

**Министерство образования и науки
Донецкой Народной Республики**
ГОУВПО «Донецкий национальный технический университет»
Факультет «Компьютерные науки и технологии»
Кафедра «Программная инженерия»

**Программная инженерия:
методы и технологии разработки информационно-
вычислительных систем
(ПИИВС-2018)**

**Сборник материалов II Международной научно-практической
конференции**

**г. Донецк
14-15 ноября 2018 года**

Донецк – 2018

УДК 004.4
ББК 32.81
П78

Программная инженерия: методы и технологии разработки информационно-вычислительных систем (ПИИВС-2018): сборник научных трудов II Международной научно-практической конференции, Том. 1. 14-18 ноября 2018 г. – Донецк, ГОУВПО «Донецкий национальный технический университет», 2018. – 308 с.

Цель научно-практической конференции: создание условий для обмена новыми идеями, технологиями и результатами между профессионалами программной инженерии, принимающими непосредственное участие в деятельности по анализу, спецификации, проектированию, разработке, сертификации, сопровождению и тестированию программного обеспечения компьютерных систем различного назначения, а также для расширения сотрудничества специалистов в области индустриального программирования с коммерческими структурами.

Основные направления работы конференции:

- современные языки и технологии программирования;
- методы и средства автоматизированного проектирования;
- информационная безопасность и защита данных;
- информационные системы и базы данных;
- компьютерная графика: синтез и обработка изображений;
- управление требованиями, тестирование и обеспечение качества ПО;
- программирование мобильных, распределённых и сетевых систем;
- интеллектуальный анализ данных, распознавание и классификация;
- прикладные интеллектуальные системы управления и принятия решений;
- компьютерное моделирование в фундаментальных и прикладных исследованиях.

Конференция организована Донецким национальным техническим университетом Министерства образования и науки ДНР. В организации конференции приняли участие: Донецкий национальный университет, Министерство связи ДНР, Ульяновский государственный технический университет (г.Ульяновск), Образовательно-научный центр «Кибернетика» ФГБОУ ВО «РЭУ им. Г.В.Плеханова» (г.Москва), Полоцкий государственный университет (Республика Беларусь, г.Полоцк).

В первом томе сборника научных трудов представлено 48 докладов сотрудников академических институтов и высших учебных заведений, а также специалистов других учреждений из Москвы, Санкт-Петербурга, Кемерово, Гродно, Нижнего Новгорода, Иркутска, Таганрога, Ижевска, Донецка, Луганска, Макеевки.

УДК 004.4
ББК 32.81

ГОУВПО «Донецкий национальный технический университет», 2018

ОГЛАВЛЕНИЕ

СЕКЦИЯ 1. «СОВРЕМЕННЫЕ ЯЗЫКИ И ТЕХНОЛОГИИ ПРОГРАММИРОВАНИЯ»⁸ Семенова Т.И., Загвоздкина А.В., Загвоздкин В.А. (Москва, Московский технический университет связи и информатики) ГРАФИЧЕСКОЕ ПРЕДСТАВЛЕНИЕ РЕЗУЛЬТАТОВ РАСЧЕТОВ В SCILAB.....	8
Винокуров С.А., Семенова Т.И. (Москва, Московский технический университет связи и информатики) ТЕСТИРУЮЩАЯ ПРОГРАММА ДЛЯ ОПРЕДЕЛЕНИЯ УРОВНЯ ЗНАНИЙ И ОБУЧЕНИЯ	15
Страхович Э.В. (Санкт-Петербург, Высшая школа менеджмента Санкт-Петербургского государственного университета) ПРИМЕНЕНИЕ МЕТОДА ДИЗАЙН-МЫШЛЕНИЯ В ОБУЧЕНИИ УПРАВЛЕНИЮ ИТ-ПРОЕКТОМ.....	20
Суворова Е.Ю., Гвоздюкова С.Н. (Луганск, Луганский национальный университет имени Тараса Шевченко) ЭКСПЕРИМЕНТАЛЬНОЕ ИССЛЕДОВАНИЕ ЗАВИСИМОСТИ ПОКАЗАНИЙ MEMS-ДАТЧИКОВ ОТ КОЛЕБАНИЙ ТЕМПЕРАТУРЫ	26
Диденко А.А., Пимонов А.Г. (Кемерово, Кузбасский государственный технический университет имени Т.Ф. Горбачева) ИСПОЛЬЗОВАНИЕ ТЕХНОЛОГИИ MAPREDUCE ДЛЯ ПАРАЛЛЕЛЬНОГО ПОИСКА В ШИРИНУ КРАТЧАЙШЕГО ПУТИ В ГРАФАХ БОЛЬШОЙ РАЗМЕРНОСТИ, АЛГОРИТМ BFS: WEIGHTED EDGES	34
Демчук В.Ю. (Донецк, Министерство связи) ОДНОСТРАНИЧНЫЕ ПРИЛОЖЕНИЯ (SPA).....	38
Рычка О.В. (Донецк, ДонНТУ) УЛУЧШЕНИЕ ПРОГНОЗНЫХ ЗНАЧЕНИЙ С ИСПОЛЬЗОВАНИЕМ МЕТОДА, ОТБРАСЫВАНИЯ ДАННЫХ	44
Шевелёв И.А., Федяев О.И., Коломойцева И.А. (Донецк, ДонНТУ) СИСТЕМА ТЕСТИРОВАНИЯ ПРИКЛАДНОГО ПРОГРАММНОГО ОБЕСПЕЧЕНИЯ ДЛЯ ОБРАЗОВАТЕЛЬНОГО ПРОЕКТА “ЭЛЕКТРОННЫЙ УЧЕБНИК LANG-BOOK”	51
Филипишин Д.А., Серёженко О.А., Коломойцева И.А., Павлюк Е.Н. (Донецк, ДонНТУ) АНАЛИЗ ИЗМЕНЕНИЙ ФУНКЦИОНАЛЬНОСТИ САЙТА ДЛЯ СОПРОВОЖДЕНИЯ КОНФЕРЕНЦИИ ПИИВС-2018	57
Медведева В.Ю., Гоманчук М.В. (Беларусь, Гродненский государственный университет имени Янки Купалы) РАЗРАБОТКА ПРИЛОЖЕНИЯ ДЛЯ РАСПОЗНАВАНИЯ ЛИЦ	65
СЕКЦИЯ 2. «ИНФОРМАЦИОННЫЕ СИСТЕМЫ, БЕЗОПАСНОСТЬ И ЗАЩИТА ДАННЫХ».....	70

Бутко А.В. (Донецк, Министерство связи) ПРИМЕНЕНИЕ ТЕХНОЛОГИИ БЛОКЧЕЙН В ЗЕМЕЛЬНОМ КАДАСТРЕ И РЕЕСТРЕ НЕДВИЖИМОСТИ	70
Гущин В.В. (Донецк, Министерство связи Донецкой Народной Республики) ГОСУДАРСТВЕННАЯ ИНФОРМАТИЗАЦИЯ С НУЛЯ.....	75
Джура Е.С. (Донецк, Министерство связи) ИНФОРМАЦИОННАЯ БЕЗОПАСНОСТЬ. ПЕРСПЕКТИВЫ И ВЫЗОВЫ	82
Пьянков К.А. (Донецк, Министерство связи) АВТОМАТИЗИРОВАННАЯ ИНФОРМАЦИОННАЯ СИСТЕМА ОБРАБОТКИ ВЫЗОВОВ СКОРОЙ МЕДИЦИНСКОЙ ПОМОЩИ	89
Vabicheva M.V. (Донецк, ДонНУ) THE AUTHENTICITY OF DIGITAL DOCUMENTS CHECKING AND PROTECTION BY OWN STEGANOGRAPHY ALGORITHM	96
Грищенко О.С., Чернышова А.В. (Донецк, ДонНТУ) АНАЛИЗ АЛГОРИТМОВ АУТЕНТИФИКАЦИИ В РАСПРЕДЕЛЕННЫХ ПРОГРАММНЫХ СИСТЕМАХ	102
Кубашевский Д.В., Чернышова А.В. (Донецк, ДонНТУ) ИССЛЕДОВАНИЕ МЕТОДОВ И СРЕДСТВ ЗАЩИТЫ АВТОРСКОГО ПРАВА В ОБЛАСТИ РАЗРАБОТКИ ПРОГРАММНОГО ОБЕСПЕЧЕНИЯ ДЛЯ РАСПРЕДЕЛЁННЫХ СИСТЕМ	109
Толстых В.К., Лихолетов А.В. (Донецк, ДонНУ) УПРАВЛЕНИЕ ТЕХНОЛОГИЧЕСКИМИ ОБЪЕКТАМИ ПОСРЕДСТВОМ МОБИЛЬНОГО УСТРОЙСТВА С ДОПОЛНЕННОЙ РЕАЛЬНОСТЬЮ	115
Чередникова О.Ю., Щедрин С.В., Ногтев Е.А., Исаков А.Ю. (Донецк, ДонНТУ) ВЫБОР СИСТЕМЫ УПРАВЛЕНИЯ БАЗАМИ ДАННЫХ ДЛЯ ПРОГРАММНОГО КОМПЛЕКСА ПРИЕМНОЙ КОМИССИИ ДОННТУ	120
СЕКЦИЯ 3. «ИНТЕЛЛЕКТУАЛЬНЫЙ АНАЛИЗ ДАННЫХ, РАСПОЗНАВАНИЕ И КЛАССИФИКАЦИЯ».....	126
Аверкин А.Н., Савинова В.М., Ярушев С.Я., Угнячев А.А. (Москва, Российский экономический университет им. Г.В. Плеханова) ПРОГНОЗИРОВАНИЕ ПОКАЗАТЕЛЕЙ СФЕРЫ ИНВЕСТИЦИЙ И ИННОВАЦИЙ С ИСПОЛЬЗОВАНИЕМ ГИБРИДНЫХ МОДЕЛЕЙ	126
Жуков А.И. (Санкт-Петербургский национальный исследовательский университет информационных технологий, механики и оптики), Федяев О.И. (Донецк, ДонНТУ) РАЗРАБОТКА АЛГОРИТМОВ И ПРОГРАММ ПАРАМЕТРИЗАЦИИ РЕЧЕВОГО СИГНАЛА НА ОСНОВЕ ВЕЙВЛЕТ-АНАЛИЗА	134
Тырса С.В. (Санкт-Петербургский национальный исследовательский университет информационных технологий, механики и оптики), Федяев О.И. (Донецк, ДонНТУ)	

РАЗРАБОТКА ПРОГРАММНОЙ СИСТЕМЫ РАСПОЗНАВАНИЯ ЛИЦ НА ОСНОВЕ НЕЙРОСЕТИ.....	140
Бакаленко В.С., Федяев О.И. (Донецк, ДонНТУ) ПРОЕКТИРОВАНИЕ ИНТЕЛЛЕКТУАЛЬНОГО РЕЧЕВОГО ИНТЕРФЕЙСА ДЛЯ МЕДИЦИНСКОЙ СПРАВОЧНОЙ СИСТЕМЫ.....	146
Пикалёв Я.С., Веретельник Д.М., Добренко Д.С. (Донецк, ДонНТУ) О МЕТОДАХ ДЕТЕКЦИИ ОБЪЕКТОВ ИЗ ИЗОБРАЖЕНИЙ В РЕАЛЬНОМ ВРЕМЕНИ, ОСНОВАННЫХ НА ГЛУБОКОМ ОБУЧЕНИИ	152
Сторожук Н.О., Коломойцева И.А. (Донецк, ДонНТУ) РЕАЛИЗАЦИЯ ПРОТОТИПА АНАЛИЗАТОРА ЖАНРОВОЙ ПРИНАДЛЕЖНОСТИ ПРОИЗВЕДЕНИЙ НА ОСНОВАНИИ ВЕКТОРНО – ПРОСТРАНСТВЕННОЙ МОДЕЛИ ПРЕДСТАВЛЕНИЯ ДОКУМЕНТОВ	161
Титаренко М.Г., Коломойцева И.А. (Донецк, ДонНТУ) ОБЗОР И АНАЛИЗ АЛГОРИТМОВ ДЛЯ ОСУЩЕСТВЛЕНИЯ БИНАРНОЙ КЛАССИФИКАЦИИ ИНФОРМАЦИИ О ВНЕШНЕТОРГОВОЙ ДЕЯТЕЛЬНОСТИ ГОСУДАРСТВ	167
Федяев О.И., Медгаус С.В. (Донецк, ДонНТУ) МУЛЬТИАГЕНТНОЕ ИМИТАЦИОННОЕ МОДЕЛИРОВАНИЕ СИСТЕМЫ ТРУДОУСТРОЙСТВА СТУДЕНТОВ	172
Чернышов Б.С., Фоминых И.Д., Рудак Л.В., Федяев О.И. (Донецк, ДонНТУ) СОВРЕМЕННОЕ РЕШЕНИЕ ПРОБЛЕМЫ РАСПОЗНАВАНИЯ ЛИЦ НА ОСНОВЕ НЕЙРОННЫХ СЕТЕЙ	180
Шевцов Д.В. (Донецк, ДонНУ) СИСТЕМЫ РАСПОЗНАВАНИЯ ИЗОБРАЖЕНИЙ КАК СРЕДСТВО АВТОМАТИЗАЦИИ ПРОЦЕССОВ ДОКУМЕНТООБОРОТА.....	185
СЕКЦИЯ 4. «ПРИКЛАДНЫЕ ИНТЕЛЛЕКТУАЛЬНЫЕ СИСТЕМЫ УПРАВЛЕНИЯ, ПРОЕКТИРОВАНИЯ И ПРИНЯТИЯ РЕШЕНИЙ»	194
Савкова Д.Г. (Москва, Компания «Интеллектуальные системы автоматизации и технологии» ИнСАТ) ОБЗОР ПРОБЛЕМ РЕАЛИЗАЦИИ ЧЕЛОВЕКО-МАШИННОГО ИНТЕРФЕЙСА В СИСТЕМАХ SCADA.....	194
Гвоздюкова С.Н., Суворова Е.Ю. (Луганск, Луганский национальный университет имени Тараса Шевченко) ПРИМЕНЕНИЕ ERP-ТЕХНОЛОГИЙ В СТРАТЕГИЧЕСКОМ УПРАВЛЕНИИ ПРЕДПРИЯТИЕМ	199
Евгеньев Г.Б. (Москва МГТУ им. Н.Э. Баумана) МНОГОАГЕНТНЫЕ СИСТЕМЫ ПОЛУАВТОМАТИЧЕСКОГО 3D ПРОЕКТИРОВАНИЯ ИЗДЕЛИЙ МАШИНОСТРОЕНИЯ.....	206

Жевнерчук Д.В. (Нижний Новгород, Нижегородский государственный технический университет им. Р.Е. Алексеева) КОНЦЕПТУАЛЬНЫЕ КАРКАСЫ ОНТОЛОГИЙ САМОСТРУКТУРИРУЕМЫХ ОТКРЫТЫХ ИНФОРМАЦИОННЫХ СИСТЕМ	212
Захаров А.С., Жевнерчук Д.В. (Нижний Новгород, Нижегородский государственный технический университет им. Р.Е. Алексеева) АНАЛИЗ СУЩЕСТВУЮЩИХ ПОДХОДОВ И ПЕРСПЕКТИВ РАЗВИТИЯ ИНТЕЛЛЕКТУАЛЬНЫХ ГЕНЕРАТОРОВ ПРОГРАММНОГО КОДА	221
Ройзензон Г.В. (Москва, Институт системного анализа ФИЦ ИУ РАН) СТАНДАРТЫ ЭТИКИ В ИСКУССТВЕННОМ ИНТЕЛЛЕКТЕ.....	227
Скрыпникова М.К. (Москва, Московский технический университет связи и информатики (МТУСИ)) ЭКСПЕРТНО-АНАЛИТИЧЕСКИЙ МЕТОД ОЦЕНКИ КАЧЕСТВА КОНКУРИРУЮЩЕЙ ПРОДУКЦИИ.....	237
Юрин А.Ю., Берман А.Ф., Николайчук О.А., Дородных Н.О. (Иркутск, Институт динамики и теории управления им. В.М. Матросова СО РАН (ИДСТУ СО РАН), Иркутский национальный исследовательский технический университет (ИрННТУ)) СОЗДАНИЕ БАЗ ЗНАНИЙ ДЛЯ ПОДДЕРЖКИ ПРИНЯТИЯ РЕШЕНИЯ В ЗАДАЧАХ ЭКСПЕРТИЗЫ ПРОМЫШЛЕННОЙ БЕЗОПАСНОСТИ.....	242
Юсков И.О. (Москва, Московский технический университет связи и информатики) НЕЙРОСЕТЕВОЙ ПОДХОД К МОДЕЛИРОВАНИЮ КОРПОРАТИВНЫХ СЕТЕЙ.....	250
Джура С.Г., Чурсинов В.И., Якимишина В.В. (Донецк, ДонНТУ) ВНЕДРЕНИЕ ПРОГРАММНЫХ ПРОДУКТОВ ГРУППЫ КОМПАНИЙ CSOFT В УЧЕБНЫЙ ПРОЦЕСС ДОННТУ И ПРОМЫШЛЕННОСТЬ ДНР	256
Олейник А.Н. (Донецк, Министерство образования и науки Донецкой Народной Республики), Горун В.В., Гуридов А.Г. (Республиканская служба по контролю и надзору в сфере образования и науки) ПРОБЛЕМАТИКА ОРГАНИЗАЦИИ ПРОЕКТИРОВАНИЯ СИСТЕМ ПРИ РЕАЛИЗАЦИИ МЕТОДОВ ПРОГРАММНО-ИНФОРМАЦИОННОГО ОБЕСПЕЧЕНИЯ УПРАВЛЕНИЯ ПРОЕКТАМИ, ОЦЕНКА РИСКОВ И СПОСОБЫ ИХ УСТРАНЕНИЯ.....	263
Григорьев А.В. (Донецк, ДонНТУ) МОДУЛЬНЫЙ ПОДХОД К ПРЕДСТАВЛЕНИЮ МОДЕЛЕЙ МЕТОДИК ПРОЕКТИРОВАНИЯ В ИНТЕЛЛЕКТУАЛЬНЫХ САПР	269
Морозова О.В. (Донецк, ДонНТУ) МОДЕЛЬ ПРЕДСТАВЛЕНИЯ ЗНАНИЙ В ИНТЕЛЛЕКТУАЛЬНОЙ НАДСТРОЙКЕ НАД ПРОБЛЕМНО-ОРИЕНТИРОВАННЫМИ САПР	274
СЕКЦИЯ 5. «КОМПЬЮТЕРНОЕ МОДЕЛИРОВАНИЕ, ГРАФИКА И ОБРАБОТКА ИЗОБРАЖЕНИЙ».....	280
Тайлакова А.А. (Кемерово, Кузбасский государственный технический университет имени Т.Ф. Горбачева)	

РАЗРАБОТКА ПРОГРАММНОГО КОМПЛЕКСА ДЛЯ АНАЛИЗА ЭФФЕКТИВНОСТИ АЛГОРИТМОВ ОПТИМИЗАЦИИ ПРИ РЕШЕНИИ ЗАДАЧИ ПОДБОРА КОНСТРУКЦИИ НЕЖЕСТКИХ ДОРОЖНЫХ ОДЕЖД.....	280
Капустин Д. А., Сентяй Р.Н., Швыров В.В. (Луганск, Луганский национальный университет имени Тараса Шевченко) ИССЛЕДОВАНИЕ ВОЗМОЖНОСТИ ОЦЕНКИ ЭФФЕКТИВНОСТИ РАБОТЫ ПРОМЫШЛЕННОЙ ГИДРОТРАНСПОРТНОЙ СИСТЕМЫ (ПГТС) СИСТЕМАМИ ИСКУССТВЕННОГО ИНТЕЛЛЕКТА	286
Пирская Л.В. (Таганрог, Южный федеральный университет) ИССЛЕДОВАНИЕ ФУНКЦИОНИРОВАНИЯ СПЕЦИАЛИЗИРОВАННОГО ВЫЧИСЛИТЕЛЬНОГО УСТРОЙСТВА НА ОСНОВЕ АЛГОРИТМОВ ДЕЛЬТА- ПРЕОБРАЗОВАНИЙ ПЕРВОГО ПОРЯДКА С ПЕРЕМЕННЫМ КВАНТОМ	292
Зори С.А., Аноприенко А.Я. (Донецк, ДонНТУ) ОСОБЕННОСТИ ОРГАНИЗАЦИИ ОБЪЕМНОЙ ВИЗУАЛИЗАЦИИ В СИСТЕМАХ ТРЕХМЕРНОЙ КОМПЬЮТЕРНОЙ ГРАФИКИ.....	302
Конопацкий Е.В. (Макеевка, Донбасская национальная академия строительства и архитектуры) ПРИНЦИПЫ ПОСТРОЕНИЯ КОМПЬЮТЕРНЫХ МОДЕЛЕЙ МНОГОФАКТОРНЫХ ПРОЦЕССОВ И ЯВЛЕНИЙ МЕТОДОМ МНОГОМЕРНОЙ ИНТЕРПОЛЯЦИИ.....	309
Лазебная Л.А., Едемская Е.Н., Бельков Д.В. (Донецк, ДонНТУ) ДИНАМИЧЕСКОЕ РАСПРЕДЕЛЕНИЕ ФАЙЛОВ ПО УЗЛАМ КОМПЬЮТЕРНОЙ СЕТИ	319

СЕКЦИЯ 1. «СОВРЕМЕННЫЕ ЯЗЫКИ И ТЕХНОЛОГИИ ПРОГРАММИРОВАНИЯ»

УДК 004.94

Графическое представление результатов расчетов в Scilab

Семенова Т.И., Загвоздкина А.В., Загвоздкин В.А.
Московский технический университет связи и информатики
sematata@yandex.ru, ann-zag@mail.ru, acc-zag@yandex.ru

Семенова Т.И., Загвоздкина А.В., Загвоздкин В.А. Графическое представление результатов расчетов в Scilab. Свободно распространяемый пакет Scilab включает в себя современную систему программирования и является мощным средством проведения математических расчетов и представления полученных результатов как в традиционном числовом, так и в графическом виде. В статье представлены возможности программных и графических средств пакета Scilab, а также примеры их использования для визуализации расчетов в различных дисциплинах, изучаемых студентами.

***Ключевые слова:** пакет SCILAB, современные системы программирования, библиотечные функции и программные средства, графическое представление расчетов.*

Введение

В настоящее время трудно представить проведение сложных математических расчётов без использования возможностей современных математических пакетов. Они содержат достаточный набор математических и программных средств для решения вычислительных задач практически любой степени сложности, а также, что немаловажно, средства визуализации полученных результатов [1]. Однако большинство математических пакетов являются коммерческими, поэтому появление в настоящее время такого пакета, как Scilab, вызвало у пользователей несомненный интерес. Этот пакет является не только свободно распространяемым, но и универсальным, поскольку содержит объемную библиотеку встроенных функций, мощную систему программирования, а также графический редактор, что делает его необходимым при подготовке современного специалиста программной инженерии [2] и других специалистов [3].

Назначение пакета Scilab

Пакет Scilab построен как программная система, ориентированная на работу с матрицами, поэтому содержит большое количество матричных функций. Возможности пакета Scilab 6.01 вполне достаточны для использования его в учебном процессе при изучении методов вычислительной математики [4], а также других дисциплин. Так, например, в системе Scilab реализованы следующие методы решения вычислительных задач [5]: задачи линейной алгебры; методы решения нелинейных уравнений и систем линейных и нелинейных уравнений; обработка экспериментальных данных; задачи оптимизации; интегрирование и дифференцирование; решение обыкновенные дифференциальных уравнений и их систем [6]. Кроме того, Scilab позволяет работать с большим числом специальных функций (Бесселя, Неймана и т.д.), имеет средства для построения, настройки и работы с графиками в графических окнах, а для выполнения численных расчётов могут использоваться библиотеки Lapack, LINPACK, Atlas и другие. В Scilab могут конвертироваться программы, созданные в других языках программирования, в том числе и в пакете Matlab [7].

Для решения нестандартных задач в Scilab имеется встроенный объектно-ориентированный язык программирования, sci-язык, с помощью которого пользователь может создавать свои визуальные приложения в виде отдельной программы.

Графическое представление результатов

В графическом редакторе имеется ряд инструментов, позволяющих представить зависимости, полученные в ходе проведения расчетов, в отдельных масштабируемых и перемещаемых графических окнах в виде 2D и 3D графиков. С помощью средств панели графического окна, на которую традиционно выносятся самые используемые команды, можно легко управлять параметрами графиков и наносить на них текстовые комментарии.

В простейшем случае для построения графика с использованием средств пакета Scilab, функцию необходимо предварительно описать внутри командного окна. В случае возможного многократного использования функции, ее удобно описать в текстовом редакторе SciNotes, и сохранить ее в файле. Все способы описания функции представлены в [8]. Функция, сохраненная на внешнем носителе, перед ее использованием должна быть подключена командой `exes`. Если файл находится в текущем каталоге, то достаточно указать имя файла, в противном случае указывается полный путь к файлу.

Все графики в Scilab выводятся в отдельных окнах, называемых графическими. В одном графическом окне допускается построение нескольких графиков, как разных функций, так и одной и той же функции, полученной при различных значениях какого-либо ее параметра. Кроме того, графическое окно может быть разбито на ряд областей, в каждой из которых может быть проведено построение своего графика.

В этой статье представлены графические средства пакета Scilab и примеры их использования для визуализации расчетов нескольких технических дисциплин.

Самый простой способ для построения графика функции одной переменной $y=f(x)$ можно произвести с использованием команд `plot`. Для этого предварительно необходимо сформировать два вектора одинаковой длины: вектор значений аргументов (x) и вектор соответствующих им значений функции (y). Выполнение команды `plot(x,y)` открывает графическое окно и отображает в нем график функции $y(x)$. Создать новое графическое окно можно и путем выполнения команды `scf(n)`, где n – номер графического окна.

Рассмотрим пример построения графика функции $y = e^{-x^2}$ на отрезке $[-1.5; 1.5]$ (рис. 1).

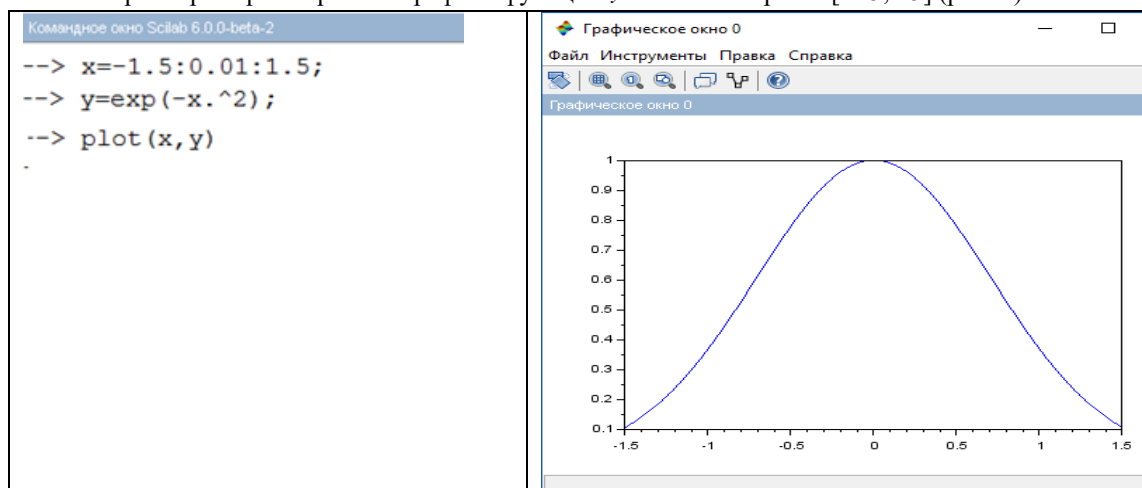


Рисунок 1 – График, полученный с использованием команды `plot(x,y)`

На рисунке 2 приведены команды, необходимые для построения в одном окне трех графиков: $y1(x) = \sin(x)$, $y2(x) = \sin(2x)$, $y3(x) = \sin(4x)$. Сначала показано, как можно вывести несколько графиков в одном графическом окне одной командой `plot`. В этом случае каждый график автоматически различается своим цветом. Однако в черно-белом варианте такой способ неприменим и ниже приведен другой способ. Как правило, для вывода нескольких графиков используется несколько команд `plot`, тогда после построения первого графика необходимо выполнить команду `mtlb_hold('on')`. Эта команда позволяет в данном сеансе работы расположить все выводимые в дальнейшем графики в одном окне [8]. Графики функций выведены в графическое окно, прокомментированы заголовком, названиями осей и легендой, показывающей цвет, тип точек и линий, с помощью которых рисуется каждый график.

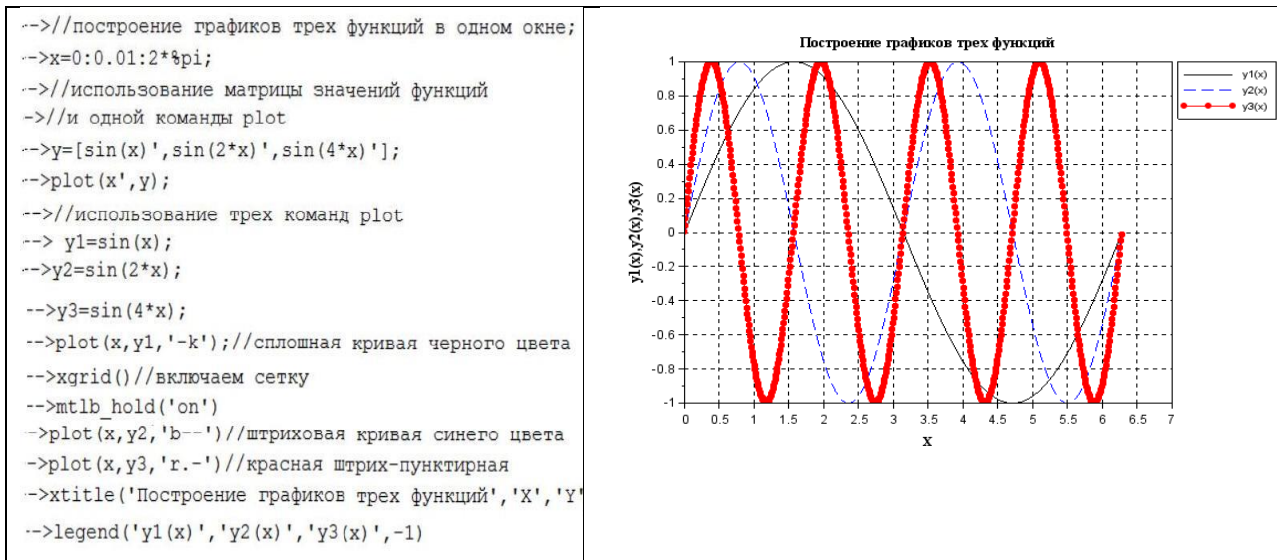


Рисунок 2 – Графики трех функций в одном графическом окне, полученные с использованием команды plot()

Для построения трехмерных графиков используются несколько команд: mesh() - сетчатый график; contour() – график контурных линий; surf() – график сплошной поверхности; plot3d() – точки графика соединены отрезками прямых и другие. Перед созданием любого трехмерного графика необходимо сформировать матрицу значений функции от двух аргументов командой meshgrid(n1:h1:m1,n2:h2:m2), где n1, h1, m1 – начальное значение, шаг изменения и конечное значение первого аргумента, а n2, h2, m2 – соответственно, второго аргумента. Для компактного расположения нескольких графиков в одном следует указать область построения текущего графика с помощью команды subplot(n,m,k). На рисунке 3 графическое окно разбито на четыре части (строки n=2 и столбцы m=2), а последний параметр (k) указывает номер области построения текущего графика.

Следует отметить, что при описании математического выражения функции, для которой происходит построение графика, используются математические операции с точкой (*, ./, .^), поскольку здесь необходимо проводить вычисления с использованием поэлементных операций.

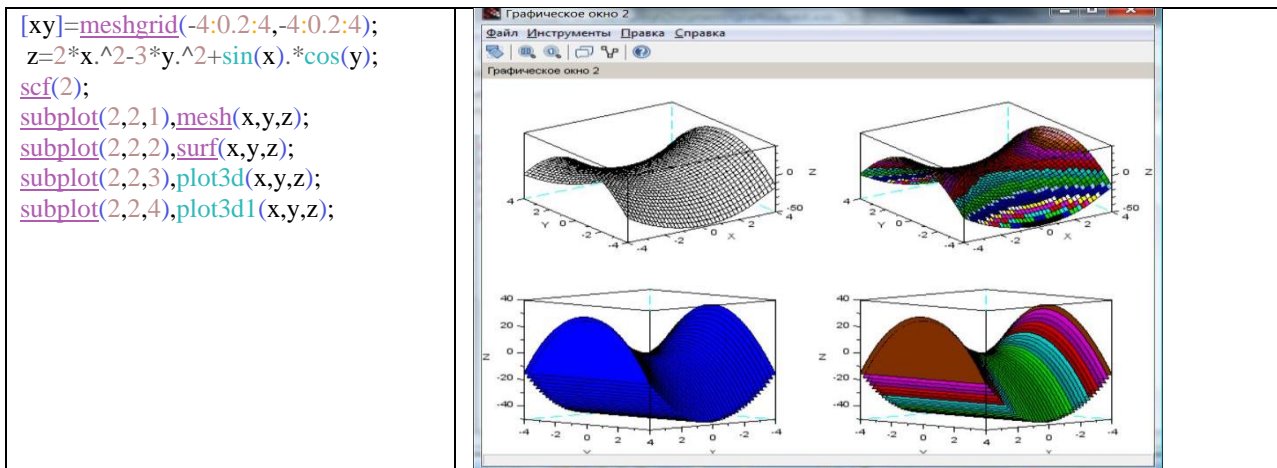


Рисунок 3 – Трехмерные графики функции, полученные с использованием различных команд в одном графическом окне

Для построения графика в полярных координатах применяется команда polarplot с двумя обязательными параметрами; первый параметр определяет диапазон значений угла, второй – саму функцию. На рисунке 4 приведен пример построения в полярных координатах графика функции $\rho = 4\cos(3\phi)$, где ϕ (полярный угол) изменяется на интервале $[0; 2\pi]$ с шагом $0,01$.

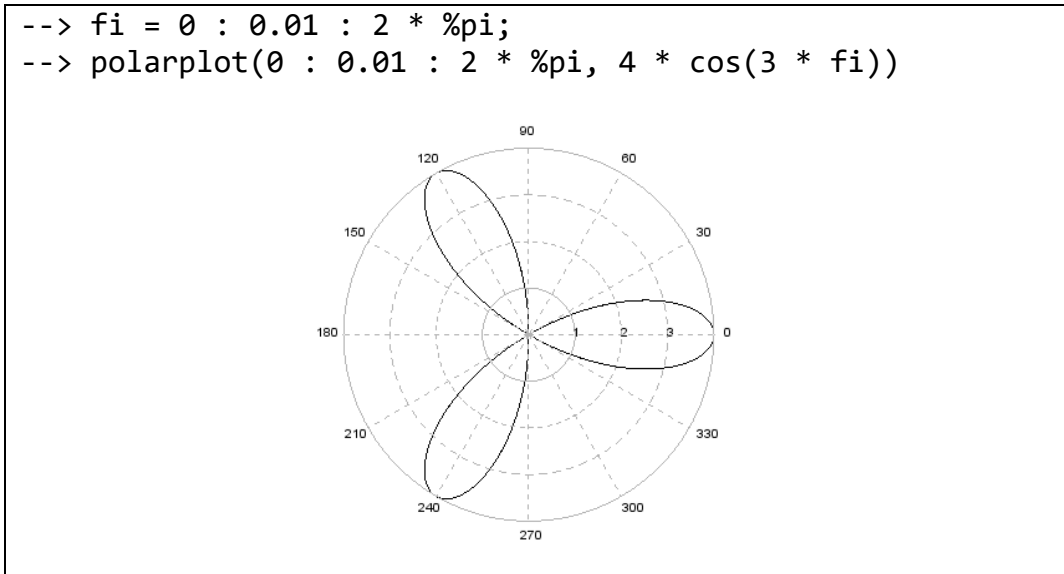


Рисунок 4 – Построение графика в полярной системе координат

Далее приведены несколько примеров графических окон, содержащих графики функций, отражающих специфику дисциплин: высшая математика (рис. 5), физика (рис. 6), методы оптимизации (рис. 7), теория электрических цепей (рис. 8) и общая теория связи (рис. 9). Каждый пример пояснен комментариями ко всем проводимым вычислениям и средствам визуализации полученных результатов.

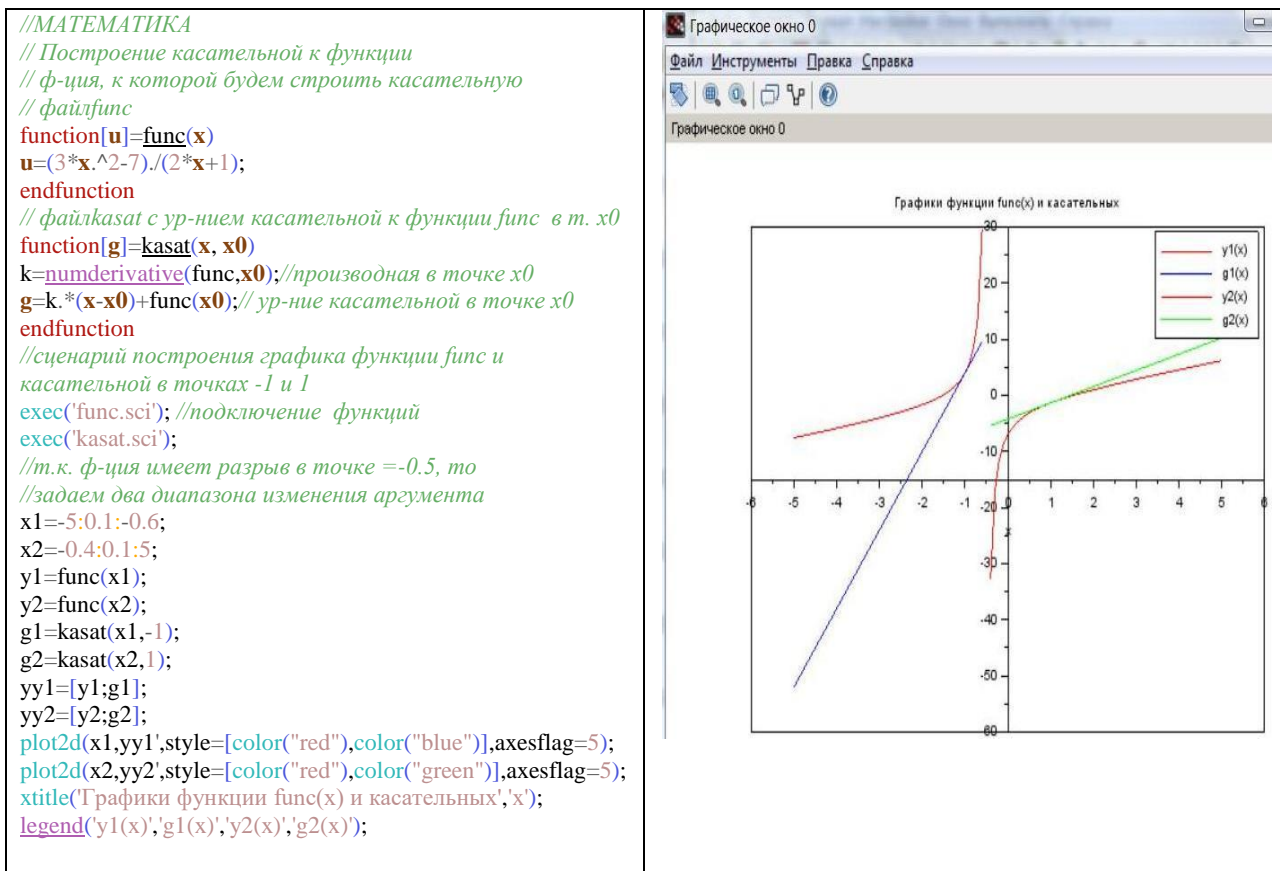


Рисунок 5 – Построение графика касательной к функции в заданной точке

```

// ФИЗИКА
// Амплитуда вынужденных колебаний
// w-частота внешней силы
// g-декремент затухания (1/c)
// m-масса маятника (кг)
// F- амплитуда внешней силы (Н)
// w0-собственная частота (1/c)
function[A]=Favk(F, m, w0, w, g)
A=(F./m)./sqrt((w0.^2-w.^2).^2+4*g.^2*w.^2);
Endfunction

//сценарий построения графиков
exec('Favk.sci'); //подключение функции
m=1; //масса маятника (кг)
F=1; //амплитуда внешней силы (Н)
w0=10; //собственная частота (1/c)
w=0:0.01:20;
g1=0.1; g2=0.5; g3=1;
A1=Favk(F,m,w0,w,g1);
A2=Favk(F,m,w0,w,g2);
A3=Favk(F,m,w0,w,g3);
plot(w,A1,'-',w,A2,'--',w,A3,'-');
xlabel('Амплитуды вынужденных колебаний','w','A1(w), A2(w), A3(w)');
legend('A1(w)', 'A2(w)', 'A3(w)');

```

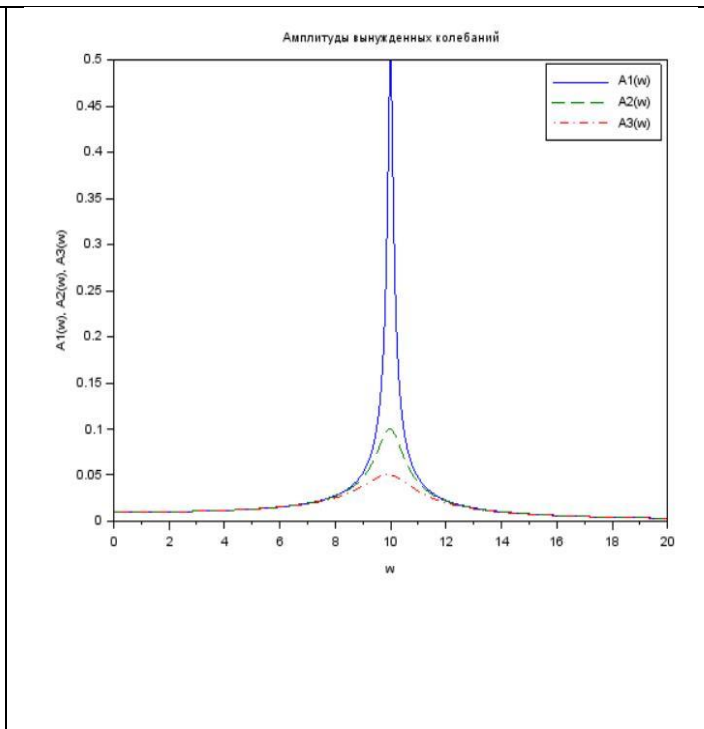


Рисунок 6 – Построение графиков амплитуды вынужденных колебаний

```

// МЕТОДЫ ОПТИМИЗАЦИИ

[x y]=meshgrid(-10:10,-10:10);
deff('k=f(x,y)', 'k=2*x.^2+y.^2');
z=f(x,y);
scf(2);
xlabel('Графики функций от 2-х переменных и контурных линий');
x1=-10:10; y1=-10:10;
subplot(2,1,1), surf(x,y,z);
subplot(2,1,2); contour(x1,y1,z,5);

```

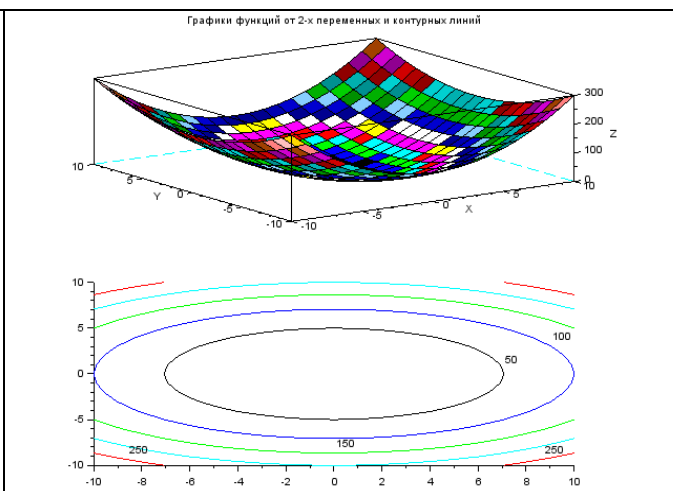
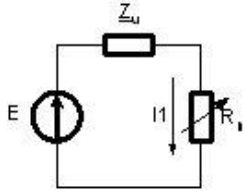


Рисунок 7 – Построение в одном графическом окне графика функции от 2-х переменных и графика контурных линий

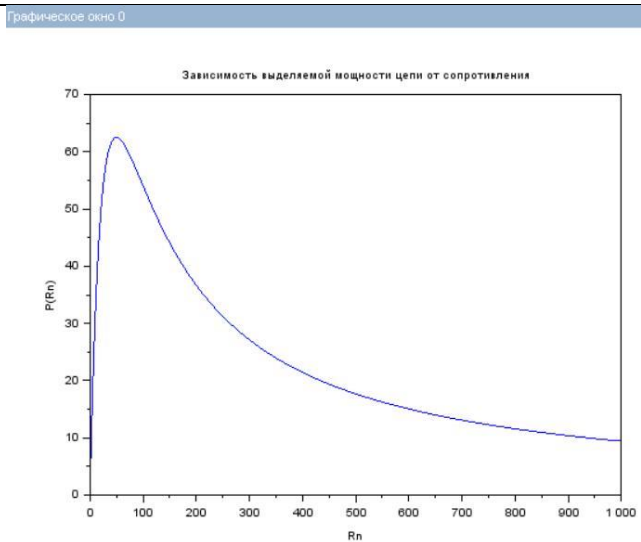

```

//ТЭЦ
// определение величины резисторного
//сопротивления нагрузки цепи, при котором
//выделяется максимальная мощность



// Вычисление мощности цепи для заданных
// значений E и Z_u
function[P]=FP(Rn)
E=100;//(В)
Zu=30-%i*40;//(Ом)
I1=E./(Zu+Rn);
P=(abs(I1).^2).*Rn;
endfunction
//Производная от мощности
function[dP]=FPdif(Rn)
dP=10000.*(Rn.^2-2500)./(2500+60.*Rn+Rn.^2).^2;
endfunction
//сценарий
exec('C:\Users\anna\Documents\FP.sci');
exec('C:\Users\anna\Documents\FPdif.sci');
//построение графика выделяемой мощности цепи
//от сопротивления
Rn=0:0.1:1000;
P=FP(Rn);
plot(Rn,P)
xtitle('Зависимость выделяемой мощности цепи от
сопротивления','Rn','P(Rn)')
//максимум мощности найдем как корень
//производной от мощности
R0=30;//начальное значение для поиска корня
Rnmax=fsolve(R0,FPdif)
Pmax=FP(Rnmax)

```



```

Командное окно Scilab 5.5.2
Файл Правка Управление Инструменты Справка
Командное окно Scilab 5.5.2

Rnmax =
    50.
Pmax =
    62.5

```

Рисунок 8 – Построение графика зависимости выделяемой мощности цепи от сопротивления и определение величины резисторного сопротивления нагрузки цепи, при котором выделяется максимальная мощность

```

//ОТС
//Функция плотности вероятности
//гауссовского случайного процесса,
//нормальное распределение
//m- среднее значение, D- дисперсия
function[w]=FPV(m, D, x)
w=(1/sqrt(2*%pi.*D)).*exp(-(x-m).^2./(2*D));
endfunction
//сценарий построения графиков ф-ции
плотности //вероятности
exec('FPV.sci'); //подключение функции
x=-4:0.01:4;
m1=0;m2=2;m3=2;
D1=1;D2=0.5;D3=3;
w1=FPV(m1,D1,x);
w2=FPV(m2,D2,x);
w3=FPV(m3,D3,x);
plot(x,w1,'-',x,w2,'--',x,w3,'-.-')
xgrid()
xtitle('Функция плотности вероятности
гауссовского случайного процесса','x','w1(x),
w2(x), w3(x)')
legend('w1(x)','w2(x)','w3(x)');

```

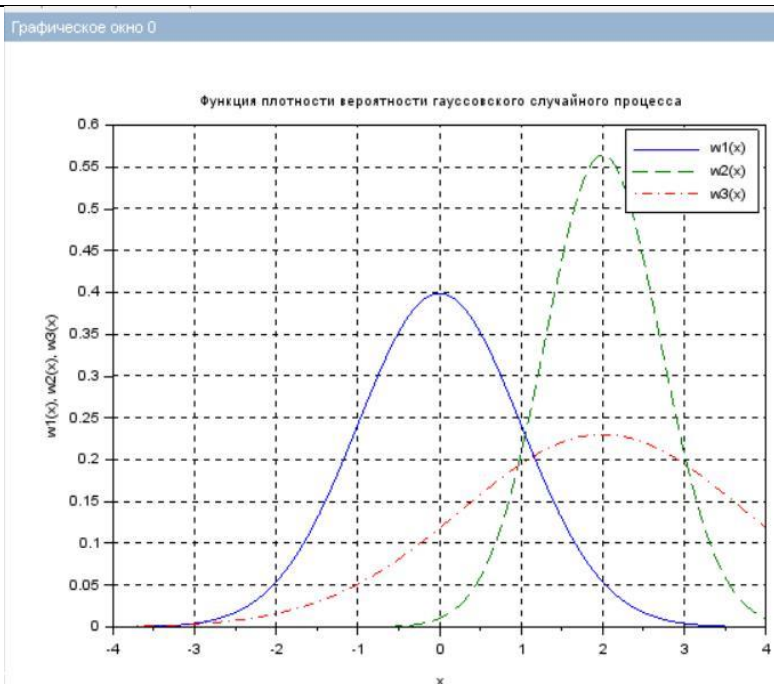


Рисунок 9 – Построение графиков функции плотности вероятности гауссовского случайного процесса

Данные для построения графиков могут быть получены самым разными способами: путем ввода исходных данных, в результате проведения расчетов и, наконец, путем передачи данных из других пакетов и программ. Scilab допускает использование всех перечисленных способов, обладая возможностью конвертирования данных, созданных в других системах программирования.

Выводы

В заключение хочется сказать, что свободно распространяемый пакет Scilab, являясь по сути дела незначительно упрощенным аналогом системы Matlab, обладает достаточными возможностями для решения многих инженерных и научных задач. Пакет Scilab постоянно совершенствуется, каждая его версия пополняется новыми функциями и возможностями, что безусловно, способствует росту его популярности в России.

Литература

1. Шакин, В.Н. Основы работы с математическим пакетом Matlab: учебное пособие/ В.Н. Шакин, Т.И. Семенова. – Москва: МТУСИ, 2016. – 133 с.
2. Семенова, Т.И. Роль математического образования при подготовке специалиста в области информационных технологий. / Т.И. Семенова, А.В. Загвоздкина // Сборник тезисов конференции «Мобильный бизнес: перспективы развития и реализации систем радиосвязи в России и за рубежом Российской академии естественных наук «Экономика и качество систем связи» // ЗАО «Национальный институт радио и инфокоммуникационных технологий», 2018. – С.136-138.
3. Клесарева, Е. Ю. Применение интерактивных технологий при реализации компетентностного подхода к подготовке бакалавров рекламы и связи с общественностью в техническом ВУЗе / Е. Ю. Клесарева, А.Р. Каберова // Экономика и качество систем связи, 2017 - №4 (6), С. 56-60.
4. Семенова, Т.И. Использование пакета Scilab при изучении методов вычислительной математики. / Т.И. Семенова, А.В. Загвоздкина, В.А. Загвоздкин // Международный сборник научных трудов «Новые технологии в науке, образовании и производстве» по материалам международной научной производственной конференции 10-13 ноября 2017, РИБиУ (Региональный институт бизнеса и управления) – г. Рязань, 2017. – С. 471-482.
5. Семенова, Т.И. Вычислительные модели и алгоритмы решения задач численными методами: учебное пособие/ Т.И. Семенова, О.М. Кравченко, В.Н. Шакин. – Москва: МТУСИ, 2017. – 84 с.
6. Семенова, Т.И. Изучение численных методов с использованием средств пакета Scilab / Т.И. Семенова, А.В. Загвоздкина, В.А. Загвоздкин // Экономика и качество систем связи, 2017 - №4 (6), С. 60-69.
7. Ерин, С.В. Автоматизация инженерных расчетов с использованием пакета Scilab: практическое пособие / С.В. Ерин, Ю.Л. Николаев. – Москва: «Русайнс», 2017. – 280 с.
8. Семенова Т.И., Шакин В.Н. Основы работы с математическим пакетом Scilab. Учебное пособие и практикум: – Москва : МТУСИ, 2017. - 129 с.[Электронный ресурс] – Режим доступа: http://www.mtuci.ru/structure/library/catalogue/download.php?book_id=1834

Семенова Т.И., Загвоздкина А.В., Загвоздкин В.А. Графическое представление результатов расчетов в Scilab. Свободно распространяемый пакет Scilab включает в себя современную систему программирования и является мощным средством проведения математических расчетов и представления полученных результатов как в традиционном числовом, так и в графическом виде. В статье представлены возможности программных и графических средств пакета Scilab, а также примеры их использования для визуализации расчетов в различных дисциплинах, изучаемых студентами.

Ключевые слова: пакет SCILAB, современные системы программирования, библиотечные функции программные средства, графическое представление расчетов.

Semenova Tatiana, Zagvozdki Anna, Zagvozdkin Vladimir. Graphical presentation of calculations' results in Scilab. Free provided package Scilab is a modern programming system with powerful possibilities for making calculations and representing results in numerical and graphical forms. The article presents programmatic and graphical possibilities of Scilab package and showcases some examples of using them in different disciplines being studied by students.

Key words: package Scilab, modern programming systems, library functions and programmatic instruments, graphical presentation of calculations' results.

Тестирующая программа для определения уровня знаний и обучения

Винокуров С.А., Семенова Т.И.

Ордена Трудового Красного Знамени федерального государственного бюджетного образовательного учреждения высшего образования «Московский технический университет связи и информатики»
sergey.vinokurow@yandex.ru

Винокуров С.А., Семенова Т.И. Тестирующая программа для определения уровня знаний и обучения. Программа тестирования проста в применении и удобна для любого пользователя. Ее редактор позволяет производить выбор тем, количество вопросов или задач, проводить настройку порогового уровня и критерия оценки. Можно создать текстовый файл, в который будут автоматически занесены данные и результаты тестируемых пользователей. Произведя замену базы данных тестирующую программу можно использовать для оценки знаний практически по любой дисциплине.

Ключевые слова: программа тестирования, настройка режимов тестирования, критерий оценки уровня знаний, универсальность использования программы в учебных процессах, помощь пользователю.

В данной статье представлена программа тестирования, предназначенная для решения актуальных проблем, связанных, как с сокращением времени, которое ежедневно тратится на проверку степени усваивания изучаемого материала, так и на упрощение этого процесса, за счет использования разработанной программы тестирования. Использование в учебном процессе тестирующих программ способствует повышению личной компетентности пользователя, а также развитию навыков самоконтроля.

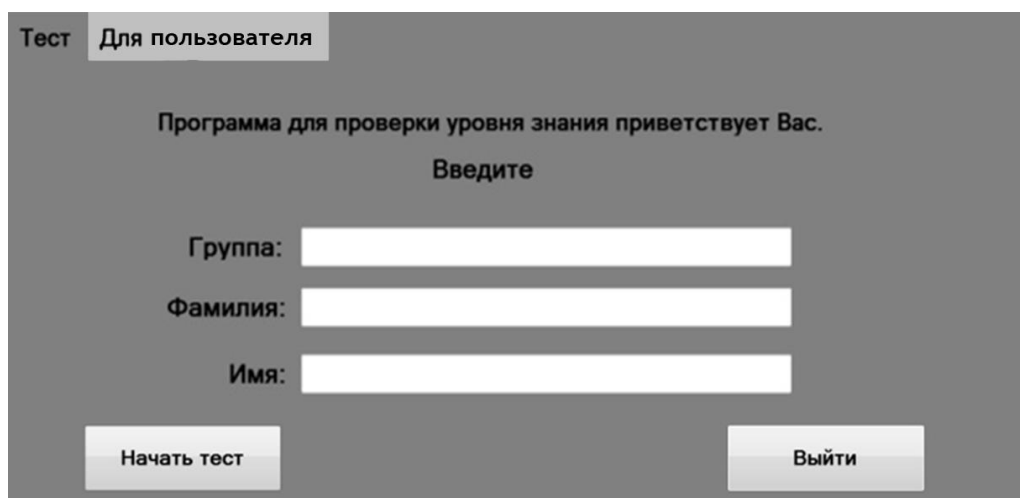
Программа была разработана на основе анализа последних публикаций и современных методов электронного обучения [1].

Эта программа позволяет практически реализовать многие возможности, которые раньше были не учтены в подобных программах [2].

Целью данной статьи является ознакомление с данной программой и демонстрации ее возможностей и достоинств: простота, универсальность, и возможность дистанционного использования [3].

Задачей является внедрение программы в учебный процесс для увеличения качества самообразования, самоконтроля и уменьшение времени на проверку уровня знаний [4].

При запуске программы открывается стартовое окно (см.рис.1). Цветовая гамма выбрана в спокойных тонах. В центре - строки ввода. Сверху расположены вкладки. Внизу - кнопки для запуска теста и для выхода из программы.



Тест Для пользователя

Программа для проверки уровня знания приветствует Вас.

Введите

Группа:

Фамилия:

Имя:

Начать тест Выйти

Рисунок 1 - Стартовое окно

Перед началом тестирования необходимо произвести ее настройку: выбрать тему (одну или несколько), а также количество вопросов или задач по данным темам, время на прохождения теста и пороговый процент правильных ответов, или выбрать готовый тест.

Программа имеет дополнительные функции, которые можно подключить, когда они необходимы пользователю:

- автоматическое оценивание теста (по параметрам установленными по умолчанию), или ручное (параметры оценки задает пользователь);
- запись результатов пройденного теста в файл.

Настройку можно сохранить и использовать в дальнейшем. Для настройки теста нужно войти в меню настроек, активизацией вкладки «для пользователя». Для защиты от несанкционированных входов в меню настроек используется пароль, который при желании пользователя можно изменить или отключить.(см.рис.2).

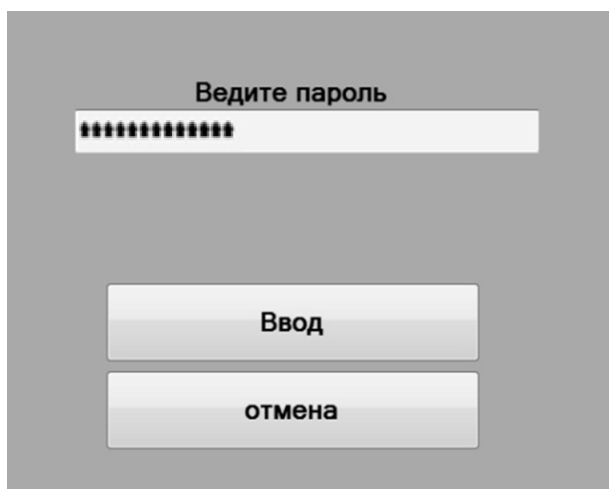


Рисунок 2 - Защитное окно

Если пароль введён правильно, то защитное окно исчезнет, и Вы перейдете в меню настроек (см.рис.3).



Рисунок 3 - Меню настроек

В верхнем правом углу окна расположена настройка времени и порогового процента правильных ответов. Ниже - дополнительная функция: оценивание пройденного теста. Еще ниже располагаются строки, где задается необходимое количество вопросов или задач. Ниже располагаются названия тем, настройка включения вопросов или задач для каждой темы. Для примера была выбрана дисциплина «Вычислительные модели»[5].

Внизу окна может быть подключена еще одна дополнительная функция: запись результатов тестирования в файл. В левой части окна располагается значок глобальных настроек. При нажатии на него появится окно настройки самой программы. Можно изменить количество тем, названия тем, количество вопросов и задач для каждой темы (см.рис.4).

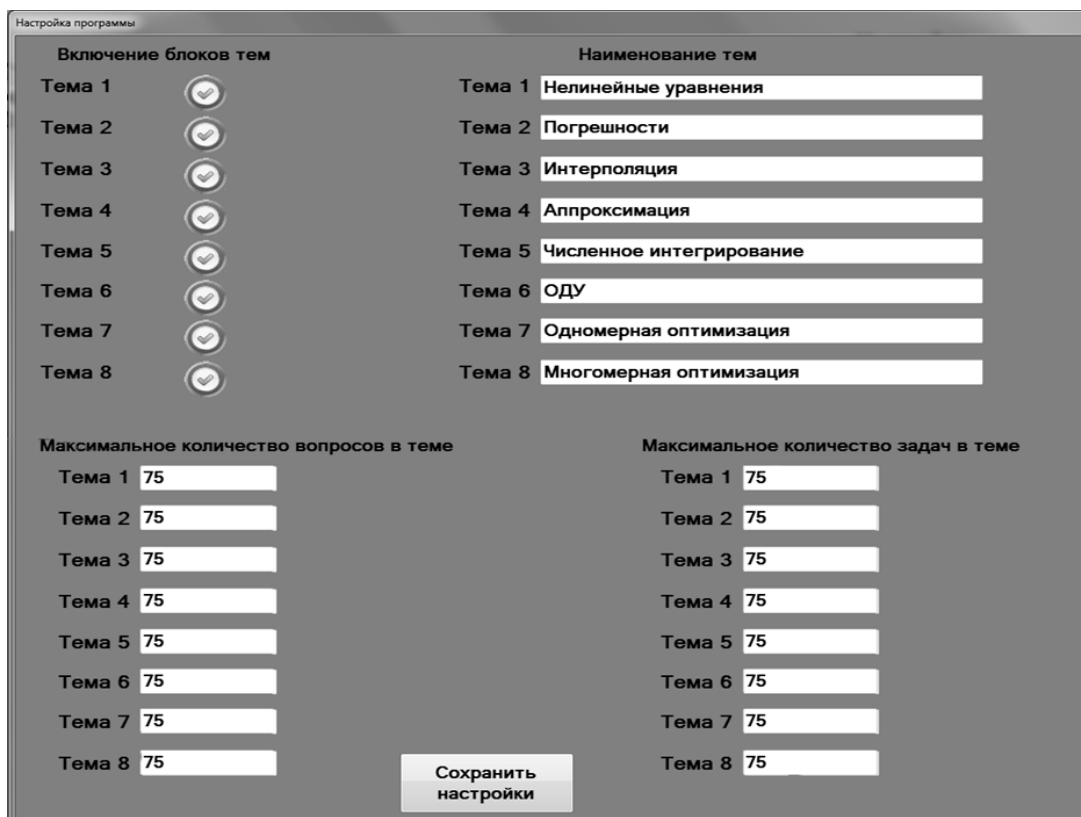


Рисунок 4 - Глобальные настройки

Для демонстрации программы было выбрано несколько тем с вопросами и задачами. После произведения всех настроек их необходимо сохранить, для чего достаточно просто нажать на кнопку «сохранить настройки». Вернувшись в стартовое окно (нажатием на вкладку «тест»), можно ввести личные данные и начать тестирование, нажав на кнопку «начать тест».

Перед Вами откроется окно с тестом, где будут расположены вопросы и задачи (см.рис.5).

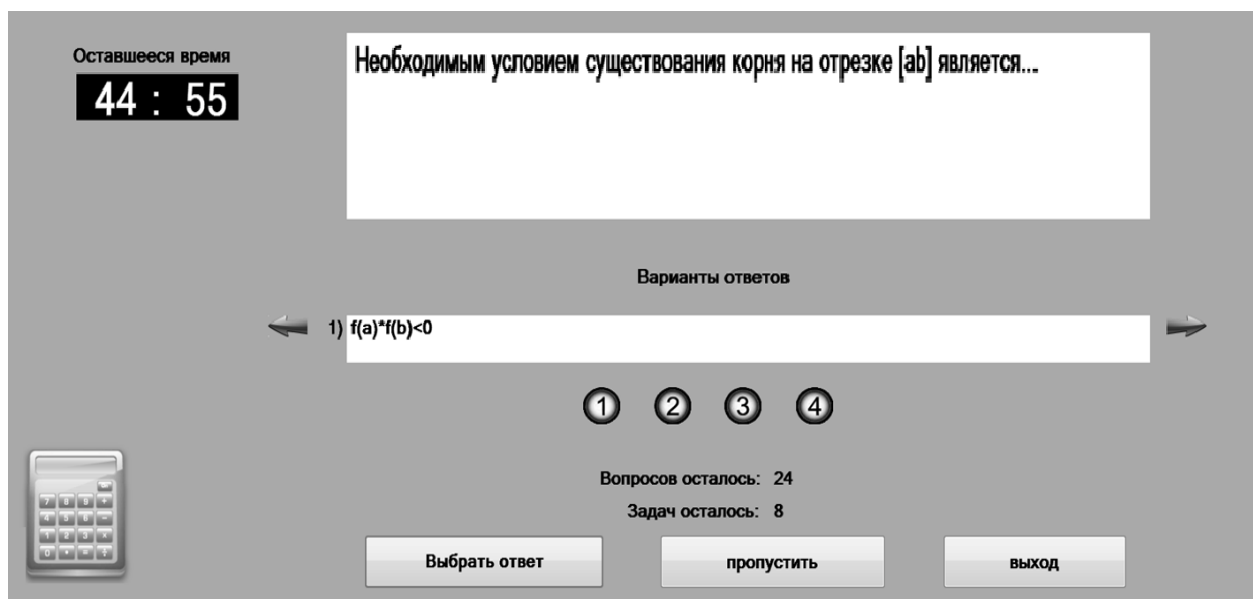


Рисунок 5 - Окно теста

Слева от окна с тестами находится большой таймер, который всегда напомнит об оставшемся времени. Ниже располагается строка, где можно просматривать варианты ответов с помощью стрелок, находящихся по обе стороны строки и произвести выбор правильного ответа (или ввести самому правильный ответ). Под строкой с ответами располагается строка с номерами вариантов ответов, а еще ниже строка с указанием количества оставшихся вопросов и задач. В помощь пользователю в нижнем левом углу экрана предусмотрен калькулятор, который может быть вызван щелчком по его изображению.

При выборе ответа нужно нажать на кнопку с его номером, после чего она загорится зеленым цветом (можно выбрать как один вариант ответа, так и несколько), после чего нужно нажать кнопку «выбрать ответ». Если необходимо пропустить вопрос или задачу, то нужно нажать кнопку «пропустить» и тогда загрузиться следующий вопрос или задача.

Если необходимо раньше времени завершить тест, то просто нужно нажать кнопку «выход». В программе предусмотрены три возможности завершения тестирования: ответив на все вопросы и задачи, либо закончится время, либо будет нажата кнопка «выход». По окончании тестирования на экране появляется окно с результатами тестирования (см.рис.6).

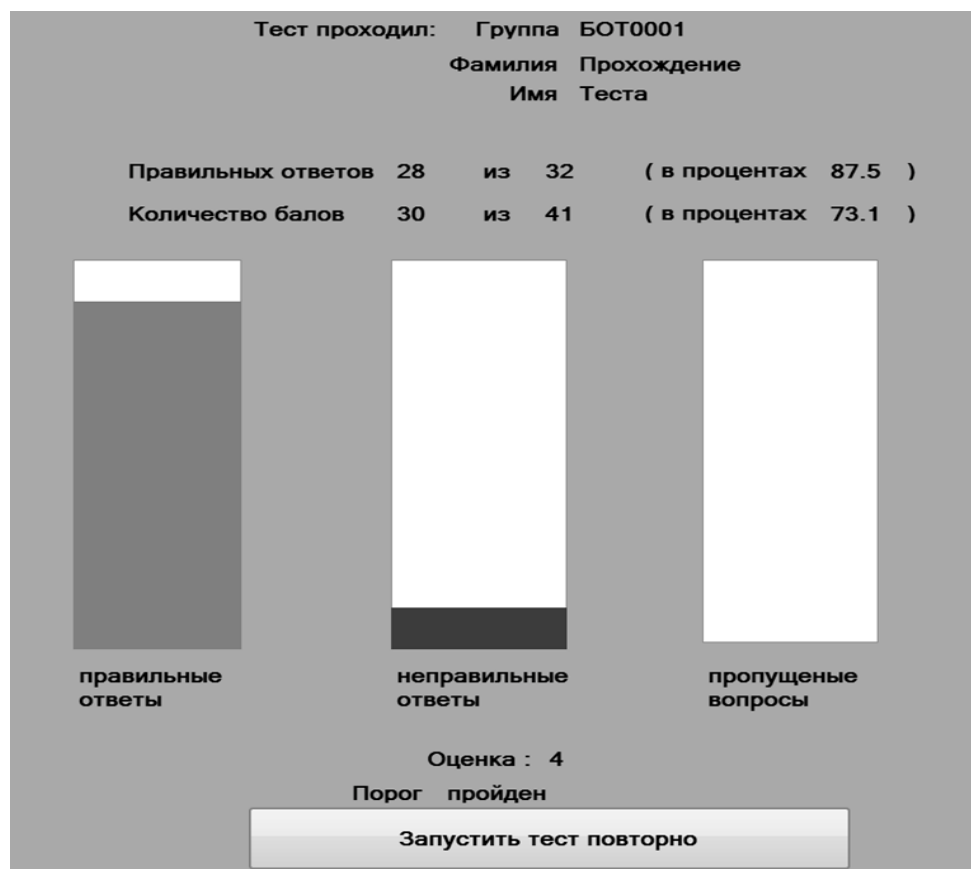


Рисунок 6 - Окно с результатами теста

Перед нами появятся: личные данные до этого введенные пользователем, количество правильных ответов из возможных и их процентное соотношение. Количество полученных баллов из возможных и их процентное соотношение[6]. Графики, визуально демонстрирующие соотношение правильных и неправильных ответов. Также предусмотрен график для вопросов и задач, на которые не успели дать ответ. Внизу будет указана оценка, если функция оценивания была включена и будет указан, прошел пороговый процент правильных ответов пользователь или нет.

При включении функции записи результатов тестирования в файл, результаты автоматически уже были записаны. Для того, чтобы следующий пользователь мог пройти тест, достаточно просто нажать на кнопку «запустить тест повторно». Следующий пользователь для прохождения теста должен проделать всю ту же процедуру, что и предыдущий.

Из вышесказанного следует, что программа очень проста, эффективна и многофункциональна. Надеюсь, она найдет широкое применение в учебном процессе и поможет многим людям в образовательной сфере.

Данная программа была разработана, чтобы помочь решить актуальность проблемы нашего времени [7]. Она решается, но не в таком объеме как хотелось бы. Но есть надежда, что при учете ошибок, анализов

предыдущих попыток сможем справиться с проблемой[8].С учетом огромного количества программ и материалов необходимо проводить детальный анализ для выявления правильных способов решения, реализации разрешения актуальной проблемы[9].В будущем программы требуют улучшения, расширения, ускорения обработки данных и проверки знаний, а также универсальности для различных сфер образования [10].

Литература

1. Ким, В.С.Разработка тестов по физике: учеб. пособие/ В.С. Ким. – Владивосток: Дальневосточный федеральный университет, 2015. – 228 с.
2. Сергеев, С.Ф. Методы тестирования и оптимизации интерфейсов информационных систем: учеб. пособие / С.Ф. Сергеев. – СПб: НИУ ИТМО, 2013. – 117 с.
3. Батулин, Н.А. Пособие по разработке тестов достижений: учеб. пособие/ Н.А. Батулин, Н.А. Курганский. – Челябинск, - Санкт-Петербург, 2001. –75с.
4. Гданский, Н.И. Адаптивные методы тестирования знания при компьютерных формах обучения: монография / Н.И. Гданский, Д.Ю. Альтиментова. – М.: РГСУ, 2015. – 220 с.
5. Денисова, Э.В. Основы вычислительной математики: учеб.-метод. пособие / Э.В. Денисова, А.В. Кучер. – СПб.: СПбГУ ИТМО, 2010. – 164 с.
6. Копченова, Н.В. Вычислительная математика в примерах и задачах/ Н.В. Копченова, И.А. Марон. – М.: Лань, 2008. –367с.
7. Семенова, Т.И. Вычислительные модели и алгоритмы решения задач численными методами: учеб. пособие / Т.И. Семенова, О.М. Кравченко, В.Н. Шакин. –М.:МТУСИ, 2017. – 84с.
8. Абрамов, В.Г.Введение в язык Паскаль: учеб. пособие / В.Г. Абрамов, Н.П. Трифионов, Г.Н. Трифионова. – М.: КноРус, 2017. –380 с.
9. Касторнова, В.А.Структуры данных и алгоритмы их обработки на языке программирования Паскаль:учеб. пособие. – СПб.: БХВ-Петербург, 2016, 304 с.
10. Агуров, П.В. Последовательные интерфейсы ПК. Практика программирования. – СПб.: БХВ-Петербург, 2005. – 496 с.

Винокуров С.А., Семенова Т.И. Тестирующая программа для определения уровня знаний и обучения. Программа тестирования проста в применении и удобна для любого пользователя. Ее редактор позволяет производить выбор тем, количество вопросов или задач, проводить настройку порогового уровня и критерия оценки. Можно создать текстовый файл, в который будут автоматически занесены данные и результаты тестируемых пользователей. Произведя замену базы данных тестирующую программу можно использовать для оценки знаний практически по любой дисциплине.

Ключевые слова: программа тестирования, настройка режимов тестирования, критерий оценки уровня знаний, универсальность использования программы в учебных процессах, помощь пользователю.

Vinokurov S. A., Semenova T. I. Testing program to determine the level of knowledge and learning. The testing program is easy to use and convenient for any user. Its editor allows you to select topics, the number of questions or tasks, to configure the threshold level and evaluation criteria. You can create a text file in which data is automatically entered and the results of the tested users. By replacing the database testing program can be used to assess the knowledge of almost any discipline.

Keywords: program testing, setup of test modes, a criterion for assessing the level of knowledge, universality of use of the program in educational processes, user assistance.

Применение метода дизайн-мышления в обучении управлению ИТ-проектом

Страхович Э.В.

Высшая школа менеджмента Санкт-Петербургского государственного университета
e.strakhovich@spbu.ru

Страхович Э.В. Применение метода дизайн-мышления в обучении управлению ИТ-проектом. В докладе обсуждается опыт применения дизайн-мышления в изучении курса дисциплины управления ИТ проектами на программе бакалавриата в бизнес-школе. Использование метода дизайн-мышления при создании бизнес-кейса проекта позволило студентам лучше понять фазу сбора требований проекта и определение объема работ. Обсуждаются результаты экспериментов с использованием всего цикла процесса дизайн-мышления и использования отдельных его фаз в разных студенческих группах. Обсуждается значение фазы эмпатии в процессе обучения.

Ключевые слова: дизайн-мышление, бизнес-кейс, требования к проекту, управление ИТ-проектом, обучение.

Введение

Управление проектами - это дисциплина, которая входит в обучение каждого менеджера. Студенты бакалавриата, изучающие менеджмент и специализирующиеся в области информационного менеджмента, изучают специфику управления информационно-технологическими (ИТ) проектами. Изучения процессов всего жизненного цикла проекта осуществляется на сквозном учебном проекте, для которого задается бизнес-кейс, определяющий целесообразность выполнения проекта. В зависимости от глубины понимания начального бизнес-кейса студенты с тем или иным успехом справляются со сбором требований проекта, определением объема работ и планированием проекта.

Одной из ключевых областей управления проектом является сбор и управление требованиями. Студенты, изучающие менеджмент, в основном являются пользователями ИТ-систем и продуктов, не имеют специального образования или опыта работы в области ИТ и нуждаются в дополнительном разъяснении конкретных требований к продукту и проекту для этой области. Согласно РМВОК (Project Management Body of Knowledge) [1], каждый проект начинается с определения бизнес-кейса, подтверждающего целесообразность и обоснованность выгод исполнения данного проекта, и в соответствии с бизнес-кейсом уточняются основные функции и функциональные возможности результата проекта. На основании собранных требований к проекту и его результату определяется объем работ проекта. В предыдущие годы на курсе управления ИТ-проектом учащиеся обычно получали готовые бизнес-кейсы для учебного проекта, для которых собирали требования с использованием известных им методов, таких, как интервьюирование, мозговой штурм, анализ документации и т.д. Получая готовый бизнес-кейс из области ИТ, студенты не всегда полностью понимали его специфику, так как он мог быть не связан с их предыдущим персональным опытом, и в результате собранные требования оказывались неполными, отражали функциональности, знакомые студентам по другим программным продуктам и часто не были специфичны для конкретного бизнес-кейса.

Студенты, изучающие информационный менеджмент, в своей будущей профессиональной деятельности будут выступать либо как представители бизнеса, формирующие бизнес-кейсы ИТ-проектов, либо как бизнес-аналитики, принимающие участие в формировании бизнес-кейсов и сборе требований для проектов на стороне исполнителей. Поэтому для этих студентов в первую очередь важно понимание бизнес-составляющей и той ценности, которую можно получить из каждого конкретного проекта. Вовлеченность студентов в понимание конкретного примера из реального бизнеса для будущих менеджеров важнее, чем технологический аспект. Поэтому было принято решение в рамках курса управления ИТ-проектом предоставить возможность самим студентам формировать бизнес-кейсы для сквозного учебного проекта. Для эксперимента была определена проблемная область разработки ИТ-продукта, учащимся предстояло определить тип ИТ-продукта и бизнес-кейс его применения. Для формирования бизнес-кейса было решено применить метод дизайн-мышления.

Метод дизайн-мышления и его применение в обучении

Образование в 21 веке ищет новые способы и пути и один из них – дизайн-мышление. Дизайн-мышление – это инновационный, творческий и ориентированный на человека процесс, который

используют междисциплинарные команды для создания продуктов или услуг, ориентированных на пользователя [2]. Дизайн-мышление изначально широко использовалось в технических областях при разработке информационно-технологических продуктов. Продвижение этого метода происходит во многом благодаря разработкам компании SAP. Компания SAP рассматривает последовательность шагов метода дизайн-мышления, принятого в школе дизайна Стэнфордского университета (см. рисунок 1).

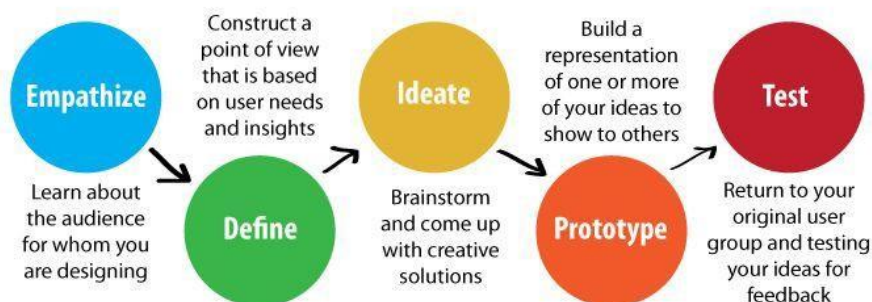


Рисунок 1 – Метод дизайн-мышления, принятый в школе дизайна Стэнфордского университета

На представленном рисунке выделены пять этапов или шагов метода дизайн-мышления:

- на этапе эмпатии (Empathize) следует как можно больше узнать о тех, для кого разрабатывается проект, т.е. о пользователях результатов проекта,
- затем составить (Define) представление о существе проекта, сформировать понимание проекта на основании потребностей пользователей,
- собрать идеи (Ideate) для реализации проекта,
- подготовить представление (Prototype) своих идей для обсуждения,
- обсудить свои предложения с потенциальными пользователями (Test), узнать их мнение о представленной идее.

После каждого из пройденных этапов возможно возвращение к предыдущим этапам, если есть необходимость. Так, как правило, после последнего этапа Test появляются дополнения или корректировки на основании отзывов потенциальных пользователей. Особенностью метода является его концентрация на интересах потенциального пользователя результатов проекта. Следует подчеркнуть развитие интегративного мышления при применении метода дизайн-мышления, так как по условию надо рассматривать все возможные предъявленные или озвученные требования, или пожелания потенциальных пользователей. Следуя методу, надо стремиться поддержать даже противоположные потребности пользователей, что не нарушает свойство непротиворечивости требований при разработке информационных систем. При применении метода дизайн-мышления приветствуются к использованию такие инструменты, как визуализация информации, мозговой штурм, интервью потенциальных пользователей и т.д.

У разных авторов можно встретить разную степень детализации шагов или этапов метода. Так, например в работе [3] предлагается следующая последовательность этапов:

- определение проблемы,
- наблюдение,
- визуализация и осмысление (понимание),
- формирование идеи или концепции,
- прототипирование и тестирование,
- дизайн бизнес-модели, внедряющей инновации.

При этом содержательная составляющая метода и общий смысл его применения остаются теми же. По-прежнему основными характеристиками метода являются человеко-центричность, интегративное мышление, управление проектами и дизайн как стратегия.

Вслед за успешным применением метода дизайн-мышления в технических областях, он был успешно применен в гуманитарных науках, включая образование [4]. Прежде всего исследователи в области образования предлагают разработать учебные планы на основе этого метода, учитывая потребности учащихся в освоении тех или иных предметных областей. Особенно активно метод продвигается университетами и школами, преподающими этот метод как предмет, как например, школа дизайна Стенфордского университета. Отдельно можно выделить рекомендацию исследователей в области образования использовать дизайнерское мышление в бизнес-школе [5]. Эта рекомендация основана на том, что, как правило, учащиеся бизнес-школы изучают

бизнес-кейсы, нацеленные на результат, в то время как дизайн-мышление ориентировано на процесс и определяет процесс, который также очень важен в управлении. Особое внимание уделяется использованию метода в обучении управлению проектами [5, 6]. При этом отмечается, что последовательность и содержание этапов метода дизайн-мышления близки к фазам жизненного цикла проекта. В исследованиях отмечается, что этот метод успешно используется для развития таких навыков и умений, как общение, переговоры, понимание потребностей других людей, работа в команде, принятие решений и т.д., что особенно важно для работы менеджера. Некоторые исследователи предлагают не внедрять весь цикл дизайн-мышления в процесс обучения, а включать некоторые из этапов, утверждая, что это также способствует развитию соответствующих навыков и умений [6].

Рассматривая соотношения метода дизайн-мышления и методов и подходов, применяемых в управлении знаниями, следует отметить близость этих методов [7]. Дизайн-мышление и управление знаниями хорошо дополняют друг друга, при этом дизайн-мышление является более практико-ориентированным подходом. С другой стороны, такие методы работы со знаниями, как мозговой штурм, интервью, обсуждения, наблюдение, документирование и т.д. применяются практически на всех этапах метода дизайн-мышления [3].

Особая роль в методе дизайн-мышления уделяется визуализации информации. Важность использования визуального представления в обучении обсуждалась в работе [8], отмечено, что визуализация информации является мощным инструментом в содействии пониманию предмета. В методе дизайн-мышления помимо задачи представления, передачи и хранения информации, визуализация работает как инструмент, активизирующий ассоциативные связи при обработке информации, что в свою очередь способствует поиску новых и нестандартных решений.

Создание бизнес-кейса для сбора требований в управлении ИТ-проектом

Как уже упоминалось выше, не все студенты-бакалавры бизнес-школы имеют опыт проектирования ИТ-продуктов или знания архитектуры ИТ-систем, в основном студенты имеют опыт применения программных продуктов в качестве пользователей и как пользователи они опробовали множество различных ИТ-систем, включая различные бизнес-системы.

В управлении проектами большое внимание уделяется управлению требованиями, особенно в части сбора требований. Особо надо выделить бизнес-требования, на основании которых определяется и оценивается создаваемая для бизнеса ценность продукта проекта. Основой для определения бизнес-требований может послужить бизнес-кейс проекта. Хорошо построенный процесс управления требованиями и сбор информации позволяет правильно определить объем работы проекта. Что касается курса управления ИТ-проектами для студентов, изучающих менеджмент, то для изучения процесса сбора требований был выбран следующий сценарий: студенты должны определить, какая система необходима в процессе обучения с точки зрения учащихся и что является бизнес-кейсом ее использования. Студенты должны были сформировать бизнес-кейс в определенной предметной области. Таким образом, этот сценарий позволил студентам определить привычную и понятную ситуацию для использования системы, т. е. бизнес-кейс, а затем позволил им обсудить и собрать правильные требования к системе, основываясь на полном понимании предметной области и самой системы. Как указано в [9], чтобы лучше понять бизнес пример, он должен основываться на предыдущей практике и опыте студентов (знания, проблемная область). Для определения бизнес-кейса был применен метод дизайн-мышления. Пошагово следуя методу, студенты определили необходимые функции будущей системы. Для работы над бизнес-кейсами студенты были организованы в группы по пять человек, что представляется оптимальным для подобных задач. Поскольку сами студенты были аудиторией, которая будет использовать эту систему в будущем, каждая команда формировала свой бизнес-кейс, исходя из потребностей, высказанных членами команды. Таким образом, студенты могли полностью применить метод до этапа тестирования, работая только внутри команды. По условиям поставленной перед ними задачи, студенты проводили десятиминутные обсуждения (мозговые штормы) на каждом этапе выполнения задания и имели время между обсуждениями на их подготовку и оформление результатов предыдущего этапа. Время, отводимое на обсуждения на каждом этапе, было специально ограничено, чтобы стимулировать участников команд к динамичной работе над принятием решений. Подойдя к этапу тестирования, каждая группа имела сформированный бизнес-кейс, описывающий информационную систему, сценарий ее использования, профиль потенциального пользователя и доводы в пользу ценности предлагаемой системы. На этапе тестирования каждая команда представляла свой бизнес-кейс на обсуждение студентам-сокурсникам, выясняя полезность предлагаемых функций информационной системы и варианты ее использования. Целью этапа тестирования было получение оценки и рекомендаций по дополнительным функциям. Следует отметить, что во время проверки идеи прототипа и обсуждения его с другими студентами создатели бизнес-кейса лучше понимали основную идею и объем работы для ими же предлагаемого прототипа, что в дальнейшем было полезно при сборе функциональных требований к информационной системе. Основываясь на полученной обратной связи, студенты обновили информацию о своем бизнес-кейсе и представили уже финальные варианты бизнес-кейсов всех групп, среди которых был выбран самый интересный, по мнению студентов. Выбранный бизнес-кейс использовался на занятиях как

сквозной пример учебного проекта. Схематичное описание выбранного бизнес-кейса представлено на рис.2. Представленный пример касается создания электронного расписания внеурочных мероприятий школы с онлайн регистрацией на мероприятия и оповещением об изменениях. Создание такого расписания позволит, по мнению студентов, учесть расписание занятий как магистратуры, так и бакалавриата, занятия которого проводятся в кампусе на значительном расстоянии от основного здания. Кроме того, студенты предложили создание локального сервиса типа BlaBlaCar среди студентов.

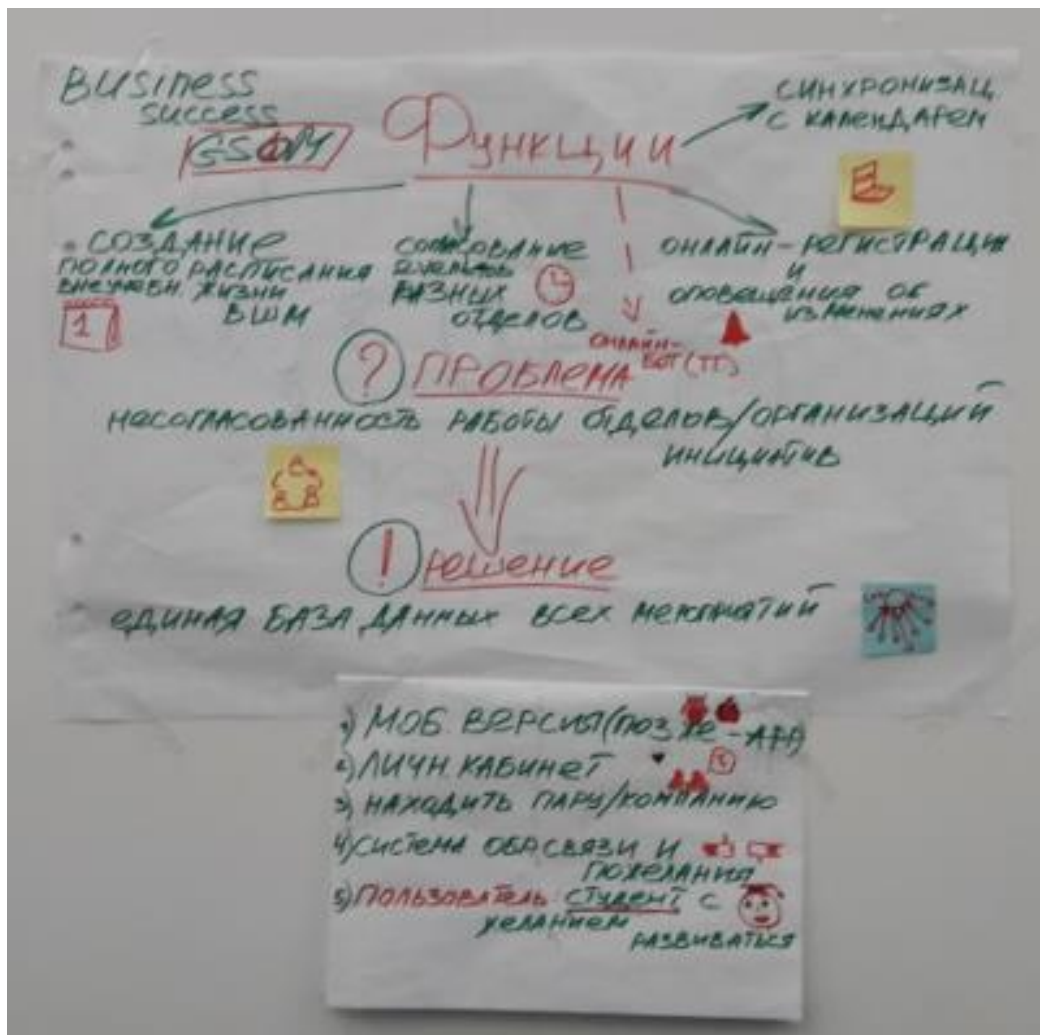


Рисунок 2 – Бизнес-кейс, созданный студентами

Следующей темой в изучении курса управления ИТ-проектом был сбор функциональных требований для информационной системы в сквозном учебном проекте. Этот этап был проще реализовать по сравнению с тем, как это было сделано в предыдущие годы, изучая ту же тему по сбору требований, поскольку сам бизнес-кейс был хорошо понят студентами и не вызывал вопросов. Повысилось качество собираемых требований, они отражали специфику именно рассматриваемого продукта. Можно привести требования, которые относились, а персонализации расписания, к учету расписания обязательных и элективных курсов и т.д.

Говоря о результатах применения метода дизайн-мышления на курсе управления ИТ-проектом, стоит напомнить, что метод был применен к отдельной задаче, изучаемой на курсе, и результат оказал положительное влияние на изучение последующих тем курса.

Метод дизайн мышления с включением разных его этапов был применен на других учебных курсах. В разработке проекта в рамках изучения темы «Умный город» на курсе магистратуры по использованию информационных технологий в государственном муниципальном управлении участвовали студенты, сознательно подошедшие к выбору специализации и имеющие некоторый опыт работы в государственных и муниципальных органах. Приступая к проекту, они имели представление о том, какие информационные технологии и как могут применяться в городской среде. На первом этапе эмпатии была сформулирована тема проекта, здесь как раз состоялись самые горячие дискуссии. На этапе прототипирования была разработана некоторая схема, представляющая, как будет использован предлагаемый проект, и кто будет его пользователем.

В этом случае применения метода дизайн-мышления не был выполнен этап тестирования и соответственно обновления проекта после его тестирования в виду ограниченности времени выполнения проекта. Тем не менее, проект был успешно представлен и можно сказать, что этап тестирования и финального представления проекта были совмещены.

Другой случай применения метода дизайн-мышления состоялся в рамках изучения курса «Информационные технологии в управлении человеческими ресурсами» на бакалавриате. Студенты предварительно изучили теорию предмета, но не имели опыта в данной области управления. Поэтому был пропущен этап эмпатии, а собственно описание информационной системы и пользовательские сценарии ее использования были представлены студентам в готовом виде. Задача, поставленная перед студентами, заключалась в определении функций информационной системы, которые позволят организовать систему мотивации сотрудников некой компании. К сожалению, уже на этапе тестирования прототипа было понятно, что студенты не до конца понимают все многообразие вариантов использования системы и основными назывались функции, связанные с материальной мотивацией сотрудников компании. Таким образом, предлагаемое решение оказалось ограниченным в сравнении с возможным, вопросы, заданные потенциальным пользователем системы на этапе тестирования, были неспецифическими и не отражали особенности пользовательского сценария, а студенты выступили пассивно. На основе этого наблюдения можно резюмировать, что включение шага эмпатии открывает возможности для творческого подхода к решению проблемы и позволяет получить наиболее интересные случаи.

Преимущества применения метода дизайн-мышления в обучении

Использование метода дизайн-мышления для разработки бизнес-кейса для сбора требований к проекту показало следующие преимущества:

- студенты, работающие над бизнес-кейсом, поняли сущность планируемой системы и хорошо сформулировали понятные им функциональные требования к системе,
- сотрудничество для выполнения задачи и человеко-ориентированность при следовании методу дизайн-мышления, позволяют развивать умения и навыки, т.н. soft skills, включая эмоциональный интеллект, возрастающая роль которого отмечается в литературе [10, 6],
- учитывая разные уровни знания информационных систем, совместная работа требовала от студентов обмена информацией, обучения друг у друга,
- четкое понимание проблемной области и бизнес-кейса позволило лучше выполнять последующие задания в разработке учебного проекта,
- возрастает вовлеченность студента в процесс обучения, вырастает его интерес к результату через участие в процессе создания этого результата, студенты активно подходят к обучению,
- работая над бизнес-кейсом, студенты осваивают метод дизайн-мышления, который становится все более популярным в сфере бизнеса.

Заключение

Подводя итог можно отметить, что применение метода дизайн-мышления в изучении управления проектами оказало положительное влияние на понимание и освоение материала, на активность студентов и на развитие их сотрудничества в рамках выполнения поставленных задач. Представляет определенный интерес развитие метода для применения в образовательных задачах.

Литература

1. A guide to the project management body of knowledge (PMBOK® guide). 6th ed. 2017, Newtown Square, Pennsylvania : Project Management Institute, Inc., 2017. – 756 p.
2. Lor R.R. Design Thinking in Education: A Critical Review of Literature // Proceedings of Asian Conference on Education and Psychology, Bangkok, Thailand, May 24-26, 2017. - P. 36-68
3. Glen R., Suci C., Baughn C., Anson R. Teaching design thinking in business schools // The International Journal of Management Education, № 13, 2015. - P. 182-192.
4. Melles G., Anderson N., Barrett T., Thompson-Whiteside S. Problem Finding through Design Thinking in Education. // In: Inquiry-Based Learning for Multidisciplinary Programs: A Conceptual and Practical Resource for Educators (Innovations in Higher Education Teaching and Learning, Volume 3), ed. Blessinger P., Carfora J.M. Emerald Group Publishing Limited, 2015. – P.191 – 209.
5. Matthews J., Wrigley C. Design and Design Thinking in Business and Management Higher Education. // Journal of Learning Design, Vol. 10, № 1. - 2017. – P. 41-54.
6. Ewin N., Luck J., Chugh R., Jarvis J. Rethinking Project Management Education: A Humanistic Approach based on Design Thinking. // Procedia Computer Science, № 121. – 2017. – P. 503-510.

7. Boersma, A. 2017. The Future of KM? Design Thinking! So here are 5 Elements to 'ECHO'... // Knowledge Management for Development Journal, Vol. 13, № 1. 2017. – P. 25-30. – Режим доступа: <http://journal.km4dev.org/>- Загл. с экрана.
8. Гаврилова Т.А., Лещева И.А., Страхович Э.В. Об использовании визуальных концептуальных моделей в преподавании. // Вестник Санкт-Петербургского университета. Серия 8. Менеджмент. Вып. 4., 2011. – С. 124-150.
9. Wang S., Wang H. A Design Thinking Approach to Teaching Knowledge Management // Journal of Information Systems Education, Vol. 19, № 2. – 2008. – P. 137-139.
10. Хомутский Д.Ю., Андреев Г.С. 2016 Дизайн-мышление как инструмент профессионального обучения персонала компании. // Управление развитием персонала. Том 2, № 46. – 2016. - С. 116-120

Страхович Э.В. Применение метода дизайн-мышления в обучении управлению ИТ-проектом. В докладе обсуждается опыт применения дизайн-мышления в изучении курса дисциплины управления ИТ проектами на программе бакалавриата в бизнес-школе. Использование метода дизайн-мышления при создании бизнес-кейса проекта позволило студентам лучше понять фазу сбора требований проекта и определение объема работ. Обсуждаются результаты экспериментов с использованием всего цикла процесса дизайн-мышления и использования отдельных его фаз в разных студенческих группах. Обсуждается значение фазы эмпатии в процессе обучения..

Ключевые слова: дизайн-мышление, обучение, бизнес-кейс, требования к проекту, управление ИТ-проектом.

Strakhovich Elvira Use of design thinking method in teaching IT project management. The paper discusses the experience of use the design thinking method in teaching the course of the discipline of IT project management in the undergraduate program at a business school. Using the design thinking method to create a business case for a project, the students better understand the phase of project requirements gathering and of definition of project scope of work. The paper presents the results of experiments when the entire cycle of the design thinking method or only its separate phases were used in different student groups. The importance of the empathy phase in the learning process is highlighted.

Key words: design thinking, teaching, business case, project requirements, IT project management.

Экспериментальное исследование зависимости показаний MEMS-датчиков от колебаний температуры

Суворова Е.Ю., Гвоздюкова С.Н.

Луганский национальный университет имени Тараса Шевченко

suvorova.itstep@gmail.com, hvozdiukova@mail.ru

Суворова Е.Ю., Гвоздюкова С.Н. Экспериментальное исследование зависимости показаний MEMS-датчиков от колебаний температуры. Статья посвящена анализу влияния температурных флуктуаций на точность показаний периферийных датчиков контроллера. Описано проведение эксперимента. Реализована программно-алгоритмическая компенсация влияния температуры на выходные показания микромеханических датчиков с использованием коэффициента корреляции Пирсона для коррекции показаний датчиков в зависимости от окружающей температуры.

Ключевые слова: мультикоптер, температурные флуктуации, точность показаний датчиков, программно-алгоритмическая компенсация.

Введение

Активно развивающиеся технологии микроэлектромеханических систем (MEMS) сделали доступными для применения в недорогих массовых устройствах датчики и системы измерения параметров, которые ранее использовались лишь в дорогой авиационной, космической и военной технике. Особенно это относится к датчикам для измерения параметров движения – инерциальным датчикам (гироскопам, акселерометрам) [1]. MEMS-сенсоры, используемые в беспилотных летательных аппаратах, обладают отличными характеристиками по параметрам энергопотребления, цене, габаритным размерам и массе.

Датчики должны надежно функционировать. Грамотный подбор датчиков дает гарантию на жизнеспособность системы, а значит необходимо их дополнительно верифицировать на предмет соответствия заявленным эксплуатационным характеристикам, особенно в экстремальных внешних условиях. Известно, что существенным недостатком систем на основе микромеханических датчиков является относительно невысокая точность измерения параметров, а также существенная зависимость этой точности от нестабильности напряжения питания и температурных флуктуаций [2,3].

Таким образом, исследование температурной зависимости показаний периферийных датчиков и их работоспособность является актуальным направлением исследования.

Целью работы является исследование влияния температурных флуктуаций на точность показаний датчиков контроллера беспилотного летательного аппарата.

Наиболее точно учесть погрешность показаний можно, сняв экспериментальную тарировочную характеристику акселерометра и датчика давления в лабораторных условиях и затем вносить поправки в получаемые результаты [4]. Задачами экспериментальных исследований являлись определение зависимости погрешности барометрического датчика давления MS5611-01BA и акселерометра - датчика перемещения MPU-6050, применяемого с полетным контроллером ArduPilot Mega, от температуры и нахождение их аналитических выражений.

MPU-6050 - интегрированное 6-осевое MotionTracking устройство, которое сочетает в себе 3-осевой гироскоп, 3-осевой акселерометр и Digital Motion Processor™ (DMP) в небольшом пакете 4x4x0,9 мм. Благодаря специальной сенсорной шине I2C он напрямую принимает входные сигналы от внешнего 3-осевого компаса, чтобы обеспечить полный выход на 9-позиционный MotionFusion™. Устройство MotionTracking MPU-6050 с его 6-осевой интеграцией, встроенным программным обеспечением MotionFusion™ и встроенной программой калибровки времени выполнения позволяет производителям устранять дорогостоящую и сложную выборку, квалификацию и системную интеграцию дискретных устройств, гарантируя оптимальную производительность [5-7]. MPU-6050 также предназначен для взаимодействия с несколькими неинерциальными цифровыми датчиками, такими как датчики давления, на его вспомогательном порту I2C. MPU-6050 оснащен тремя 16-разрядными аналого-цифровыми преобразователями (АЦП) для оцифровки выходов гироскопа и тремя 16-разрядными АЦП для оцифровки выходов акселерометра. Для точного отслеживания как быстрых, так и медленных движений детали имеют программируемый пользователем гироскоп с полным диапазоном ± 250 , ± 500 , ± 1000 и $\pm 2000^\circ/\text{сек}$ (dps) и полномасштабный пользовательский программируемый акселерометр, диапазон $\pm 2g$, $\pm 4g$, $\pm 8g$ и $\pm 16g$.

MS5611-01BA – новое поколение датчиков высотомера высокого разрешения от MEAS Switzerland с интерфейсом шины SPI и I2C. Этот барометрический датчик давления оптимизирован для высотомеров и вариометров с разрешением высоты 10 см. Модуль датчика включает в себя датчик высокого линейного давления и 24-битный АЦП сверхнизкой мощности с внутренними заводскими калиброванными коэффициентами. Он обеспечивает точное цифровое значение давления и различные режимы работы, которые позволяют пользователю оптимизировать скорость преобразования и потребление тока. Протокол связи прост, без необходимости программирования внутренних регистров в устройстве. Небольшие размеры: 5,0 x 3,0 x 1,0 мм позволяют интегрировать его в мобильные устройства. Используемый принцип измерения приводит к очень низкому гистерезису и высокой стабильности сигнала давления и температуры. MS5611-01BA используют в диапазоне от 10^{-4} до 10^{10} Па; они имеют приведенную погрешность $\pm (0,1-5) \%$ [8].

Методика проведения эксперимента

Все исследования проводились на экспериментальном стенде, в состав которого входит: испытуемый акселерометр, эталонный акселерометр, датчик давления, барометр, термометр, термокамера, диафрагменный вакуумный насос. Абсолютная погрешность испытуемого датчика определялась как разница между полученными показаниями и показаниями эталонного прибора.

При испытании барометрического датчика давления в термической камере при фиксированной температуре создавалось разрежение вакуумным насосом в диапазоне давлений, указанных в паспорте датчика. Испытания акселерометра проводились в разных положениях относительно оси чувствительности. При этом ось X сонаправлена с осью чувствительности акселерометра, а оси Y и Z взаимно перпендикулярны. Измерения выполнены в положительную и отрицательную стороны относительно осей. Эталонный и поверяемый акселерометры устанавливались соосно, что обеспечивает идентичность ускорений, испытываемых акселерометрами. Все испытания проводились при значениях температур в диапазоне от -50°C до 70°C .

Расчет ошибок измеряемых величин проводился с использованием статистических методов. Аппроксимация зависимостей проводилась методом регрессионного анализа, основанном на методе наименьших квадратов с допущением, что распределение случайных ошибок носит нормальный (Гаусовский) характер.

Расчеты ошибок измерения проводились по следующей зависимости (1):

$$\sigma = \sqrt{\sigma_S^2 + \sigma_R^2}, \quad (1)$$

где σ_R - систематическая погрешность прибора;

σ_S - по теории малых выборок вычислялась по зависимости (2):

$$\sigma_S = \sqrt{\sigma_A^2 + \sigma_O^2} \quad (2)$$

В приведенной выше формуле: σ_A – ошибка прибора; σ_O – ошибка считывания.

Для расчета доверительного интервала с учетом поправки Стьюдента t_{st} на конечное число опытов, в границах которого находятся все (случайные) ошибки измерений среднего значения фактора, использовалось соотношение (3):

$$|\Delta \bar{y}| = t_{st} \sigma \quad (3)$$

Окончательный результат параллельного измерения представлялся в следующем виде (4):

$$y = \bar{y} \pm \Delta y, \quad (4)$$

с обязательным указанием доверительной вероятности, для которой учитывалась поправка Стьюдента.

Во всех наших измерениях доверительная вероятность была принята постоянной и равной 0,95.

Определение случайной погрешности при параллельных испытаниях производилось по следующей методике.

Для конкретного фактора x_i , есть наиболее вероятное значение измеренной величины \bar{y} из n измерений, являющееся среднеарифметическим значением для принятого («нормального») закона распределения ошибки (5):

$$\bar{y} = \sum_{i=1}^{i=n} y_i / n \quad (5)$$

Ошибка одиночного измерений оценивалась его дисперсией (6):

$$S^2 = \frac{\sum_{i=1}^{i=n} (\bar{y} - y_i)^2}{n-1} \quad (6)$$

По критерию Стьюдента, вид которого приведен ниже (7):

$$t_s = \frac{|\bar{y} - y_i|}{S} \quad (7)$$

определялись грубые измерения, для которых расчетное значения критерия превышает таблично. Потом грубые измерения исключались и заодно определялось значение \bar{y} и оценивалась его дисперсия воспроизводимости (8):

$$\sigma_R^2 = \frac{\sum_{i=1}^{i=n} (\bar{y} - y_i)^2}{n(n-1)} \quad (8)$$

Доверительный интервал для всех измеряемых величин рассчитывался исходя из доверительной вероятности, равной $\alpha = 0,95$, с учетом поправки Стьюдента на конечное число опытов.

Для исследования точности показаний датчика ускорения использовался лабораторный эталон AP10. Технические характеристики: чувствительность 1 пКл/г, деформационная чувствительность 5000 Мом, электрическая ёмкость 36 пФ, температурный диапазон – 60...+ 200°С, допускаемая погрешность 0,01%. В исследовании точности показаний барометрического датчика использовался эталонный прибор для высокоточного измерения значений абсолютного и избыточного давления жидкостей и газов ПДЭ-010, к основным техническим характеристикам которого относится предел допускаемой основной относительной погрешности $\pm 0,02$ %. В качестве эталонного измерителя температуры применялся ЭТС-100/2, пределы допускаемой доверительной абсолютной погрешности при доверительной вероятности 0,95 составляет 0,02%.

Как варьируемые факторы выберем: давление p и температуру T для датчика давления и ускорение a и температуру T для акселерометра датчика перемещения. Итог варьирования и основной уровень каждого фактора приведены в таблице 1, а в таблицах 2-5 приведены матрицы планирования эксперимента для акселерометра датчика перемещения и для барометрического датчика давления, соответственно.

Таблица 1 – Исходные данные. Кодирование факторов

	-1.41	-1	0	1	1.41
p, Pa	$4,05 \cdot 10^4$	$5,2 \cdot 10^4$	$8 \cdot 10^4$	$1,08 \cdot 10^5$	$1,195 \cdot 10^5$
$a_{x, y, z}, m \varphi^2$	-15,51	-11	0	11	15,51
$T, ^\circ C$	-56,4	-40	10	50	66,4

Таблица 2 – Матрица планирования для акселерометра датчика перемещения для оси X

№ опыта	X ₁	X ₂	Y
1	-1	-1	0,076
2	-1	1	0,073
3	1	-1	0,074
4	1	1	0,072
5	-1,41	0	0,0727
6	1,41	0	0,07196
7	0	-1,41	0,07448

8	0	1,41	0,07376
9	0	0	0,077

Таблица 3 – Матрица планирования для акселерометра датчика перемещения для оси Y

№ опыта	X ₁	X ₂	Y
1	-1	-1	0,075
2	-1	1	0,073
3	1	-1	0,074
4	1	1	0,074
5	-1,41	0	0,074
6	1,41	0	0,072
7	0	-1,41	0,076
8	0	1,41	0,075
9	0	0	0,072

Таблица 4 – Матрица планирования для акселерометра датчика перемещения для оси Z

№ опыта	X ₁	X ₂	Y
1	-1	-1	0,075
2	-1	1	0,073
3	1	-1	0,074
4	1	1	0,074
5	-1,41	0	0,074
6	1,41	0	0,073
7	0	-1,41	0,075
8	0	1,41	0,074
9	0	0	0,071

Таблица 5 – Матрица планирования для барометрического датчика давления

№ опыта	X ₁	X ₂	Y
1	-1	-1	212
2	-1	1	204
3	1	-1	206
4	1	1	201
5	-1.41	0	202
6	1.41	0	200
7	0	-1.41	207
8	0	1.41	205
9	0	0	214

Используя полученные кодированные факторы были проведены физические эксперименты и получены уравнения регрессии, а также трехмерные изображения функций отклика (рис. 1 – 4).

Уравнение регрессии для акселерометра датчика перемещения (ось X):

$$\Delta a = 0.07699 - 0.0005069x_1 - 0.0007541x_2 - 0.002208x_1^2 - 0.001308x_2^2 + 0.0002500, r^2 = 0.91.$$

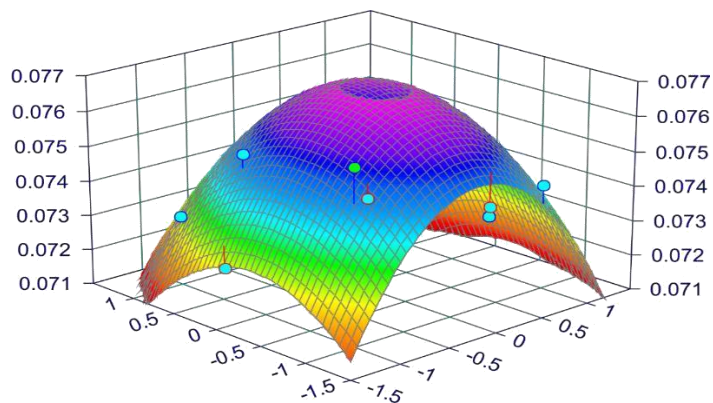


Рисунок 1 – Трехмерное изображение функции отклика для акселерометра датчика перемещения (ось X)

Уравнение регрессии для акселерометра датчика перемещения (ось Y):
 $\Delta a = 0.07200 - 0.0003536x_1 - 0.0004275x_2 + 0.0004352x_1^2 + 0.001693x_2^2 + 0.0005000$, $r^2 = 0.92$.

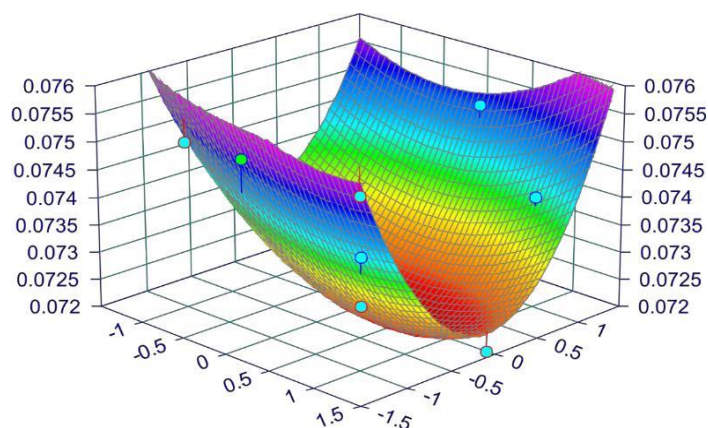


Рисунок 2 – Трехмерное изображение функции отклика для акселерометра датчика перемещения (ось Y)

Уравнение регрессии для акселерометра датчика перемещения (ось Z):
 $\Delta a = 0.07100 - 0.0001768x_1 - 0.0004275x_2 + 0.001253x_1^2 - 0.001756x_2^2 + 0.0005000$, $r^2 = 0.98$

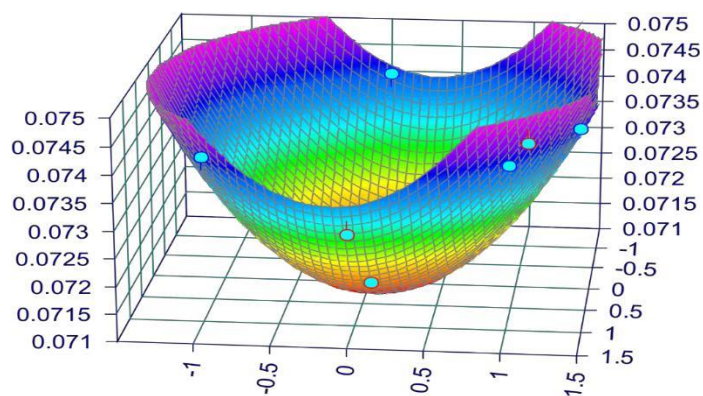


Рисунок 3 – Трехмерное изображение функции отклика для акселерометра датчика перемещения (ось Z)

Уравнение регрессии для барометрического датчика давления:
 $\Delta p = 202.0 + 0.6043x_1 - 0.1768x_2 + 5.204x_1^2 - 1.084x_2^2 - 1$, $r^2 = 0.98$.

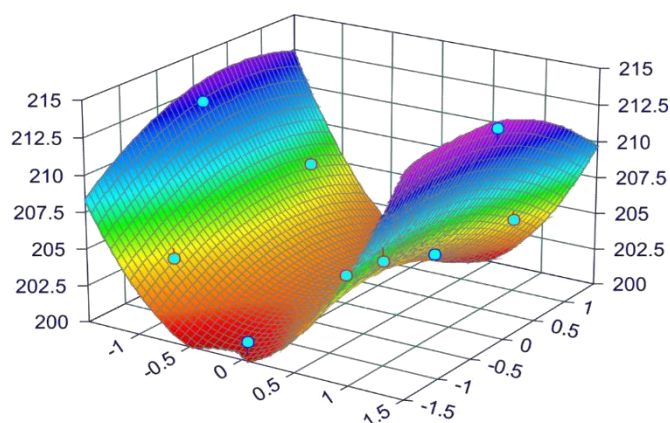


Рисунок 4 – Трехмерное изображение функции отклика для барометрического датчика давления

Выполненные расчеты показали, что для данного случая квадрат коэффициента корреляции Пирсона равен 0.91, 0.92, 0.98 для оси X, оси Y и оси Z акселерометра датчика перемещения, 0.98 для барометрического датчика давления, что говорит об адекватности полученной аппроксимационной модели и возможности ее использования для коррекции показаний датчика в зависимости от окружающей температуры.

Реализация коррекции датчиков

На основании экспериментально полученных данных была установлена необходимость внести в конструкцию совместимый с платформой датчик температуры, подключаемый по стандартному интерфейсу. Встроенный датчик температуры позволит в случае необходимости производить алгоритмическую компенсацию температурных погрешностей. Программное обеспечение реализовано на языке программирования C++ и содержит следующие функции:

- проводит инициализацию блока исполнительных устройств при запуске летательного аппарата;
- обеспечивает постоянное взаимодействие с периферийными модулями такими, как датчики, беспроводной интерфейс;
- осуществляет непрерывное дистанционное соединение и синхронизацию с пультом управления базовой станции;
- поддерживает алгоритмы для аварийных ситуаций (потеря сигнала пульта управления, сильный разряд бортовых батарей, избежание столкновения с боковыми препятствиями на низких скоростях, необходимость экстренного отключения двигателей);
- при необходимости осуществляет фильтрацию информации, получаемой с периферийных датчиков и производить коррекцию показаний.

Таким образом, применяется способ коррекции датчиков по температуре, при котором характеристика датчика аппроксимируется полиномом. На вход процессора подаются сигналы с датчика температуры и пересчитываются в соответствии с указанной аппроксимацией. Это имеет большое значение при обработке сигнала ограниченными вычислительными ресурсами микроконтроллера, так как степенные многочлены легко вычислить за конечное число шагов и проводить операции дифференцирования и интегрирования с применением только основных арифметических операций.

Летные испытания

Точность датчика должна быть тщательно проверена, иначе всякая модернизация теряет смысл.

На рисунке 5 представлены зафиксированные значения параметров высоты до коррекции показаний барометра, $Alt (m) = 22,36$. В пределах зоны физического нахождения коптера выходные показания барометра некорректны.

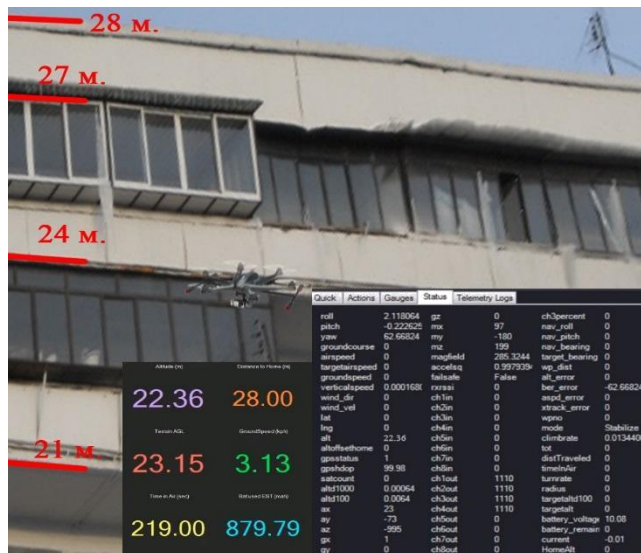


Рисунок 5 – Некорректные показания барометра

После проведения программной компенсации показаний датчика по температуре, произведено следующее полетное испытание. Как видно на рисунке 6, погрешность компенсирована, показания корректны, Alt (m) = 26,45).



Рисунок 6 – Корректные показания барометра после компенсации

Вывод

Установлено, что зависимости $U = f(T, ^\circ\text{C})$ выходных информационных сигналов с акселерометра и датчика давления от температуры носят нелинейный характер и при их математическом описании целесообразно применять полиномиальную аппроксимацию третьего порядка.

Анализ полученных экспериментальных данных показывает, что в диапазоне температур от -50°C до $+70^\circ\text{C}$, линия тренда температурного дрейфа меняет свое положение, что соответствует увеличению ошибки значений выходных сигналов микромеханических датчиков.

Использован коэффициент корреляции Пирсона для коррекции показаний датчика в зависимости от окружающей температуры. Применена алгоритмическая компенсация влияния температуры на выходные показания микромеханических датчиков. Введение коррекции приводит к увеличению стабильности работы прибора, а именно уменьшению среднеквадратической ошибки. Компенсация позволяет уменьшить нестабильность смещения нуля, исключить дрейф, вызванный температурой, и тем самым увеличить долговременную стабильность выходного сигнала. Конечным результатом работы является повышение точности показаний периферийных датчиков за счет предложенного метода уменьшения погрешностей, использующего комплексный подход алгоритмической компенсации.

Литература

1. Аврутов В.В. Надежность и диагностика / В.Аврutow, Н. Бурау. - Laplambert Academic Publishing, 2015. – 164 с.
2. Аврутов В.В. Интеркардинальная девиация гироскопа с динамически настраиваемым гироскопом // *Прикладная механика*. – 1991. – т. 27, № 6. – С.99-106.
3. Антонова М.В. Модель погрешностей волоконно-оптического гироскопа / Автореферат диссертации на соискание уч. степ. канд. техн. наук / - М.: МГТУ им. Н.Э. Баумана, 2014. – 18 с.
4. Бурмистров В.П. Автоматизированная система испытаний прецизионных гироскопических приборов / В.П. Бурмистров, Н.А. Темляков, А.А. Хасанов // *Гироскопия и навигация*. – 1995. – №3 (10). – с. 24-29.
5. Бутенин Н.В. Курс теоретической механики: учебник в 2-х томах. Т. II: Динамика. – М.: Наука, 1979. – 544 с.
6. Измайлов Е.А. Скалярный способ калибровки и балансировки бесплатформенных инерциальных навигационных систем / Е.А. Измайлов, С.Н. Лепе, А.В. Молчанов, Е.Ф. Поликовский // XV Международная конференция по интегрированным навигационным системам. / ГНЦ РФ ЦНИИ «Электроприбор». – СПб., 2008. – С.145-154.
7. Avrutov V. On Scalar Calibration of an Inertial Measurement Unit / V. Avrutov, S. Golovach, T. Mazepa // 19-th St. Petersburg International Conference on Integrated Navigation Systems. St. Petersburg, Russia, 2012. - State Research Center (CSRI) Elektropribor, 2012, pp.117-121.
8. Ideal-Aerosmith. The Model 9790-BVR Scorsby Motion Test Table [Электронный ресурс]. – Режим доступа: <http://www.ideal.aerosmith.com/motion/scorsby-motion-test-table-9790-bvr>

Суворова Е.Ю., Гвоздюкова С.Н. Экспериментальное исследование зависимости показаний MEMS-датчиков от колебаний температуры. Статья посвящена анализу влияния температурных флуктуаций на точность показаний датчиков контроллера. Описано проведение эксперимента. Реализована программно-алгоритмическая компенсация влияния температуры на выходные показания микромеханических датчиков с использованием коэффициента корреляции Пирсона для коррекции показаний датчика в зависимости от окружающей температуры.

Ключевые слова: мультикоптер, температурные флуктуации, точность показаний датчиков, программно-алгоритмическая компенсация.

Suvorova Eugenia, Hvozdiukova Svetlana Experimental study of the dependence of MEMS sensors readings from temperature fluctuations. The article is devoted to the analysis of the influence of temperature fluctuations on the accuracy of the sensor readings of the controller. An experiment has been described. Implemented algorithmic compensation of the effect of temperature on the output readings of micromechanical sensors using the Pearson correlation coefficient to correct the sensor readings depending on the ambient temperature.

Keywords: multicopter, temperature fluctuations, accuracy of sensor readings, algorithmic compensation.

ИСПОЛЬЗОВАНИЕ ТЕХНОЛОГИИ MAPREDUCE ДЛЯ ПАРАЛЛЕЛЬНОГО ПОИСКА В ШИРИНУ КРАТЧАЙШЕГО ПУТИ В ГРАФАХ БОЛЬШОЙ РАЗМЕРНОСТИ, АЛГОРИТМ BFS: WEIGHTED EDGES

Диденко А.А., магистрант
Научный руководитель: А.Г. Пимонов, д.т.н., профессор
Кузбасский государственный технический университет
имени Т.Ф. Горбачева г. Кемерово
mozggg1@gmail.com, pag_vt@kuzstu.ru

Диденко А.А., Пимонов А. Г. Использование технологии MapReduce для параллельного поиска в ширину кратчайшего пути в графах большой размерности, алгоритм BFS: weighted edges. Статья посвящена проблеме анализа информации, хранящейся в виде графов, с помощью алгоритма BFS. В статье рассматриваются способы решения данной проблемы.

Ключевые слова: MapReduce, Hadoop, графы, обход графа, поиск кратчайшего пути, BFS, BFS: weighted edges.

Введение

В современном мире достаточно много информации хранится в виде графов, например: сообщества пользователей в социальных сетях, интернет-трафик, структура связей компьютеров в сети интернет, энергетические и телекоммуникационные вышки. При хранении данных в таком формате возникает ряд проблем, таких как: поиск кратчайшего пути, поиск minimum spanning tree, поиск максимального потока, bipartitematching, поиск групп вершин графа. С решением данных задач может помочь алгоритм распределенных вычислений MapReduce [1]. В статье рассматривается решение задачи нахождения кратчайшего пути с помощью алгоритма параллельного поиска в ширину (Breadth-First Search).

Постановка задачи

Найти кратчайший путь от начальной вершины до заданной или заданных вершин. На рисунке 1 изображен пример графа и его обхода по ширине (см.рис.1).

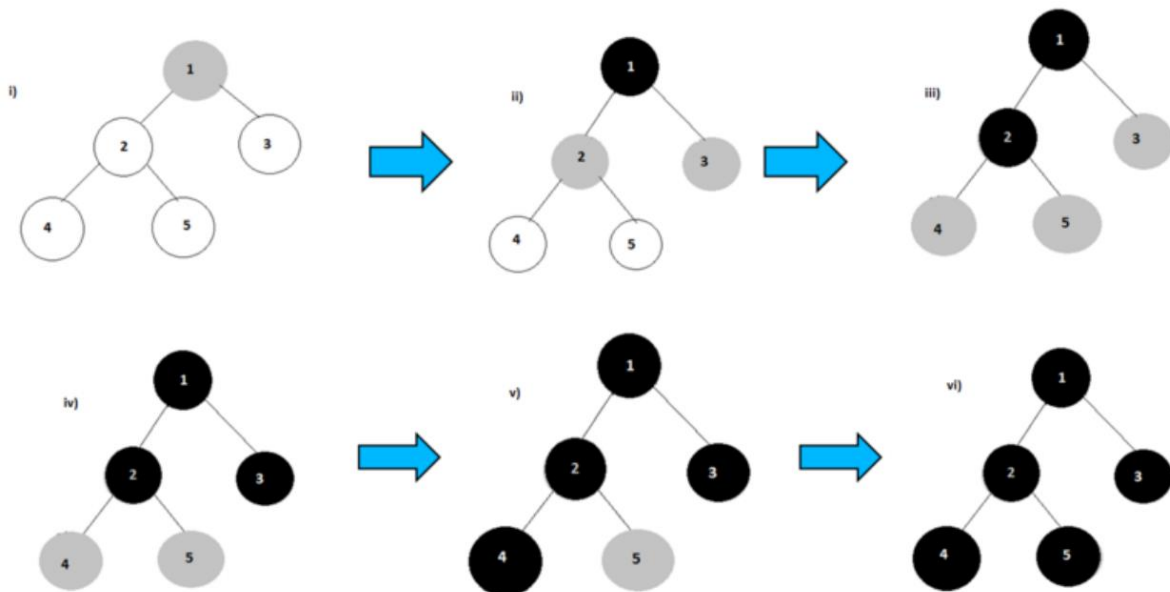


Рисунок 1 — Обход графа в ширину

Алгоритм параллельного поиска в ширину (Breadth-First Search)

Вершины графа будут представлены структурой, состоящей из трех величин: n – номер вершины, d – расстояние от начала, al – список вершин, доступных из текущей вершины. При инициализации для всех вершин кроме начальной значение d будет установлено как `integer.max_value`.

Mapper

$$\forall m \in al: emit(m, d + 1)$$

Sort/Shuffle

На данном этапе будет происходить группировка расстояний по достижимым вершинам.

Reducer

В этой части алгоритма будет происходить выборка пути с кратчайшим расстоянием для каждой достижимой вершины, а также некоторые дополнительные проверки по отслеживанию актуального пути [2].

Псевдокод алгоритма представлен на рисунке 2 (см.рис.2). Пример обхода графа с помощью параллельного алгоритма BFS представлен на рисунке 3 (см.рис.3).

```
1 class Mapper
2     method Map(nid n, node N)
3         d <- N.Distance
4         Emit(nid n, N) //Прход по всей структуре графа
5         for all nodeid m c N.AdjacencуList do
6             Emit(nid m, d + 1) //Отображение дистанций до достижимых вершин
7
8 class Reducer
9     method Reducer(nid m, [d1, d2, ...])
10        dmin <- inf
11        M <- null
12        for all d c counts [d1, d2, ...] do
13            if IsNode(d) then
14                M <- d //Восстановление структуры графа
15            else if d < dmin then Look for shorter distance
16                dmin <- d
17        M.Distance <- dmin //Обновление минимальной дистанции
18        Emit(nid n, node M)
```

Рисунок 2. Псевдокод алгоритма BFS

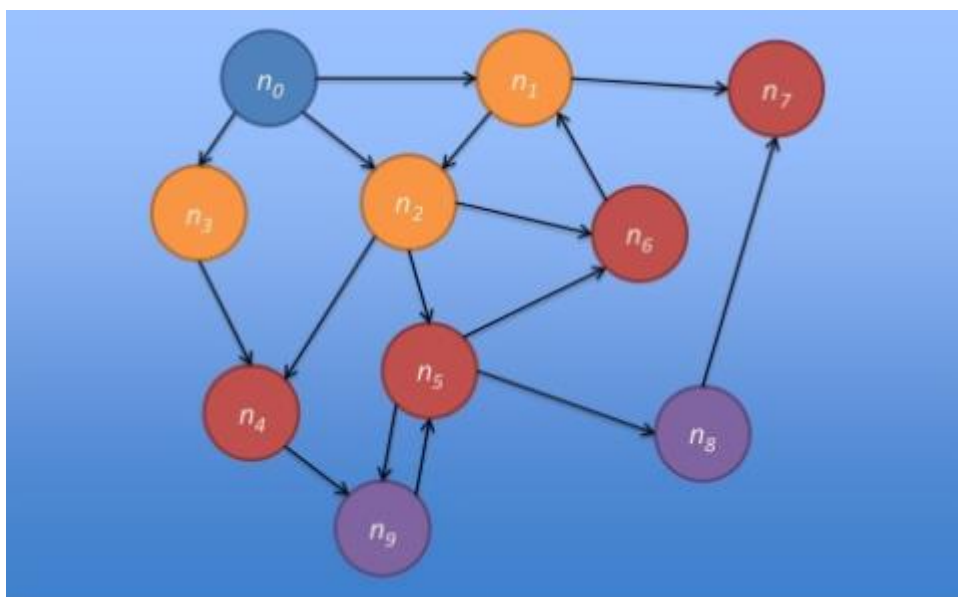


Рисунок 3. Обход графа алгоритмом BFS

Итерации BFS

Каждая итерация MapReduce смещает границу продвижения по графу или frontier на один hop. Последующие операции включают в себя все увеличивающееся количество посещенных вершин, потому что frontier расширяется. Для обхода всего графа требуется множество итераций. В ходе выполнения Mapper и Sort/Shuffle части алгоритма возникает проблема с сохранением структуры графа, в частности списка смежных вершин al . Для решения данной проблемы на шаге Mapper также пишется (n, al) [3]. Далее будет представлен лог решения задачи поиска кратчайшего пути для графа, изображенного на рисунке 1 (см.рис.1), а также будут отображены все итерации на рисунке 4 (см.рис.4).

Лог итераций алгоритма BFS

Input 1: 1 2,3|0|V, 2 1,3,4,5|mv|N, 3 1,4,2|mv|N, 4 2,3|mv|N, 5 2|mv|N;
Output 1/Input 2: 1 2,3|0|S, 2 1,3,4,5|1|V, 3 1,4,2|1|V, 4 2,3|mv|N, 5 2|mv|N;
Output 2/Input 3: 1 2,3|0|S, 2 1,3,4,5|1|S, 3 1,4,2|1|S, 4 2,3|2|V, 5 2|2|V;
Output 3: 1 2,3|0|S, 2 1,3,4,5|1|S, 3 1,4,2|1|S, 4 2,3|2|S, 5 2|2|S|.

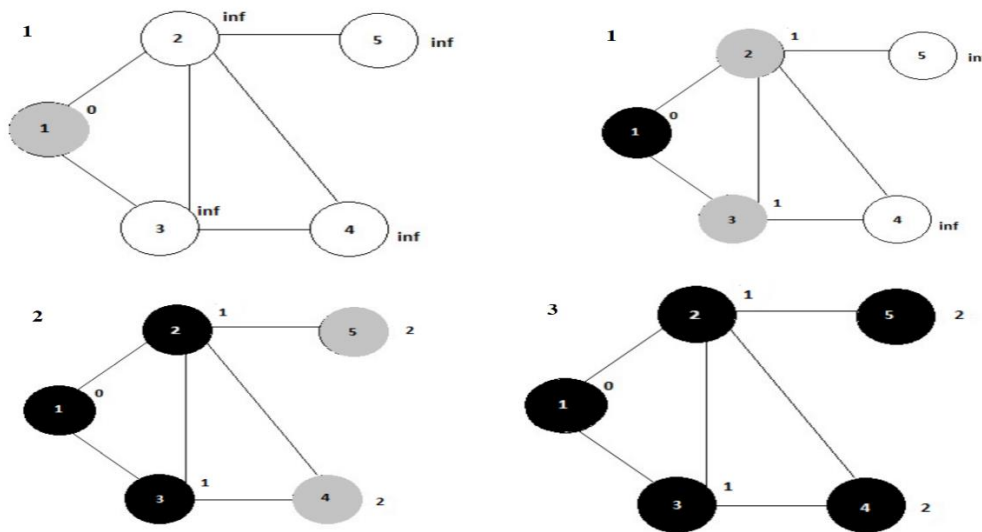


Рисунок 4. Итерации BFS

Анализ результатов работы алгоритма BFS

Данный алгоритм хорошо себя показывает в задачах на графах очень большого размера, что не позволяет их эффективно решать с помощью линейных алгоритмов. Однако у данного алгоритма есть 2 проблемы:

- 1) В ходе выполнения алгоритма выполняется много лишней работы, из-за использования подхода прямого перебора;
- 2) Полезная часть алгоритма выполняется только на текущей границе обхода.

Из-за чего возникает вопрос: Можно ли использовать MapReduce более эффективно?

BFS: WeightedEdges

Для улучшения эффективности работы алгоритма добавим параметр w , отвечающий за вес ребра. Так же на этапе Mapper вместо функции $\forall m \in al: emit(m, d + 1)$ будем использовать $\forall m \in al: emit(m, d + w_p)$ для каждой вершины m [4].

Критерий завершения

Предыдущий критерий завершения: первое посещение искомой вершины, не подойдет в данном случае, так как может быть найден более короткий путь с точки зрения веса, пример такой ситуации изображен на рисунке 5 (см.рис.5).

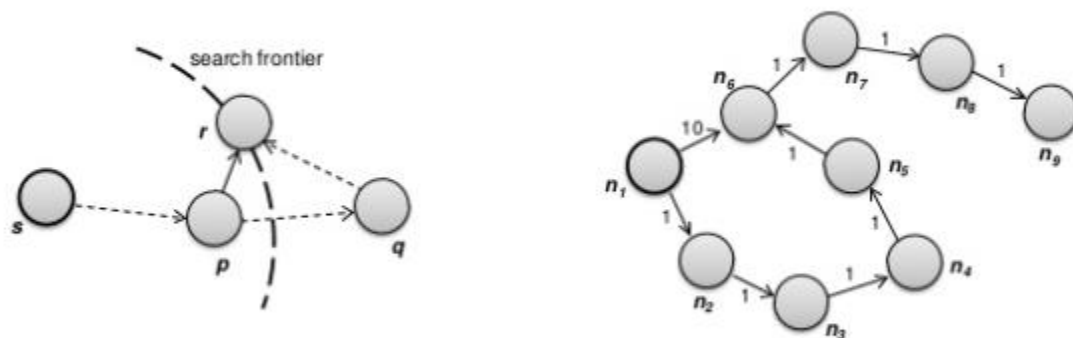


Рисунок 5. Сложности алгоритма BFS Weighted Edges

В худшем случае необходимо будет обойти $N-1$ вершин графа, однако в большинстве реальных задач количество будет примерно равно диаметру графа. Критерием завершения станет ситуация, когда минимальный путь у каждой вершины перестает изменяться. Для реализации этого критерия можно использовать счетчики MapReduce [5].

Заключение

В данной статье был рассмотрен алгоритм параллельного поиска в ширину (Breadth-First Search) с применением технологии MapReduce на примере задачи поиска кратчайшего пути в графе. Данный алгоритм хорошо себя показывает в задачах на графах очень большого размера, что не позволяет их эффективно решать с помощью линейных алгоритмов, но у него так же есть 2 проблемы, которые можно решить с помощью модификации алгоритма добавлением параметра веса вершины, в результате чего мы получаем алгоритм *BFS Weighted Edges*, который позволяет использовать технологию MapReduce более эффективно при анализе графов.

Литература

- 1) MapReduce [Электронный ресурс]. – Режим доступа: <https://ru.wikipedia.org/wiki/MapReduce>, свободный (дата обращения: 29.10.2018).
- 2) Breadth-firstGraphSearchinMapReduce [Электронный ресурс]. – Режим доступа: <http://puffsun.iteye.com/blog/1905524>, свободный(дата обращения: 29.10.2018).
- 3) Parallelbreadth-firstsearchusinghadoop (IterativeMapReduce) [Электронный ресурс]. – Режим доступа: https://www.cdac.in/index.aspx?id=ev_hpc_bfs-iterative1, свободный (дата обращения: 29.10.2018).
- 4) Parallelbreadth-firstsearchalgorithmsfortreesandgraphs [Электронный ресурс]. – Режим доступа: <https://www.tandfonline.com>, свободный (дата обращения: 29.10.2018).
- 5) Parallel Breadth-First Search on Distributed Memory Systems [Электронныйресурс]. – Режимдоступа: https://www.researchgate.net/publication/51944584_Parallel_Breadth-First_Search_on_Distributed_Memory_Systems, свободный (дата обращения: 29.10.2018).

Диденко А.А., Пимонов А. Г. Использование технологии MapReduce для параллельного поиска в ширину кратчайшего пути в графах большой размерности, алгоритм BFS: weighted edges. Статья посвящена проблеме анализа информации, хранящейся в виде графов, с помощью алгоритма BFS. В статье рассматриваются способы решения данной проблемы.

Ключевые слова: MapReduce, Hadoop, графы, обход графа, поиск кратчайшего пути, BFS, BFS: weighted edges.

Одностраничные приложения (SPA)

Демчук В.Ю.

Министерство связи Донецкой Народной Республики
demchuk.vyu@minsvyazdnr.ru

Демчук В.Ю. Одностраничные приложения (SPA). В статье рассмотрена архитектура одностраничного приложения, озвучены основные преимущества использования компонентного подхода. Описана взаимосвязь виртуального DOM дерева с построением пользовательского интерфейса. SPA приложения весьма распространены за счет своей производительности, высокого уровня интерактивности и обновления данных в режиме реального времени. За счет положительной обратной связи от пользователей, в перспективе, SPA приложения могут частично заменить классические мультистраничные приложения.

Ключевые слова: одностраничное приложение, SPA, компонентный подход, виртуальный DOM, пользовательский интерфейс и взаимодействие.

Введение

При использовании классических мультистраничных веб-сайтов наблюдается ряд существенных проблем:

- обновление данных в режиме реального времени;
- потребность перезагружать всю страницу при изменении одного или нескольких элементов представления;
- ограниченные функции пользовательского интерфейса;
- производительность при рендеринге представления (длительная генерация DOM дерева);
- потребность в специалистах, которые обладают знаниями как в backend, так и во frontend части приложения.

Для повышения скорости получения информации и быстроты реагирования на события, улучшения удобства и простоты пользовательского интерфейса, используется подход одностраничных приложений (SPA).

Одностраничное приложение (англ. single page application, SPA) — это веб-приложение или веб-сайт, использующий единственный HTML-документ как оболочку для всех веб-страниц и организующий взаимодействие с пользователем через динамически подгружаемые HTML, CSS, JavaScript, обычно посредством AJAX [1].

Цель статьи – продемонстрировать возможности SPA приложений, а также показать пути решения проблем, которые возникают в мультистраничных сайтах.

Архитектура SPA приложения

Классическая архитектура SPA приложения представлена на рисунке 1.

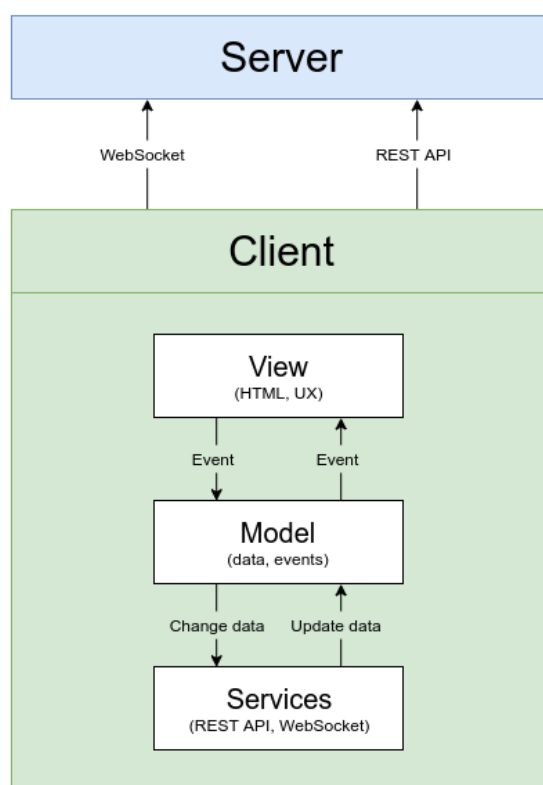


Рисунок 1 – Классическая архитектура SPA приложения.

SPA приложение в основном состоит из представления (view), модели (model) и сервисов (services).

Представление (view) предназначено для отрисовки пользовательского интерфейса. Через представление осуществляется взаимодействие пользователя с функциями системы. Результатом работы представления является обычный HTML. При помощи событий, изменения в представлении сразу же обновляют данные в модели. Представление так же предоставляет пользователю элементы управления для взаимодействия с приложением.

Модель (model) предназначена для хранения данных приложения и работы с событиями. При помощи событий, изменения в модели сразу же обновляют данные в представлении. Модель организует связь с сервером через REST API или WebSocket соединение.

Сервисы (services) предназначены для получения данных через REST API и WebSocket, а также для реализации логики отдельных компонентов приложения.

Для SPA приложений взаимодействие с backend частью в большей степени сводится к использованию REST API и WebSocket соединения. За счет этого, разработчику SPA приложения не требуется реализовывать функционал на backend, и он может полностью сфокусироваться на своей части. Также, процесс создания приложения в целом, может протекать параллельно – backend и frontend части пишутся одновременно.

Компонентный подход

Хорошей практикой при проектировании SPA приложений является использование компонентного подхода. Максимальная декомпозиция частей приложения на компоненты и слабая связанность между ними обеспечивает расширяемость приложения, простоту поддержки и предотвращает повторную реализацию имеющихся функций. Так же, при помощи компонентного подхода становится возможной разработка и тестирование проекта по частям (между разными участниками команды), а не целиком, как при монолитной архитектуре.

В свою очередь компоненты могут быть как простые — содержащие минимум логики и отвечающие за отрисовку конкретных элементов пользовательского интерфейса, так и сложные — содержащие в себе несколько простых компонентов, и обеспечивающие взаимодействие между ними (см. рис. 2).

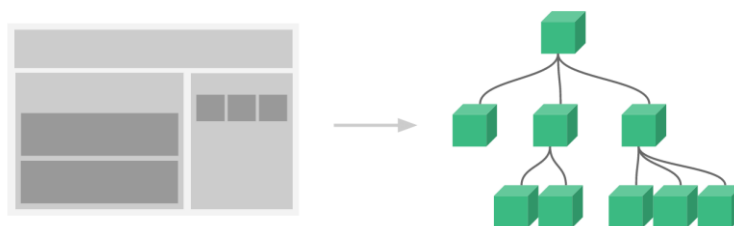


Рисунок 2 – Компонентный подход в SPA приложении.

Основным преимуществом использования компонентного подхода является то, что при изменении данных на одном из элементов страницы, нам не требуется перезагружать всю страницу. Достаточно обновить модель данных конкретного компонента. За счет этого, компонентный подход эффективно решает проблемы обновления данных в режиме реального времени и повышает скорость показа новой информации.

Использование реактивности и виртуальный DOM

Реактивность в приложении – это обновление слоя представления при изменении модели данных. За счет реактивности управление состоянием приложения становится простым и очевидным. Реактивность положительно влияет на удобство пользовательского интерфейса и взаимодействия.

В качестве примера использования реактивности данных для улучшения пользовательского взаимодействия можно рассмотреть прогрессивную загрузку. Прогрессивная загрузка – это способ поэтапной загрузки страницы, за счет которого, у пользователя складывается впечатление о быстроте работы приложения. На начальном этапе рендерится каркас будущих элементов страницы. Это может быть текст или изображение. Пример начального экрана прогрессивной загрузки изображен на рисунке 3.

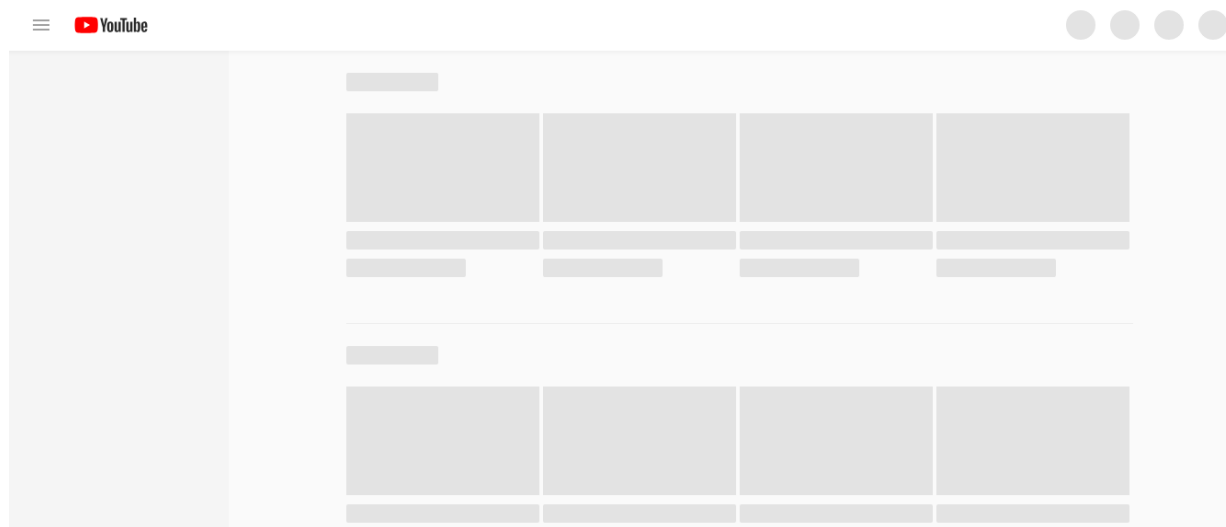


Рисунок 3 – Пример начального экрана прогрессивной загрузки Youtube.

Как только приходит ответ от REST API, представление перерисовывается и экран каркаса заменяется реальными данными. За счет использования прогрессивной загрузки пользователь осведомлен, что приложение запущено, и страница скоро будет обновлена, вследствие чего время ожидания не воспринимается таким долгим.

Во время изменения модели данных, выполняется изменение представления, которое в свою очередь влечет за собой перестраивание виртуального DOM дерева.

Преимущество использования виртуального DOM дерева перед нативным DOM деревом заключается в скорости рендеринга элементов страницы.

DOM (аббревиатура от Document Object Model) — способ представления структурного документа с помощью объектов. Это кроссплатформенное и языко-независимое соглашение для представления и

взаимодействия с данными в HTML, XML и т.д. [2]

При использовании обычного DOM дерева и при наличии большого количества узлов, изменение структуры DOM и дальнейший рендеринг на странице, является весьма медленной операцией. То есть, например, перемещение нескольких тысяч блоков <div> в DOM дереве, будет занимать весьма продолжительное время.

Виртуальный DOM является легковесной копией нативного DOM. При изменении представления, значения сначала отражаются в виртуальном DOM. Затем, происходит сравнение виртуального DOM с нативным, определяется разница и запускается рендеринг той части, которая была изменена.

За счет использования виртуального DOM, представление рендерится намного быстрее, обеспечивая тем самым высокую производительность приложения. Так же, это положительно влияет на пользовательский опыт, ведь анимации обновления частей страницы происходят плавно и быстро.

Примеры SPA приложений

Grafana – дашборд для отображения метрик (рисунок 4)[3]. Frontend часть написана на AngularJs – javascript фреймворк для разработки SPA приложений. Используется для различных целей, в том числе для мониторинга серверов, баз данных, сервисов и служб. Также может использоваться для построения отчетов из базы данных. Ключевой особенностью Grafana является демонстрация данных в реальном времени и система оповещения по различным метрикам.



Рисунок 4 – Grafana.

Google calendar – сервис для планирования мероприятий и создания напоминаний (рисунок 5)[4].

Позволяет создавать и редактировать мероприятия и напоминания без перезагрузки страницы, переключаться между периодами календаря с динамической загрузкой данных. Предусмотрена синхронизация данных между устройствами в режиме реального времени.

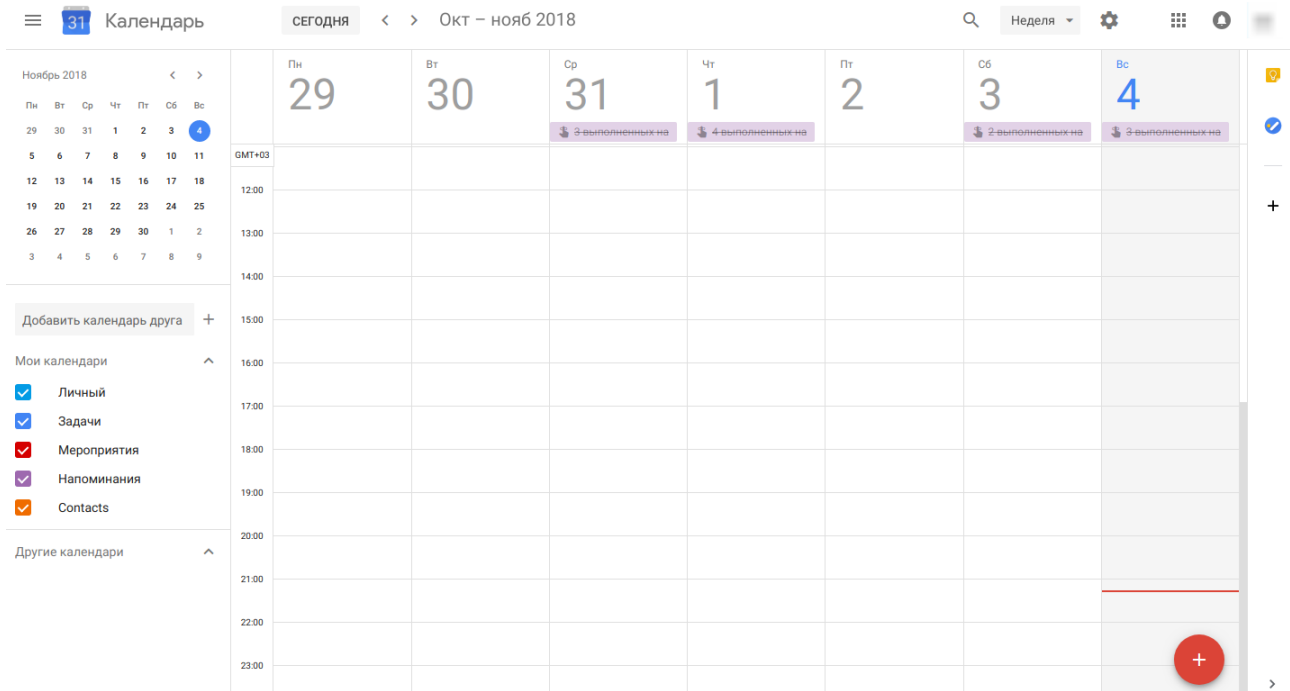


Рисунок 5 – Google calendar.

Trello – сервис для управления проектами (рисунок 6)[5]. Интерфейс построен по принципу досок и карточек, которые обновляются в режиме реального времени. Так же предусмотрена синхронизация данных между устройствами и участниками доски. За счет реактивности элементов достигается обновление информации в режиме реального времени без перезагрузки страницы.

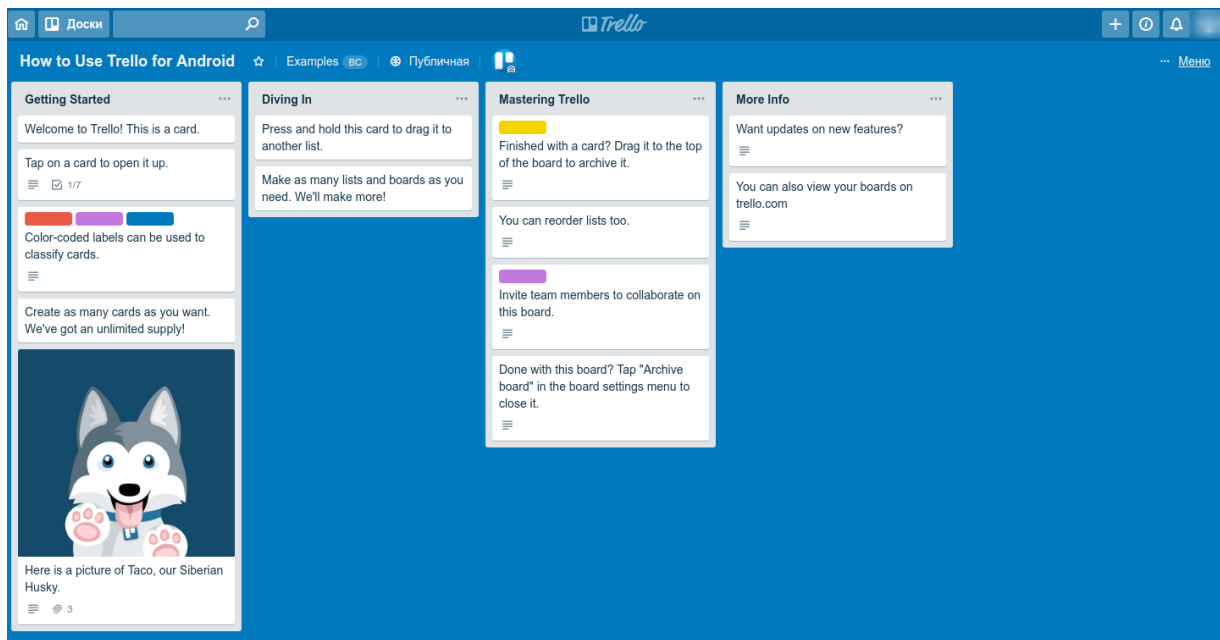


Рисунок 6 – Trello.

Выводы

SPA приложения весьма распространены за счет своей производительности и высокого уровня интерактивности. Они позволяют отслеживать данные в реальном времени, быстро реагировать на события системы и способствуют оперативному принятию решений. Отличительной особенностью SPA

приложений является скорость обновления данных и быстрые переходы между разделами приложения. За счет своего многофункционального интерфейса SPA приложения увеличивают вовлеченность пользователей, что способствует увеличению популярности данного подхода. За счет положительной обратной связи от пользователей, в перспективе, SPA приложения могут частично заменить классические мультистраничные приложения.

Литература

1. Одностраничное приложение [Электронный ресурс] Режим доступа: https://ru.wikipedia.org/wiki/Одностраничное_приложение(дата обращения 04.11.2018)
2. Что такое Virtual DOM? [Электронный ресурс] Режим доступа: <https://habr.com/post/256965>(дата обращения 04.11.2018)
3. Grafana [Электронный ресурс] Режим доступа: <https://grafana.com> (дата обращения 04.11.2018)
4. Google calendar [Электронный ресурс] Режим доступа: <https://calendar.google.com> (дата обращения 04.11.2018)
5. Trello [Электронный ресурс] Режим доступа: <https://trello.com> (дата обращения 04.11.2018)

Демчук В.Ю. Одностраничные приложения (SPA). В статье рассмотрена архитектура одностраничного приложения, озвучены основные преимущества использования компонентного подхода. Описана взаимосвязь виртуального DOM дерева с построением пользовательского интерфейса. SPA приложения весьма распространены за счет своей производительности, высокого уровня интерактивности и обновления данных в режиме реального времени. За счет положительной обратной связи от пользователей, в перспективе, SPA приложения могут частично заменить классические мультистраничные приложения.

Ключевые слова: одностраничное приложение, SPA, компонентный подход, виртуальный DOM, пользовательский интерфейс и взаимодействие.

Demchuk Vitaliy Single page application (SPA). The article describes the architecture of a one-page application, sounded the main advantages of using the component approach. The relationship of the virtual DOM tree with the construction of the user interface is described. SPA applications are quite common due to their performance, high level of interactivity and real-time data updates. Due to the positive feedback from users, in the long run, SPA applications can partially replace the classic multi-page applications.

Key words: single page application, SPA, component approach, virtual DOM, user interface and interaction.

Улучшение прогнозных значений с использованием метода, отбрасывания данных

Рычка О.В.

Донецкий национальный технический университет

olga_rychka@mail.ru

Рычка О.В. Улучшение прогнозных значений с использованием метода, отбрасывания данных. В статье представлен краткий обзор существующих методов прогнозирования. Предложен новый метод (и его модификации) для линейных регрессионных моделей, позволяющий повысить точность прогнозирования. Описаны основные преимущества программного приложения, разработанного для нового метода.

Ключевые слова: прогноз, регрессионный анализ, метод, аномальные данные, программное приложение, точность прогнозирования

Введение

Прогнозирование позволяет подготавливать и принимать обоснованные решения, обеспечивающие эффективность управления, предупреждает возможные срывы и сбои в работе, создавая тем самым условия для нормального функционирования системы. На сегодняшний момент существует огромное количество (около 200) методов и средств прогнозирования. Однако широко применяются около 15-20 методов.

Большинство методов прогнозирования можно разбить на две группы: эвристические методы и формализованные. Эвристические методы прогнозирования включают в себя прямые экспертные оценки и экспертные оценки с обратной связью.

Формализованные методы делятся на:

- методы аналогий;
- опережающие методы;
- статистические методы (экстраполяция, корреляция, дисперсионный анализ, регрессионный анализ, многомерный анализ).
- методы моделирования.

Одним из наиболее применяемых среди статистических методов является прогнозирование с использованием регрессионного анализа. Частным случаем регрессионного анализа является метод линейной регрессии. Он заключается в анализе взаимосвязи двух переменных – влияние вариации факторного показателя X на результирующий показатель Y. Уравнение линейной регрессии имеет вид (1):

$$\hat{Y} = A \cdot x + B \quad (1)$$

Коэффициенты регрессионного уравнения, определяются методом наименьших квадратов.

Для качественного прогноза важное значение имеет точность измерений или исходных данных, на основании которых, осуществляется дальнейшее прогнозирование. Однако существуют данные, которые представляют собой "аномальные" и ненадежные измерения. Аномальным измерением (выбросом), считается наблюдение, выделяющееся по величине и статистическим свойствам из общей выборки. Игнорирование их наличия может привести к неверным результатам прогнозирования.

Существует большое количество различных методов для нахождения и устранения ненадежных измерений [1-5]. К таким методам относятся – методы Эктона, Титьена-Мура-Бекмана, и Прескотта-Лунда. Их основными преимуществами являются простота применения и понимания, а также то, что при определении значения остатков, легко обнаруживается существование аномального значения.

К недостаткам данных методов относятся:

- для применения методов необходимым условием является нормальное распределение случайных величин невязок;
- минимальный объем статистики предполагает более 30 измерений;
- значительное время анализа при большом количестве измерений, поскольку осуществляется полный перебор выборки.

Выявленные недостатки методов Эктона, Титьена-Мура-Бекмана, и Прескотта-Лунда способствовали разработке принципиально новых методов, таких как – метод Кука, методы Белсли-Ку-Уэлша и Аткинсона. Однако методы Белсли-Ку-Уэлша и Аткинсона представляют собой незначительную модификацию метода Кука.

Расстояние Кука – величина, показывающая разницу между вычисленными коэффициентами уравнения регрессии и значениями, которые получились бы при исключении соответствующего наблюдения.

Сущность метода Кука заключается в нахождении при отбрасывании уравнивающих измерений, которые стабилизируют параметры, нового регрессионного уравнения по отношению к исходному. Это приводит к тому, что даже при отбрасывании множества невязок, эффект смещения оценок прогнозного значения, рассчитанного с использованием уравнения регрессии, будет минимальным. В случае применения метода Кука или его модификаций отбрасываются не аномальные невязки, а любые невязки, которые удовлетворяют минимуму критерия Кука.

Благодаря значительному сокращению величины смещения прогнозного значения метод Кука считается достаточно эффективным при ручном анализе. Однако можно выделить ряд недостатков:

- искусственность и непрозрачность при выборе отбрасываемых невязок;
- сложность вычислений (особенно при расчете статистики Кука высокого порядка);
- большое время анализа;
- весомые измерения могут быть отброшены совместно с аномальными и малозначимыми, поскольку основной задачей критерия Кука является минимизация суммарной величины смещения при отбрасывании.

Цель работы. В связи с выявленными выше недостатками существующих методов преобразования исходных данных, необходима разработка новых методов, позволяющих избежать основные недостатки, что приведёт к увеличению качества прогнозов. Простота реализации данных методов позволит применять их в современных компьютерных технологиях.

1 Описание метода, основанного на отбрасывании данных

Смысл метода заключается в том, что среди исходных статистических данных находятся измерения, которые выходят за пределы определенной области. Эти измерения исключаются из выборки.

Работа данного метода, осуществляется по следующему алгоритму:

- 1) по всем исходным статистическим данным, с помощью метода наименьших квадратов, находятся коэффициенты линейного уравнения (1);
- 2) определяются невязки e_i (2), как разница между исходными данными (Y_i) и расчетными (\hat{Y}), т.е. полученными с помощью построенного выше линейного уравнения:

$$e_i = Y_i - \hat{Y} \quad (2)$$

- 3) находится среднеквадратическое отклонение (СКО) (3):

$$\sigma_e = \sqrt{\frac{\sum_{i=1}^n e_i^2}{n-2}} \quad (3)$$

где n – количество элементов в выборке.

- 4) рассчитываются параметры уравнения $\hat{Y}' = A'X + B'$, описывающего линию перпендикулярную линии исходного регрессионного уравнения. Коэффициенты нового регрессионного уравнения, определяются по формулам (4) и (5)

$$A' = -\frac{1}{A} \quad (4)$$

$$B' = \frac{\sum_{i=1}^n Y_i}{n} - A' \cdot \frac{\sum_{i=1}^n X_i}{n} \quad (5)$$

- 5) определяется новое значение невязок e'_i , как разность исходных значений Y_i и \hat{Y}' , рассчитанных по уравнению $\hat{Y}' = A'X + B'$. СКО невязок определяется по формуле (6):

$$\sigma'_e = \sqrt{\frac{\sum_{i=1}^n e_i'^2}{n-2}} \quad (6)$$

- 6) строится прямоугольная область со сторонами $2k \cdot \sigma_e$ и $2k \cdot \sigma'_e$ (рис. 1), где k – коэффициент (обычно $0,6 \leq k < 3$), соответствующий вероятности попадания в заданную область. Вероятность попадания в область рассчитывается следующим образом:

$$P_0 = 2 \cdot \frac{1}{\sqrt{2\pi}} \int_0^k e^{-t^2/2} dt \quad (7)$$

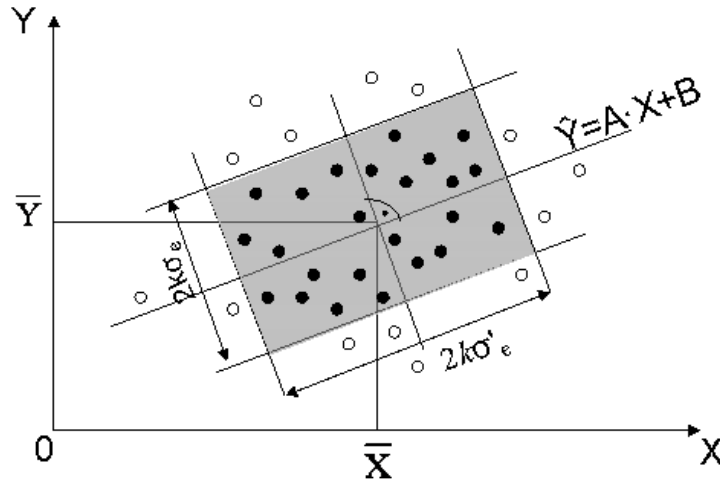


Рисунок 1 – Метод повышения точности регрессионной модели, основанный на отбрасывании "аномальных" и ненадежных данных

Эффективность метода определяется следующими критериями:

1) коэффициентом детерминации R^2 (8). Данный коэффициент определяет силу взаимосвязи между зависимыми переменными Y и независимыми X .

$$R^2 = \frac{\sum_{i=1}^m (\hat{Y}_i - \bar{Y})^2}{\sum_{i=1}^m (Y_i - \bar{Y})^2} \quad (8)$$

где \bar{Y} – математическое ожидание случайной величины Y_i

2) модуль величины смещения результата прогноза (9)

$$\Delta_{\text{прогн}} = |(A \cdot X_{\text{прогн}} + B) - (A_n \cdot X_{\text{прогн}} + B_n)|,$$

где: $A \cdot X_{\text{прогн}} + B$ – линейное регрессионное уравнение, построенное по исходным данным;

$A_n \cdot X_{\text{прогн}} + B_n$ – линейное регрессионное уравнение после отбрасывания части исходных данных.

A, B и A_n, B_n находятся с использованием традиционного метода наименьших квадратов;

3) доверительный интервал прогнозных значений $Y_{\text{прогн}}$.

Используемая статистика невязок имеет распределение отличное от нормального. Усечение аномальных и значительных по величине невязок еще более усугубляет эту проблему. В этой связи использовать квантили Стьюдента ($P_{\text{дов}}; k=n-2$) не корректно. Поэтому использовались доверительные интервалы свободные от закона распределения случайных величин невязок [6].

Пусть $c_1, c_2, c_3, \dots, c_i$ порядковые статистики и $c_1 < c_2 < c_3 \dots c_{i-1} < c_i$. При этом доверительный интервал (c_1, c_i) симметричен относительно своего медианного значения e_k и определяется доверительной вероятностью [7]:

$$P_{\text{дов}} = I_{1/2}(1, i) - I_{1/2}(i, 1), \quad (10)$$

$$P_{\text{дов}} = 1 - 2 \cdot I_{1/2}(b, a), \quad (11)$$

$$I_x(a, b) = \frac{1}{B(a, b)} \int_0^x t^{a-1} \cdot (1-t)^{b-1} dt, \quad (12)$$

$$B(a, b) = \int_0^1 t^{a-1} \cdot (1-t)^{b-1} dt, \quad (13)$$

где: $I_x(a, b)$ – неполная Бета-функция;

$B(a, b)$ – Бета-функция.

Для вычисления доверительного интервала подбирались пары 1, i ; 2, $i-1$; 3, $i-2$ и т.д. симметрично расположенные относительно медианного значения c_k , обеспечивающие по (10-13) заданное значение $P_{\text{дов}}$. Найденная пара и представляет размах доверительного интервала. Естественно, что точное значение величины доверительного интервала таким способом найти не удастся из-за дискретного характера порядковой статистики, но с этим приходится мириться.

4) точность.

Необходимо определить до каких пор следует уменьшать размеры прямоугольника, внутри которого исходные статистические данные считаются надежными, т.к. неограниченное уменьшение сторон этого прямоугольника может привести к абсурдному научному результату. Для этого определяется точность, которая рассчитывается по формуле (14):

$$T = R^2 \cdot \frac{m}{n} \quad (14)$$

где n – количество исходных данных;

m – количество данных оставшихся после отбрасывания.

При этом наилучшим вариантом, считается вариант, при котором величина коэффициента детерминации R^2 является максимальной, при обязательном условии, что $T \geq 0.5$;

5) количество элементарных операций ЭВМ, которое необходимо для реализации отбрасывания аномальных и ненадежных измерений. Определяется по формуле (15):

$$K = 26 \cdot n + 9 \cdot m + 7 \quad (15)$$

где n – количество исходных данных;

m – количество данных, оставшихся после отбрасывания.

2 Модификации метода отбрасывания данных

Для определенных случаев были разработаны две модификации, описанного выше метода.

Первую модификацию можно использовать, если существует уверенность в исходных значениях независимой переменной X (например, значения x_i известны заранее и в них не может быть случайных ошибок), Она состоит в том, что область, в пределах которой, располагаются "надежные" данные представляет собой не прямоугольник, а так называемый «коридор», пределы которого равноудалены от линии исходного регрессионного уравнения. Расстояние между границами данного «коридора» составляет $2k\sigma_e$ (рис. 2). Данная модификация схематически изображена на рисунке 2.4.

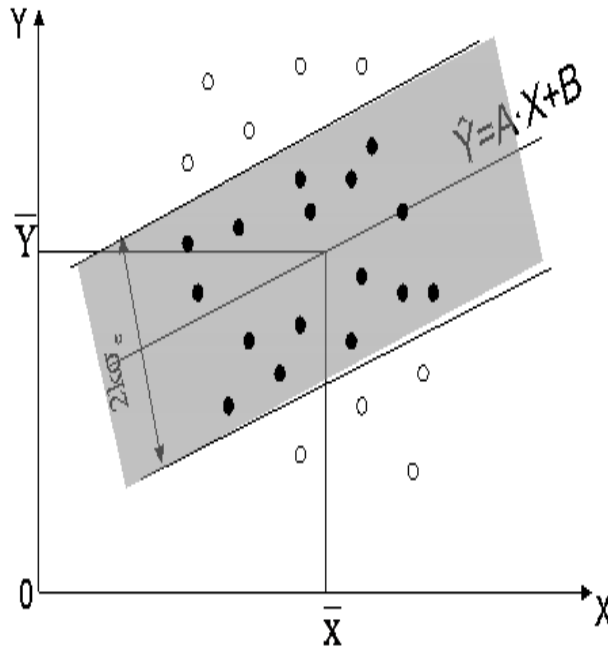


Рисунок 2 – Первая модификация метода отбрасывания данных

Количество элементарных операций в данном случае будет значительно меньше, чем при использовании основного метода (16):

$$K = 16 \cdot n + 9 \cdot m + 1 \quad (16)$$

Вторая модификация метода позволяет отбрасывать данные, выходящие за пределы надёжности, по значениям переменной X . В отличие от первой модификации метода, в данном случае, определяются 2 линии параллельные, найденному уравнению перпендикуляр (рис. 3). Расстояние между границами «коридора», в этом случае, составляет $2k\sigma'_e$.

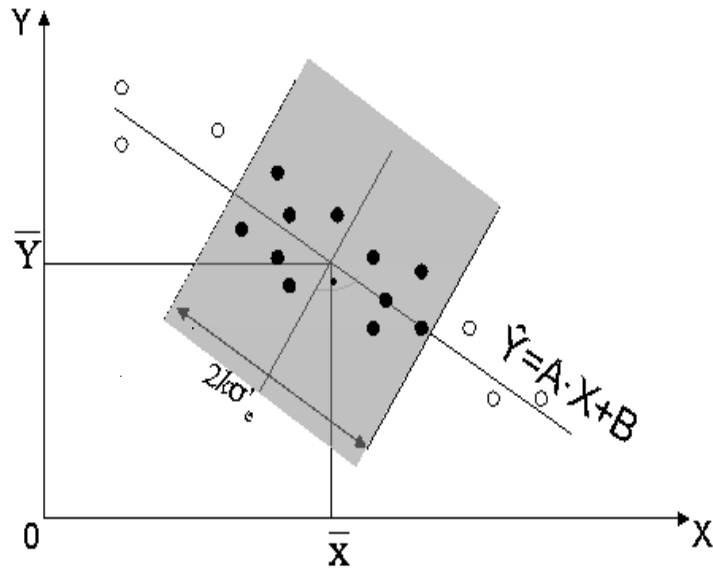


Рисунок 3 – Вторая модификация метода отбрасывания данных

Количество элементарных операций для данной модификации, больше чем для первой модификации, но меньше чем в основном (17):

$$K = 19 \cdot n + 9 \cdot m + 4 \quad (17)$$

В некоторых случаях, при применении второй модификации, величина коэффициента детерминации R^2 уменьшается. Это объясняется тем, что крайние точки являются определяющими для уравнения регрессии, и если они удалены на значительное расстояние от остальных точек, то оказывают большое влияние на вид исходного уравнения.

Применение описанных выше модификаций проще и быстрее, чем реализация основного метода, однако использовать их следует, только в конкретных случаях, описанных выше.

3 Программная реализация метода и его модификаций

Для простоты применения, изложенного выше метода и его модификаций, было разработано программное приложение с использованием языка программирования VisualBasicforApplications для MicrosoftExcel.

Данное приложение имеет ряд достоинств:

- 1) простота использования;
- 2) небольшое время обработки данных;
- 3) не требует дополнительных ресурсов памяти компьютера, поскольку не нужна дополнительная установка программного обеспечения;
- 4) объём программного приложения не превышает 1 МБ;
- 5) приложение позволяет осуществлять расчеты для различного количества исходных данных;
- 6) значение Xпрогн можно устанавливать произвольно, по желанию пользователя;
- 7) существует возможность выбора метода или одной из его модификаций (рис.4);

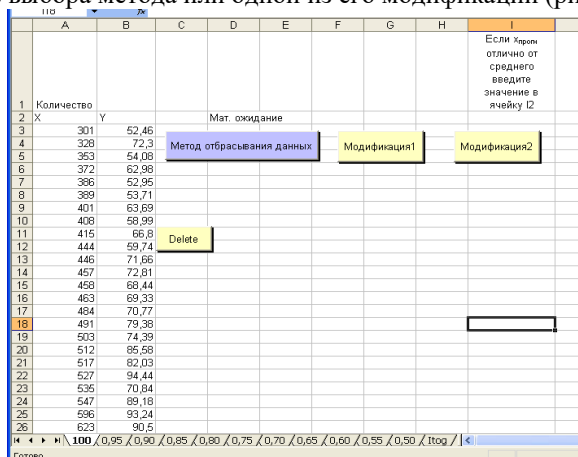


Рисунок 4 – Выбор метода отбрасывания данных

8) возможность просмотра вспомогательных значений (рис. 5)

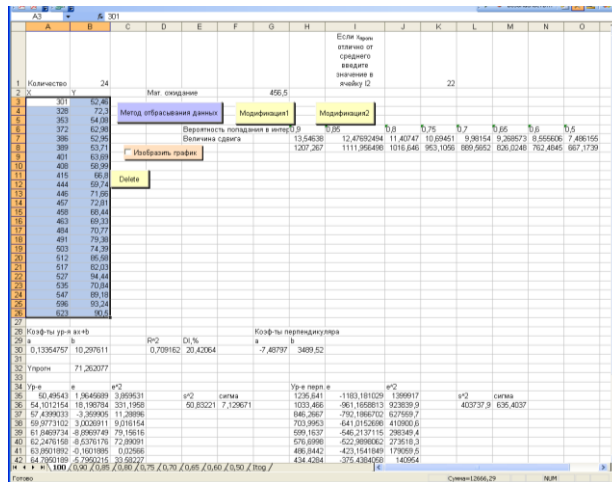


Рисунок 5 – Отображение результатов, после выбора метода

- 9) вывод результатов расчетов для каждой вероятности на отдельных листах;
- 10) расширение стандартных возможностей MSExcel путём добавления возможности построения на графике уравнения перпендикулярного исходному;
- 11) наглядное изображение на графиках соответствующих областей с надёжными данными, при этом ненадёжные значения, не попадающие в заданную область, окрашиваются другим цветом (рис.6)

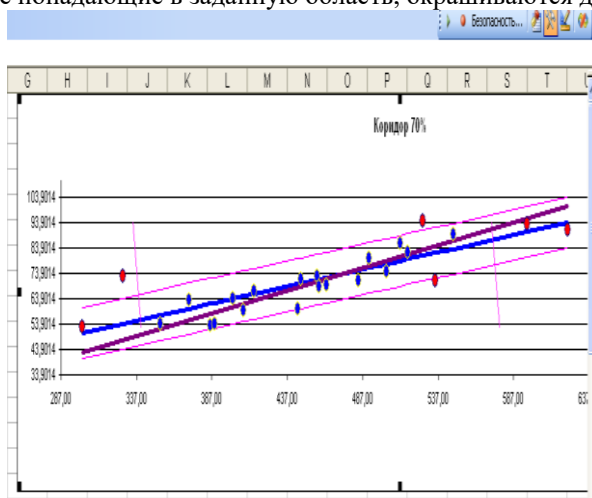


Рисунок 6 – Графическое изображения области попадания надёжных измерений

12) формирование итоговой сводной таблицы результатов.

Выводы

В результате рассмотрения наиболее известных методов нахождения аномальным данных было выявлено, что данные методы являются трудоемкими, т.к. при их использовании нахождение аномальных точек происходит методом перебора.

Главным недостатком методов отбраковки аномальных результатов измерений при непрерывном наблюдении является сложность их применения для решения большинства практических задач. Поэтому был предложен и проанализирован новый метод, позволяющий улучшить качество прогноза, полученного с использованием регрессионных моделей. Данный метод заключается в нахождении и последующем исключении из выборки аномальных и ненадежных измерений. Также для нового метода, предлагается два упрощения. Были разработаны критерии оценки эффективности метода и его модификаций. На основе проведенного анализа можно сделать следующие выводы: 1. Использование предложенного методов повышения точности прогнозирования приводит к увеличению коэффициента детерминации R^2 . Рост коэффициента детерминации может достигать 40% при использовании

метода на основе отбрасывания данных. Это происходит благодаря обнаружению и устранению ненадежных измерений.

3. Так невязки e_i не подчиняются нормальному закону распределения, использовались доверительные интервалы свободные от закона распределения.

4. Предложенный в данной работе метод и его модификации можно применять для нелинейных регрессионных прогнозных уравнений с внутренней нелинейностью. При этом исходное нелинейное регрессионное уравнение путем специальных преобразований приводится к линейному виду. Производится отбрасывание части статистики и далее путем обратного преобразования возвращаются к исходному нелинейному уравнению.

5. Разработанное программное приложение позволяет существенно сократить время анализа, а также наглядно отображает все результаты, проводимого анализа.

Литература

1. Дрейпер Н.Р., Смит Г. Прикладной регрессионный анализ. 3-е изд.: Пер. с англ. – М.: Вильямс, 2007. – 912 с.
2. Кобзарь А.И. Прикладная математическая статистика. Для инженеров и научных работников. – М.: ФИЗМАТЛИТ, 2006. – 816 с.
3. Лемешко Б. Ю, Лемешко С. Б. Расширение области применения критериев типа Граббса, используемых при отбраковке аномальных измерений// Измерительная техника. – 2005. – № 6 – С.13-20
4. Rawlings, John O. Applied regression analysis: a research tool. — 2nd ed. / John O. Rawlings, Sastry G. Pentula, David A. Dickey – USA.: Springer, 1998.
5. Мудров В.И., Кушко В.Л. Методы обработки измерений: Квазиправдоподобные оценки. – Изд. 2-е, перераб. и доп. – М.: Радио и связь, 1983. – 304 с.
6. Ллойд Э. Справочник по прикладной статистике. В 2-х т. Т.1: Пер. с англ./ Под ред. Э. Ллойда, У. Ледермана, Ю.Н. Тюрина. – М.: Финансы и статистика, 1989. – 510 с.
7. Абрамовиц М. Справочник по специальным функциям с формулами, графиками и математическими таблицами./ Под. ред. М.Абрамовица и Н. Стигана: Пер. с англ. под ред. В.А. Диткина и Л.И. Кармазиной. – М.: Наука, 1979. – 830 с.

Рычка О.В. Улучшение прогнозных значений с использование метода, отбрасывания данных. В статье представлен краткий обзор существующих методов прогнозирования. Предложен новый метод (и его модификации) для линейных регрессионных моделей, позволяющий повысить точность прогнозирования. Описаны основные преимущества программного приложения, разработанного для нового метода.

Ключевые слова: прогноз, регрессионный анализ, метод, аномальные данные, программное приложение, точность прогнозирования

Rycka Olga. The improvement of forecasting values with use the method of removing data. The article provides a brief overview of existing forecasting methods. A new method (and its modifications) for linear regression models is proposed to improve the prediction accuracy. The main advantages of the software application developed for the new method are described.

Key words: forecast, regression analysis, method, anomalous data, software application, forecasting accuracy

Система тестирования прикладного программного обеспечения для образовательного проекта “Электронный учебник Lang-book”

Шевелёв И.А., Федяев О.И., Коломойцева И.А.
Донецкий национальный технический университет
rayan-cooper@mail.ru, fedyayev@donntu.org, bolatiger@mail.ru

Шевелёв И.А., Федяев О.И., Коломойцева И.А. Система тестирования прикладного программного обеспечения для образовательного проекта “Электронный учебник Lang-Book”. В работе рассматриваются особенности процессов ручного и автоматизированного тестирования и автоматизации образовательного проекта для московских школ, который предоставит возможность улучшить систему образования в России, помочь абитуриентам успешно подготовиться к ЕГЭ, поступить в ВУЗ. Анализ проекта даёт возможность составить и поддерживать процессы комплексного тестирования проекта и обеспечить качество тестирования за счёт покрытия интерфейса автоматизированными программными тестами. Для решения этой задачи используется анализ существующих методов ручного тестирования и автоматизации тестирования пользовательских интерфейсов, средствами Selenium, Jasmine, Protractor, которые используют язык JavaScript.

Ключевые слова: образовательный проект, электронный учебник, автоматизация тестирования

Введение

Развитие информационного и программного обеспечения (ПО) изменяет способ приобретения и распространения знаний. В качестве средства передачи информации и обучения в системе образования применяется специальное обучающее ПО. IT-технологии являются инструментами, которые помогают школьнику расширить возможности памяти, способность решать проблемы и сложные логически задачи. Она позволяет школьнику активно обучаться и в значительной мере компенсировать дефицит общения с преподавателем. Компьютерная обучающая программа как программное средство учебного назначения можно представить в качестве системы, состоящей из двух подсистем:

- контента (материалы системы);
- программного компонента [1].

Чтобы получить качественную образовательную систему необходимо тщательно её протестировать, начиная от анализа бизнес-идей, заканчивая проверкой правильности решения задач после стадии разработки. Тестирование программного обеспечения является большой IT-отраслью, которая позволяет поддерживать качество программного продукта на высоком уровне и не допускать наличия дефектов на этапе пользования продуктом конечными пользователями. Само тестирование включает в себя:

- процесс исследования ПО с целью получения информации о качестве продукта;
- процесс проверки соответствия заявленных к продукту требований и реально реализованной функциональности, осуществляемый путем наблюдения за его работой в искусственно созданных ситуациях и на ограниченном наборе тестов, выбранных определенным образом;
- оценку системы с тем, чтобы найти различия между тем, какой система должна быть и какой она есть.

В широком смысле, тестирование – это одна из техник контроля качества (Quality Control), которая включает планирование, составление тестов, непосредственно выполнение тестирования и анализ полученных результатов.

Тестирование ПО включает не только собственно проведение тестов, но и многие другие действия, связанные с процессом обеспечения качества:

- анализ и планирование;
- разработку тестовых сценариев;
- оценку критериев окончания тестирования;
- написание отчетов;
- формальные инспекции документации (в том числе и исходного кода);
- проведение статического анализа.

Известные компании и коммерческие системы тестирования

Известно, что для каждого проекта необходимо подобрать индивидуальные процессы и инструменты для тестирования, чтобы удовлетворить требования заказчиков. Эти процессы могут кардинально различаться в зависимости от выбранной методологии команды разработчиков или личных предпочтений заказчика.

Существуют несколько коммерческих универсальных зарубежных компаний, которые предоставляют свои услуги в области тестирования веб-сайтов, оконных и мобильных приложений. Среди наиболее интересных стоит отметить следующих представителей:

– AlpanaQualityServices – компания, которая предоставляет услуги внешнего тестирования продуктов методами чёрного ящика, написания документации, тест-планов и тест-кейсов, автоматизации интерфейса приложения, своей технической поддержки;

– QATestLab – независимое объединение тестировщиков со всего мира, которое предоставляет услуги тестирования приложений, анализа кода методами белого ящика, подсчёта показателей тестирования и его окупаемости;

– QAWerk – компания, которая занимается тестированием веб-приложений, составляет техническую документацию, по которой пишутся автоматизированные тесты. Под каждый проект подбирается отдельная команда специалистов с индивидуальной специализацией.

Постановка задачи разработки системы тестирования

Несмотря на то, что коммерческие услуги тестирования предлагают большой спектр услуг по обеспечению качества программного обеспечения, любому заказчику дешевле и продуктивнее нанять отдельную QA-команду в свой штат сотрудников, которые будут обеспечивать качество с вниманием только на один продукт, а не на многие, как в отдельных QA-компаниях.

Чтобы обеспечить максимальное качество продукту необходимо:

– знать продукт, понимать его бизнес-цель и иметь представления о планах его развития на ближайший квартал-год;

– составить полный тест-план работ на ближайший промежуток времени (например, квартал);

– составлять тестовую документацию, чек-листы, тест-кейсы по критическому функционалу;

– проводить постоянное регрессионное тестирование продукта силами ручных тестировщиков или заскриптованными авто-тестами;

– общаться с заказчиками, понимать их проблемы и уметь переводить их в технический язык;

– всегда иметь качественную тестовую инфраструктуру (например, окружение, максимально приближенное к пользовательскому);

– покрывать функционал автоматизированными тестами для облегчения работы ручных тестировщиков;

– составлять списки уязвимых мест продукта, чтобы следить за ними после каких-либо правок со стороны разработки.

Таким образом, чтобы обеспечить в конкретном программном продукте все вышеуказанные признаки качества, нанимают отдельных людей: ручных тестировщиков, автоматизаторов и тест-менеджеров. Следовательно задача тестирования сложного ПО для образовательного проекта “Электронный учебник Lang-book” является актуальной.

Выбор архитектуры автоматизированной системы тестирования проекта

Процессы тестирования рассмотрены на гибких моделях разработки программного обеспечения, такой как Scrum и Kanban [2]. В Kanban над задачей может работать несколько узкопрофильных команд. К примеру, сначала работает отдел аналитики, потом дизайнеры рисуют интерфейсы задачи, а на третьем этапе включаются разработчики и тестировщики. В Kanban внутри команды нет ролей. Каждый разработчик может быть ответственным за любую задачу.

В Scrum над проектом работает одна универсальная команда. В ней столько разноплановых специалистов, сколько нужно для решения любой задачи проекта. Поскольку команда самоорганизуется, у специалистов scrum-команды нет формальной компетенции. Когда необходимо, тестировщик помогает дизайнеру, а аналитик – разработчику. В scrum-команде помимо собственно специалистов есть две роли: scrum-мастер и владелец продукта.

Scrum-мастер – человек, который организует работу. Это не управленческая должность, и он не раздает указания. Его задачи:

– вести собрания;

– устранять препятствия в работе (если команде мешает перфоратор в соседнем офисе, мастер ищет выход);

– замечать и вытаскивать на поверхность скрытые проблемы;

- отвечать за соблюдение методологии;
- следить за статусом задач.

В свободное от этих задач время скрам-мастер работает так же, как другие члены команды.

Владелец продукта определяет ход проекта, он может представлять внешнего заказчика. Владелец знает все о рынке и целевой аудитории, расставляет приоритеты задачам. Примеры каждой методологии представлены на рисунках 1 и 2 соответственно.

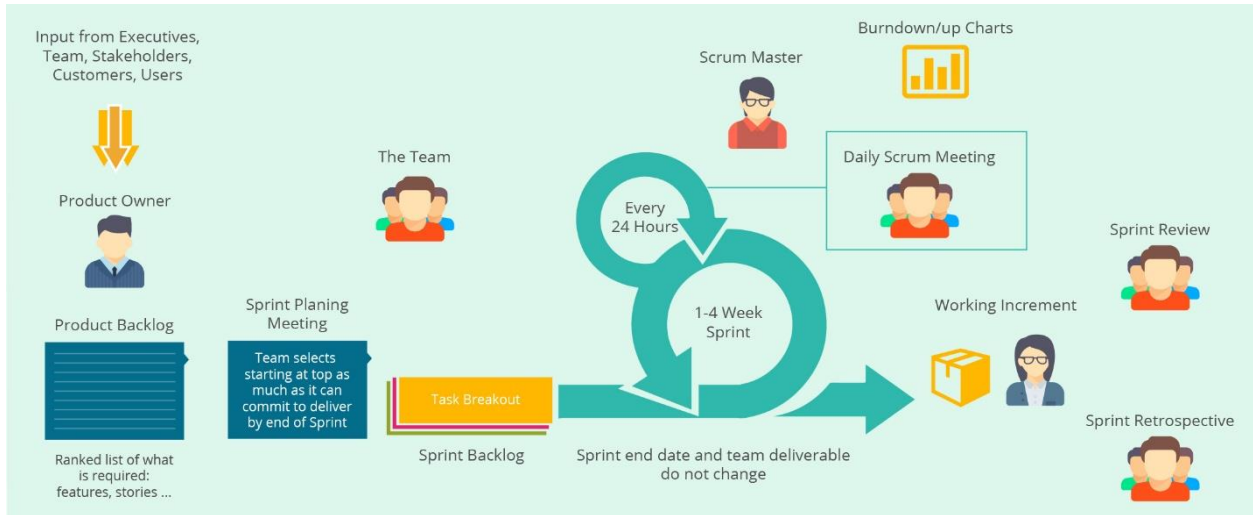


Рисунок 1 – Методология разработки Scrum

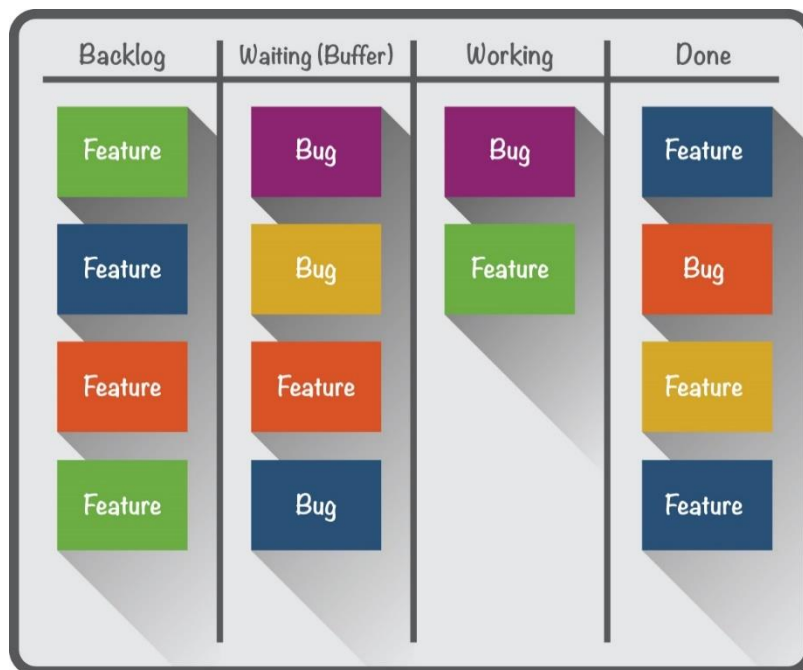


Рисунок 2 – Методология разработки Kanban

Из анализа гибких методологий разработки было установлено, что в команде тестировщиков (QA-команда) сотрудники делятся на 3 вида:

– ручной тестировщик, который выполняет тест-кейсы вручную без использования средств автоматизации. Несмотря на то что это звучит очень просто, от тестировщика в те или иные моменты времени требуются такие качества, как терпеливость, наблюдательность, креативность, умение ставить нестандартные эксперименты, а также умение видеть и понимать, что происходит «внутри системы», т.е. как внешние воздействия на приложение трансформируются в его внутренние процессы. Обычно тестировщики помимо выполнения тест-кейсов пишут их, тестовую документацию, общаются с разработчиками по задачам;

– автоматизатор, который тестовую документацию в виде тест-кейсов преобразовывает в сценарии использования продукта программным кодом при помощи соответствующих средств разработки на основе BDD (разработка на основе поведений);

– тест-менеджер, обычно он же руководитель отдела QA, который распределяет нагрузку в команде, общается с заказчиками и руководителями разработки, составляет тест-планы.

Наиболее распространенная форма автоматизации тестирования – это тестирование приложений через графический пользовательский интерфейс. Популярность такого вида тестирования объясняется двумя факторами: во-первых, приложение тестируется тем же способом, которым его будет использовать человек, во-вторых, можно тестировать приложение, не имея при этом доступа к исходному коду.

В качестве средства автоматизации тестирования выступает SeleniumWebdriver – драйвер браузера, то есть программную библиотеку, которая позволяет разрабатывать программы, управляющие поведением браузера. По своей сущности Selenium WebDriver представляет собой: спецификацию программного интерфейса для управления браузером, референсные реализации этого интерфейса для нескольких браузеров, набор клиентских библиотек для этого интерфейса на нескольких языках программирования.

Драйвер – не имеющая пользовательского интерфейса программная библиотека, которая позволяет различным другим программам взаимодействовать с браузером, управлять его поведением, получать от браузера какие-то данные и заставлять браузер выполнять какие-то команды [3].

Фреймворк для автоматизации тестирования является Jasmine, который использует веб-драйвер Selenium. Jasmine – это BDD фреймворк для тестирования JavaScript кода.

Главной особенностью проекта Lang-book является то, что он написан на Angular 5 – одна из самых популярных библиотек для создания сложных Frontend-приложений. Для автоматизации интерфейсов, которые написаны на Angular используется библиотека автоматизации тестирования Protractor – end-to-end фреймворк для автоматизации Angular-приложений, очень широко распространёна в бизнес-ИТ направлении. Protractor довольно легко портируется на другие языки – он использует небольшое подмножество протокола JsonWire на котором основан SeleniumWebDriver, а именно всего один интерфейс. Принцип работы автоматизированных тестов представлен на рис. 3 [4].

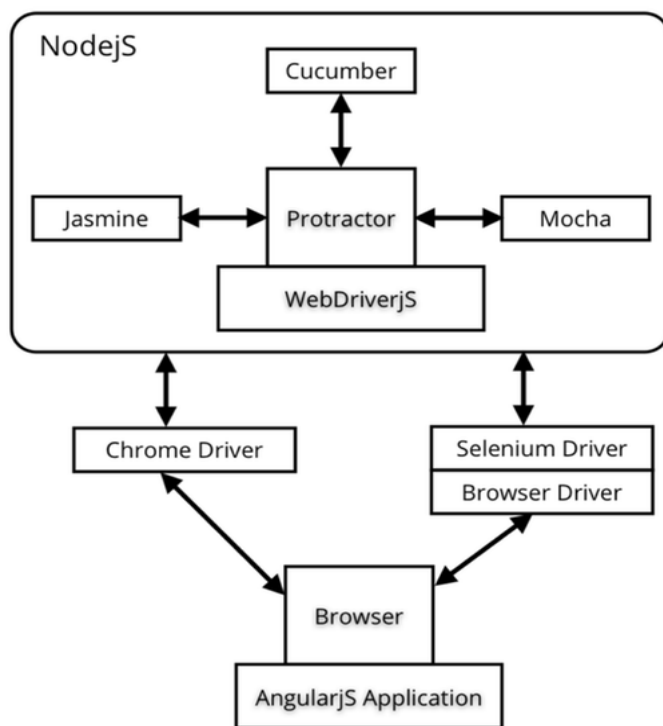


Рисунок 3 – Структура работа автоматизированных тестов

В качестве шаблона проектирования используется Pageobject. Это шаблон проектирования, который широко используется в автоматизированном тестировании и позволяет разделять логику выполнения тестов от их реализации. PageObject как бы моделирует страницы тестируемого приложения в качестве объектов в коде. В результате его использования получаются отдельные классы, отвечающие за работу с DOM-элементами каждой конкретной веб-страницы. Такой подход значительно уменьшает объем повторяющегося кода, потому что одни и те же объекты страниц можно использовать в различных тестах.

Основное преимущество PageObject заключается в том, что в случае изменения пользовательского интерфейса, можно выполнить исправление только в одном месте, а не исправлять каждый тест, в котором этот интерфейс используется.

Класс PageObject не обязательно должен представлять собой всю страницу. Он может быть частью страницы, которая часто используется на сайте (или даже на одной странице). Основной принцип состоит в том, что есть только один класс в проекте, который знает о структуре HTML конкретной страницы или ее части [5].

Существует большая разница между логикой тестирования (что проверить) и его реализацией (как проверить). Пример тестового сценария: пользователь вводит неверный логин или пароль, нажимает кнопку входа, получает сообщение об ошибке. Этот сценарий описывает логику теста, в то время как реализация содержит в себе такие действия как поиск полей ввода на странице, их заполнение, проверку полученной ошибки. И если, например, изменится способ вывода сообщения об ошибке, то это никак не повлияет на сценарий теста, все также нужно будет ввести неверные данные, нажать кнопку входа и проверить ошибку. Но это напрямую затронет реализацию теста – необходимо будет изменить метод получающий и обрабатывающий сообщение об ошибке. При разделении логики теста от его реализации авто-тесты становятся более гибкими и их, как правило, легче поддерживать.

PAGE OBJECT MODEL

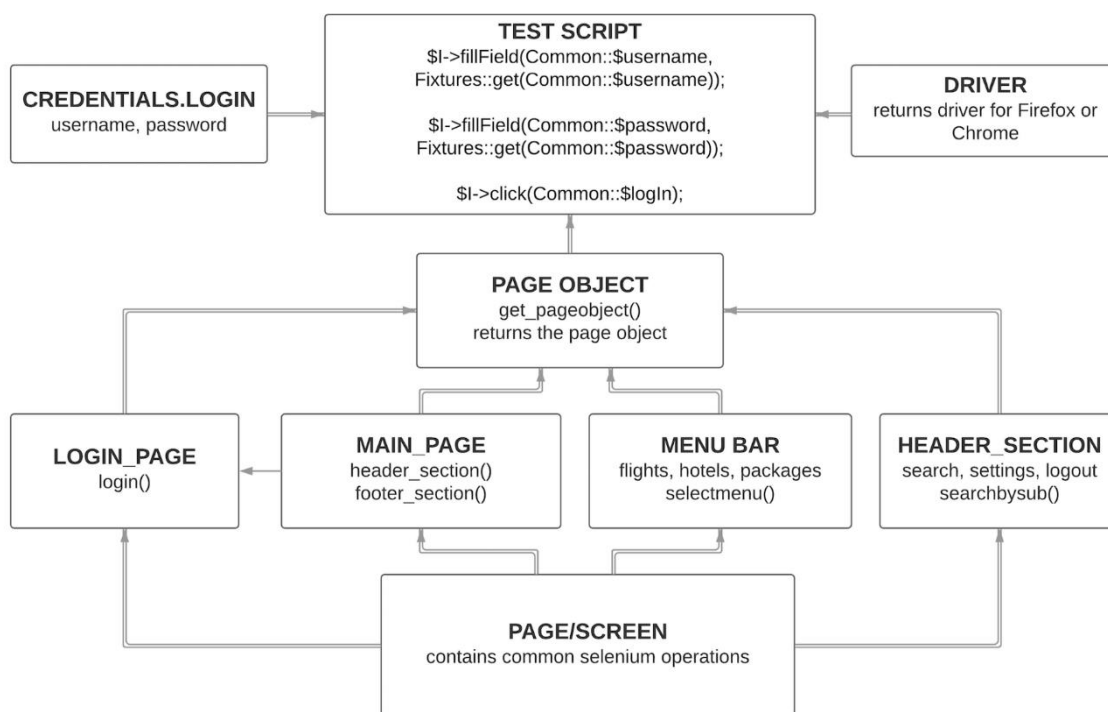


Рисунок 4 – Шаблон проектирования Pageobject

Автоматизация тестирования обучающей системы

Программная реализация заключалась только в покрытии автоматизированными тестами пользовательских сценариев использования проекта Lang-Book.

Для программной реализации выбранного процесса тестирования данной системы использованы следующие инструментальные средства: JavaScript, Xpath, CSS, SeleniumWebDriver, Jasmine, Protractor, NodeJS, Yarn, REST API Requestnpmframework.

JavaScript – язык программирования, который поддерживает объектно-ориентированный, императивный и функциональный стили. Является реализацией языка ECMAScript.

Jasmine – это BDD фреймворк для тестирования JavaScript кода.

Xpath – язык запросов к элементам XML-документа. Разработан для организации доступа к частям документа XML в файлах трансформации XSLT. XPath призван реализовать навигацию по DOM в XML. В XPath используется компактный синтаксис, отличный от принятого в XML.

CSS – каскадные таблицы стилей, используются для описания внешнего вида документа, написанного языком разметки.

SeleniumWebDriver – это драйвер браузера, который позволяет разрабатывать программы, управляющие поведением браузера.

Protractor – e2e тест-фреймворк сделанный на основе WebDriverJS, который приходит на смену первоначальным e2e тестам AngularJS.

Node.js – это среда выполнения JavaScript, построенная на JavaScript-движке V8 из Chrome.

Yarn – это менеджер пакетов. Его целью является решение целого ряда проблем, с которыми столкнулись разработчики при использовании npm, а именно: установка пакетов не была достаточно быстрой и последовательной; существовали проблемы с безопасностью, так как npm позволяет пакетам запускать код при установке.

Заключение

При анализе актуальных вопросов разработки и создания системы тестирования прикладного программного обеспечения для образовательного проекта “Электронный учебник Lang-book” были рассмотрены вопросы использования современных интернет-технологий в разработке автоматизированных тестов на языке программирования JavaScript. В качестве рассматриваемой методологии разработки была выбрана гибкая методология разработки Agile. На её основе были составлены роли для QA-команды, чтобы соответствовать требованиям методологии разработки. Были выбраны инструментальные средства для автоматизации пользовательских интерфейсов.

Литература

1. Технология создания компьютерных обучающих программ - Международный журнал экспериментального образования – Режим доступа: <https://www.expeducation.ru/ru/article/view?id=6702> - Загл. с экрана;
2. Разбираемся в Scrum и Kanban – Режим доступа: <https://netology.ru/blog/scrum-kanban> - Загл. с экрана;
3. Введение – Selenium – автоматизация веб-приложений – Режим доступа: <https://selenium2.ru/docs/introduction.html> - Загл. с экрана;
4. Скринкаст по Angular – Режим доступа: <https://learn.javascript.ru/screencast/angular> - Загл. с экрана;
5. Selenium и Page Object паттерн — Блог вебразработчика – Режим доступа: <http://internetka.in.ua/selenium-page-object/> - Загл. с экрана.

Шевелёв И.А., Федяев О.И., Коломойцева И.А. Система тестирования прикладного программного обеспечения для образовательного проекта “Электронный учебник Lang-Book”. В работе рассматриваются особенности процессов тестирования и автоматизации образовательного проекта для московских школ, который предоставит возможность улучшить систему образования в России, помочь абитуриентам успешно подготовиться к ЕГЭ, поступить в ВУЗ. Анализ проекта даёт возможность составить и поддерживать процессы комплексного тестирования проекта и обеспечить качество тестирования за счёт покрытия интерфейса автоматизированными программными тестами. Для решения этой задачи используется анализ существующих процессов ручного тестирования и автоматизации пользовательских интерфейсов, средства для автоматизации, как Selenium, Jasmine, Protractor, которые будут использованы на языке JavaScript.

Ключевые слова: образовательный проект, электронный учебник, автоматизация тестирования

Shevelov I.A., Fedyaev O.I., Kolomoitseva I.A. System testing application software for the educational project "Electronic Lang-Book". The paper discusses the following processes of testing and automating an educational project for Moscow schools, providing an opportunity to improve the education system in Russia, help applicants to successfully prepare for the USE, and enter the university. Project analysis: creation and support of integrated project testing processes and quality assurance of testing due to interface coverage with automated software tests. To solve this problem, we use analysis of processes like Selenium, Jasmine, Protractor, which will be used in the JavaScript language.

Keywords: educational project, electronic textbook, test automation

Анализ изменений функциональности сайта для сопровождения конференции ПИИВС-2018

Филипишин Д.А., Серёженко О.А., Коломойцева И.А., Павлюк Е.Н.

Донецкий национальный технический университет

domaco@mail.ru, seryozhenkoolga94@gmail.com, bolatiger@mail.ru, luxury.l@yandex.ru

Филипишин Д.А., Серёженко О.А., Коломойцева И.А., Павлюк Е.Н. Анализ изменений функциональности сайта для сопровождения конференции ПИИВС-2018. В работе рассматриваются вопросы использования каждого из модулей, а также второй глобальной реконструкции и рефакторинга сайта для сопровождения конференции ПИИВС-2018.

Ключевые слова: Joomla, веб-сайт, электронное хранилище, конференция, научный доклад.

Введение

Целью данной работы является описание функциональности и проведение рефакторинга модулей сайта для сопровождения конференции ПИИВС-2018.

Актуальность работы: в ходе проведения конференции, у преподавателей недостаточно много времени на разбор функциональности и назначения всех возможностей сайта, что создаёт трудности и ошибки при подаче доклада. Последующая работа обеспечит удобное и быстрое добавление статей на сайт конференции с помощью веб форм с возможностью подачи доклада через мобильное устройство.

Краткий обзор общего назначения сайта и подготовка к проведению конференции

Сайт для сопровождения конференции ПИИВС-2018 [2, 3, 4] предназначен для удобного удалённого сбора докладов с последующим формированием научных сборников. Место для хостинга предоставляет официальный сайт Донецкого национального технического университета [1]. Управление сайтом разделено на такие роли: администратор, секретарь, рецензент и автор.

Администратор полностью управляет сайтом, может менять данные о докладах в режиме реального времени, а также введённые данные при регистрации, за исключением просмотра паролей зарегистрированных пользователей, которые хранятся в базе в зашифрованном виде и исключают возможность подсмотреть пароль. Администратор создаёт конференцию, секции конференции и студенческих секций, сопровождает конференцию.

Секретарь просматривает новые доклады, отправляет на доработку, назначив ему статус «исправить» в случае, если они не соответствуют требованиям шаблона, или назначает докладу рецензента, который рецензирует доклад, пока ему не будет назначен один из статусов «принят» или «не принят». Секретарь имеет возможность лично рецензировать доклад, в отличие от рецензента, ему доступны все доклады конференции.

Рецензент имеет возможность рецензировать предварительно назначенный ему доклад, а также как и все пользователи сайта вне зависимости от роли, подавать доклад в роли автора. В случае наличия мелких правок рецензент может, предварительно назначив докладу статус «принят», лично исправить ошибки и заменить файл доклада без уведомления авторов.

Автор доклада имеет право подавать доклад как на конференцию ПИИВС-2018, так и на студенческие секции. Количество докладов не ограничено.

Первая и вторая (текущая) версии сайта требовали предварительной регистрации каждого участвующего автора, после окончательного рефакторинга планируется внедрить систему ввода данных участвующих авторов без обязательной предварительной регистрации. Такой подход обусловлен тем, что оргкомитету конференции требуется информация обо всех участниках, будут ли они участвовать лично (очно/заочное участие), а также их персональные данные для составления программы проведения конференции. Также стоит отметить, что ранее указанный предварительно не зарегистрированный автор, в случае последующей регистрации и подачи других докладов, должен иметь возможность управлять и просматривать все доклады, в которых он указан, что вносит трудности и путаницу в реализацию данного модуля и выносятся на реализацию третьей глобальной модификации движка сайта.

При второй глобальной модификации сайта были полностью переписаны следующие модули: фотогалерея, новости, меню управления с добавлением ссылки на справочное руководство текущей роли

пользователя, упрощена регистрация пользователя (на один e-mail можно регистрировать несколько пользователей), система отчётов.

Подготовка конференции к проведению осуществляется со страницы управления ... (см. рис. 1). Система рассматривает студенческие секции, как отдельную конференцию, которой можно управлять независимо от основной. На этой странице администратор может запретить подачу докладов, что позволит авторам только просматривать свои доклады в текущем их состоянии. Также можно скрыть для просмотра конференцию целиком, в таком случае авторам сообщено, что конференция недоступна либо ещё не началась.

Краткое название	Полное название	Дата проведения	Крайний срок приёма докладов	Запрещено ли добавление докладов?	Открыта ли конференция для просмотров на сайте	
ПИИВС-2018	Программная инженерия: методы и технологии разработки информационно-вычислительных систем (ПИИВС-2018)	14.11.2018	30.10.2018	разрешено	открыта	
Студенческие секции	ПИИВС(студенческие секции)-2018	14.11.2018	30.10.2018	разрешено	открыта	

Рисунок 1 – Страница управления

После указания дат конференций и разрешения пользователям добавлять в неё доклады администратор назначает секции на странице управления секциями (см. рис. 2), которые будут доступны авторам докладов на странице создания нового доклада в графе секция в виде выпадающего списка.

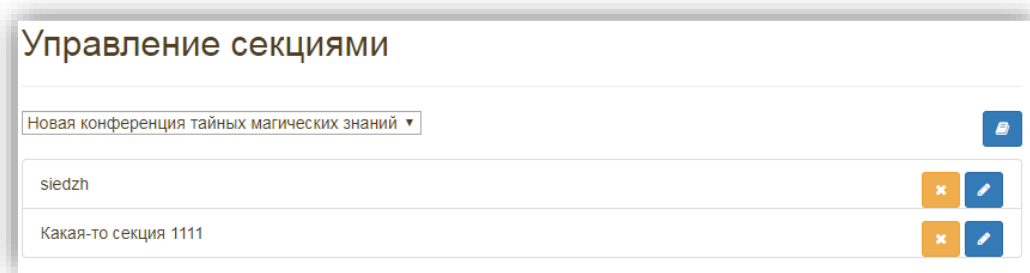


Рисунок 2 – Страница «управление секциями»

Первые доклады и начальная работа секретаря

Работа конференции начинается с добавлением первого доклада. Зарегистрировав всех авторов (до 5-ти на доклад), один из авторов добавляет доклад. Используя панель управления, автору следует пройти на страницу добавления доклада и, указав форму участия, нажать на кнопку «создать доклад» (рис. 3).

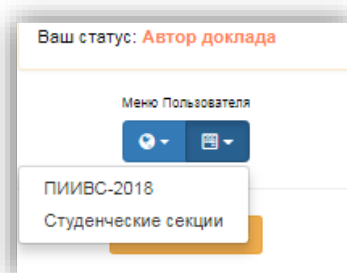


Рисунок 3 – Меню управления для роли «автор доклада»; выбор страницы добавления доклада

Далее ввести название доклада, выбрать секцию, указать авторов (для системы, главное чтобы все участвующие авторы были указаны в файле доклада), выбрать файл и отправить доклад (рис. 4). Авторов можно добавлять по ID, доступному на странице профиля каждого из авторов после авторизации конкретного пользователя, либо по фамилии и инициалам (вводить до тех пор, пока система автоматически не добавит ID в поле справа).

Рисунок 4 – Пример формы создания доклада

Файл доклада может быть размером до 5-ти мегабайт и с расширением doc или docx. Ранее допустимо было добавление и с расширением pdf, но в текущей версии оргкомитет постановил отказаться от этого расширения. После чего на странице «приём докладов» будет заполнена таблица первым докладом, а все последующие будут добавляться ниже. В таблице просмотра поданных докладов секция представлена иконкой «чашечка с кофе», по наведению курсора мыши на которую всплывает подсказка с полным названием указанной секции. В поле «действия» изначально находится только иконка-ссылка просмотра в виде стрелочки, а после того как докладу секретарём или рецензентом будет назначен статус «исправить», будет доступна иконка-ссылка в виде карандаша «редактировать» (рис. 5).

№	Название	Версия	Статус	Секция	Действия
1	Название доклада 1234	5	исправить	☕	➡ ✎
2	ty betry rt y	1	принят	☕	➡

Рисунок 5 – Таблица просмотра добавленных докладов автором; версии доклада инкрементируются с каждым изменением доклада и учитываются только рецензентом

После того как будут добавлены первые несколько докладов под панелью управления секретаря появится уведомление о количестве не назначенных докладов, которые ему следует проверить и распределить со страницы «секретариат» (рис. 6). Также система отслеживает так называемые простые рецензентов (бездействие в течение 3-х и более дней от последнего изменения доклада), которые секретарь также может просмотреть со специальной страницы «простой рецензентов».













Секретариат						
ПИИВС-2018						
Рецензент	Доклады					
неназначено	№	Название	Версия	Статус	Секция	Действия
	1	Экспертно - аналитический метод оценки качества конкурирующей продукции	1	на рецензировании	Прикладные интеллектуальные системы управления и принятия решений	   
Секретарь Секретарь Секретарь	№	Название	Версия	Статус	Секция	Действия
	1	ты бету рт у	1	на рецензировании	Интеллектуальный анализ данных, распознавание и классификация	   
	2	Название доклада	1	на рецензировании	Интеллектуальный анализ данных, распознавание и классификация	   

Рисунок 6 – Пример страницы «Секретариат»

Рецензирование и система уведомлений

Сразу после того, как секретарём рецензенту будет назначен доклад, система отправит уведомление с вложенным файлом, что позволяет рецензенту не находиться на сайте в ожидании назначенных ему докладов. Для рецензирования пользователю. Необходимо авторизоваться с помощью логина и пароля, которые отправляются на почту сразу после регистрации. После проверки вложенного в письме файла доклада рецензенту необходимо выбрать нужную страницу в панели управления (рис. 7) и указав статус докладу («исправить», «принят» или «не принят») написать комментарий к докладу, после чего закончить рецензирование кнопкой «закончить рецензирование». Всем указанным авторам доклада система вышлет на почтовые адреса уведомление с указанием названия, назначенного статуса и комментария доклада.

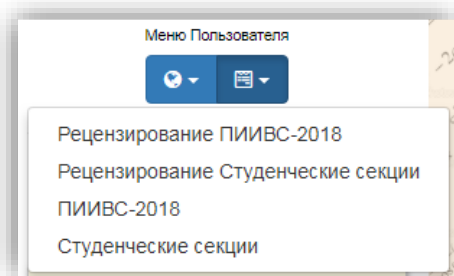


Рисунок 7 – Панель управления для пользователя с ролью рецензент

Авторы имеют право на усмотрение рецензента или секретаря несколько раз изменять (с помощью иконки-ссылки в виде карандаша, которая появляется, только если докладу назначен статус «исправить») свой доклад. Форма редактирования доклада позволяет изменять название, секцию и добавлять новый файл доклада

(рис. 8). Каждая версия сохраняется на сайте с целью учёта ранее произведённых действий над докладом.

The screenshot shows a web form for editing a report. It includes a text input for the title, a dropdown menu for the section, a list of authors, a file upload button, and a submit button.

Название:
Название доклада 1234

Секция:
Интеллектуальный анализ данных, распознавание и классификация

Авторы доклада

Дмитрий Александрович

Секретарь Секретарь Секретарь

Всем авторам и рецензенту будет отправлено уведомление об отправке данного доклада.

Файл можно добавлять с расширением .pdf / .doc / .docx и размером не больше 5MB

Выберите файл файл не выбран

Отправить доклад

Рисунок 8 – Форма редактирования доклада

На странице просмотра доклада доступна информация по всем версиям доклада, посмотреть которую можно на странице «просмотр доклада» (рис. 9). Основная информация представлена в блоке залитым серым цветом, где перечислены все указанные авторы доклада, отображён статус доклада, его название и секция. Чуть ниже с помощью кнопок пагинации выведены все версии доклада. Каждая из версий содержит название доклада, которое, по мнению разработчиков, может меняться по разным версиям, ссылку на скачивание файла доклада, который был добавлен для этой версии, а также комментарий рецензента, естественно изменяющийся от версии к версии.

The screenshot shows the 'View Report' page. It displays metadata for the report, a list of authors, a pagination control, and details for the 8th version of the report, including a download link and a reviewer's comment.

Просмотр доклада

Конференция:
ПИИВС-2018

Секция:
Интеллектуальный анализ данных, распознавание и классификация

Статус:
[исправить](#)

Авторы:

ФИО: Филиппин Дмитрий Александрович
Страна: Донецк
Город: ДНР
Учёная степень:
Учёное звание:
Должность: инженер программист
Организация: ДонНТУ

ФИО: Секретарь Секретарь Секретарь
Страна: Донецк
Город: ДНР
Учёная степень:
Учёное звание:
Должность: должность
Организация: ДонНТУ

8 7 6 5 4 3 2 1

Версия доклада №8

Название:
Название доклада 1234
Ссылка для скачивания файла:
↓

Комментарий рецензента:

Уважаемые коллеги!

Приглашаем Вас принять участие в II Международной научно-практической конференции
"Программная инженерия: методы и технологии разработки информационно-
вычислительных систем" (ПИИВС-2018)

Рисунок 9 – Страница просмотра доклада с 8-ью ранее добавленными версиями

Рецензирование производится на специально выделенной для этого странице. В случае какой-либо неспособности одного из рецензентов прорецензировать доклад, это может сделать секретарь, которому доступны все доклады на конференции. Страница «рецензирование» содержит полноценный html-редактор, который позволяет в полной мере описать ошибки авторам. Html-редактор не позволяет рецензенту загружать картинки в комментарий (см. рис. 10).

В качестве тестового примера одному из докладов был назначен комментарий рецензента полным html-скриптом главной страницы сайта конференции ПИИВС-2018, который корректно отображается на странице просмотра доклада.

В случае если доклад имеет небольшие грамматические или орфографические ошибки рецензент может, предварительно назначив докладу статус «принят», то есть одобрить его и уведомить авторов об успешной проверке доклада, заменить исправленный файл доклада самостоятельно, который из хостинга не удаляется, а хранится с суффиксом «del».

Рецензирование доклада

Конференция:
ПИИВС-2018

Название доклада:
ty betry rt y

Версия доклада:
1

Секция:
Компьютерная графика: синтез и обработка изображений

Статус:
не принят

Комментарий рецензента:
*В редакторе представлен комментарий, как шаблон для статуса "принят". Очистите поле чтобы ввести свой.
Редактор поддерживает html теги, для этого используйте ссылку "выключить редактор".*

Файл - Изменить - Вставить - Вид - Формат - Инструменты -

B *I* U ~~S~~ [буллит] [нумерация] [формат] [параграф] [шрифт] [размер ш...]

[буллит] [нумерация] [буллит] [нумерация] [undo] [redo] [ссылка] [удалить] [картинка] [code] [underline] [background-color] [table]

x₂ x² Ω -

✂ [table] [table] [table] “ ”

Начало конференции 14.11.2018

р

Количество слов: 2

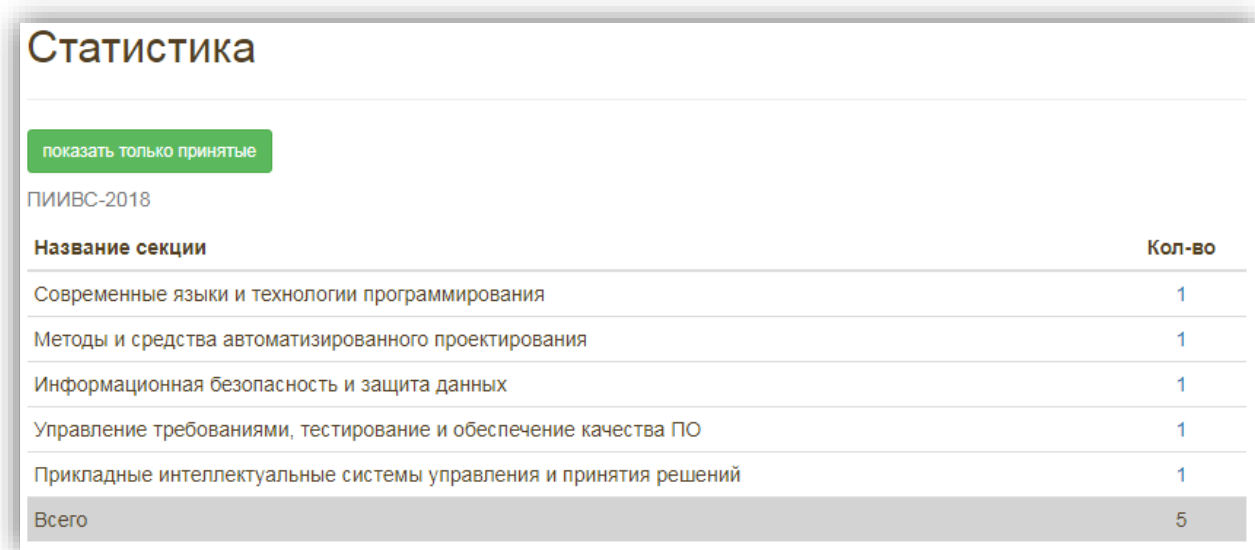
[Выключить редактор](#)

[Рецензировать](#) [Не рецензировать \(вернуться на пред. страницу\)](#)

Рисунок 10 – Страница рецензирования

Статистика конференции и формирование сборника

Система отчётов позволяет доверенным пользователям сайта конференции просматривать данные о докладах, как в пределах секции, так и в конкретном пользовательском виде. Ключевым модулем для этих целей является страница «Статистика», на которой в виде таблицы приведён список секций с указанием количества докладов (см. рис. 11).

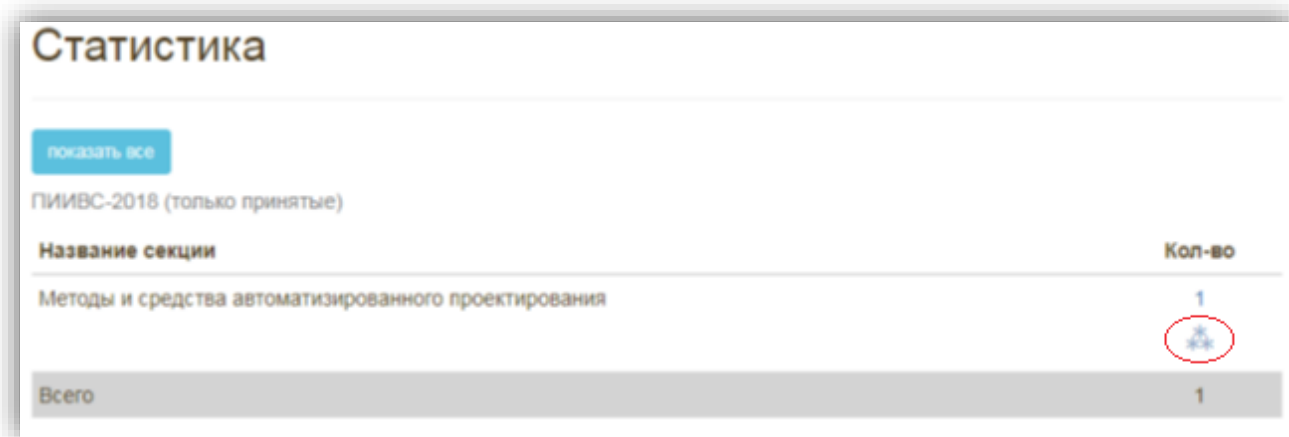


The screenshot shows a web interface titled 'Статистика'. At the top left, there is a green button labeled 'показать только принятые'. Below it, the text 'ПРИИВС-2018' is displayed. A table follows with two columns: 'Название секции' and 'Кол-во'. The table lists five sections, each with a count of 1, and a final row for 'Всего' with a count of 5.

Название секции	Кол-во
Современные языки и технологии программирования	1
Методы и средства автоматизированного проектирования	1
Информационная безопасность и защита данных	1
Управление требованиями, тестирование и обеспечение качества ПО	1
Прикладные интеллектуальные системы управления и принятия решений	1
Всего	5

Рисунок 11 – Страница статистики с общим тестовым подсчётом докладов по секциям

Страница изначально работает в режиме «всего»: секретарь видит количество докладов по секциям в целом. Просмотреть только одобренные или доклады с назначенным статусом «принят» можно с помощью кнопки «показать только принятые», которой перезагрузит страницу в режиме принятых докладов (см. рис. 12). В режиме просмотра только принятых докладов, появляется дополнительная иконка-ссылка в виде снежинки, которая позволяет скачать все файлы текущей секции одним zip-архивом.



The screenshot shows the 'Статистика' page with a blue button labeled 'показать все'. The text 'ПРИИВС-2018 (только принятые)' is displayed. The table shows only one section: 'Методы и средства автоматизированного проектирования' with a count of 1. A red circle highlights a snowflake icon next to the count '1'. The 'Всего' row shows a count of 1.

Название секции	Кол-во
Методы и средства автоматизированного проектирования	1
Всего	1

Рисунок 12 – Страница «статистика» в режиме «только принятые»

После загрузки докладов, оргкомитет конференции формирует научные сборники, которые по количеству предзаказов на печатный вариант заказываются в полиграфии и распределяются по желающим. Электронный вариант каждого сборника, также доступен в правой панели сайта в виде файлов с расширением .pdf (см. рис. 13).



Рисунок 13 – Пример обложек научных сборников конференции ПИИВС-2016

Выводы

В данной работе были перечислены и проанализированы изменения функциональности сайта, а также его работа для сопровождения конференции ПИИВС-2018 [1] в целом.

В результате работы был рассмотрен вопрос потребности упрощения работы некоторых модулей сайта конференции, а также наличия справочных руководств доступных пользователям после регистрации.

Литература

1. Официальный сайт ДонНТУ [Электронный ресурс]. — Режим доступа: <http://donntu.org/>.
2. Сайт конференции ПИИВС-2018 [Электронный ресурс]. — Режим доступа: <http://pi.conf.donntu.org/>.
3. Webstudio2u [Электронный ресурс]. — Режим доступа: <http://webstudio2u.net/ru/programming/96-cms.html>.
4. Joomla.org Русский [Электронный ресурс]. — Режим доступа: <https://www.joomla.org/>.

Филипшин Д.А., Серёженко О.А., Коломойцева И.А., Павлюк Е.Н. Анализ изменений функциональности сайта для сопровождения конференции ПИИВС-2018. В работе рассматриваются вопросы использования каждого из модулей, а также второй глобальной реконструкции и рефакторинга сайта для сопровождения конференции ПИИВС-2018.

Ключевые слова: Joomla, вебсайт, электронное хранилище, конференция, научный доклад.

Filipishin D.A., Serezhenko O.A., Kolomoitseva I.A., Pavlyuk E.N. Analysis of changes in the functionality of the site to accompany the conference PIIVS-2018. This article is about the using of the module. Also the paper describes the second global reconstruction and refactoring of site to accompany the PIIVS-2018 conference.

Keywords: Joomla, website, electronic repository, conference, scientific report.

Разработка приложения для распознавания лиц

Медведева В.Ю., Гоманчук М.В.

Гродненский государственный университет имени Янки Купалы
Medvedeva_vj_97@mail.ru , Mikhail.hamanchuk@instinctools.ru

Медведева В.Ю., Гоманчук М.В. Разработка приложения для распознавания лиц. В работе рассматривается разработка программного продукта для выделения и распознавания лиц из видеопотока. Она позволит распознавать людей по их лицамна изображениях. Для решения задачи выделения использован метод Виолы-Джонса, который позволяет надежно обнаруживать лица независимо от яркости фона и цвета лица. Для распознавания лиц используется метод LBPН, как один из наиболее эффективных на сегодняшний день.

Ключевые слова: распознавание лиц, поиск лиц, алгоритмы выделения и распознавания лиц, алгоритм Виолы-Джонса.

Введение

Задачи, связанные с компьютерным зрением получили актуальность со времён первых компьютеров. Для извлечения данных в них применяются статистические методы и используются модели, построенные с помощью геометрии, физики и теории обучения. Однако только в последние несколько лет появилась возможность решать эти задачи более эффективно. Важным этапом в создании эффективных алгоритмов стало применение в них методов машинного обучения.

Распознавание лиц является одной из основных практических проблем, решение которых способствуют развитию теории распознавания объектов. Задача крайне сложна и требует колоссальных технических средств. В случае распознавания лиц выделяют две основные задачи: задача выделения лица и задача распознавания. Но даже такое разделение сводит к двум довольно сложным подзадачам. Для их решения используют разные подходы.

В алгоритмах выделения лица выделяют два основных направления:

- 1) первый подход рассматривает лицо как набор разнообразных признаков, что характеризует лицо как объект среди множества других объектов. Такими признаками могут служить цвет кожи, брови, глаза и т.д.
- 2) второй подход основывается на рассмотрении снимка с лицом как изображения, абстрагируясь от задачи выделения лица, расширяя ее к задаче выделения какого-либо объекта на снимке.

Выбор алгоритмов и разработка приложения

Алгоритмы, реализующие первый подход, характеризуются высокой скоростью работы и относительно простой структурой, но вероятность распознавания недостаточно высока для применения в реальных условиях. Другие алгоритмы используют различные сложные структуры хранения и обработки информации, представляют собой целую систему. Один из примеров такой системы – нейронная сеть. Хотя скорость работы и вероятность правильного распознавания у таких алгоритмов гораздо выше, чем у алгоритмов первого подхода, их реализация крайне сложна. Таким образом, более эффективным является второй подход — подход, абстрагирующий от задачи выделения лица до задачи выделения любого объекта на снимке. Наиболее популярный - метод Виолы-Джонса для выделения лиц. Это алгоритм, позволяющий обнаруживать объекты в реальном времени.

Основные принципы, на которых он основан:

- а) Принцип сканирующего окна.

Сканируется изображение окном поиска (окно сканирования), а затем применяется классификатор к каждому положению. Система обучения и выбора наиболее значимых признаков полностью автоматизирована и не требует вмешательства человека, поэтому данный подход работает крайне быстро.

- б) Использование изображений в интегральном виде для быстрого вычисления.

Главным преимуществом такого представления является то, что оно позволяет быстро рассчитывать суммарную яркость произвольного прямоугольника на данном изображении, не зависимо от площади прямоугольника.

Интегральное представление изображения представляет собой матрицу, совпадающую по размерам с исходным изображением. В каждом ее элементе хранится сумма интенсивностей всех пикселей, находящихся левее и выше данного элемента. Элементы матрицы рассчитываются по следующей формуле

$$I(x, y) = \sum_{x' \leq x, y' \leq y} i(x', y'),$$

где $I(x, y)$ — значение точки (x, y) интегрального изображения; $i(x, y)$ — значение интенсивности исходного изображения. На основе применения интегрального представления изображения вычисление признаков одинакового вида, но с разными геометрическими параметрами, происходит за одинаковое время. Каждый элемент матрицы $I(x, y)$ представляет собой сумму пикселей в прямоугольнике от $i(0, 0)$ до $i(x, y)$, т.е. значение каждого элемента $I(x, y)$ равно сумме значений всех пикселей левее и выше данного пикселя $i(x, y)$. Расчет матрицы занимает линейное время, пропорциональное числу пикселей в изображении и его можно производить по следующей формуле:

$$I(x, y) = i(x, y) - I(x-1, y-1) + I(x, y-1) + I(x-1, y).$$

Интегральное представление имеет интересную особенность. По интегральной матрице можно очень быстро вычислить сумму пикселей произвольного прямоугольника.

с) Использование признаков Хаара, с помощью которых происходит поиск нужного объекта.

В стандартном методе Виолы – Джонса используются прямоугольные признаки, которые называются признаками Хаара. Признак Хаара - это число, характеризующееся разностью сумм пикселей между черной и белой областью. В рамках системы детектирования лиц множество всех признаков задается формой, размером и положением на изображении.

В расширенном методе Виолы – Джонса используются дополнительные признаки.

Для того чтобы компенсировать влияние различных условий освещенности, изображение должно быть нормализовано, т.е. иметь нулевое математическое ожидание и единичную дисперсию, также важным критерием вычисления признаков является единичное соотношение сторон изображения и его представление в сером формате.

Вычисляемым значением признаков Хаара будет по формуле $F = X - Y$, где X - сумма значений яркостей точек, закрываемых светлой частью признака, а Y - сумма значений яркостей точек, закрываемых темной частью признака. Для их вычисления используется понятие интегрального изображения. Признаки Хаара дают точечное значение перепада яркости по оси X и Y соответственно.

д) Бустинг — процедура последовательного построения композиции алгоритмов машинного обучения, когда каждый следующий алгоритм стремится компенсировать недостатки композиции всех предыдущих алгоритмов.

Так как количество признаков Хаара достигает почти 200 000, необходимо отсеять наиболее неинформативные из них. В случае лиц, можно заметить, что области глаз, носа и рта темнее, то их легко можно выделить с помощью признаков, поэтому они являются информативными.

Ускорение (усиление) - это подход в машинном обучении, который основывается на создании высокоточного правила предсказания, путем комбинирования относительно слабых и неточных правил. Алгоритм AdaBoost, предложенный Йоавом Фройндом и Робертом Шапиром был первым алгоритмом усиления, который до сих пор используется и изучается в различных областях.

Данный алгоритм является бинарным, поскольку он может разделять множество только на два класса. Основной задачей адаптивного ускорения является построение сильного классификатора на основе слабых. Слабым классификатором, в данной работе, является признак Хаара. Сильный классификатор – это множество слабых классификаторов, которые наилучшим образом определяют лицо на изображении.

Особенностью применения алгоритма обучения в алгоритме Виолы-Джонса является отбор сильных классификаторов. Эта цель достигается путем оценивания вероятности правильно классифицированных изображений. Например, вероятность правильной классификации лиц из выборки не должна быть ниже 95%.

Для разработки приложения были выбраны два алгоритма: метод Виолы-Джонса [1] (для поиска лиц), и метод LVRN [2] (для распознавания лиц), как наиболее эффективные среди имеющихся на сегодня методов. Метод LVRN позволяет извлекать признаки из окрестностей изображения, поэтому он является устойчивым к масштабу, поворотам и т.д. Этот набор признаков имеет маломерную структуру, что является положительной чертой данного метода.

Сначала изображение делится на одинаковые блоки, которые образуют сетку. Далее для каждого блока строится гистограмма кодов, которые вычисляются таким образом: берется пиксель, который сравнивается с соседями, если интенсивность центрального пикселя больше или равна интенсивности соседа, то он обозначается 1, иначе — 0. В итоге каждому пикселю будет соответствовать двоичное число, состоящее из результатов сравнений. Эти преобразования наглядно показаны на рисунке 1. Полученные гистограммы объединяются в одну общую, которая является итоговым дескриптором, который используется для классификации лица.

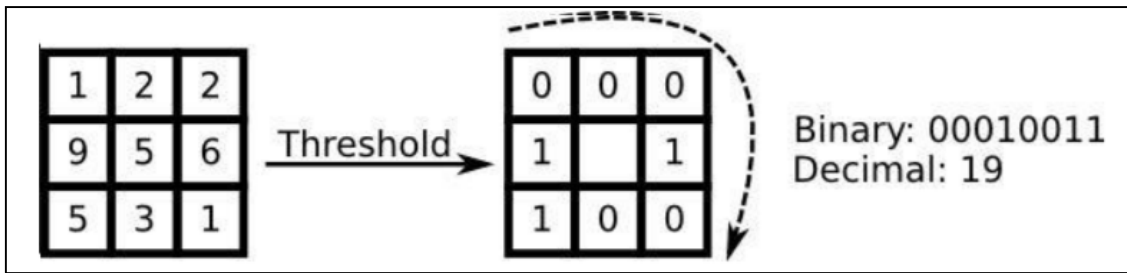


Рисунок 1 – Преобразование области в двоичный код

Описанный подход позволяет захватывать очень мелкозернистые детали. Сначала LBPH не мог кодировать детали разного масштаба, из-за этого он был расширен, и теперь число соседей может варьироваться. Полная схема работы изображена на рисунке 2:

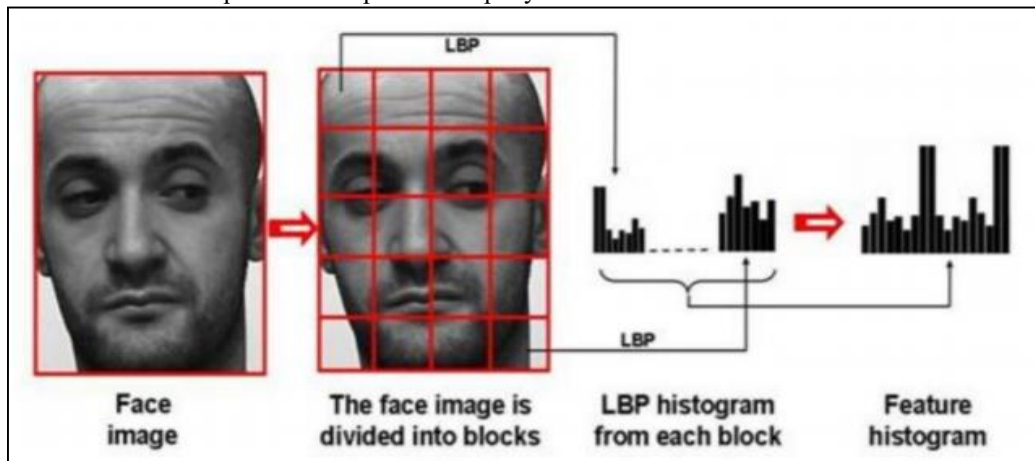


Рисунок 2 – Схема работы алгоритма LBPH

Идея расширения заключалась в том, чтобы расположить соседей по кругу определенного радиуса, и, таким образом, выделялись следующие признаки окрестностей (рисунок 3):

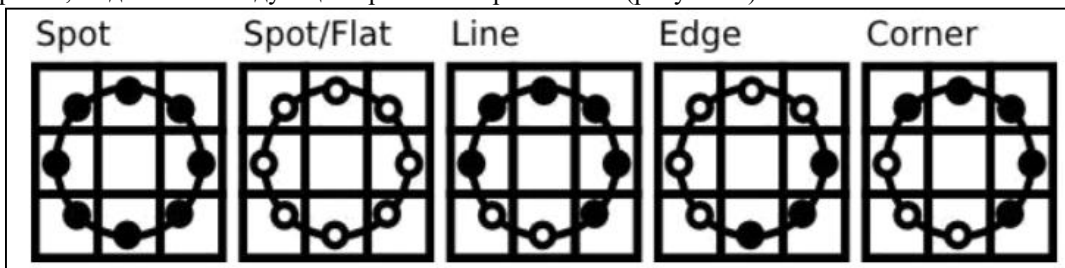


Рисунок 3 – Признаки круговой окрестности

Формулы для расчета координат соседей:

$$x_p = x_c + R \cos \frac{2\pi\rho}{P}; \quad y_p = y_c + R \sin \frac{2\pi\rho}{P}$$

где x_p, y_p — координаты соседа; x_c, y_c — координаты центра; R — радиус окружности; ρ — номер соседа; P — количество соседей. Если координаты точек не соответствуют координатам изображения, к ним будет применен метод интерполяции.

Тестирование программы - это этап, на котором проверяется, как ведет себя программа на как можно большем количестве входных наборов данных, в том числе, и на заведомо неверных.

При запуске приложения пользователь видит окно, оформление и размеры которого могут изменяться непосредственно программистом в процессе создания программы и в соответствии с его вкусами, предпочтениями и дизайнерскими идеями (рисунок 4).

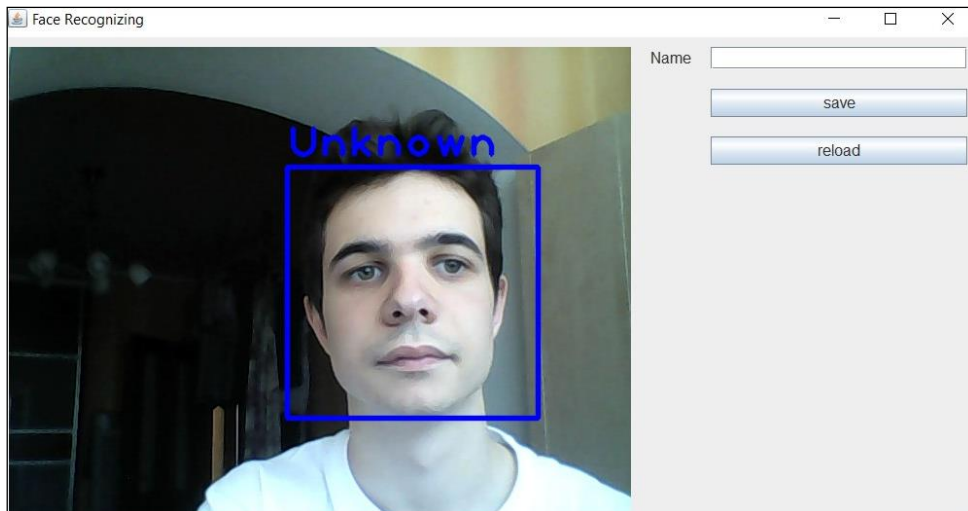


Рисунок 4 – Начальное окно приложения

При запуске web-камеры разработанное приложение выделяет лицо, которое присутствует на изображении, в синий квадрат. С помощью поля «Name» в правом верхнем углу окна пользователю на экране можно присвоить имя, которое после нажатия кнопки «Save» появится на экране (рисунок 5).

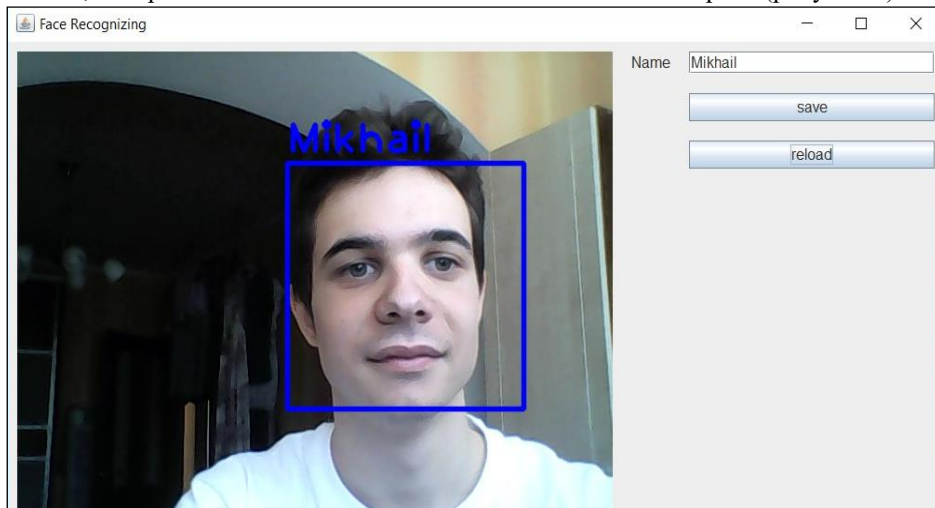


Рисунок 5 – Пользователю присвоено имя

Рисунок 6 демонстрирует ещё одну возможность приложения: определение и подпись на изображении двух и более лиц. Добавление последующего подписанного лица возможно только после сохранения предыдущих.

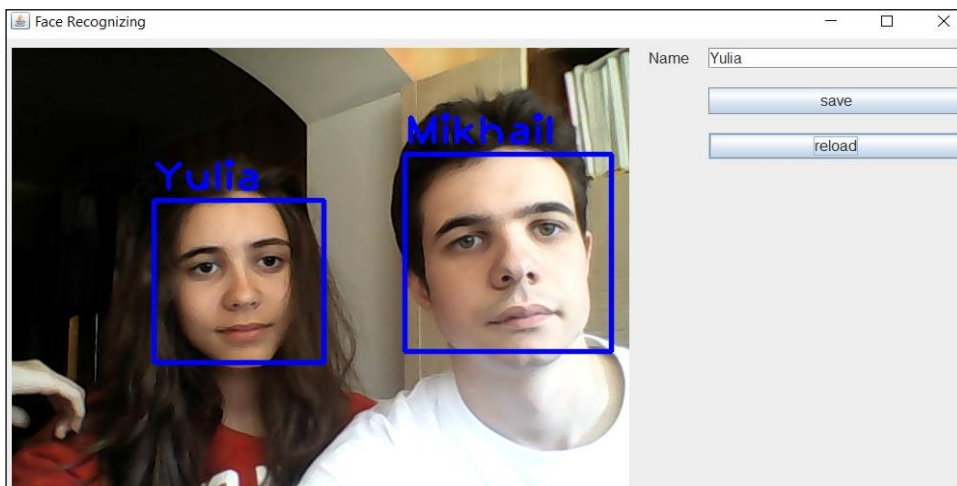


Рисунок 6 – Определение двух и более лиц

Код данного программного обеспечения выложен в общем доступе на GitHub [3]. Программа не занимает много места, не требовательна к установленному программному обеспечению и работает корректно. В ходе написания приложения были углублены знания в языке Java, популярных фреймворках и библиотеках, таких как SpringFramework, Hibernate и др. При разработке программы был использован объектно-ориентированный язык программирования Java. Разработанная программа является Web-приложением.

Заключение

Сейчас набирает популярность приложения с так называемой «дополненной реальностью». Созданная программа является только начальным этапом в разработке приложения, которое сможет пополнить коллекцию популярного сегодня направления, а также стать неотъемлемым атрибутом современного развлечения для человека. Эта цель будет достигнута по средствам улучшения:

- распознавание лиц, находящихся под углом;
- увеличения скорости обработки.

Литература

1. P.Viola, M.Jones. Rapid object detection using a boosted of simple features / 2001 IEEE Computer Society Conference on Computer Vision and Pattern Recognition Vol. 1. 8-14 December 2001 / The Institute of Electrical and Electronics Engineers, Inc. С. 511-518.
2. Ojala, T., Pietikainen, M., Maenpaa, T. “Multi resolution gray-scale and rotation invariant texture classification with local binary patterns.” IEEE Transactions on Pattern Analysis and Machine Intelligence. Vol. 24, № 7, 2002. P. 971–987.
3. Программный код приложения. GitHub [Электронный ресурс]. – Режим доступа: <https://github.com/mishagrodno/FaceRecognizer/>. – Дата доступа: 14.10.2018.

Медведева В.Ю., Гоманчук М.В. Разработка приложения для распознавания лиц. В работе рассматривается разработка программного продукта для выделения и распознавания лиц из видеопотока. Она позволит распознавать людей по их лицам на изображении. Для решения этой задачи использован метод Виолы-Джонса для поиска лиц. Он позволил надежно обнаруживать лица независимо от яркости фона и цвета лица.

Ключевые слова: распознавание лиц, поиск лиц, алгоритмы выделения и распознавания лиц, алгоритм Виолы-Джонса.

Medvedeva Viktoria, Gomanchuk Mikhail Developing an application for face recognition. The paper discusses the development of a software product for the selection and recognition of persons from the video stream. It will allow people to be recognized by their faces in the image. To solve the problem of selection, the method of Viola-Jones was used, which allows you to reliably detect faces, regardless of the brightness of the background and complexion. For facial recognition, the LBPH method is used as one of the most effective to date.

Key words: face recognition, face search, face detection and recognition algorithms, Viola-Jones algorithm.

СЕКЦИЯ 2. «ИНФОРМАЦИОННЫЕ СИСТЕМЫ, БЕЗОПАСНОСТЬ И ЗАЩИТА ДАННЫХ»

УДК 004.9

Применение технологии блокчейн в земельном кадастре и реестре недвижимости

Бутко А.В.

Министерство связи Донецкой Народной Республики
butko.av@minsvyazdnr.ru

Бутко А.В. Применение технологии блокчейн в земельном кадастре и реестре недвижимости. В статье обзревается практическое применение технологии блокчейн в Российской Федерации в области управления недвижимостью.

Ключевые слова: блокчейн, смарт-контракт, Eхom, Ethereum, Мастерчейн, земельный кадастр, реестр недвижимости, Российская Федерация, Росреестр, договор долевого участия, ипотека.

Введение

Технология блокчейн чаще всего ассоциируется с криптовалютами. Однако, эта технология также может применяться и в других сферах деятельности. Перспективными направлениями являются логистика, регистрация прав собственности (земельные участки, недвижимость), здравоохранение (защита от подделок медикаментов, мониторинг передвижения, обмен медицинскими знаниями, ведение медицинской карты пациента), государственный и корпоративный документооборот, голосования, защита авторских прав. В рамках этого исследования мы рассмотрим применение технологии в области земельного кадастра и реестра недвижимости.

Что такое блокчейн?

В основе технологии блокчейн находится хранение информации в децентрализованной распределенной базе данных. Все участники системы при помощи пиринговой сети имеют актуальные копии этой базы данных.

Данные, которые хранятся в блокчейне, защищены криптографически. Любой участник может проверить подлинность данных. При этом, чтобы получить доказательство их истинности не нужно видеть сами данные в открытом виде. Для этого используется механизм шифрования при помощи публичных и частных ключей. Защита от несанкционированных изменений основана на хеш функции – способе преобразования данных произвольной длины в цифровую строку фиксированной длины, называемой хеш. Особенность этой функции в том, что вычислить хеш легко, а обратно восстановить исходные данные по заданному хешу очень трудно.

В последовательной цепочке блоков хранится информация о транзакциях, сделках и контрактах либо любая другая информация, которую обычно записывают в виде последовательных реестров. Также каждый блок содержит хэш предыдущего блока. При любом изменении блока пересчитывается его хеш, поэтому невозможно удалить или провести изменение блока без изменения всех последующих блоков в цепочке.

Любой участник может создать новый блок, включив в него ещё не записанные в блокчейн транзакции. Для записи новых блоков необходим консенсус узлов блокчейна. Это позволяет фильтровать операции и записывать только легитимные транзакции, что гарантирует достоверность содержимого. Консенсус обеспечивают алгоритмы Proof of Work (PoW, проделанной работы) и Proof of Stake (PoS, подтверждение доли). Для создания нового блока, необходимо включить в него ссылку на предыдущий блок (хеш блока) и выполнить подбор случайного модификатора, чтобы хеш, данного блока начинался с определенным количеством нулевых битов. Для подбора случайного модификатора требуются вычислительные мощности и время, т. к. эта задача решается перебором пока в результате не будет подобран нужный формат. Подбор производится всеми участниками на конкурентной основе. За подбор модификатора и создание блока участник получает вознаграждение. Если несколько участников подберут модификатор и сделают независимые блоки, то происходит ветвление, и новые блоки присоединяются к нескольким веткам. На этот случай в блокчейне есть правило, что в системе остаётся только самая длинная ветвь, поэтому блоки из коротких ветвей расформируются, а транзакции из них вновь считаются неподтверждёнными и используются в создании нового блока.

Опыт разных стран по внедрению блокчейна

Внедрение технологии блокчейн в области земельного кадастра востребовано как в развитых, так и в развивающихся странах, хотя обусловлено разными причинами. Развитые страны совершенствуют операционные процессы, сокращая в несколько раз время совершения сделки, повышают надёжность и прозрачность процессов, тем самым снижая риск мошенничества и ошибок в документах и транзакциях. В развивающихся странах есть проблемы с плохим документированием права собственности на землю, в результате чего у собственников ограничены операции с землей и возможны злоупотребления со стороны чиновников.

В 2016 году Национальная земельная служба Швеции, заявила об успешной работе с блокчейн стартапом ChromaWay, консалтинговой фирмой Kairos Future и поставщиком услуг телефонной связи Telia. Было разработано и протестировано решение для покупателей и продавцов. В тестировании участвовали покупатели, продавцы, агенты недвижимости, банк и государственная служба кадастровой регистрации.

В 2016 году правительство Ганы выдало официальное разрешение проекту Bitland, основанному на платформе Graphene на составление земельного кадастра на основе блокчейна Bitshares. Bitland выпустил базовые токены CADASTRAL с помощью которых можно будет регистрировать права на землю, решать спорные вопросы землевладения, продавать и покупать землю. В случае успеха проект планируется распространить его на другие африканские страны, в первую очередь Нигерию и Кению.

В 2016 году правительство Гондураса приняло решение использовать распределенный реестр для регистрации прав на землю, чтобы бороться с мошенничеством. Исторически в странах Латинской Америки есть проблема с правами на землю, которая связана с высоким уровнем коррупции. Разработка проводится при сотрудничестве с компанией Factom. В Гондурасе уже несколько десятков лет существует конфликт, связанный с проблемой прав собственности.

В 2016 году Национальное агентство публичного реестра Грузии и компания BitFury объявили о запуске пилотного проекта по кадастровому учету на базе блокчейна. Практическая реализация проекта началась в 2017 году. Первый этап внедрения завершился, далее планируется использовать смарт-контракты для оптимизации таких процессов как продажа имущества, передача права собственности и других.

«Блокчейн-коммуна» в России

В 2017 году в Москве открылся первый в России специализированный экспертный центр по внедрению блокчейн-технологий в сфере государственного и корпоративного управления «Блокчейн-коммуна». Организован Внешэкономбанком и Национальным исследовательским технологическим университетом «МИСиС». Партнерами стали международные компании Ethereum, Bitfury, Waves, E&Y, PwC, представительства свыше двадцати компаний располагаются на территории Центра. Также «Блокчейн-коммуна» стала участником независимой ассоциации Crypto Valley, крупнейшего мирового центра развития блокчейн-технологий.

В рамках «Блокчейн-коммуны» разработчики технологий и продуктов, основанных на блокчейне могут заниматься разработкой и перенимать современный опыт. Были проведены крупные хакатоны с международным участием и действуют образовательные программы, начато несколько пилотных государственных проектов на технологии блокчейн. Среди них перенос на блокчейн данных Федеральной службы государственной регистрации, кадастра и картографии (далее – Росреестр) и Федеральной службы по интеллектуальной собственности, отслеживание цепочек поставок лекарственных препаратов, мониторинг лесов и другие.

Внешэкономбанк совместно с Росреестром запустил федеральный проект Blockchain Pitch, в рамках которого происходит поиск инновационных компаний, готовых создавать проекты на основе технологии блокчейн и внедрять их совместно с государственными учреждениями. В 2017 году в рамках проекта Blockchain Pitch был осуществлен конкурсный отбор компаний имеющих прототипы решений для рынка недвижимости и тестирование их разработок совместно с представителями государственного заказчика. Запущено два пилотных проекта, относящихся к использованию технологии блокчейн в операциях с недвижимостью:

- Ведение реестра взаимодействия основных участников процесса регистрации договоров долевого участия в строительстве и регистрация фактов, предшествующих внесению изменений в информационные системы участников. Разрабатывается в сотрудничестве с Росреестром и Агентством ипотечного жилищного кредитования.
- Создание решения, позволяющего фиксировать изменение сведений в Едином государственном реестре недвижимости (далее – ЕГРН) об объекте недвижимости, зарегистрированном праве и правообладателе по г. Москва в распределенной сети, построенной на технологии блокчейн и доступной для внешнего аудита и контроля.

Блокчейн для регистрации договоров долевого участия в строительстве

Ленинградская область выбрана регионом, где тестируется использование технологии блокчейн для регистрации договоров долевого участия в строительстве. Для системы регистрации договоров использован фреймворк для создания блокчейн-проектов Eхonum компании Bitfury. В проекте участвуют Росреестр, Агентство ипотечного жилищного кредитования, Внешэкономбанк и публично-правовая компания «Фонд защиты прав граждан – участников долевого строительства».

Первый договор долевого участия с применением технологии распределенного реестра был зарегистрирован 26 января 2018 года. По состоянию на 1 октября 2018 года в рамках проекта было зарегистрировано 894 сделки. Если проект будет признан удачным, то его распространят на другие регионы.

Согласно законодательства, для компенсационных выплат участникам долевого строительства застройщики обязаны перечислять страховые взносы от суммы каждого договора в Фонд защиты прав участников долевого строительства. Деньги дольщиков могут быть использованы для строительства только после регистрации сделки и ранее приходилось ждать несколько месяцев, т.к. Росреестр не может зарегистрировать договор, пока не получит от Фонда уведомление о выплате застройщиком страхового взноса.

Теперь благодаря блокчейн-решению участники взаимодействия имеют доступ к актуальной информации по процессу регистрации каждого договора. Кроме того, информационные системы Росреестра и Фонда теперь постоянно синхронизированы, а также при отсутствии связи между ними могут локально работать с полными.

Использование блокчейна позволяет существенно сократить сроки взаимодействия между Агентством ипотечного жилищного кредитования, Росреестром и Фондом защиты прав участников долевого строительства. При проведении регистрации с применением этой системы информация из Фонда в Росреестр поступает в течение трех-пяти минут.

Подтверждать, соответствует ли новая транзакция уже записанной информации, должны несколько участников – верифицирующих узлов. У Росреестра, Агентства ипотечного жилищного кредитования и Фонда функционируют по два верифицирующих узла. Позже узлы могут появиться и у застройщиков. Копия на блокчейне гарантирует полную неизменность данных т.к. для подделки записи придется изменить её во всех организациях, управляющих верифицирующими узлами. Если какой-то узел подтвердит несанкционированное изменение информации, и она попадет в блокчейн, это будет видно всем другим участникам.

Блокчейн для фиксации изменений ЕГРН в г. Москва

В Москве в 2018 году также проводятся два эксперимента: первый — создание блокчейн-реестра сделок с недвижимостью как «зеркала» ЕГРН, второй — автоматизация правовой экспертизы документов по сделкам с недвижимостью при помощи смарт-контрактов.

В рамках первого эксперимента параллельно с фиксацией сведений в ЕГРН вносятся изменения в реестр на основе блокчейн-платформы Ethereum. Держателями блокчейн-узлов в этом проекте являются Росреестр, департамент информационных технологий Москвы, Федеральная налоговая служба, Сбербанк, «Ростелеком» и Агентство ипотечного жилищного кредитования. Первая стадия эксперимента включает в себя накопление неизменной истории владения объектом. На второй стадии будет осуществляться стресс-тестирование реестра для чего в него внесут до 300 миллионов дополнительных записей. По окончании эксперимента будет оценена неизменность данных блокчейн-реестра, производительность и потребность в вычислительных мощностях, а также необходимость дополнительной защиты данных.

Эксперимент по использованию смарт-контрактов

В рамках второго эксперимента на основе смарт-контрактов проверяется система правовой экспертизы документов, поданных на регистрацию перехода прав собственности. Исследуется насколько возможна передача прав собственности на недвижимость без посредников. Конечная цель проекта – автоматизация с помощью смарт-контрактов всех этапов подготовки к сделке, которые требуют длительного согласования: регистрации права собственности у нотариуса, оформления справки об отсутствии задолженности за жилищно-коммунальные услуги и т. п.

Смарт-контракт — компьютерный алгоритм, предназначенный для заключения и поддержания коммерческих контрактов в технологии блокчейн. Впервые начали применяться на практике в проекте Ethereum.

Автоматизированная логика смарт-контрактов, построенная по правилам работы регистраторов, будет принимать решения о регистрации сделки на основе анализа доверенных внешних данных и данных Росреестра. В качестве вариантов реализации рассматриваются блокчейн-платформы Eхonum, Corda или Hyperledger. Результаты автоматической экспертизы будут заноситься в блокчейн-реестр.

Параллельно те же операции будут проводить сотрудники. Если мнения программы и человека не совпадут, то это несоответствие будет исследоваться для доработки

Оба эксперимента в Москве должны закончиться 14 декабря 2018 года. По их результатам Росреестр

оценит целесообразность внедрения технологии и внесения необходимых изменений в законодательство. Если в будущем информация, записанная с помощью блокчейна, будет признана юридически значимой, система фактически сделает ненужным получение выписок из Росреестра: покупатель квартиры или судебные органы смогут посмотреть запись об объекте недвижимости сразу в блокчейне.

Блокчейн для ипотеки

Ещё одна область применения блокчейна – ипотека. В настоящее время для получения ипотеки нужно оформить существенное количество бумажных документов при участии многих посредников: оценщиков, юридических компаний, кредитных агентств, что повышает время и сумму расходов на оформление сделки.

При помощи блокчейна для всех объектов недвижимости можно создать цифровой идентификатор, для хранения информации о праве собственности и рыночной стоимости, что позволит банкам быстрее определять статус владения или подтверждать рыночную стоимость и сократит участие оценочных и юридических компаний. Также при помощи смарт-контрактов можно обрабатывать условия заключения сделки, например, переводить денежные средства продавцу только после одобрения банком ипотеки и цифрового подписания документов.

В сентябре 2018 года Райффайзенбанк (Россия) провел первую сделку по выпуску электронной ипотечной закладной с хранением и учетом в децентрализованной депозитарной системе на платформе Мастерчейн.

Инициация сделки происходит после заполнения клиентом банка полей электронной формы закладной. Далее документ заверяется усиленной квалифицированной электронной подписью, дополняется другими заверенными документами по сделке и направляется по цифровому каналу в виде xml-файла в Росреестр. Росреестр проверяет и регистрирует закладную, после чего через СМЭВ отправляет файл в депозитарий хранения (в данном случае — Райффайзенбанка).

Мастерчейн – это российская блокчейн-сеть, предназначенная для передачи цифровых ценностей и информации о них между участниками; разрабатывается с 2016 года в рамках ассоциации «Финтех» совместно крупнейшими банками под эгидой Центробанка

После присоединения Росреестра к платформе Мастерчейн станет возможным проведение операций с электронными ипотечными закладными в децентрализованной депозитарной системе.

Используемые инструменты – Exonum

Exonum — фреймворк для создания блокчейн-проектов. При помощи этого фреймворка можно создавать блокчейны, в которых блоки добавляют только заранее известные узлы и регулировать уровень открытости блокчейна для конечных пользователей. В Exonum разработан собственный алгоритм консенсуса, который гарантирует, что в любой момент существует только одна согласованная версия блокчейна и устойчивый к «византийскому» поведению узлов. Поэтому отдельные узлы не могут изменить историю транзакций или добавить неподтвержденные данные. Алгоритм консенсуса не требует майнинга, не требует специального оборудования и рассчитан на небольшие вычислительные мощности.

Лёгкий клиент позволяет пользователям отправлять транзакции и запрашивать данные из блокчейна. Клиент называется «легким» потому, что не имеет полной версии блокчейна и содержит только те данные, которые важны для узла, что требует меньше вычислительных ресурсов и объемов памяти. Клиентом может быть браузер на смартфоне или ноутбук, или серверное приложение. Скорость Exonum при применении в глобальной сети составляет 5000 транзакций в секунду. В локальной сети с небольшим количеством узлов — до 15000 транзакций в секунду. Получая данные, лёгкий клиент также получает и проверяет криптографические доказательства их истинности. Такая реализация защищена от MITM атаки и позволяет обмениваться данными даже используя открытый Wi-Fi в публичном месте.

Exonum — это проект с открытым исходным кодом. Его репозиторий есть на GitHub (<https://github.com/exonum>). Ядро Exonum написано на языке системного программирования Rust.

Используемые инструменты – Ethereum и смарт-контракты

Ethereum (Эфириум) — платформа для создания децентрализованных онлайн-сервисов на базе блокчейна, работающих на базе смарт-контрактов. Реализована как единая децентрализованная виртуальная машина.

Изначально была создана как криптовалюта, но в отличие от других криптовалют, авторы не ограничивают роль Ethereum платежами, а предлагают его, например, в качестве средства для обмена ресурсами или регистрации сделок с активами.

Технология Ethereum даёт возможность регистрации любых сделок с любыми активами на основе распределенной базы смарт-контрактов, не прибегая к традиционным юридическим процедурам. Эта возможность является конкурентной по отношению к существующей системе регистрации сделок.

Стороны подписывают смарт-контракт, используя методы, аналогичные подписанию отправки средств в

действующих криптовалютных сетях. После подписания сторонами смарт-контракт вступает в силу. Для обеспечения автоматизированного исполнения обязательств контракта непременно требуется среда существования, которая позволяет полностью автоматизировать выполнение пунктов контракта. Это означает, что смарт-контракты смогут существовать только внутри среды, имеющей беспрепятственный доступ исполняемого кода к объектам смарт-контракта. Все условия смарт-контракта должны иметь математическое описание и ясную логику исполнения. В связи с этим первые смарт-контракты имеют задачу формализации наиболее простых взаимоотношений, состоящих из небольшого количества условий. Имея беспрепятственный доступ к объектам контракта, смарт-контракт отслеживает по указанным условиям достижения или нарушения пунктов и принимает самостоятельные решения, основываясь на запрограммированных условиях. Таким образом, основной принцип смарт-контракта состоит в полной автоматизации и достоверности исполнения договорных отношений.

Смарт-контракты в Ethereum представлены в виде классов, которые могут быть реализованы на различных языках, включая визуальное программирование и компилируются в байт-код для виртуальной машины Эфириума (Ethereum Virtual Machine) перед отправкой в блокчейн.

Ethereum — это проект с открытым исходным кодом. Его репозитории есть на GitHub (<https://github.com/ethereum>).

Заключение

Технология блокчейн сейчас является одним из трендов информационных технологий и обладает большим потенциалом. В ближайшее время в Российской Федерации ожидается завершение экспериментов внедрения пилотных проектов, по которым можно будет судить насколько успешно продвигается внедрение и какие изменения нормативно-правовой базы потребуются для обеспечения юридической значимости операций в распределённой сети.

Литература

1. Артем Генкин, Алексей Михеев. Блокчейн. Как это работает и что ждет нас завтра. — М.: Альпина Паблишер, 2017. — 592 с. — ISBN 978-5-9614-6558-7.
2. Мелани Свон. Блокчейн: Схема новой экономики. — Олимп-Бизнес, 2016. — 240 с. ISBN 978-5-9693-0360-7
3. Лоран Лелу. Блокчейн от А до Я. Все о технологии десятилетия. — М.: Эксмо, 2018. — 256 с. — ISBN 978-5-699-98942-3.
4. Алекс Тапскотт, Дон Тапскотт. Технология блокчейн - то, что движет финансовой революцией сегодня. — М.: Эксмо, 2017. — 448 с. — ISBN 978-5-699-95092-8.
5. Уильям Могайар, Виталик Бутерин. Блокчейн для бизнеса. — М.: Эксмо, 2017. — 224 с. — ISBN 978-5-699-98499-2.

Бутко А.В. Применение технологии блокчейн в земельном кадастре и реестре недвижимости. В статье обзревается практическое применение технологии блокчейн в Российской Федерации в области управления недвижимостью.

Ключевые слова: блокчейн, смарт-контракт, Exonum, Ethereum, Мастерчейн, земельный кадастр, реестр недвижимости, Российская Федерация, Росреестр, договор долевого участия, ипотека.

Butko A.V. The use of blockchain technology in the land cadastre and real estate registry. The article reviews the practical use of blockchain technology in the Russian Federation in the field of real estate management.

Keywords: blockchain, smart contract, Exonum, Ethereum, Masterchain, land cadastre, real estate register, Russian Federation, Rosreestr, equity agreement, mortgage

Государственная информатизация с нуля

Гущин В.В.

Министерство связи Донецкой Народной Республики
samokat@minsvyazdnr.ru

Гущин В.В. Государственная информатизация с нуля. В статье рассмотрено состояние и развитие информатизации государственных органов власти Донецкой Народной Республики. Освещен тот факт, что перспективы будут строиться вокруг концепций в области информационных технологий, принятых в Российской Федерации и современных подходов в производстве программного обеспечения.

Ключевые слова: Информатизация, государственные органы власти, централизованное хранилище данных, электронный документооборот, единый репозиторий исходного кода, система межведомственного электронного взаимодействия, контактная панель, государственный удостоверяющий центр.

Введение

В современных реалиях уровень развития информационного пространства самым решающим образом влияет на экономику, обороноспособность и политику страны именно поэтому планируя архитектуру пространства информационного взаимодействия государственных органов мы принимали во внимание тенденции развития информационно-коммуникационных технологий (далее — ИКТ) в целом и планы России в построении информационного общества, в частности.

Проблемы, с которыми столкнулось Министерство связи ДНР, такие как состояние межведомственного информационного пространства, а точнее его отсутствие как в техническом, так и в нормативно-правовом плане, отсутствие финансирования и кадровый дефицит в ИТ-сфере побудили к началу работы одновременно по всем направлениям от построения центров обработки данных до обучения специалистов.

Управленческие решения

Принимая во внимание объективные сложности при формировании межведомственного информационного пространства, Министерством связи ДНР была принята концепция Единой Информационной Системы, состоящая из объединения современных подходов в государственном управлении и передовых технических решений, соответствующих общим тенденциям импортозамещения в государственных органах власти (рисунок 1).



Рисунок 1 – Взаимодействие государственных информационных систем и органов государственной власти

Для реализации основных модулей Единой Информационной Системы были задействованы ресурсы государственных хозрасчетных предприятий, так же на базе гос. предприятий проводилась опытная эксплуатация этих решений что позволило реализовать комплексный подход к проектированию, созданию и внедрению информационных систем.

Очевидный положительный эффект, который дал централизованный подход, трудно переоценить потому, как в результате работы над проектами сформировались общие принципы разработки/внедрения информационных технологий и команды ИТ-специалистов, были определены основные сферы государственного управления, в которых функционируют существующие и перспективные информационные системы.

Государственный Центр Обработки Данных

Одной из основных частей ядра Единой Информационной Системы стал Государственный Центр Обработки Данных, реализованный на площадке государственного предприятия «Углетелеком». На базе ЦОД разрабатываются/тестируются ключевые модули информационных систем и ведутся научно-исследовательские работы в сфере информационных технологий, а также подготовка кадров для обеспечения всё возрастающей потребности в ИТ-специалистах.

Репозиторий исходного кода и конфигураций

Проект этой подсистемы, в рамках развития пространства информационного взаимодействия, призван систематизировать процесс производства программного обеспечения на основе открытых стандартов и технологий.

Сама по себе идея объединить кодовую базу, сценарии развертывания и прочие процедуры необходимые в разработке информационных систем, довольно тривиальна, но процесс её реализации на государственном уровне требует не только современного производительного оборудования, но и беспрецедентного уровня координации всех субъектов информационного взаимодействия (рисунок 2).



Рисунок 2 – Репозиторий исходного кода

Для осуществления координации такого уровня, требуется инструмент, регламентирующий все участки необходимой совместимости, а именно нормативно-правовой, организационной, семантической и технической, мы используем инструментарий, входящий в состав бесплатного и свободно распространяемого продукта GitLab, на базе этого продукта реализован базовый функционал для тестирования и развертывания, включая стандартные Linux-репозитории.

Создание такой системы на уровне любого государства, какое бы оно не было, это вопрос времени. Предпосылки к этому имеются в многих современных государствах и союзах государств, как раз для повышения уровня координации. Данный проект задумывается как большая площадка координации взаимодействия, которая будет включать в себя не только стандартные функционал любой облачной инфраструктуры, но поможет спроецировать объекты облачной инфраструктуры на нормативно-правовую, организационную и семантическую плоскость.

Все исходные коды разработок объединяются в единый реестр. Репозиторий исходного кода позволяет

вести совместную разработку сложных систем, а также открывает серьезные возможности экономии ресурсов на разработку информационных систем за счет повторного использования и совершенствования ранее разработанных систем

Государственный удостоверяющий центр

Следующей задачей в процессе развития информационного взаимодействия является обеспечение надежной аутентификации субъектов взаимодействия, то есть обеспечение доверия, для этого российские криптографические алгоритмы были интегрированы в фреймворк с открытым исходным кодом — BouncyCastle (рисунок 3). Архитектурно мы отказались от пространства дочерних удостоверяющих центров, которые неоправданно усложняют схему взаимодействия, государственный удостоверяющий центр, в нашем случае, выполняет только одну функцию — гарантирует что обладатель данной электронной подписи действительно является тем, кем является.

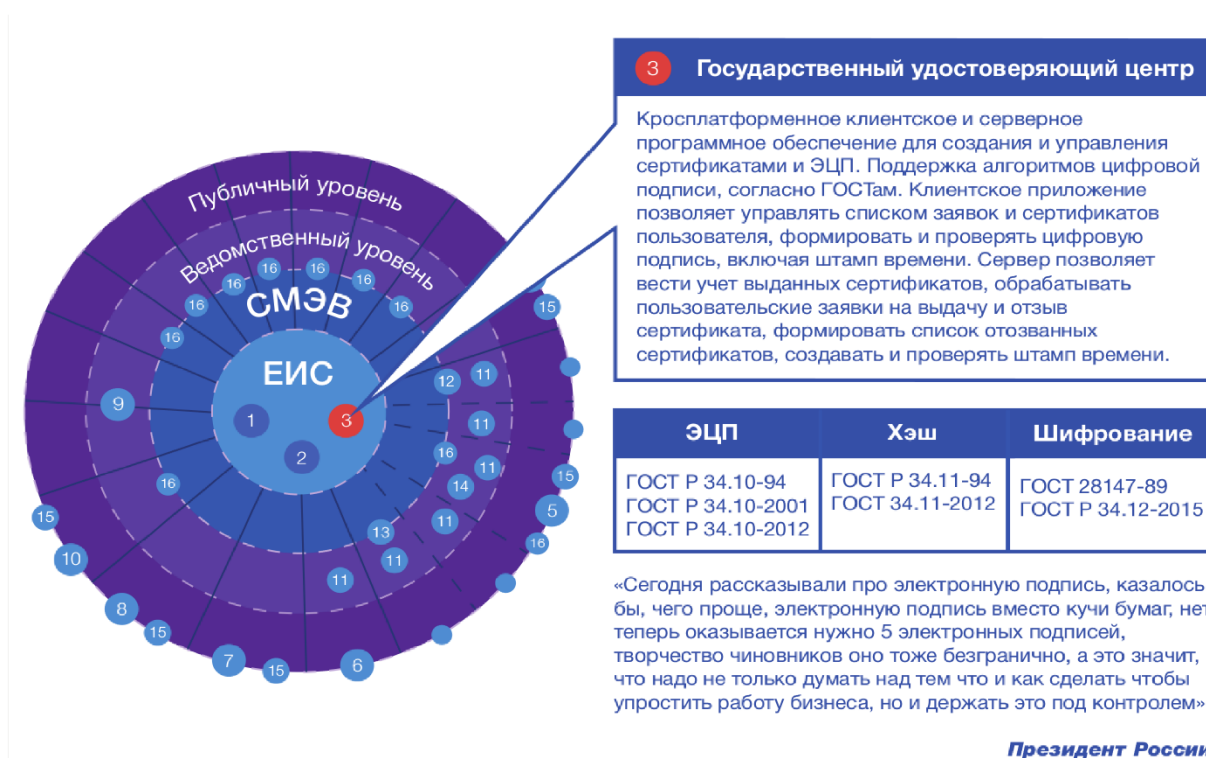


Рисунок 3 – Государственный удостоверяющий центр

Такая реализация требует минимальных ресурсов на нормативно-правовое и организационное обеспечение, при этом в совокупности с процессом производства данного продукта мы получаем широчайшие возможности по интеграции электронной подписи во все государственные органы, популяризации путем удешевления процесса её «производства» и агрегацией типовых решений обеспечивая её повсеместное применение.

Стоит отдельно отметить, что все модули задействованные в процессе создания и эксплуатации данного аппаратно-программного комплекса являются профессиональными, хорошо себя зарекомендовавшими, динамично развивающимися OpenSource-решениями.

Кроссплатформенное клиентское и серверное программное обеспечение для создания и управления сертификатами и ЭЦП. Поддержка алгоритмов цифровой подписи, согласно ГОСТам. Клиентское приложение позволяет управлять списком заявок и сертификатов пользователя, формировать и проверять цифровую подпись, включая штамп времени. Сервер позволяет вести учет выданных сертификатов, обрабатывать пользовательские заявки на выдачу и отзыв сертификата, формировать список отозванных сертификатов, создавать и проверять штамп времени.

Электронное сопровождение документооборота

Подсистема на базе Руби он Рэйлс, с определенным набором разработанных нами и открытых модулей, полностью соответствует методическим рекомендациям по ведению документооборота и дополнительно является хорошим инструментом для управления проектами, а также контролем времени. Так как сама система

модульная, а способ доставки обновлений автоматизирован и полностью реализован, процесс разработки нового функционала и скорость развертывания постоянно увеличивается (рисунок 4).



Рисунок 4 – Электронное сопровождение документов

Call-центр гуманитарной программы

Так же разработан на основе предыдущей подсистемы, и представляет из себя сервис обработки обращений и приема звонков. К системе подключено 10 ведомств которые задействованы в работе над обращениями граждан. В целом система является высоко производительной платформой на OpenSource-решениях.

Единая контактная панель

Единая контактная панель явилась частью воплощения мечты о государственном портале как о едином «мега» сервисе, есть мнение что предел развития информационных систем — это создание платформы для их объединения, о платформе я скажу в следующем слайде, а этот «виджет» должен объединить государственные ресурсы повысив уровень предоставления информации на порядок (рисунок 5). Этот инструмент гарантирует что гражданин всегда попадет на рабочий сайт с актуальной информацией.



Рисунок 5 – Единая контактная панель для государственных органов власти

Целями Единой контактной панели для государственных органов власти являются: централизованный сбор и обработка обращений граждан, прием звонков, текстовых обращений с веб-сайта и дальнейшее распределение заявок по ответственным участникам «Гуманитарной программы по воссоединению народа Донбасса», контактную информацию о министерствах и ведомствах. Данная система предлагается, как централизованное решение по распространению открытой информации о государственных органах власти.

Проект - Linux Cluster

Необходимой частью государственной инфраструктуры являются ведомственные информационные системы. Которым в свою очередь нужны стандартные возможности аутентификация, авторизация и управление конфигурациями (рисунок 6).

Для обеспечения этого нами разработаны и внедрены два варианта Линукс-Кластера с полным набором инструментов по администрированию, конфигурированию и мониторингу.

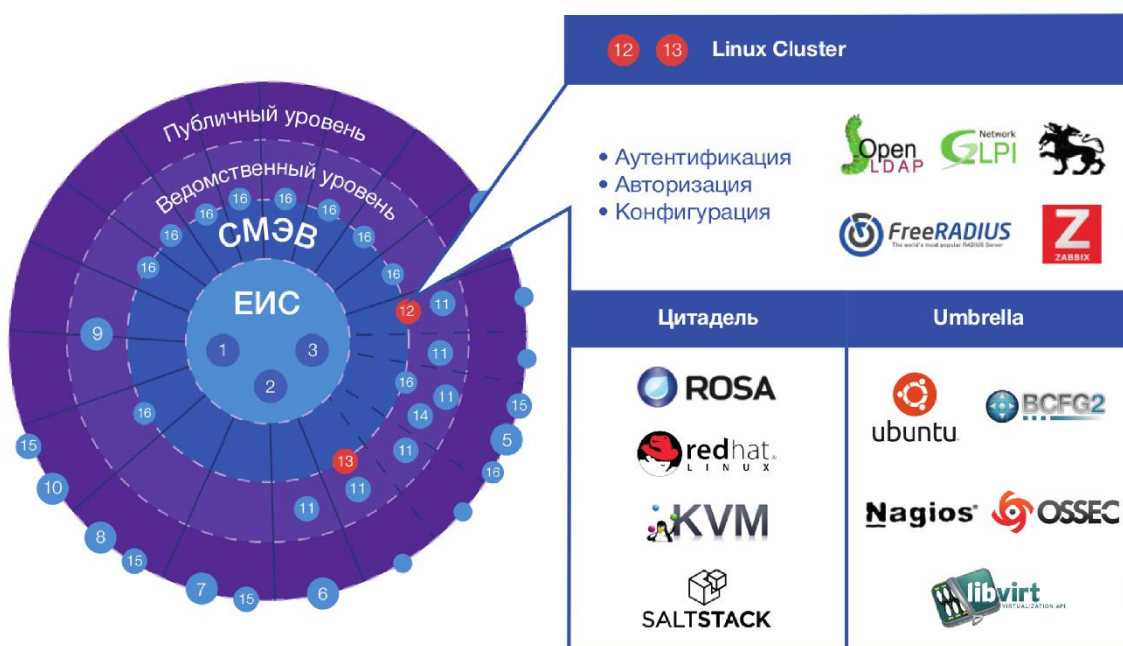


Рисунок 6 – Линукс-Кластер

Архитектура, заложенная в этих двух системах, позволяет глубоко модернизировать не только внутриведомственные сервисы, но повысить уровень информационной безопасности в целом. Базовые решения уже обеспечивают «Единый вход», журналирование и, например, контроль подключаемых устройств на уровне Кластера.

Государственный реестр общеобязательного социального страхования ДНР

Данная система разработана по указу Главы Донецкой Народной Республики в целях снижения административной нагрузки на граждан и ведомства, автоматизировав обработку документов, которые подаёт работодатель в Министерство Налогов и сборов, а Пенсионный фонд получает данные по единому социальному взносу.

Целью системы является синхронизация данных Пенсионного фонда с Министерством доходов и сборов ДНР и последующая их обработка. Учет проверок. Организация реестров по предпринимателям, застрахованным лицам, страхователям-должникам, проверкам и запросам. Формирование отчетов (рисунок 7).



Рисунок 7 – Государственный реестр общеобязательного государственного социального страхования ДНР

Система межведомственного электронного взаимодействия (далее-СМЭВ)

Подходя к решению задачи обеспечения высокопроизводительного, надежного и легко масштабируемого сервиса обмена сообщениями между всеми участниками СМЭВ вне зависимости от технологии реализации информационных систем, мы остановились на решении брокера сообщений RabbitMQ, защита каналов связи OpenVPN + ГОСТ-криптография, данная платформа, по нашему мнению, сможет обеспечить необходимую скорость обработки и масштабируемость системы (рисунок 8).



Рисунок 8 – Система межведомственного электронного взаимодействия

Данный проект находится на стадии научно-исследовательских работ, построения моделей, выбор формата обмена и т.п., но существует уверенность, что с запуском удостоверяющего центра Министерство связи далеко шагнет и в рамках данного проекта, выведя информационное взаимодействие на новый уровень.

Заключение

Информационно – технологические нововведения в государственном и муниципальном управлении имеют комплексный характер, связаны с одновременным и согласованным использованием информационных, организационных, правовых, социально–психологических, кадровых, технических, логико – семантических факторов. Все это требует комплексного подхода, качественного изменения как системы работы с информацией, так и функциональной и организационной структур управления, состава и структуры всей управленческой деятельности.

«Для выхода на новый уровень развития экономики, социальных отраслей нам нужны собственные передовые разработки и научные решения». (с) Президент России.

Литература

1. Голенищев Э. П. Информационное обеспечение систем управления. — Ростов-на-Дону: Феникс, 2003.
2. Иванова Е.И. Информатизация государственного управления и проблемы национальной безопасности // Угрозы и безопасность. — 2008. — № 2.
3. Королев А. Н., Плешакова О. В. Об информации, информационных технологиях и о защите информации. Постатейный комментарий к Федеральному закону. — М.: Юстицинформ, 2007. — 128 с. — (Библиотека журнала «Право и экономика». Комментарий специалиста). — ISBN 5-7205-0791-4.
4. Федеральный закон Российской Федерации от 27 июля 2010 года № 210-ФЗ «Об организации предоставления госуслуг».
5. Постановление Правительства Российской Федерации № 697 «О единой системе межведомственного электронного взаимодействия».
6. Приказ Минкомсвязи РФ от 27.12.2010 года № 190 «Об утверждении технических требований к взаимодействию информационных систем в единой системе межведомственного электронного взаимодействия».

Гущин В.В. Государственная информатизация с нуля. В статье рассмотрено состояние и развитие информатизации государственных органов власти Донецкой Народной Республики Освещен тот факт, что перспективы будут строиться вокруг концепций в области информационных технологий, принятых в Российской Федерации и современных подходов в производстве программного обеспечения.

Ключевые слова: Информатизация, государственные органы власти, централизованное хранилище данных, электронный документооборот, единый репозиторий исходного кода, система межведомственного электронного взаимодействия, контактная панель, государственный удостоверяющий центр.

Gushchin V. V. State informatization from scratch. The article considers the state and development of Informatization of public authorities of the Donetsk people's Republic Highlights the fact that the prospects will be built around the concepts in the field of information technology adopted in the Russian Federation and modern approaches in the production of software.

Keywords: Informatization, state authorities, centralized data storage, electronic document flow, single source code repository, system of interdepartmental electronic interaction, contact panel, state certification center.

Информационная безопасность. Перспективы и вызовы.

Джура Е.С.

Министерство связи Донецкой Народной Республики
dzhura.gs@minsvyazdnr.ru

Джура Е.С. Информационная безопасность. Перспективы и вызовы. В статье рассмотрены проблемы и тенденции обеспечения информационной безопасности в современных реалиях. Приведены междисциплинарные связи. Рассмотрены наиболее чувствительные сферы с точки зрения государственного управления. Приведена структура Государственной системы информационной безопасности в Российской Федерации, изложены методы и практики обеспечения информационной безопасности в России, а также расставлены акценты в проблематике и путях решения выявленных проблем.

Ключевые слова: Сетевые атаки, безопасность информационных систем, безопасность данных, нормативно-правовое обеспечение, государственное управление, защита информации.

Введение

Уровень информационно-аналитического обеспечения современного государства является важнейшей характеристикой его деятельности и оказывает фундаментальное влияние на все процессы развития общества. Использование информационных технологий имеет огромное значение для повышения конкурентоспособности государства и повышения эффективности работы органов государственной власти. Ввиду роста охвата и развития технологий во всех сферах деятельности государства и общества, обеспечение безопасности государственных информационных ресурсов является стратегической и фундаментальной задачей любого современного государства.

Сетевые атаки сегодня становятся несложными в реализации и крайне разнообразными. Этому способствуют такие факторы, как всестороннее и повсеместное проникновение Интернета и широкое распространение простых в использовании операционных систем и программного обеспечения [1]. Сетевые атаки настолько же разнообразны, как и информационные системы (далее – ИС), против которых они направлены. Нарушитель, в процессе осуществления атаки, может нарушить конфиденциальность, целостность или доступность передаваемой информации, а также нанести урон как общей работоспособности ИС, так и отдельным ее компонентам.

Компьютерные атаки сегодня могут провоцировать техногенные аварии, экологические катастрофы, социальные потрясения и наносить непоправимый ущерб государству.

Существуют различные подходы к пониманию термина «информационная безопасность». Многогранность и многоуровневость данного направления обуславливает необходимость рассмотрения данного термина с многих ракурсов. Важность безопасности информации сегодня очевидна для большинства отраслей и сфер деятельности. Этот фактор обуславливает ее разветвленность и интегрированность с другими сферами и науками. Сегодня информационная безопасность – это и юриспруденция (нормативное регулирование) и психология (социальная инженерия) и безусловно такие направления, как системное администрирование и программирование. Говорить о важности информационной безопасности для военно-промышленного комплекса государства сегодня излишне.

Меры и средства обеспечения информационной безопасности

Наряду с безусловными преимуществами от внедрения различных ИС в органы и структуры государственного управления, возникают обоснованные проблемы с безопасностью информационных ресурсов. По этой причине создание эффективной системы обеспечения информационной безопасности является одним из наиболее приоритетных направлений развития государства в целом.

Система информационной безопасности – это совокупность мер, средств, органов и лиц, обеспечивающих безопасность информации на основе единых нормативно-правовых норм.

Среди основных мер для защиты информации можно выделить следующие:

- идентификация и аутентификация субъектов и объектов доступа;
- управление доступом субъектов к объектам доступа;
- защита программной среды;
- защита машинных носителей информации;

- регистрация событий безопасности;
- обнаружение (предотвращение) вторжений;
- комплексная антивирусная защита;
- контроль (анализ) защищенности информации;
- обеспечение целостности информации в ИС;
- обеспечение доступности информации;
- защита среды виртуализации;
- защита технических средств;
- защита средств и систем связи и передачи данных ИС.

Для реализации вышеназванных мер обычно используются методы и средства обеспечения информационной безопасности (далее – ИБ), приведенные на рисунке 1 [2]:



Рисунок 1 - Методы и средства обеспечения ИБ

Все рассматриваемые средства защиты разделены на формальные (выполняющие защитные функции по заранее строго предусмотренной процедуре без непосредственного участия человека) и неформальные (определяются целенаправленной деятельностью человека либо регламентируют эту деятельность). Так как в использовании неформальных методов и средств государственные органы должны выполнять ведущую роль, в данном исследовании особое внимание уделено их применению. Рассмотрим данные методы и средства:

1. Регламентация – создание таких условий автоматизированной обработки, хранения и передачи защищаемой информации, при которых нормы и стандарты по защите выполняются в наибольшей степени.

2. Принуждение – метод защиты, при котором пользователи и персонал сопровождающий ИС вынуждены соблюдать правила обработки, передачи и использования защищаемой информации под угрозой материальной, административной или уголовной ответственности.

3. Побуждение – метод защиты, побуждающий пользователей и сопровождающий ИС не нарушать установленные порядки за счет соблюдения сложившихся моральных и этических норм.

4. Организационные средства – осуществляют регламентацию производственной деятельности в ИС и взаимоотношений исполнителей на нормативно-правовой основе таким образом, что разглашение, утечка и несанкционированный доступ к конфиденциальной информации становятся невозможными или существенно затрудняются за счет проведения организационных мероприятий. Комплекс данных мер реализуется группой ИБ, но должен находиться под контролем первого руководителя.

5. Законодательные средства – определяются нормативно-правовыми актами государства, регламентирующими правила пользования, обработки и передачи информации ограниченного доступа и устанавливаются меры ответственности за нарушение этих правил.

6. Морально-этические средства защиты – включают всевозможные нормы поведения, которые традиционно сложились ранее, складываются по мере распространения ИС и технологий в государстве и в мире или специально разрабатываются. Морально-этические нормы могут быть неписанные (например, честность) либо оформленные в некий свод (устав) правил или предписаний. Эти нормы, как правило, не являются законодательно утвержденными, но поскольку их несоблюдение приводит к падению престижа организации, ввиду чего они считаются обязательными для исполнения.

7. Противодействие атакам вредоносных программ предполагает комплекс мер организационного характера и использование антивирусных программ. Цели принимаемых мер – это уменьшение вероятности инфицирования ИС, выявление фактов заражения системы; уменьшение последствий информационных инфекций, локализация или уничтожение вирусов; восстановление информации в ИС.

Таким образом, решение проблем обеспечения безопасности государственных информационных систем принято вести в трех направлениях:

1. Унификация и централизация ИС государственных органов.
2. Переход государственных ИС на единые нормы и стандарты.
3. Повышение технологического уровня средств обеспечения ИБ.

Проблематика, анализ опыта

К основным проблемам в Донецкой Народной Республике с субъективной точки зрения можно подойти, выделив следующие:

1. проблемы нормативно-правового регулирования (отсутствие требований и явных критериев соответствия для государственных органов и бизнеса);
2. отсутствие централизованных подходов (централизация рождает управляемость, а управляемость в свою очередь порядок);
3. проблемы кадрового характера (отрицательно влияет как фактор утечки кадров, так и сложности с поддержанием необходимого уровня образования);
4. отсутствие площадок для обмена опытом (необходимость постоянного обучения и переобучения).

Ввиду наличия стратегических планов развития в Донецкой Народной Республике, опыт Российской Федерации для нас является приоритетным, в том числе по причине того, что с многих точек зрения является передовым. На рисунке 2 изображена обобщенная государственная система обеспечения информационной безопасности РФ.

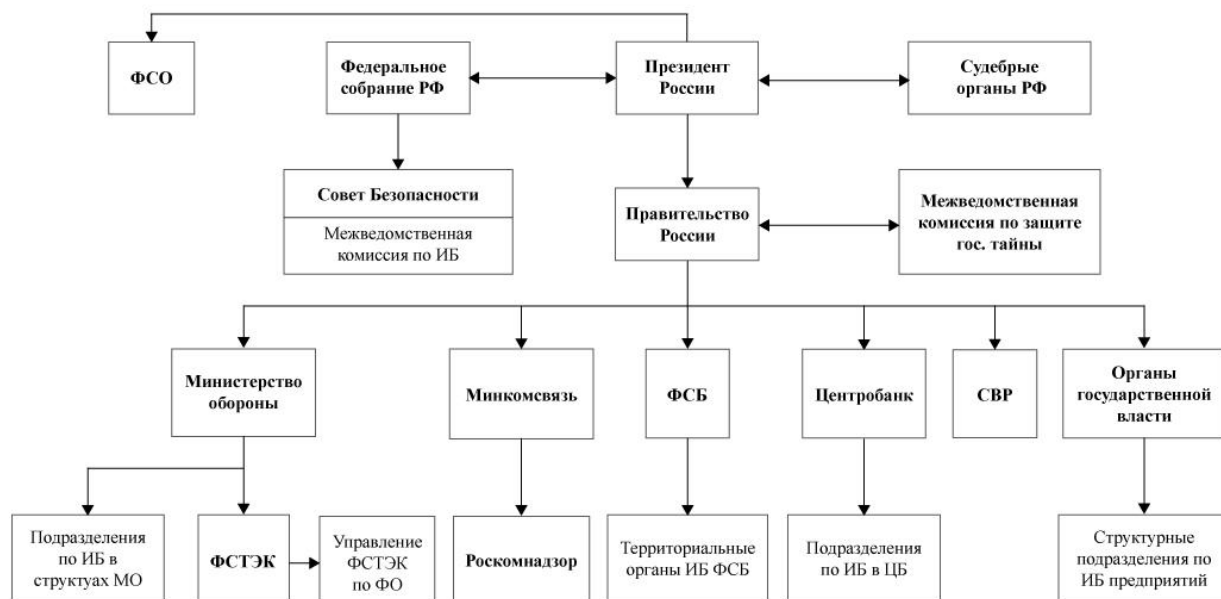


Рисунок 2 - Государственная система обеспечения информационной безопасности Российской Федерации

Сравнение с системой обеспечения информационной безопасности в Донецкой Народной Республике пока уместным назвать сложно, ввиду чего стоит отметить не только целесообразность, но и необходимость пересмотра структуры органов государственного управления как с точки зрения распределения компетенций, так и относительно имплементации изложенных в законодательстве Донецкой Народной Республики положений.

Главными тенденциями развития информационной безопасности в Российской Федерации сегодня можно назвать:

- поступательный и планомерный рост надзора и контроля в области ИБ;
- формирование требований к объектам критической информационной инфраструктуры (далее – КИИ);
- интеграция органов государственной власти и корпораций с государственной системой обнаружения, предупреждения и ликвидации последствий компьютерных атак (далее – ГосСОПКА) [3].

Отраслями, относящимися к КИИ, являются: здравоохранение, наука, транспорт, связь, энергетика, банки, финансы, топливно-энергетический комплекс, атомная энергетика, оборонная, ракетно-космическая, горнодобывающая, металлургическая и химическая отрасли промышленности.

Обращая внимание на тенденции развития нормативно-правового регулирования информационной безопасности в РФ, стоит обратить внимание на некоторые вошедшие в обиход термины [4]:

- 1) Субъект КИИ – любое лицо (орган власти, юридическое лицо, индивидуальный предприниматель), которому принадлежит хотя бы одна ИС, используемая в одной из вышеперечисленных отраслей.
- 2) Объект КИИ – любая ИС, принадлежащая субъекту КИИ.
- 3) НКЦКИ – Национальный координационный центр по компьютерным инцидентам.
- 4) ГосСОПКА – глобальная система сбора и обмена информацией о компьютерных атаках на территории РФ.

Законодательство РФ в сфере обеспечения безопасности КИИ представлено на рисунке 3 [5].

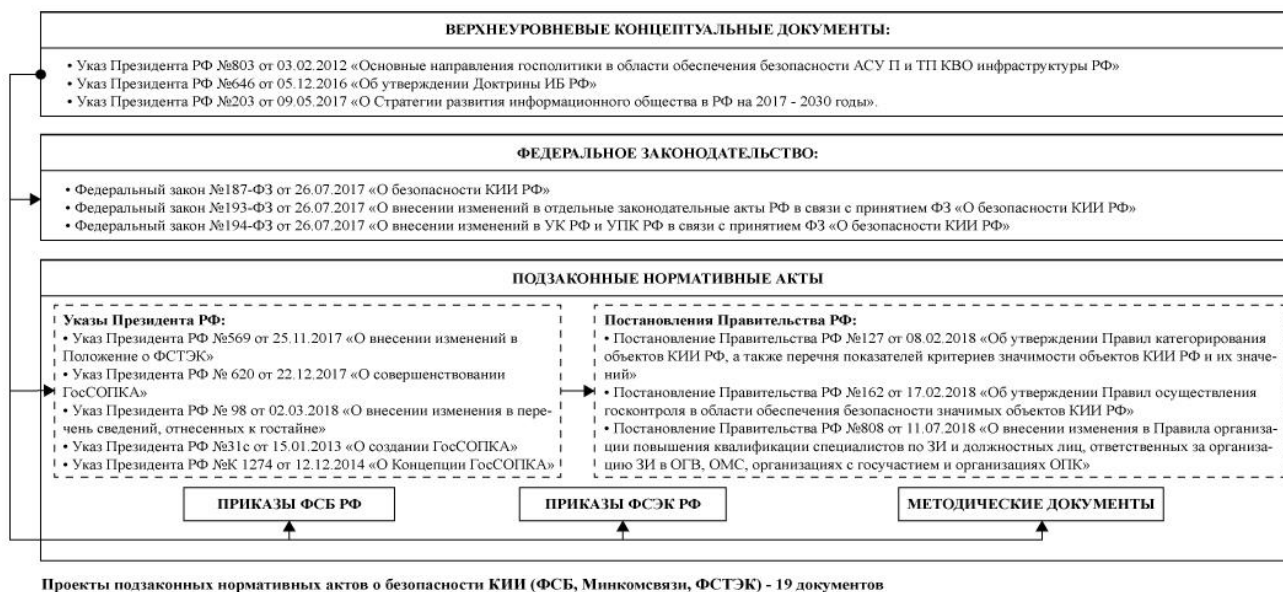


Рисунок 3 - Нормативные документы о безопасности КИИ РФ

Разветвленность и развитость нормативного регулирования и контроля за его исполнением в Российской Федерации является передовым с точки зрения перенимания опыта. Из рисунка 3 видно, что платформа для тенденций сегодняшнего развития была заложена еще в 2012 году, а в 2017 году достигло кульминации.

К основным задачам ГосСОПКи относят:

- прогнозирование инцидентов в области обеспечения информационной безопасности РФ;
- обеспечение взаимодействия владельцев информационных ресурсов РФ, операторов связи, иных субъектов, осуществляющих лицензируемую деятельность в области защиты информации, при решении задач, касающихся обнаружения, предупреждения и ликвидации последствий компьютерных атак;
- осуществление контроля степени защищенности информационных ресурсов РФ;
- установление причин компьютерных инцидентов, связанных с функционированием информационных ресурсов РФ.

На рисунке 4 изображены категории компьютерных атак и инцидентов, охватываемые системой ГосСОПКа [6].



Рисунок 4 - Категории компьютерных атак и инцидентов, охватываемые средствами ГосСОПКи

Средствами ГосСОПКи являются:

- инвентаризация информационных ресурсов;
- выявление уязвимостей;
- анализ угроз;
- повышение осведомленности персонала и пользователей;
- прием сообщений о возможных инцидентах;
- обнаружение компьютерных атак;
- анализ данных о событиях безопасности;
- регистрация инцидентов;
- реагирование на инциденты и ликвидация их последствий;
- расследование инцидентов;
- анализ результатов устранения последствий инцидентов.

На рисунке 5 изображена организационная иерархическая структура системы ГосСОПКА [6].

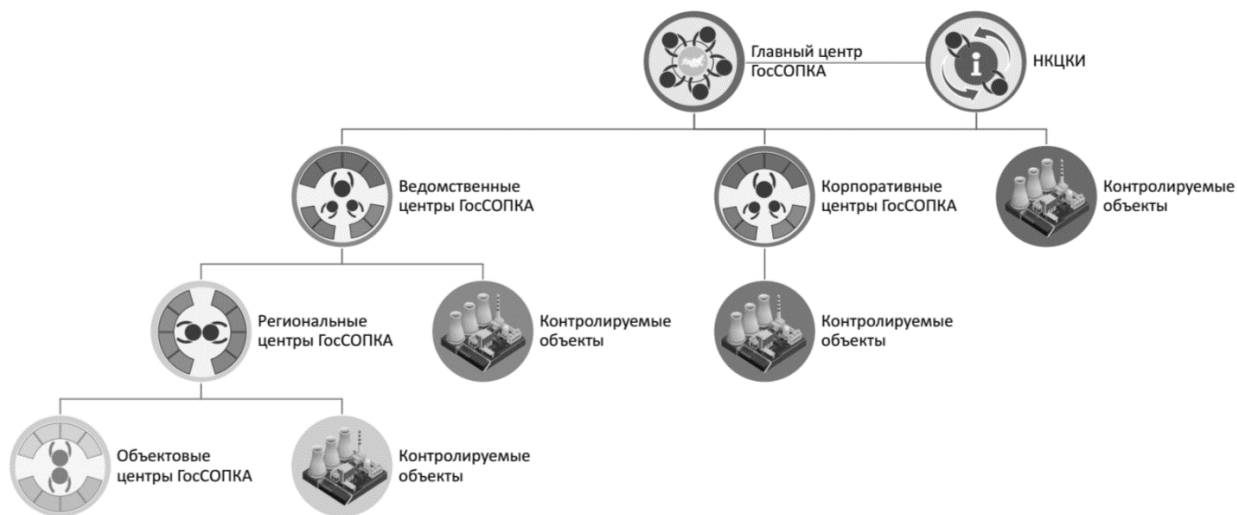


Рисунок 5 – Организационная иерархическая структура системы ГосСОПКА

Масштаб и охват сфер и областей является крайне внушительным с той точки зрения, что в настоящее время создаются сегменты более чем в 20 министерствах и государственных предприятиях. Сегодня уже подписаны соглашения и проходят интеграцию с системой ГосСОПКА такие организации, как Минобороны и Минэнерго России, «Роскосмос», «Росатом», «Ростех» и многие другие.

На рисунке 6 изображены компоненты безопасности КИИ и сферы компетенции [7].

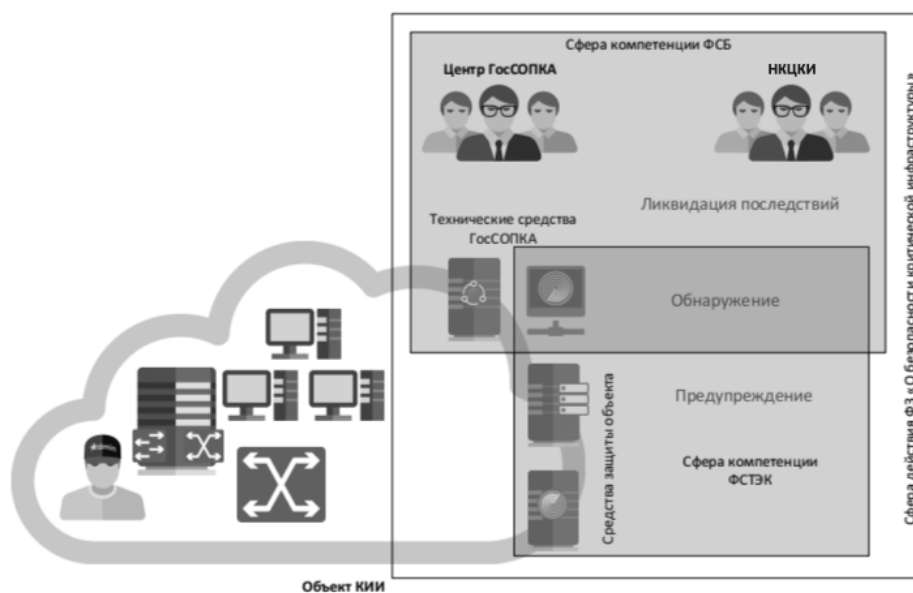


Рисунок 6 - Компоненты безопасности КИИ и основные зоны ответственности

Основные направления выполнения требований в КИИ это:

- обеспечение безопасности государственной тайны;
- обеспечение безопасности государственных информационных систем;
- обеспечение безопасности информационных систем персональных данных;
- обеспечение безопасности информационно-телекоммуникационных сетей.

Вышеприведенные перечни подтверждают актуальность проведенного исследования и с субъективной точки зрения могут быть использованы как стратегические ориентиры для развития сферы информационной безопасности в Донецкой Народной Республике.

Подводя итоги можно расставить следующие акценты:

1. Актуальность защиты информации для Республики растет с каждым днем и требует незамедлительного пересмотра подходов к ее обеспечению.
2. Уровень нормативного регулирования в Республике минимален, без необходимых нормативно-правовых актов невозможна ни реализация требований, ни надзор, ни контроль за их исполнением.
3. С учетом максимальной чувствительности сферы ИБ, рассмотрение возможности создания профильных межведомственных комиссий и управлений для государственных органов власти может стать не только решением многих задач, а еще и платформой для формирования полноценной стратегии как в области информационной, так и в целом национальной безопасности Республики.

Заключение

Ввиду интенсивного перехода всех родов взаимодействий государственного уровня на инфокоммуникационную платформу, задача обеспечения безопасности информации приобретает общегосударственный масштаб. Пути решения этой задачи лежат в трех направлениях: повышение технологического уровня средств обеспечения ИБ, унификация и централизация ИС государственных органов и переход государственных ИС на единые стандарты.

Мировой опыт совершенствования государственного управления подтверждает, что информационно-аналитическое обеспечение следует рассматривать как одно из стратегических направлений повышения эффективности деятельности государства во всех сферах, а защита государственных информационных систем и ресурсообеспечение рассматривать как стратегический вопрос обеспечения национальной безопасности.

Литература

7. Эволюция угроз информационной безопасности в бизнес-среде [Электронный ресурс]. – Режим доступа: https://cdn.securelist.ru/files/2015/12/KSB_2015_business_threats_ru.pdf.
8. Глобальное исследование утечек конфиденциальной информации в I полугодии 2018 года Методы и средства обеспечения информационной безопасности [Электронный ресурс]. – Режим доступа: <http://900igr.net/prezentacija/informatika/metody-obespechenija-bezopasnosti-v-operatsionnykh-sistemakh-264547/metody-i-sredstva-obespechenija-informatsionnoj-bezopasnosti-8.html>.

9. Все нормативные документы по КИИ [Электронный ресурс]. – Режим доступа: https://zlonov.ru/kii/all_laws/.
10. СОИБ. Анализ. ГосСОПКА [Электронный ресурс]. – Режим доступа: <https://www.securitylab.ru/blog/personal/sborisov/131813.php#>.
11. Средства ГосСОПКА. Переводим терминологию [Электронный ресурс]. – Режим доступа: <https://habr.com/company/solarsecurity/blog/422819/>.
12. ГосСОПКА: основные задачи, функции и реализуемые мероприятия [Электронный ресурс]. – Режим доступа: <http://it-security.admin-smolensk.ru/files/367/gossopka.pdf>.
13. Безопасность КИИ: что нас ждет в 2018 году [Электронный ресурс]. – Режим доступа: <https://www.ptsecurity.com/ru-ru/research/webinar/284799/>.

Дзюра Г.С. Информационная безопасность .Перспективы и вызовы. В статье рассмотрена актуальность информационной безопасности в реалиях сегодняшнего дня. Приведены междисциплинарные связи. Рассмотрены наиболее чувствительные сферы и области с точки зрения государственного управления. Приведена структура Государственной системы информационной безопасности в Российской Федерации, а также тенденции и практики обеспечения информационной безопасности в РФ.

Ключевые слова: Сетевые атаки, безопасность информационных систем, безопасность данных нормативно-правовое обеспечение, государственное управление, защита информации.

Dzhura G. S. Information security. Prospects and challenges. The article considers the relevance of information security in the realities of today. Interdisciplinary connections are given. The most sensitive areas and areas from the point of view of public administration are considered. The structure of the State system of information security in the Russian Federation, as well as trends and practices of information security in the Russian Federation.

Keywords: Network attacks, security of information systems, data security, regulatory and legal support, public administration, information protection

Автоматизированная информационная система обработки вызовов скорой медицинской помощи

Пьянков К.А.

Министерство связи Донецкой Народной Республики
pyankov.k@minsvyazdnr.ru

Пьянков К.А. Автоматизированная информационная система обработки вызовов скорой медицинской помощи. В статье представлен краткий обзор выполненного проекта по автоматизации деятельности станции скорой медицинской помощи. Описаны функциональность системы и основные компоненты, использовавшиеся для построения системы.

Ключевые слова: информационная система, микросервисная архитектура, облачная инфраструктура, скорая медицинская помощь, диспетчер, территориально-распределенная система.

Введение

Современные информационные технологии и развитие средств связи позволяют значительно улучшить основные показатели деятельности службы скорой медицинской помощи.

Работа станций скорой медицинской помощи нуждается в автоматизации для обеспечения эффективной информационной поддержки процессов деятельности, координации и управления силами и средствами служб с использованием передовых информационных и навигационных технологий, а именно:

- повышения оперативности и эффективности работы за счет обеспечения диспетчерского персонала полной и достоверной информацией о состоянии выездных бригад на любой момент времени;
- повышения производительности работы выездных бригад и сокращение времени прибытия выездных бригад на место вызова за счет возможности оперативного выбора ближайшей к месту происшествия бригады необходимого профиля и оптимизации маршрута ее следования;
- повышения производительности работы выездных бригад за счет уменьшения средней дальности поездки при обслуживании вызовов;
- повышения скорости поиска и предоставления информации согласно поступившим запросам;
- совершенствования технологии работы и информационного взаимодействия между различными функциональными подразделениями службы скорой медицинской помощи, задействованных в процессе оказания экстренной медицинской помощи населению.

Задачи управления службой скорой медицинской помощи

Основными функциями станции скорой медицинской помощи являются:

- круглосуточное оказание экстренной и качественной медицинской помощи, в соответствии с медицинскими стандартами, заболевшим и пострадавшим, находящимся вне лечебного учреждения, при катастрофах и стихийных бедствиях;
- осуществление своевременных транспортировок больных, в том числе инфекционных, пострадавших и рожениц, которые нуждаются в экстренной стационарной помощи;
- информирование муниципальных органов управления здравоохранением обо всех чрезвычайных ситуациях и несчастных случаях в зоне обслуживания;
- обеспечение равномерного комплектования выездных бригад медицинским персоналом по всем сменам и полное их обеспечение медицинским оснащением.

Эффективное управление станции скорой медицинской помощи в современных условиях является сложной организационно-технической проблемой, т.к. станция относится к классу территориально-распределенных систем, в которой основные функции в пространстве населенного пункта выполняет медицинский персонал, объединенный в подразделения станции скорой медицинской помощи. Также сложность моделирования добавляет нерегулярность поступления вызовов от населения.

Для своевременной передачи информации о вызовах необходима разработка эффективной технологии взаимодействия станции скорой медицинской помощи с поликлиниками и другими лечебно-профилактическими учреждениями, а также с другими оперативными службами (милиция, пожарная часть,

горноспасательная служба и другие) для сокращения времени оказания медицинской помощи во время ликвидации аварийных и чрезвычайных ситуаций.

Оптимизацию среднего времени регистрации вызова требуется проводить за счет совершенствования методик приема вызова диспетчерами и использования при современных средств связи.

При планировании обслуживания вызовов бригадами скорой медицинской помощи необходимо определить минимальный по времени путь движения бригады, приоритет вызова и соответствие профиля заболевания больного со специализацией бригады. При планировании госпитализации больных требуется определить минимальный по времени путь движения бригады к лечебному учреждению, наличие свободных мест и соответствие профиля заболевания больного профилю этого учреждения.

Также важной компонентой оптимизации процессов на станции скорой медицинской помощи является введение электронного документооборота, что позволит осуществлять регистрацию и контроль исполнения документов в электронном виде, оперативно обмениваться этими документами, автоматически формировать необходимые отчеты.

Краткая характеристика объекта автоматизации

Служба экстренной (скорой) медицинской помощи Республиканского Центра экстренной медицинской помощи и медицины катастроф Донецкой Народной Республики (далее – РЦЭМП и МК ДНР) включает в себя 6 обособленных структурных подразделений – станций скорой медицинской помощи (далее – СМП) в городах: Донецк, Макеевка, Горловка, Енакиево, Торез и Новоазовск; 35 подстанций скорой медицинской помощи; 23 пункта постоянного и временного базирования.

В составе РЦЭМП и МК ДНР функционируют следующие подразделения:

- служба выездной консультативной специализированной экстренной медицинской помощи;
- служба экстренной (скорой) медицинской помощи.

Выездная консультативная специализированная экстренная медицинская помощь предоставляется пациентам, находящимся в стационарах учреждений здравоохранения городов и районов Донецкой Народной Республики по 27 лечебным специальностям круглосуточно.

Учреждения и подразделения скорой медицинской помощи Донецкой Народной Республики оказывают круглосуточную скорую медицинскую помощь больным и пострадавшим, находящимся вне медицинской организации, в пути следования в медицинское учреждение, при непосредственном обращении в учреждения и подразделения СМП, по заявкам медицинских работников СМП, оказывающих первичную медико-санитарную помощь.

Объектом автоматизации являются процессы приема, обработки и накопления информации о вызовах, поступающих на номер «103».

АИС «СМП»

Министерством связи Донецкой Народной Республики для Министерства здравоохранения Донецкой Народной Республики разработана автоматизированная информационная система обработки вызовов скорой медицинской помощи (далее – АИС «СМП»).

АИС «СМП» предназначена для автоматизации деятельности диспетчерской службы в обособленном структурном подразделении Республиканского центра экстренной медицинской помощи и медицины катастроф Донецкой Народной Республики, а именно обеспечения реализации следующих процессов:

6. приём и обработка телефонных звонков, поступающих на номер «103» по г. Донецку фельдшером по приему вызовов;
7. направление бригад скорой медицинской помощи на поступивший вызов на номер «103»;
8. автоматическое формирование формы первичной учетной документации № 109/у «Карта вызова скорой медицинской помощи».

Внедрение АИС «СМП» основано на комплексном использовании современных средств разработки, в том числе и сторонних систем мониторинга, средств связи, обеспечивающих передачу данных от контролируемых транспортных средств, сбор навигационных данных о движении транспортных средств, обмен оперативными данными между элементами и структурными компонентами системы. Вся работа пользователей с модулями осуществляется через клиент-приложение.

АИС «СМП» имеет микросервисную архитектуру и построена на облачной инфраструктуре (IaaS) к которой подключаются все клиентские автоматизированные рабочие места. Общая инфраструктура включает в себя следующие компоненты:

- Kubernetes – открытое программное обеспечение для автоматизации развёртывания, масштабирования и управления контейнеризированными приложениями. Поддерживает основные технологии контейнеризации, включая Docker, kkt, также возможна поддержка технологий аппаратной виртуализации;
- Docker – программное обеспечение для автоматизации развёртывания и управления

приложениями в среде виртуализации на уровне операционной системы. Позволяет «упаковать» приложение со всем его окружением и зависимостями в контейнер, который может быть перенесён на любую Linux-систему, а также предоставляет среду по управлению контейнерами;

- PG SQL – компонент основной СУБД;
- Zabbix – свободная система мониторинга и отслеживания статусов сервисов компьютерной сети;
- Jenkins – компонент автоматизации процессов управления компонентами системы;
- Nexus – хранилище артефактов проекта;
- GitLab – хранилище исходных кодов проекта, система управления репозиториями кода для Git;
- Backup Server – сервер для хранения суточных копий базы данных АИС «СМП» и архивов записей звонков;
- Asterisk – сервис компьютерной телефонии. Обеспечивает коммутацию, запись и управление телефонными вызовами;
- Open Street Map – картографический сервис, содержащий подробную свободную карту мира;
- Geo Ritm – специальное программное обеспечение для решения задач мониторинга транспорта. Служит источником данных о местоположении (координатах) машин скорой медицинской помощи;
- VPN – точка входа для всех подключений.

Общая схема взаимодействия компонент АИС «СМП» приведена на рисунке 1.

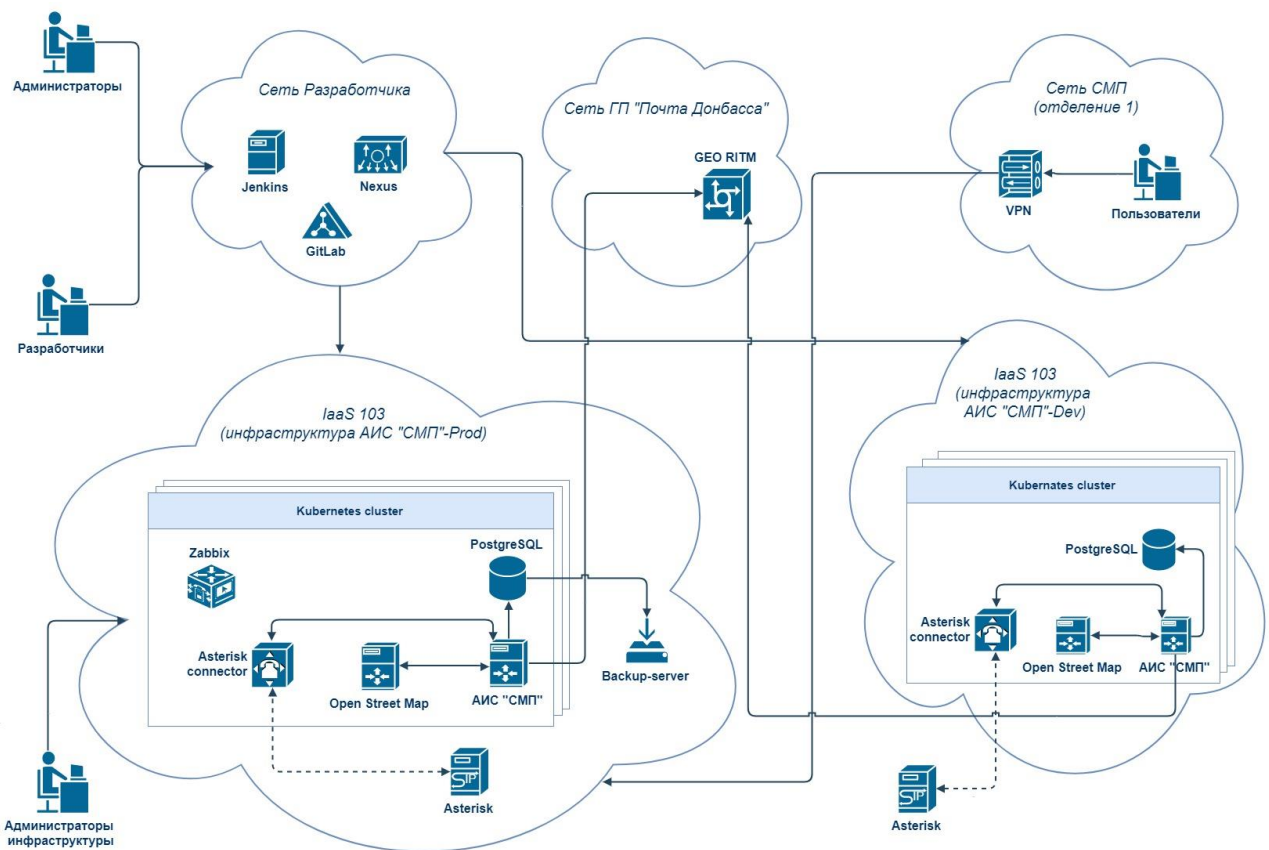


Рисунок 1 – Общая схема взаимодействия компонент АИС «СМП»

Автоматизация процессов приёма и обработки телефонных вызовов, поступающих на номер «103» включает в себя:

- системную и визуальную сортировку обрабатываемых вызовов (последний поступивший вызов попадает в конец очереди на обработку вызова);
- визуализацию обрабатываемых вызовов (экстренные вызовы помечаются в более яркий цвет и перемещаются в первую очередь обработки вызова);
- автоматическое формирование формы 109/у;
- поиск бригады для обработки вызова;
- сбор и накопление информации о вызовах в электронном виде;

- запись и повторное прослушивание всех принятых и совершенных вызовов на номер «103».

Для взаимодействия с АИС «СМП» созданы типовые (базовые) конфигурации профилей пользователей (автоматизированные рабочие места, далее – АРМ):

- АРМ администратора;
- АРМ оператора;
- АРМ старшего врача смены;
- АРМ диспетчера;
- АРМ медицинского инженера.

АИС «СМП» обеспечивает пользователям возможность выполнять определенную последовательность действий при обработке вызова, поступившего на номер «103» (диаграмма деятельности обработки вызова приведена на рисунке 2, более расширенная схема последовательности действий обработки вызова, поступившего на станцию скорой медицинской помощи приведена на рисунке 3).

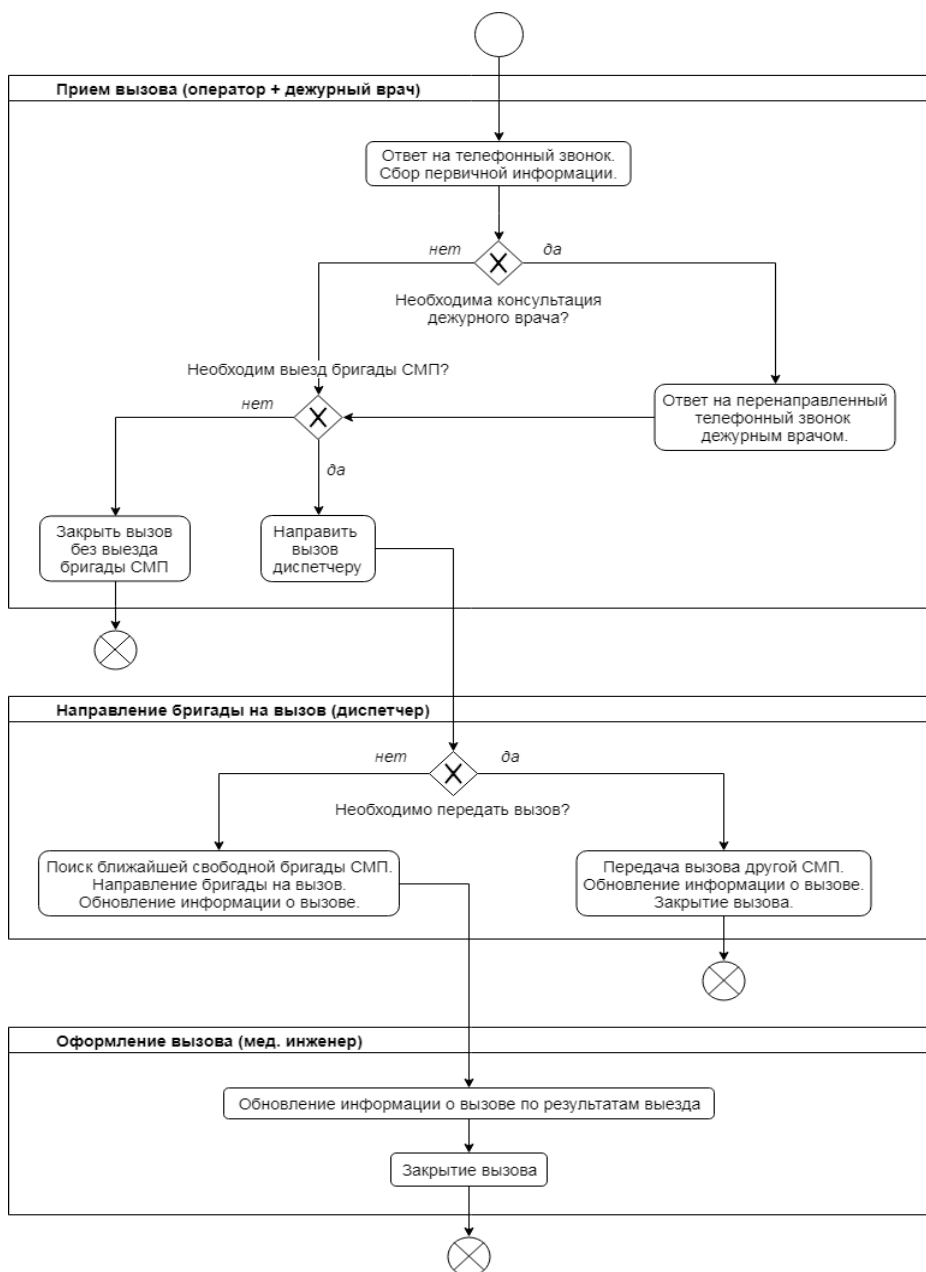


Рисунок 2 – Диаграмма деятельности обработки вызова

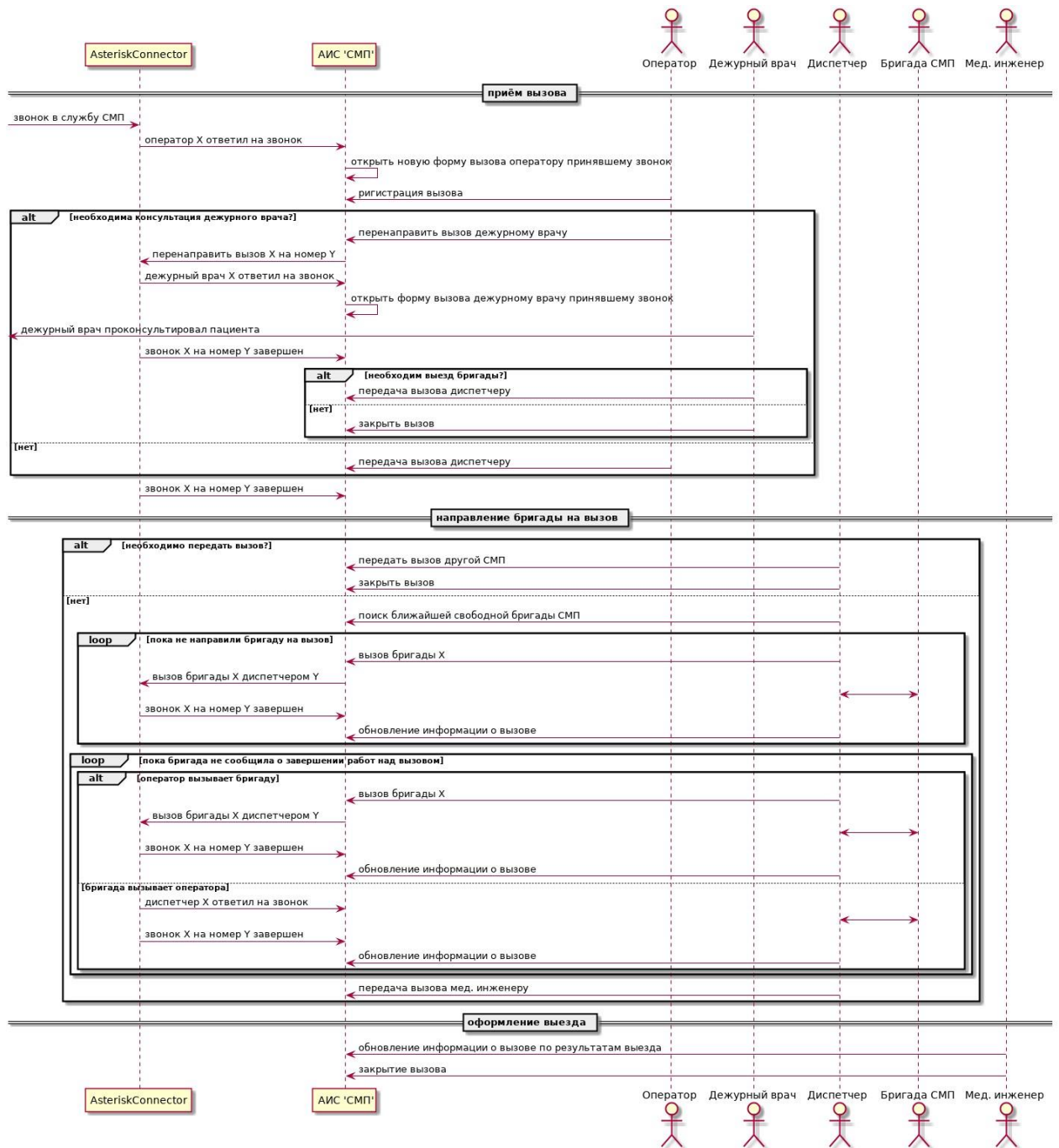


Рисунок 3 – Диаграмма последовательности действий обработки вызова, поступившего на станцию скорой медицинской помощи

При приеме вызова Оператор имеет возможность:

- Оформить вызов и передать Диспетчеру;
- Ввести первичную информацию о происшествии;
- Перенаправить вызов старшему врачу смены;
- Аннулировать вызов, удалить сорвавшийся звонок;
- Закрыть вызов, если не требуется выезд бригады скорой медицинской помощи;
- Искать зарегистрированные карты вызова.

При приеме вызова от Оператора Старший врач смены имеет возможность:

- Принять переадресованный вызов от Оператора;
- Закрыть вызов без направления бригады;
- Оформить вызов и передать диспетчеру.

При приеме вызова Диспетчер имеет возможность:

- Обработать вызов;
- Управлять дублями форм 109/у (просматривать, связывать с формой главного вызова, перемещать из дублей и присваивать соответствующий номер);
- Управлять бригадами СМП (формировать и расформировывать их);
- Направить на вызов определенную бригаду либо несколько бригад скорой медицинской помощи;
- При приеме звонков на номер «103» от бригады скорой медицинской помощи обновить информацию о вызове;
- Найти свободную бригаду скорой медицинской помощи с использованием картографического сервиса.

Управление звонками осуществляется посредством IP-телефонии с использованием технологии VoIP без отключения существующей телефонной сети. Коммутацию, запись и управление телефонными вызовами обеспечивает сервис компьютерной телефонии Asterisk.

При входящем звонке на службу 103 программный коммутатор (Softswitch) по исходящему номеру определяет группу, к которой относится абонент («потерпевшие» или бригады СМП) и, в зависимости от этой группы, осуществляет соединение с соответствующей группой абонентов (звонок «потерпевших» адресуется всем свободным операторам, звонок бригад СМП адресуется всем свободным диспетчерам).

Если все линии операторов/диспетчеров заняты, то формируется очередь звонков, а абоненту посредством системы голосовых сообщений воспроизводится сообщение с периодичностью 25 секунд до тех пор, пока линия оператора/диспетчера не освободится. При освобождении линии происходит соединение с оператором/диспетчером. Оператор имеет возможность перевести звонок на группу старших врачей смены посредством нажатия кнопки перевода звонка на телефоне, либо через АИС «СМП» (связь с Asterisk по интерфейсу АМІ). Все разговоры записываются и доступны для дальнейшего прослушивания в случае расследования инцидентов.

Заключение

Созданная АИС «СМП» находится в стадии доработки и тестирования и будет опытно эксплуатироваться на одном из отделений скорой медицинской помощи г. Донецка. В случае успешной эксплуатации систему возможно масштабировать на республиканский уровень. Также заложена возможность дальнейшей модернизации программного обеспечения, базирующегося на расширении функционала и добавления программных и инфраструктурных компонент.

Литература

1. Бутузова А. В., Моисеев В. С., Тутубалин П.И. Теоретические основы информатизации службы скорой медицинской помощи. — РИЦ «Школа»: РИЦ, 2011. — 242 с. — (Современная прикладная математика и информатика).
2. Никитина М. И. Применение информационных технологий в работе скорой медицинской помощи // Материалы VIII Всероссийской научно-практической конференции по теоретическим основам проектирования и разработке распределенных информационных систем (ПРИС - 2010). — Красноярск : Сибирский государственный технологический университет, 2010. — С. 196-206.
3. Моисеев В. С., Бутузова А. В. Основные задачи разработки автоматизированной системы управления скорой медицинской помощью // Исследования по информатике. — 2006. — № 10. — С. 141–150.
4. Бутузова (Сбоева) А.В. Задача планирования перемещения бригад скорой помощи // XIV Туполевские чтения: Материалы Межд. молод. науч. конф. 10 – 11 ноября 2006г., Т.IV. Казань, 2006 г. С. 4 – 6.
5. Сороковая А.В., Лыченко Н.М. Разработка управляющей информационной системы диспетчеризации скорой медицинской помощи // Проблемы автоматизации и управления. — Институт автоматизации и информационных технологий Национальной академии наук Кыргызской Республики (Бишкек), 2017. — № 2. — С. 72-81.
6. Дмитриева В.Г., Печеник Е.В. Внедрение автоматизированной информационной системы в службу скорой медицинской помощи в городе Южно-Сахалинск // Вестник общественного здоровья и здравоохранения Дальнего Востока России. — Дальневосточный государственный медицинский университет (Хабаровск), 2017. — № 1. — С. 6.
7. Мавлиханов Р.Ш., Пономарева О.А. Автоматизированная система диспетчеризации санитарного автотранспорта // Советник бухгалтера в здравоохранении. — 2017. — № 5.
8. Шестаков Г.С. Современное состояние организации и оценки деятельности станций и отделений

- скорой медицинской помощи // Менеджер здравоохранения. — 2007. — № 2.
9. Закон Донецкой Народной Республики «Об информации и информационных технологиях» № 71-ІНС от 07.08.2015;
 10. Положение о Министерстве связи Донецкой Народной Республики, утвержденное Постановлением Совета Министров Донецкой Народной Республики от 10.01.2015г. №1-17»;
 11. Положение о Министерстве здравоохранения Донецкой Народной Республики, утвержденное Постановлением Совета Министров Донецкой Народной Республики от 10.01.2015г. №1-33»;

Пьянков К.А. *Автоматизированная информационная система обработки вызовов скорой медицинской помощи.* В статье представлен краткий обзор выполненного проекта по автоматизации деятельности станции скорой медицинской помощи. Описаны функциональность системы и основные компоненты, использовавшиеся для построения системы.

Ключевые слова: информационная система, микросервисная архитектура, облачная инфраструктура, скорая медицинская помощь, диспетчер, территориально-распределенная система.

Pyankov K.A. *Automated information system for emergency call service.* The article presents a brief overview of the completed project on automating the activities of the ambulance station. The system functionality and the main components used to build the system are described.

Keywords: information system, microservice architecture, cloud infrastructure, emergency medical service, dispatcher, geographically distributed system

«THE AUTHENTICITY OF DIGITAL DOCUMENTS CHECKING AND PROTECTION BY OWN STEGANOGRAPHY ALGORITHM»

Babicheva M.V.

DonetskNationalUniversity
mv.babicheva60@gmail.com

Babicheva M.V. The authenticity of digital documents checking and protection by own steganography algorithm. The author's method of digital documents authenticity determining presented. Own steganography algorithm, based on a text message hidden in a container image by the hash digest matrix key, to protect digital documents from falsification is offered. The console application to demonstrate the capabilities of this method is developed.

Keywords: digital documents authenticity, steganography algorithm, a hidden message, an image container, hash digest.

Introduction

With the development of information and communication technologies, digital documents (photographs, scans) become increasingly important. A lot of enterprises and institutions use electronic document circulation. Document scans send when applying for a job, confirming an employee's qualifications, submitting offers for tenders at electronic platforms. Digital photos confirm the completed work amount, the state of the insurance objects, are widely used in the criminal expertise. No wonder, the number of different fraud with digital images is constantly growing. Image editing tools have become widely available. Therefore, it is very important to distinguish the original photo from the one modified in the graphical editor.

The primary purpose of this work is to develop a methodology for recognizing the editing fact of a digital image and some software to protect images from editing.

Forensic examination of documents

The theoretical and methodological basis of the study was the well-known methods for determining the authenticity of documents and the available research by the scientists in this field [1].

After the existing methods of recognizing the editing fact analysis and their possible vulnerabilities study, the base of real photographs and document scans was prepared. These images were loaded in Adobe Photoshop CC (2017) and edited in different ways, then saved in the most popular web-formats jpg and png. In addition, the author collected real fake documents and included into the database.

The following forensics methods were considered to determine the facts of editing:

1. Analysis of the JPEG structure (ELA, Error Level Analysis).
2. Analysis of noise (NLA, Noise Level Analysis).
3. Analysis of the brightness gradient (Luminance Gradient Analysis).
4. Principal Component Analysis (PCA).
5. Search for clones (Clone detector).
6. Analysis of quantization tables JPEG (JPEG analysis).
7. Meta data analysis (EXIF check).
8. Analysis of changes in the parity brightness of pixels.

These methods were applied to the created database of authentic and fake digital documents trying to detect the fact of editing. The researchers were not aware authenticated or fraudulent documents verified. Figure 1 shows the original and faked document difference detected by ELA method. Areas of editing seen due to the changes in the specific jpeg structure.

Competent approach in carrying out activities for the collection of digital evidence will not only provide an opportunity to detect the fact of editing, but also allow identifying ways and prerequisites for the incident occurrence. However, all the methods turned out to be possible to "deceive", that is, to hide the fact of editing by repeated re-saving, compression, and editing metadata. The fact of pixel-pixel editing no method could identify. One the scan of the passport had the person photo complete replaced and the registration number changed, but no method detected such nasty fake. The criminalists providing digital expertise should consider discovered vulnerabilities of forensic methods.

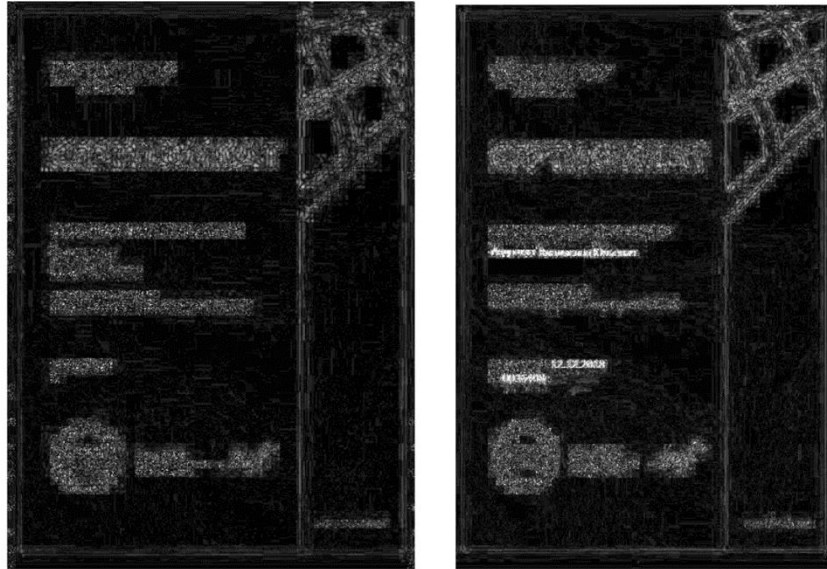


Figure 1- Original and fake certificate «Dr.Web» image (Error Level Analysis)

For a digital image suitable for identification research in the course of examination it is possible to obtain the following identification data: the establishment of shooting conditions, determination of the time image creation, restore the original image, type of digital conversion that was used to create the image, availability of further digital image processing (by means of graphic editors and software used in image processing), assessment of the change degree in the image during graphic processing, frame size created, EXIF data. This information can be given by different methods mentioned in the paper. However if the forger knows the essence of these methods, he or she can cheat the checker using pixel-by-pixel editing and EXIF modification. Figure 2 on the left shows the original scan of 1991 USSR passport. On the right - edited pixel-by-pixel image using Adobe Photoshop CC (2017) graphic editor, modified areas are selected. No one forensic method detected the fake.

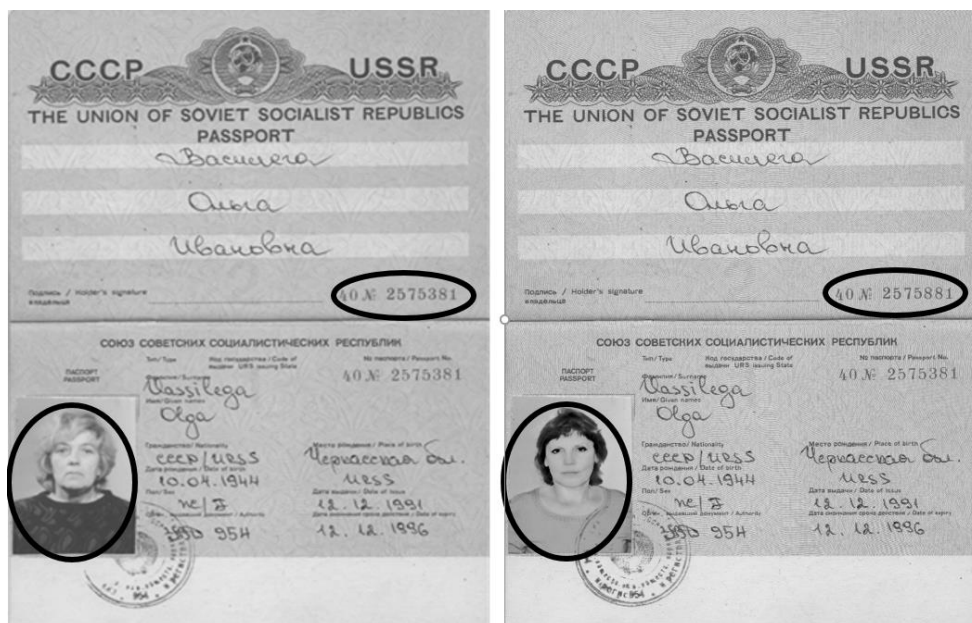


Figure 2- The original passport scan and the fake made by pixel-by-pixel editing

The table 1 summarizes the results of computer experiments. The cells indicate what percentage of images from the database made by various gadgets and edited in different ways, was identified as fake.

Table 1 - The number of recognized samples depending on the editing method and recognition method

№	Editing method	Method of recognizing the editing fact (%)							
		ELA	NLA	Gradient	PCA	Clone Det.	JPEG analysis	Meta data	the parity brightness of pixels
1	Correction of brightness, color, saturation	80	75	73	63	2	90	100	100
2	Insert areas cut from other images	96	93	56	70	0	90	100	100
3	Insert cut from the same image areas	0	0	10	0	90	90	100	0
4	Contours change, drawing	98	92	36	27	0	95	100	100
5	Tools "plastic", "blur", "retouch"	62	60	100	24	0	100	100	96
6	Download to the editor	0	0	0	0	0	100	100	0
7	Pixel-by-pixel editing	0	0	0	0	0	0	0	0

The literature describes a method for determining, whether a digital image in a png format is edited in Photoshop, using an artifact detected by researchers: when an image is placed in an editor, pixel brightness values that do not exceed the middle of the dynamic range are rounded to the nearest even values, and higher to the nearest odd [2].

Versions of graphic editors used in the research, starting with Adobe Photoshop 5.5 and ending with Adobe Photoshop CS3. Several new versions of Photoshop have been released since then, it was necessary to check for the presence of this artifact and the ability to use it as one of the tools for determining whether an image was edited.

For computer experiments, 10 images of different sizes, obtained from different sources, were saved in the formats png, bmp and tif. Then, in Adobe Photoshop CC (2017), the following manipulations were carried out:

- the image was loaded into the editor and its brightness was changed with the help of the tool "image brightness".
- the contours were edited with various tools (brushes, pencil, eraser, stamp, deformation), fragments were copied and moved.
- the image was loaded into the editor and then simply saved in the same format.

Then a program was written in MatlabImage Processing, which counted the number of pixels with even and odd values of brightness and gave the difference, determined the middle of the dynamic range of the image and calculated the number of pixels with even brightness values, below the middle of dynamic range, and the number of pixels with odd brightness values, above the middle of dynamic range, and also calculated the difference of these values. Although the values of the brightness changed when editing, the patterns discovered by the articles authors were not found.

Original steganographic algorithm

Analysis of existing steganographic methods to protect digital images from editingshows they usually have three features, good hiding capacity, good imperceptibility and robustness [3]. All of them are based on additional information embedded into the image-container, destroyed by editing. As a means of protecting digital documents from editing, were chosen well known "merging files" and "least significant bit".

Unlike hashing and cryptography algorithms, which are not recommended to be developed independently, steganographic algorithms on the contrary, recommended to develop, because exactly the fact of information transfer is disrupted. Such algorithms are used in documents protection from editing. Information is added to the document, which, if edited, may change. In this work own steganographic algorithm was developed, which allows you to enter information into the image matrix. A proprietary algorithm was developed for writing a hidden text message to the Blue matrix of a container image by key addresses hash digest matrix. The block schematic diagram of the algorithm can be seen on the figure 3. The container image split into 3 matrices, R G and B. Then the blue matrix is separated, as man's eyes are not so sensitive to notice changes in the blue color. Then the password word or a phrase changes into hash digest, the author used sha-1, but any other may be used. Changing by cyclic shift the matrix is formed with as many

symbols, as hidden message has. This is the matrix of addresses for embedding the message into the image. The message for hiding (no more than 30 percent of the image size) changes to ASCII codes and is written to the blue image matrix according to the address matrix. Three matrices merge and the image with embedded message is ready. If the intruder tries to edit the image, the hidden message is destroyed and the fact of editing becomes known.

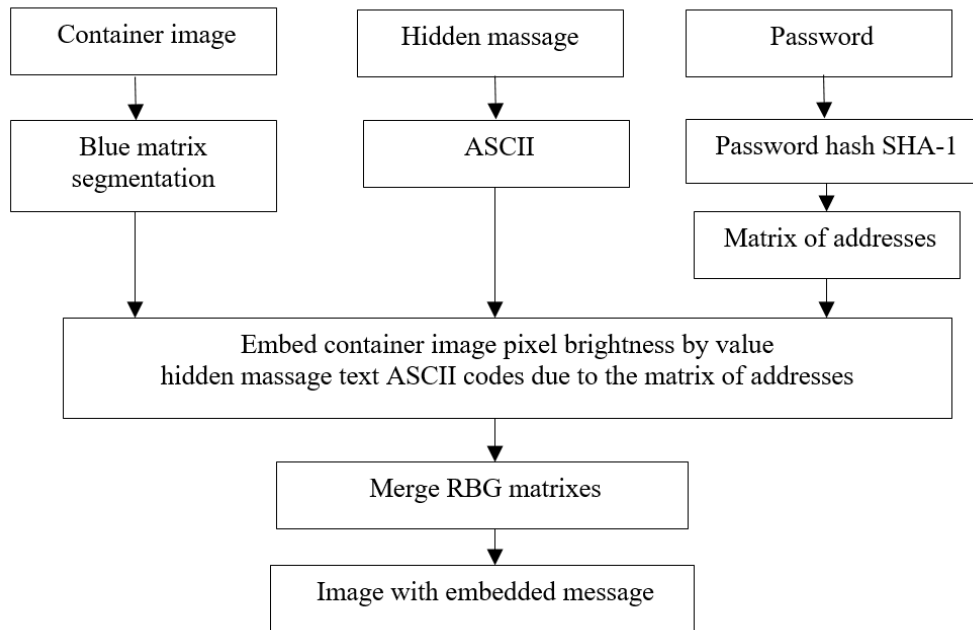


Figure 3 - Block schematic diagram of the steganographic algorithm

The figure 4 presents the blue matrix of the image with embedded message. In the left upper corner the changed pixels can be seen. After merging, they become invisible.

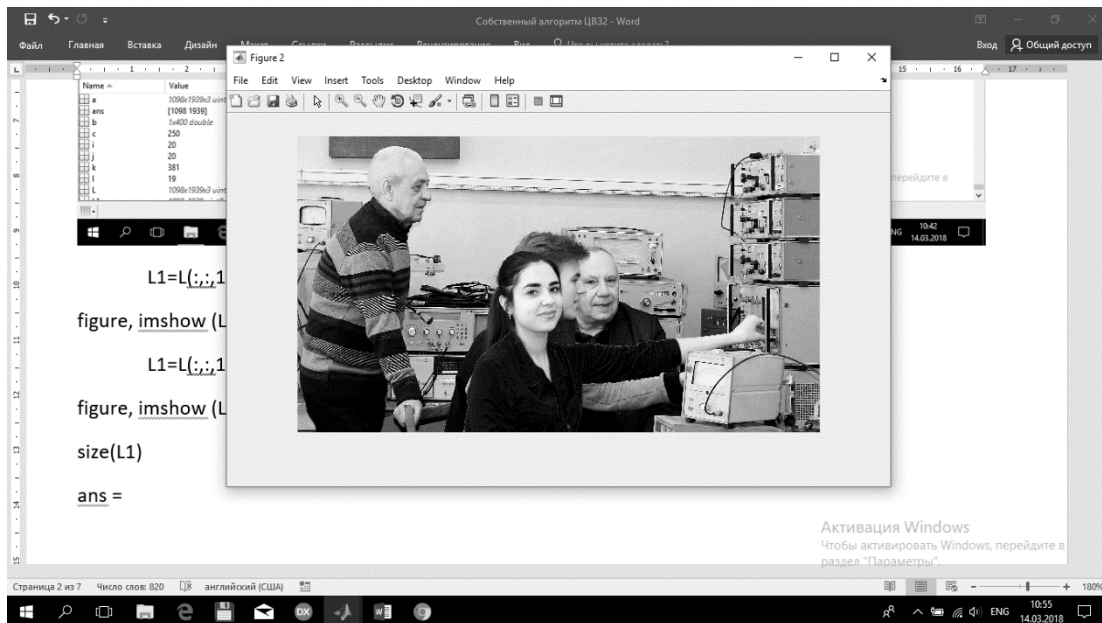


Figure 4 – The blue matrix of container image with embedded steganographic message

For three steganographic algorithms, “merging files”, “least significant bit” and new “hidden message by hash digest matrix addresses” console applications were developed in Matlab Image Processing Toolbox. The limits of applicability of each method and their advantages and disadvantages were determined. The file merging method is the easiest and quickest. At the slightest editing of an image with a hidden container, the file structure changes and the

extraction of a hidden message is impossible. However, sticking another file affects the size of the container and that may be suspicious. Another way to detect merging files is to view their headers in HEX editor.

The image is stored as a matrix of color hue values for each point of the image. If each of the RGB space components (also called color channels) is stored in one byte, it can take values from 0 to 255, inclusive, which corresponds to a 24-bit color depth. The peculiarity of a person's vision is that it weakly distinguishes minor color variations. For a 24-bit color, a change in each of the three channels of one least significant bit results in a less than 1% change in the intensity of this point, which allows you to change them inconspicuously to the eye at your discretion [4]. To calculate the bandwidth of the method, if discard in the calculations, usually insignificant relative to the size of the image the service information at the beginning of the file, then we are able to secretly transmit a message about 1/8 the size of the container ("smeared" by the last bits in each byte of the pixel color matrix) or 1 / 4 containers (respectively, when using the last 2 bits in bytes) [5].

Based on the results of the LSB method implementation, concluded that the bits of the hidden information were lost in the noise of their own image, as a result the detection of a change in the file structure does not determine the presence of hidden information. Among the weakness, we can note the long processing of the bit structure, which took a long time. Computer experiments to study the resistance of the methods to transferring files via info communication channels were done. The image container file with the embedded message saved in different formats, and different file systems was transferred by various info communication services. Then the embedded image was again extracted. Table 2 presents the results of the experiments. Different file formats were examined. Experiments conducted on sending and receiving files by the Internet services such as e-mail (mail.ru), web (VK), Skype and cloud storage (mail.ru cloud). Transferring, depend on the service capabilities, the file was directly sent or packaged by the WinRAR archiver for the Windows 10 operating system, and then sent, unpacked and checked. To send by the web, the VK file sharing service was used. After receiving the image file, it was saved on the disk, and then the embedded message was extracted.

Table 2 -The results of experiments on saving and sending images with embedded messages through info communication systems and further extracting this message

Original container image format	Record format of image container with embedded message	File system	Managed to extract the embedded message after writing to disk	Managed to extract the embedded message after writing to disk, sending and receiving over the Internet, again writing to disk
jpg	jpg	FAT32	no	no
bmp	bmp	FAT32	yes	yes
bmp	tiff	FAT32	yes	yes
bmp	tif	FAT32	yes	yes
bmp	tif	NTFS	yes	yes
bmp	jpg	NTFS	no	no
bmp	jpg	FAT32	no	no
jpg	bmp	NTFS	yes	yes
jpg	bmp	FAT32	yes	yes
bmp	png	NTFS	yes	yes
bmp	png	FAT32	yes	yes
bmp	jpeg	NTFS	no	no
bmp	jpeg	FAT32	no	no

From the results of experiments, we can conclude the algorithm is invariant to the file system type, and to the format in which the container image was originally stored. However, from the saved format in which the image with the embedded message transmitted, depends on whether we can extract the embedded message. An embedded message is well extracted if the container with the message is saved in any format without compression (for example, bmp, tiff, tif)

or lossless compression (png) and is not extracted from a file saved in jpg or jpeg format (with lossy compression). The results of the study showed to protect documents, is advisable to use embedded messages destroyed when trying to edit a container image. All three methods showed satisfactory protection. The results of this work implemented into the educational process as a laboratory work on the course "Basics of Information Security".

Conclusion

In this paper, the concepts of digital documents security are studied. The fact of image editing, with possible changes in forensically significant information, can be detected by the number methods, but with professional pixel-by-pixel editing, changes successfully hidden. The author performed a new steganography method by own algorithm. This method and two more well-known methods were implemented in a program written in Matlab language. The proposed method has proved successfully hiding text in color images to protect them from modifying. Achieved results indicate that proposed method is encouraging in terms of security, and robustness for the most file formats. As many new application areas appeared like internet banking, mobile communication security, cloud security etc., the insight into the steganographic principles will definitely guide us to identify new areas and to improve its applications in the already existing areas. The work results can be useful for criminalists, police officers, employees conducting digital examinations and specialists responsible for the digital documents authenticity.

Literature

1. Mamta Juneja, Parvinder Singh Sandhu, A New Approach for Information security using an Improved Forensic Technique, Journal of Info.Pro.Systems, – 2017. – Vol 9, No:3, -P.405-424.
2. Grebenyuk P. E., Zotov P. V., Chmutin A. M., Sharipov R. R. Programmatic Intervention in Digital Images and Its Identification // XIV Regional Conference of Young Researchers of the Volgograd Region (Volgograd, 10-13 November 2009 city): mes. report / VolgGTU [et al.]. - Volgograd, 2010. - P. 201-204.
3. M. K. I. Rahmani and N. P. Kamiya Arora, "A crypto-steganography: A survey," International Journal of Advanced Computer Science and Application, -2014. -Vol. 5, - P. 149–154.
4. J. V. Karthik and B. V. Reddy, "Authentication of secret information in image stenography," International Journal of Computer Science and Network Security (IJCSNS), -2014. - Vol. 14, no. 6, - P. 58.
5. Rajendran, S., Doraipandian, M. (2017), Chaotic Map Based Random Image Steganography Using LSB Technique, International Journal of Network Security, – 2017. – Vol.19, No.4, - P. 593-598.

Babicheva M.V. The authenticity of digital documents checking and protection by own steganography algorithm. *The author's method of digital documents authenticity determining presented. Own steganography algorithm, based on a text message hidden in a container image by the hash digest matrix key, to protect digital documents from falsification is offered. A console application to demonstrate the capabilities of this method is developed.*

Keywords: digital documents authenticity, steganography algorithm, a hidden message, an image container, hash digest.

Бабичева М.В. Проверка подлинности и защита цифровых документов с помощью собственного алгоритма стеганографии. *Представлен авторский метод определения подлинности цифровых документов. Предлагается собственный алгоритм стеганографии, основанный на текстовом сообщении, скрытом в изображении контейнера с помощью ключа матрицы хэш-таблицы, для защиты цифровых документов от фальсификации. Разработано консольное приложение, для демонстрации возможностей этого метода.*

Ключевые слова: проверка подлинности цифровых документов, алгоритм стеганографии, скрытое сообщение, изображение-контейнер, хэш-дайджест.

Анализ алгоритмов аутентификации в распределенных программных системах

Грищенко О.С., Чернышова А.В.
Донецкий национальный технический университет, г. Донецк
кафедра программной инженерии
olleg1996@mail.ru, alla@donntu.org

Грищенко О.С., Чернышова А.В. Анализ алгоритмов аутентификации в распределенных программных системах. Представлен обзор и анализ существующих подходов и алгоритмов аутентификации, используемых в современных распределенных системах. Рассмотрен принцип работы распространенных протоколов аутентификации. Проанализирован протокол аутентификации SQRL и внесены предложения по усовершенствованию данного протокола.

Ключевые слова: аутентификация, алгоритм SQRL, PKI, уязвимость, защищенность.

Постановка задачи

На современном этапе развития информационных технологий аутентификация является неотъемлемой частью распределенных программных систем. В связи с ростом ценности информации и появлением новых угроз информационной безопасности задача обеспечения надежной аутентификации пользователя является очень актуальной. Исходя из этого, было решено проанализировать существующие алгоритмы аутентификации, выделить их преимущества, недостатки и уязвимости.

Цель работы

Целью данной работы является анализ общих принципов и конкретных алгоритмов аутентификации в программных системах, изучение существующих алгоритмов и стандартов, анализ возможных уязвимостей и выработка вариантов усовершенствования одного из существующих алгоритмов.

Современные тенденции в области аутентификации

Существует множество способов программно реализовать контроль доступа к электронным ресурсам. Большинство таких способов основаны на процедурах идентификации и аутентификации. Идентификация состоит в присвоении каждому пользователю системы идентификатора. Идентификатор – это специальная информация, которая ставится в соответствие пользователю и полностью определяет его личность для системы. Аутентификация[1] – это процедура проверки принадлежности предъявленного идентификатора пользователю. Можно выделить несколько общих подходов, на которых основывается аутентификация в современных программных системах. Это проверка информации, которой должен обладать пользователь (например, парольная аутентификация[2]), проверка факта обладания пользователем материальным объектом или информационным ресурсом (например, аутентификация по usb-токену[3], криптографическому сертификату[4]), проверка уникальных свойств, присущих пользователю (сюда относится биометрическая аутентификация[5]) и проверка информации, которая имеет отношение к пользователю (на текущий момент представлена аутентификацией по местоположению[6]). Аутентификация может быть односторонней[7], когда проверяется подлинность только одной из сторон участвующих в аутентификации, и двусторонней[8], когда аутентификацию проходят обе стороны. Односторонняя аутентификация обычно используется при парольной или биометрической аутентификации. Двусторонняя аутентификация часто используется при аутентификации на основе сертификатов[9]. По количеству используемых методов аутентификация может быть однофакторной (используется только один метод) и многофакторной (используется не менее двух различных методов). Рассмотрим особенности использования и уязвимости современных методов аутентификации.

Наиболее широкое распространение на текущий момент получила парольная аутентификация, которая имеет множество реализаций, часто отличающихся по уровню защищенности. Концептуально алгоритм парольной аутентификации достаточно прост: при регистрации в системе пользователь задает свой логин (идентификатор) и пароль (аутентификатор). При последующем входе в систему пользователь передает свои логин и пароль, которые сверяются со значениями, хранящимися в системе. Главной проблемой парольной аутентификации является то, что пользователь устанавливает пароль и отвечает за его хранение. Осведомленность рядовых

пользователей в вопросах информационной безопасности обычно невысока, что приводит к созданию простых или предсказуемых паролей, которые очень уязвимы перед атаками методом прямого перебора или перебора по словарю. В современных системах этот недостаток пытаются компенсировать путем задания специальных правил и политик (например, обязательное использование цифр, букв в разных регистрах при создании пароля, ограничение по длине пароля, обязательная смена пароля по истечении определенного периода времени) [10]. Такие меры делают пароли более устойчивыми, но усложняют использование пользователем системы. Значительно повысить устойчивость паролей можно, используя генератор паролей, однако сгенерированный пароль трудно запомнить пользователю. Описанные меры не могут защитить пароль пользователя от небезопасного хранения, случайного или намеренного разглашения.

Потенциальными угрозами безопасности является перехват данных[11], используемых при парольной аутентификации, при передаче по сети и хищение базы данных пользовательских учетных записей. При реализации данных угроз множество учетных данных пользователей могут быть скомпрометированы. Очень актуальными описанные угрозы являются для распределенных систем из-за необходимости репликации данных между различными физическими устройствами. Для защиты от подобных угроз в современных системах используют шифрование данных при передаче по сети и хранении вместо паролей хеш-значений[12], сгенерированных криптографически стойкими хеш-функциями.

Один из самых распространенных способов хищения пароля пользователя – это фишинговая атака[13]. Фишинговые атаки могут иметь различные реализации, но все их объединяет то, что пользователи сами передают свои учетные данные злоумышленнику. Рассмотрим обобщенную схему наиболее популярной реализации фишинговой атаки. Пользователь получает сообщение от имени известного ресурса, в котором предлагается перейти по ссылке. Если пользователь переходит по предложенной ссылке, то он попадает на форму аутентификации, копирующую дизайн данного ресурса и вводит свои учетные данные, которые, затем отправляются на сервер злоумышленника. Фишинговые атаки могут использовать спуфинг[14], например, путем регистрации для поддельной страницы аутентификации домена, визуально похожего на домен реального ресурса. Пример реализации описанной схемы представлен на рисунке 1.



Рисунок 1 – Пример реализации фишинговой атаки

Кража учетных данных пользователя также может быть осуществлена с помощью вредоносного программного обеспечения, такого как программы кейлогеры[15].

Более безопасной по сравнению с парольной аутентификацией является аутентификация на основе инфраструктуры открытых ключей - РКІ[16]. Работа РКІ систем основана на использовании криптографического алгоритма с открытым ключом. РКІ системы позволяют проводить аутентификацию пользователей путем проверки их сертификатов. Сертификат выдается каждому пользователю в системе и содержит подписанные с помощью асимметричного криптоалгоритма данные, идентифицирующие пользователя и его открытый ключ. Выдачей сертификатов занимаются центры сертификации[17]. В реальных системах из-за большого количества пользователей используются множество центров сертификации. При этом образуется иерархическая структура, в которой нижестоящие центры сертификации имеют сертификаты, выданные вышестоящими центрами. В основании такой структуры находится корневой центр сертификации, который использует самоподписанный сертификат[18] и пользуется абсолютным доверием остальных участников системы. На рисунке 2 представлен пример аутентификации пользователя А на ресурсе Б с использованием инфраструктуры открытых ключей.

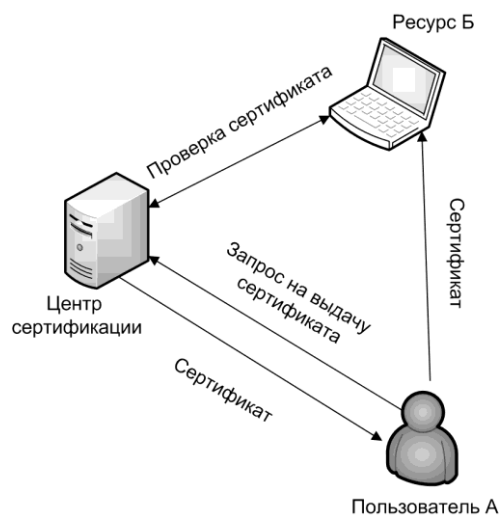


Рисунок 2 – Пример аутентификации пользователя А на ресурсе Б

PKI системы широко используются в современных программных системах. В частности инфраструктуру открытых ключей использует протокол TLS/SSL[19], который реализует аутентификацию и шифрование данных в рамках протокола HTTPS[20]. Многие компании организуют доступ к своим внутренним ресурсам на основе PKI системы с собственным центром сертификации. PKI системы являются основой для создания сервисов электронной подписи документов и могут быть использованы при реализации аутентификации по usb-токенам или смарт-картам.

Инфраструктура открытых ключей также имеет уязвимости. В частности угрозой для аутентификации, использующей PKI, представляет вирусное программное обеспечение. Например, вредоносная программа, попавшая каким-либо образом на компьютер пользователя может добавить в список доверенных сертификатов сертификат злоумышленника. Это приведет к тому, что злоумышленник и выданные им сертификаты будут использоваться полным доверием пользователя.

При небезопасном хранении приватный ключ может быть похищен и, в таком случае, злоумышленник сможет осуществлять всевозможные действия от имени пользователя.

В случае, если корневой центр сертификации будет скомпрометирован, то скомпрометированными станут и все сертификаты, которые он выдал.

Аутентификация в PKI системе невозможна, если утеряна связь с центром сертификации.

Для отзыва выданных сертификатов (например, в случае их компрометации) в PKI системах предусмотрен список аннулированных (отозванных) сертификатов. Тем не менее, из-за большого количества территориально распределенных центров сертификации может потребоваться некоторое время для обновления списка аннулированных сертификатов в рамках всей системы. В течении этого времени злоумышленник может продолжать использовать отозванный сертификат.

Все большую популярность в последнее время приобретают биометрические методы аутентификации, основанные на распознавании уникальных свойств конкретного человека. К биометрической аутентификации можно отнести методы, основанные на распознавании лица, рисунка кровеносных сосудов, радужной оболочки и сетчатки глаза, отпечатков пальцев, кисти руки. Одним из недостатков биометрической аутентификации является то, что параметры человека меняются с возрастом и в результате получения травм. Современные программные реализации аутентификации на основе распознавания лица или радужной оболочки глаза можно обмануть с помощью качественной фотографии человека. Уязвимым местом аутентификации, использующей сканирование отпечатка пальца, является то, что люди оставляют свои отпечатки, касаясь окружающих предметов. Таким образом, добыть достаточно качественную фотографию отпечатка пальца не составляет особого труда. А используя фотографию отпечатка пальца, можно обойти аутентификацию, напечатав модель пальца на 3d принтере, или на специальной бумаге с помощью струйного принтера. Аутентификацию по голосу можно обойти, используя программы, позволяющие воспроизводить голос человека по аудио записи, которую также не сложно получить в современных условиях. В реализациях биометрической аутентификации от таких компаний, как Google или Apple базы данных биометрической информации шифруются и хранят результат одностороннего преобразования биометрических данных. Однако не все существующие реализации биометрической аутентификации имеют подобные средства защиты, в связи с чем пользовательские биометрические данные могут быть скомпрометированы. Это особенно важно, учитывая то, что, в отличие от традиционных методов аутентификации, в системах биометрической аутентификацией часто невозможно заменить скомпрометированные данные.

Исследование существующих алгоритмов аутентификации

В стандарте RFC 7235[21] описывается парольная аутентификация для протокола HTTP. Стандарт определяет несколько схем аутентификации. Рассмотрим базовую схему (RFC7617[22]). При отправке от имени неавторизованного клиента запроса к защищенному ресурсу сервер вернет ответ 401 с описанием схемы аутентификации в заголовке. После получения такого ответа клиент запрашивает учетные данные у пользователя и отправляет их на сервер в поле Authorisation. После этого сервер может провести аутентификацию пользователя по полученным данным. В дальнейшем поле Authorisation добавляется к запросам пользователя на данный сервер. Главным достоинством описанного алгоритма, как и большинства алгоритмов парольной аутентификации, является простота реализации. Описанный алгоритм не обеспечивает защиту передаваемых данных, поэтому может использоваться в реальных системах только через защищенное соединение. Серьезной уязвимостью является то, что перехват и расшифровка любого запроса приведут к получению злоумышленником учетных данных пользователя. Данный алгоритм не является защищенным от атак повторного воспроизведения. Также описанная схема не предусматривает какой-либо защиты от фишинговых атак, вредоносного программного обеспечения. Данный алгоритм не может предоставлять достаточного уровня безопасности и поэтому редко используется в реальных системах.

Протокол SSL/TLS получил широкое распространение как средство защиты интернет-коммуникаций. Протокол реализует аутентификацию на основе инфраструктуры открытых ключей, а также шифрование и аутентификацию передаваемых данных.

Аутентификация и согласование ключей в TLS состоит из следующих шагов:

- установка соединения между клиентом и сервером по протоколу TCP[23];
- согласование версии протокола, алгоритмов шифрования, хеш-функции;
- проверка сертификатов (обычно происходит односторонняя аутентификация с проверкой сертификата сервера);
- обмен ключами шифрования по алгоритму RSA[24] или Диффи-Хеллмана[25];
- обмен сообщениями с проверкой MAC[26] кода.

Сам протокол является надежным и безопасным, его спецификация постоянно обновляется для устранения найденных уязвимостей. Благодаря тому, что протокол стал стандартом в области защиты интернет соединений он очень хорошо изучен. Однако многие реализации используют устаревшую спецификацию и не следят за обновлениями протокола (таким образом, не устраняются обнаруженные уязвимости). Также реализации могут использовать небезопасные криптографические алгоритмы (например, md5[27] для подписи сертификатов) или недостаточную длину ключа. Из-за большого количества используемых версий протокола появляется необходимость поддержки старых версий с известными уязвимостями. Введение многочисленных контрмер в ответ на найденные уязвимости вместо пересмотра логики протокола сделали его полную реализацию очень сложной и, в следствие этого, многие реализации протокола имеют ошибки, позволяющие использовать уязвимости, которые считаются устраненными. Из-за того, что рядовые пользователи интернет обычно не имеют сертификатов, TLS может аутентифицировать только сервер, но не клиента на сервере. Для этого приходится использовать другие алгоритмы аутентификации, которые часто становятся слабым местом всей системы. При аутентификации сервера TLS может гарантировать принадлежность домена ресурсу, однако он не может защитить пользователя от ситуации, когда сертификат выдан на домен, визуально похожий на домен другого ресурса (одна из разновидностей спуфинга). Протоколу присущи общие проблемы PKI систем, описанные ранее.

Стандарт ГОСТ Р ИСО/МЭК 9594-8-98[28] — Основы аутентификации описывает два варианта реализации аутентификации — простую и строгую. Простая аутентификация является методом аутентификации, основанным на парольной аутентификации, а строгая аутентификация использует инфраструктуру открытых ключей. Рассмотрим алгоритм строгой аутентификации, который предполагает более безопасную аутентификацию, чем простой вариант. Выдачей сертификатов занимаются уполномоченные по сертификации (аналог центра сертификации инфраструктуре открытых ключей). Уполномоченные по сертификации (УС) могут генерировать два типа сертификатов – срочные, которые выдаются другим участникам системы, и реверсивные, выданные самому себе. Сертификат содержит такую информацию, как имя владельца, публичный ключ и уникальный идентификатор. Перед началом аутентификация строится путь сертификации (список сертификатов, необходимый для получения общего ключа другого пользователя). Стандарт описывает три возможные процедуры строгой аутентификации: однонаправленную, двунаправленную и трехнаправленную. При однонаправленной аутентификации между пользователями А и В происходит одна передача информации, подписанной пользователем А и содержащей уникальный маркер (га), метку времени (та), предлагаемый к использованию ключ шифрования (зашифрованный открытым ключом В), идентификатор В и путь сертификации от В к А. Пользователь В сверяет подпись, проверяет параметры га и та для защиты от атаки повторного воспроизведения и принимает решения об установлении защищенного соединения с А, используя предлагаемый ключ шифрования. При двунаправленной аутентификации добавляется ответ от В к А, подписанный В и содержащий идентификатор А метки времени tb и ta, маркеры га и gb, и предлагаемый ключ шифрования, зашифрованный с помощью открытого ключа А. После чего пользователь А проверяет

полученные данные. Трехнаправленная аутентификация позволяет не передавать метку времени за счет использования третьей передачи подписанной А информацией, содержащей гВ идентификатор В.

Строгая аутентификация является достаточно надежной и более защищенной, чем парольная аутентификация, однако требует настройки и поддержки соответствующей инфраструктуры и более сложна в реализации. Стандарт не описывает способа хранения и генерации приватных ключей сертификатов. Если приватные ключи будут создаваться централизованно (например, администратором системы), то, в зависимости от способа передачи и уровня защищенности канала связи, приватные ключи могут быть похищены, что скомпрометирует выданные сертификаты. В случае, если приватные ключи сертификатов будут генерироваться на конечном устройстве пользователя, то при краже устройства или полном копировании носителя информации злоумышленник сможет выдавать себя за этого пользователя. При использовании в качестве хранилища ключа специализированных устройств (например usb-токенов и смарт-карт) существует угроза утери или кражи такого устройства. При неправильном администрировании системы также возможна кража приватных ключей сертификатов. Если конечный пользователь имеет доступ к своему закрытому ключу, то злоумышленник может использовать методы социальной инженерии для получения этого ключа. Однонаправленная и двунаправленная процедуры требуют синхронизации времени, в системе, иначе аутентификация будет уязвима для атак повторного воспроизведения. Реализация описанного алгоритма может быть уязвима при использовании ненадежных алгоритмов шифрования и электронной цифровой подписи или при малой длине ключа шифрования.

Анализ и варианты усовершенствования протокола SQRL

Протокол SQRL[29] разрабатывался как простая в использовании и более защищенная альтернатива парольной аутентификации. При первом запуске приложения на клиенте генерируется мастер ключ, ключ разблокировки личности и задается пароль пользователя. При первой попытке доступа к ресурсу генерируется уникальный приватный ключ на основе мастер-ключа и url службы аутентификации ресурса. На сервере публичный ключ для сгенерированного приватного ключа заносится в базу данных и ассоциируется с пользователем. На рисунке 3 представлена схема аутентификации с помощью протокола SQRL.

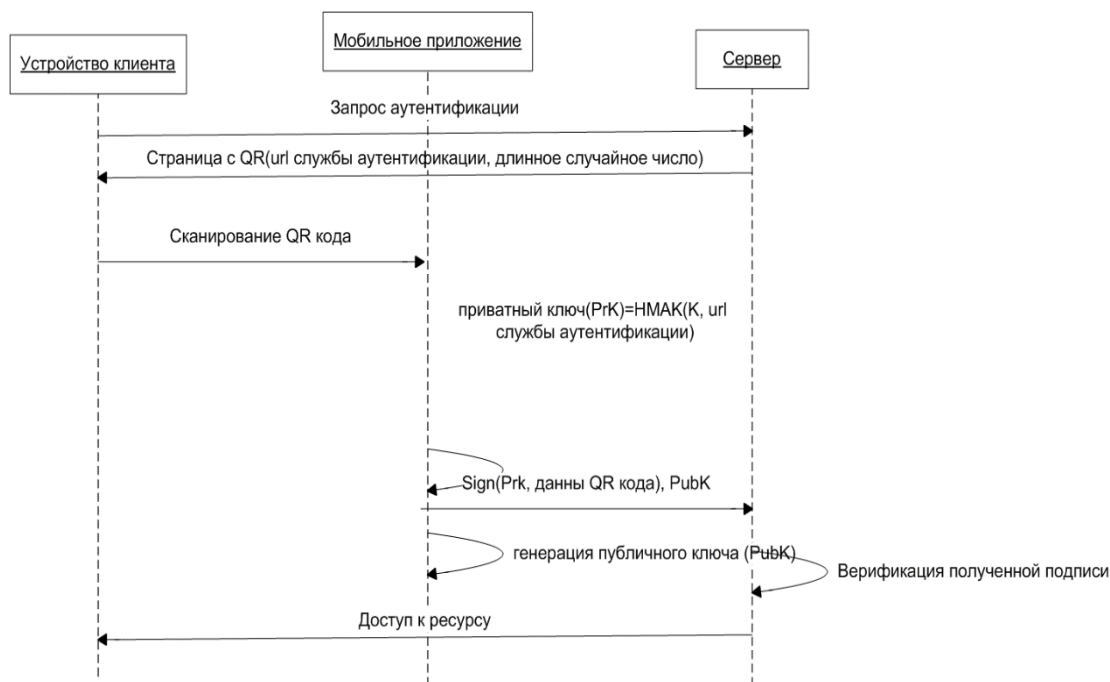


Рисунок 2 - Схема работы протокола SQRL

Для того чтобы экспортировать или заменить мастер-ключ используется ключ разблокировки личности, который хранится в зашифрованном виде в распечатываемом QR[30] коде. Расшифровать ключ разблокировки личности можно с помощью 24-х значного кода, который показывается пользователю при генерации ключа разблокировки личности. Мастер-ключ хранится в зашифрованном виде на устройстве пользователя и расшифровывается с помощью пароля пользователя.

Слабым местом в алгоритме защиты протокола SQRL является пароль пользователя. Также потенциальной уязвимостью является то, что пользователь отвечает за хранение кода разблокировки личности. Не смотря на то, что код разблокировки личности зашифрован с использованием произвольно генерируемого 24-значного ключа, пользователь, вероятно, будет хранить этот ключ вместе с кодом разблокировки личности (так как

пользователю сложно запомнить 24-х значный код, который он к тому же не использует в повседневной жизни).

Для повышения защищенности можно расширить протокол за счет использования двухфакторной аутентификации при вводе пароля пользователя и при вводе ключа разблокировки личности, например через отправку кода подтверждения на e-mail адрес. А также за счет введения в реализацию клиента правил, не позволяющих генерировать простой пароль. При этом e-mail адреса для аутентификации на ресурсе могут храниться на сервере ресурса и задаваться при регистрации на конкретном ресурсе. А e-mail, используемый для разблокировки личности, может храниться в зашифрованном виде вместе с кодом разблокировки, чтобы исключить возможность подмены e-mail адреса на адрес злоумышленника. В таком случае e-mail, необходимый для активации кода разблокировки личности, может задаваться при первом запуске приложения. Оба типа e-mail адресов могут быть изменены при активации кода разблокировки личности.

Выводы

Был проведен обзор существующих подходов в области аутентификации и проанализированы распространенные алгоритмы аутентификации, рассмотрены их уязвимости преимущества и недостатки. На основе изученных данных можно сделать вывод, что на сегодняшний день существует множество различных подходов к аутентификации, имеющих свои уязвимости и отличающихся по степени защищенности и сложности реализации. Был проведен анализ протокола SQRL и предложены возможные меры по его усовершенствованию.

Литература

1. authentication [Электронный ресурс]. // SearchSecurity. – Режим доступа: <https://searchsecurity.techtarget.com/definition/authentication> - Загл. с экрана.
2. Парольная аутентификация [Электронный ресурс]. // Volpi. – Режим доступа: <http://www.volpi.ru/umkd/zki/index.php?man=1&page=23> – Загл. с экрана.
3. CryptographicTokenKeyInitializationProtocol [Электронный ресурс]. // Tools. – Режим доступа: <https://tools.ietf.org/html/rfc4758#page-11> – Загл. с экрана.
4. The Basics of Cryptography and Digital Certificates [Электронныйресурс]. // Cryptography. – Режимдоступа: <https://whatismyipaddress.com/cryptography> - Загл. с экрана.
5. Biometric Authentication Overview, Advantages & Disadvantages [Электронныйресурс]. // Heimdal Security. – Режимдоступа: <https://heimdalsecurity.com/blog/biometric-authentication> - Загл. с экрана.
6. Location based authentication [Электронныйресурс]. // Patents. – Режимдоступа: <https://patents.google.com/patent/US20090187492> - Загл. с экрана.
7. Аутентификация односторонняя [Электронный ресурс]. // Cryptography. – Режим доступа: http://cryptography.ru/docs/аутентификация_односторонняя/ - Загл. с экрана.
8. Двусторонняя аутентификация [Электронный ресурс]. // Studopedia. – Режим доступа: <https://studopedia.org/4-95486.html> - Загл. с экрана.
9. What Is Certificate-Based Authentication [Электронныйресурс]. // Global Sign. – Режимдоступа: <https://www.globalsign.com/en/blog/what-is-certificate-based-authentication/> - Загл. с экрана.
10. Как создать надежный пароль [Электронный ресурс]. // GoogleSupport. – Режим доступа: <https://support.google.com/accounts/answer/32040?hl=ru> – Загл. с экрана.
11. Перехват данных по сети [Электронный ресурс]. // Antimalware. – Режим доступа: <https://www.anti-malware.ru/threats/network-traffic-interception> - Загл. с экрана.
12. Хэш код [Электронный ресурс]. // Academic. – Режим доступа: <https://dic.academic.ru/dic.nsf/ruwiki/1185956> - Загл. с экрана
13. Фишинг [Электронный ресурс]. // Habr. – Режим доступа: <https://habr.com/post/344066/> - Загл. с экрана.
14. Спуфинг [Электронный ресурс]. // Avast. – Режим доступа: <https://www.avast.ru/c-spoofing> - Загл. с экрана.
15. Клавиатурные шпионы [Электронный ресурс]. // Antimalware – Режим доступа: <https://www.anti-malware.ru/threats/keyloggers> - Загл. с экрана.
16. PublicKeyInfrastructure[Электронный ресурс] // Tutorialspoint. – Режим доступа: https://www.tutorialspoint.com/cryptography/public_key_infrastructure.htm- Загл. с экрана.
17. Установка центра сертификации на предприятии [Электронный ресурс]. // Habr. – Режим доступа: <https://habr.com/company/microsoft/blog/348944/> - Загл. с экрана.
18. Самоподписанный SSL сертификат [Электронный ресурс]. // Emaro-Ssl. – Режим доступа: <https://www.emaro-ssl.ru/blog/self-signed-certificate/> - Загл. с экрана.
19. TheTransportLayerSecurity (TLS) [Электронный ресурс] // Tools.– Режим доступа: <https://tools.ietf.org/html/rfc5246>- Загл. с экрана.
20. HTTPOverTLS[Электронный ресурс] // Tools.– Режим доступа: <https://tools.ietf.org/html/rfc2818> - Загл. с экрана.

21. HypertextTransferProtocol (HTTP/1.1): Authentication[Электронный ресурс] // Tools. – Режим доступа: <https://tools.ietf.org/html/rfc7235> - Загл. с экрана.
22. The 'Basic' HTTPAuthenticationScheme[Электронный ресурс] // Tools. – Режим доступа:<https://tools.ietf.org/html/rfc7617> - Загл. с экрана.
23. TRANSMISSIONCONTROLPROTOCOL[Электронный ресурс] // Tools. – Режим доступа:<https://tools.ietf.org/html/rfc793> - Загл. с экрана.
24. Public-KeyCryptographyStandards (PKCS) [Электронный ресурс]. // Tools. – Режим доступа: <https://tools.ietf.org/html/rfc3447> - Загл. с экрана.
25. Diffie-HellmanKeyAgreementMethod [Электронный ресурс]. // IEFT. – Режим доступа: <https://www.ietf.org/rfc/rfc2631.txt>- Загл. с экрана.
26. HMAC: Keyed-HashingforMessageAuthentication[Электронный ресурс]. // IEFT. – Режим доступа: <https://tools.ietf.org/html/rfc2104.html> - Загл. с экрана.
27. TheMD5 Message-DigestAlgorithm [Электронный ресурс]. // IEFT. – Режим доступа: <https://www.ietf.org/rfc/rfc1321.txt>- Загл. с экрана.
28. ГОСТ Р ИСО/МЭК 9594-8-98 Информационная технология [Электронный ресурс]. // CNDTDocs. –Режим доступа: <http://docs.cntd.ru/document/1200028710> - Загл. с экрана.
29. Simple Authentication and Security Layer [Электронныйресурс]. // Itu – Режимдоступа: [www/ URL: http://www.itu.int.](http://www.itu.int) - Загл. с экрана.
30. IntroductiontoQRCode[Электронный ресурс] // Cgi-Bin – Режим доступа: <http://twiki.org/cgi-bin/view/Blog/BlogEntry201102x2> - Загл. с экрана.

Грищенко О.С., Чернышова А.В. Анализ алгоритмов аутентификации в распределенных программных системах. Представлен обзор и анализ существующих подходов и алгоритмов аутентификации, используемых в современных распределенных системах. Рассмотрен принцип работы распространенных протоколов аутентификации. Проанализирован протокол аутентификации SQRL и внесены предложения по усовершенствованию данного протокола.

Ключевые слова: аутентификация, алгоритм, SQRL, PKI, уязвимость, защищенность.

Grishchenko OS, Chernyshova A.V. Analysis of authentication algorithms in distributed software systems. A review and analysis of existing approaches and authentication algorithms used in modern distributed systems is presented. The principle of operation of common authentication protocols is considered. The SQL authentication protocol was analyzed and suggestions were made to improve this protocol.

Keywords: authentication, algorithm, SQRL, PKI, vulnerability, security.

Исследование методов и средств защиты авторского права в области разработки программного обеспечения для распределённых систем

Кубашевский Д.В., Чернышова А.В.
Донецкий национальный технический университет
dehax12@gmail.com, chernyshova.alla@rambler.ru

Кубашевский Д.В., Чернышова А.В. Исследование методов и средств защиты авторского права в области разработки программного обеспечения для распределённых систем. В статье представлено исследование недостатков существующих методов и средств защиты авторского права в области разработки программного обеспечения для распределённых систем. Предложен новый комбинированный подход к реализации распределённой системы защиты авторского права на программное обеспечение.

Ключевые слова: авторское право, защита интеллектуальной собственности, программное обеспечение, распределённые системы, безопасность данных, ApacheHadoop.

Введение

В настоящий момент проблема нелегального копирования и распространения программного обеспечения сохраняет свою актуальность. Каждый месяц появляются новые продукты, на создание которых разработчики тратят достаточно усилий и средств. Если затраченные средства не возмещаются за счёт покупки данных программных продуктов их пользователями, разработчик несёт убытки, и это чаще всего ведёт к закрытию проекта по разработке и сопровождению такого программного продукта [1].

Для предотвращения подобных ситуаций разработчик вынужден использовать программную защиту авторского права, которая не позволит использовать функционал программного продукта пользователям, которые не приобрели у разработчика лицензию на использование такого ПО.

Анализ недостатков метода автономной активации лицензии

Применение метода автономной активации лицензии на использование программного продукта (так называемой оффлайн-активации) теперь стало неэффективно, поскольку сегодня существует огромное количество инструментов, которые позволяют исследовать и нейтрализовать защиту. Этим успешно занимаются так называемые варез-группы (англ. warez group) — группы лиц, специализирующиеся на взломе технических средств защиты авторских прав (англ. DRM — digitalrightsmanagement). Наиболее популярными на текущий момент являются 3DM, RELOADED, CPY, SKIDROW, CODEX [2].

Таким образом, какой бы сложностью ни обладала автономная система защиты, в конечном итоге она всё равно может быть взломана и устранена; вопрос лишь во времени, которое на это потребуется.

В связи с этим возникает необходимость использования другого метода, метода так называемой онлайн-активации с применением внешнего сервера лицензий.

Анализ недостатков метода онлайн-активации лицензии

Так как злоумышленник не имеет доступа к внешнему серверу, при условии его надлежащей защищённости, следовательно, подменить ответ об успешности активации лицензии по предоставленным лицензионным данным не представляется возможным. В результате такой организации системы защиты по-прежнему остаётся уязвимость, позволяющая хакеру модифицировать бинарный код исполнимого модуля системы защиты таким образом, чтобы вне зависимости от успешности результата активации лицензии программный продукт полностью функционировал.

Устранить подобного рода уязвимость можно, применив метод защиты исполнимых данных, при котором важная часть функционала программного продукта остаётся неработоспособной без наличия на клиентской машине вспомогательного (или дополнительного) исполнимого модуля до момента успешной

активации лицензии на приобретённое ПО [3].

Анализ недостатков нераспределённых систем лицензирования

Для полноценного функционирования программного продукта, защищённого с использованием подсистемы защиты авторского права с применением одного сервера лицензий, необходимо обеспечить полную отказоустойчивость такого сервера. В случае потери соединения с сервером лицензий программный продукт перестанет работать до тех пор, пока соединение не восстановится.

В распределённой системе лицензирования одновременно работают несколько серверов лицензий. При отказе одного из серверов остальные продолжают работу. И в таком случае пользователи, которые потеряли соединение с отказавшим сервером, смогут переключиться к остальным работоспособным серверам автоматически.

Кроме того, если разработчик имеет множество программных продуктов, каждый из которых в свою очередь имеет огромное количество пользователей, необходимо обеспечить надлежащий уровень производительности для сервера лицензий. Для решения данной проблемы также необходимо применение системы, в которой нагрузка будет распределяться между различными вычислительными машинами. Это касается и хранилищ данных, и вычислительных кластеров [4].

Исследование методов и средств защиты авторского права в области разработки программного обеспечения для распределённых систем

Неудивительно, что, к сожалению, в открытых источниках информации отсутствуют сведения о внутренней организации защиты в существующих решениях по защите авторского права. Однако, можно исследовать принципы работы такой системы с точки зрения обычного пользователя. Проанализировав эту информацию и сопоставив её с современными методами и средствами защиты, которые используются в распределённых системах в целом, можно сделать выводы относительно их эффективности применительно к системе защиты авторского права на использование программного обеспечения.

Обычно для различных типов программного обеспечения используются разные подходы к защите авторского права. Например, для защиты современных компьютерных игр разработчики чаще всего применяют методы онлайн-активации с ограничением функционала мультиплеерного режима, если активация не была успешно выполнена. Кроме того, часто выполняется привязка игрового продукта к аккаунту пользователя, который содержит информацию о том, был ли приобретён данный продукт.

Например, в онлайн-сервисе цифрового распространения компьютерных игр Steam (см. рис. 1) пользователь может купить компьютерную игру и после оплаты данный игровой продукт будет зарегистрирован на аккаунте пользователя Steam [5].



Рисунок 1 – Интерфейс клиента Steam

Игры, которые распространяются в Steam, тесно интегрированы с его сервисами, поэтому даже в случае, если злоумышленник сможет обойти локальную защиту, сетевые сервисы, которые необходимы для функционирования мультиплеерного режима игры, всё равно останутся недоступными по причине того, что удалённые сервера Steam содержат информацию о том, что данный пользователь не покупал запущенный программный продукт.

Сравнение существующих систем управления лицензиями

Система управления лицензиями (или менеджер лицензий) — инструмент, используемый разработчиком

программного обеспечения, который контролирует, где и как программные продукты могут быть запущены.

Keugen — облачная служба лицензирования и распространения программного обеспечения для настольных приложений, серверных приложений и других цифровых продуктов (см. рис. 2). Keugen предоставляет API для разработчиков, с помощью которого можно интегрировать данную службу в свой программный продукт. Keugen поддерживает несколько видов лицензий: бессрочная, подписка, плавающая с ограничениями на количество устройств, плавающая с ограничениями на функциональные возможности, временные и комбинированные.



Рисунок 2 – Логотип Keugen

Как заявляет разработчик, данные внутри Keugen защищены и хранятся «настолько безопасно, насколько это возможно» [6]. Пароли всегда хешируются с применением криптографической функции bcrypt [7]. Все соединения с серверами шифруются через SSL, базы данных также зашифрованы, в том числе и их резервные копии — бэкапы (англ. — backupcopy).

Nalpeiron LicensingService — облачная служба лицензирования, которая поддерживает также несколько видов лицензий: бессрочная, подписка, плавающая с ограничениями на количество устройств, с ограничениями на функциональные и вычислительные возможности, лицензия с указанным сроком действия, многопользовательская лицензия.

Разработчик утверждает, что Nalpeiron LicensingService предоставляет возможность оффлайн активации, однако на самом деле подразумевается возможность активировать программный продукт для устройства, которое не имеет доступа к интернету, с другого устройства, у которого есть доступ к интернету. Для такой активации достаточно будет обычного браузера, с помощью которого необходимо перейти на сайт Nalpeiron LicensingService и ввести необходимые лицензионные данные [8].

License4J предоставляет решения для лицензирования и защиты программного обеспечения, написанного под платформу Java. License4J включает JavaAPI, инструмент менеджера лицензий, сервера для генерации и активации лицензий. Представляет собой небольшую Java-библиотеку, предоставляющую статические методы для проверки лицензии, онлайн-активации и деактивации.

Кроме стандартных типов лицензий, которые поддерживают Keugen и Nalpeiron, License4J также поддерживает четыре различных идентификатора оборудования: имя хоста, физический адрес сетевой карты, идентификатор системного диска UUID и серийный номер жёсткого диска. Эти идентификаторы могут быть использованы для генерации лицензии (см. рис. 3). Во время проверки лицензии License4J может выполнять проверку локальной даты и времени и сравнивать значения с такими, которые предоставляет удалённый NTP-сервер или веб-сервер License4J [9].

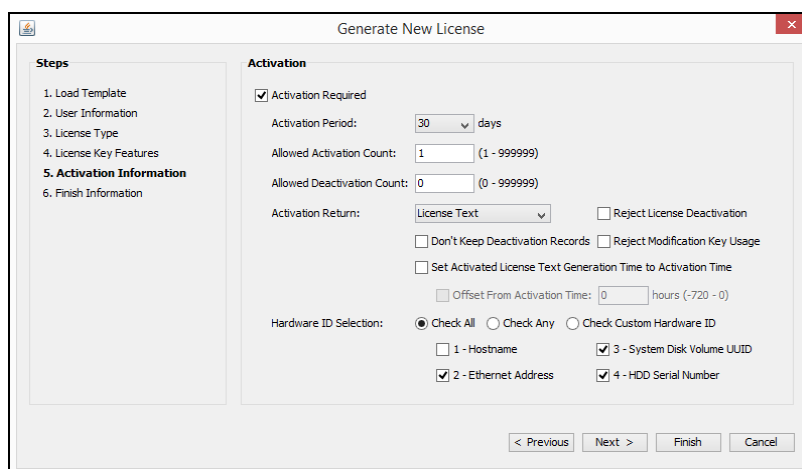


Рисунок 3 – Окно генерации лицензии License4J

SafeGuardLM — решение для лицензирования программного обеспечения, которое, подобно уже рассмотренным инструментам, предоставляет API для защиты авторского права и несколько видов системы лицензий. Для генерации лицензий существует инструмент SafeGuardManager (см. рис. 4), написанный на Java, который создаёт лицензии и сохраняет в них информацию о покупателе.

Описывая безопасность своего решения, автор SafeGuard указывает на возможность применения шифрования таким образом, что никакой другой разработчик программного продукта не будет иметь возможности сгенерировать лицензию для защищаемого программного обеспечения. Также предоставляется модуль антихакинга (с исходным кодом), который разработчик программного продукта может использовать для защиты от таких мощных инструментов исследования бинарного исполнимого кода как IDAPro [10], OllyDbg [11] и других, которые используют методы внедрения в бинарный исполнимый модуль программного обеспечения процессорных инструкций NOP или JMP [12].

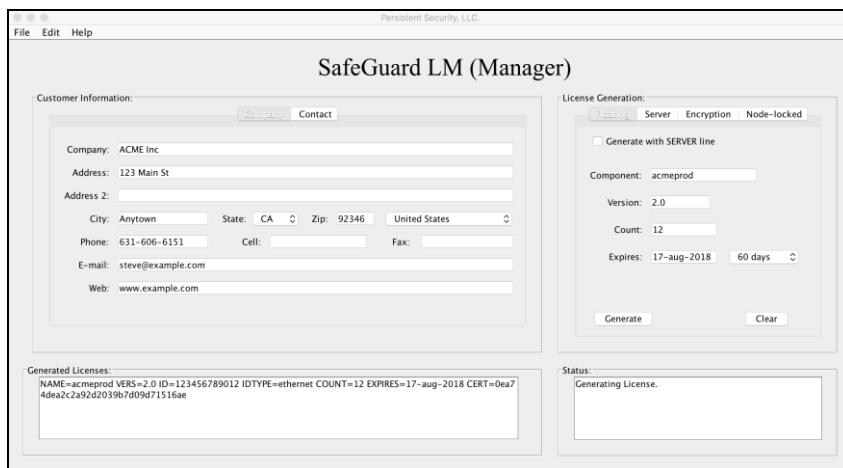


Рисунок 4 – Окно инструмента управления лицензиями SafeGuardLMManager

Для платформы WindowsSafeGuard предоставляет два уникальных сервера лицензий. Один осуществляет регулярную проверку на предмет перевода локального времени системы назад. Это самый распространённый пользовательский метод обмануть систему защиты. Другой сервер лицензий осуществляет расширенную проверку локальной даты и времени, чтобы противостоять таким хакерским инструментам, как DateCracker 2000 [13].

Исходя из описания разработчиков вышеперечисленных продуктов, характерной отличительной чертой RepriseLicenseManager (см. рис. 5) являются высокодоступные (или отказоустойчивые) распределённые сервера, которые способны одновременно обрабатывать множество запросов на активацию и проверку лицензии от пользователей. Кроме того, все лицензии подписаны электронной цифровой подписью (ЭЦП) с использованием открытого криптографического ключа, чтобы защитить доступ к лицензируемому программному продукту [14].

В описанных решениях по защите авторского права на программное обеспечение присутствует описанная выше уязвимость, связанная с возможностью полностью нейтрализовать систему защиты и получить доступ ко всем функциональным возможностям программного продукта. Для этого нужно всего лишь единожды купить одну копию защищаемого программного продукта, исследовать исполнимый бинарный код и модифицировать его таким образом, чтобы система защиты даже не осуществляла попыток проверить лицензию, либо же чтобы программное обеспечение игнорировало результат проверки и запускалось в любом случае.



Рисунок 5 – Логотип Reprise License Manager

Комбинированный подход к защите авторского права

Комбинированный подход к реализации распределённой системы защиты авторского права на программное обеспечение предполагает объединение двух основных методов:

- использование сервера лицензий, который будет передавать клиенту лицензируемый функционал в виде модуля исполнимого кода только после успешной активации лицензии;
- реализация сервера в виде распределённой системы, которая обеспечит надёжность и отказоустойчивость.

Согласно реализованному ранее алгоритму активации лицензии (см. рис. 6), после запуска клиентская часть системы защиты отправляет на сервер данные лицензии. После успешной проверки данных сервер подготавливает лицензируемый функционал программного продукта в виде модуля исполнимого кода, зашифровывает его уникальным ключом и отправляет клиенту. Клиент расшифровывает полученный модуль исполнимого кода и получает доступ к лицензируемому функционалу [4].

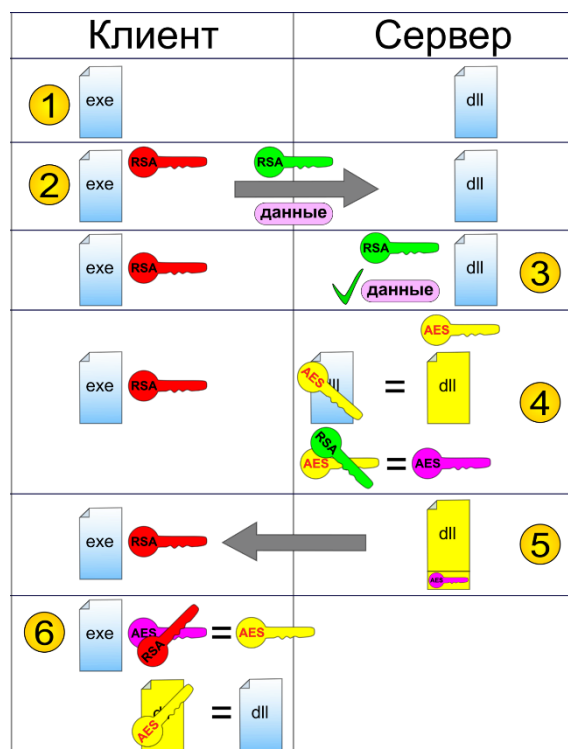


Рисунок 6 – Алгоритм активации лицензии [1]

Для реализации распределённой системы планируется использование свободно распространяемого фреймворка для разработки и выполнения распределённых программ Apache Hadoop [15]. Дистрибутив Cloudera Distribution for Hadoop (CDH) [16] включает в себя встроенные средства, которые позволяют обеспечить защиту кластеров, используя шифрование данных, аутентификацию пользователя и механизмы авторизации. Для хранения информации о лицензируемых программных продуктах, версиях, пользователях и лицензиях планируется использовать Apache HBase — нереляционную распределённую базу данных, которая также входит в состав CDH.

Выводы

Существующие распространённые методы и средства защиты авторского права в области программного обеспечения имеют ряд недостатков. Алгоритм проверки успешности активации лицензии не защищён от модификации, а серверная часть подобных систем не всегда обеспечивает требуемый уровень производительности и отказоустойчивости.

Для повышения эффективности распределённой системы защиты авторского права на программное обеспечение необходимо использовать комбинированный подход, объединив метод удалённой защиты лицензируемого функционала и метод распределённых вычислений для обеспечения отказоустойчивости системы защиты в целом.

Литература

1. Чернышова, А.В. Подсистема защиты авторского права в программном обеспечении / А.В. Чернышова, Д.В. Кубашевский // Информатика и кибернетика. – 2016. – № 2(4). – С. 68-72.
2. AllSceneGroupsandCrackers [Электронный ресурс] // CrackWatch. – Режим доступа: <https://crackwatch.com/groups> – Загл. с экрана.
3. Чернышова, А.В. Подсистема защиты авторского права в программном обеспечении [Текст] / А.В. Чернышова, Д.В. Кубашевский // Сборник материалов VIII Международной научно-технической конференции «Информатика, управляющие системы, математическое и компьютерное моделирование» (ИУСМКМ–2017). - Донецк: ДонНТУ, 2017. – С. 110-114.
4. Чернышова, А.В. Подсистема защиты авторского права в сфере программного обеспечения / А.В. Чернышова, Д.В. Кубашевский // Информатика и кибернетика. – 2017. – № 4(10). – С. 84-90.
5. Steam [Электронный ресурс] // Steam. – Режим доступа: <https://store.steampowered.com/about/> – Загл. с экрана.
6. Keygen [Электронный ресурс] // Keygen. – Режим доступа: <https://keygen.sh/> – Загл. с экрана.
7. Secure Salting Password Hashing - How to do it Properly [Электронный ресурс] // CrackStation. – Режим доступа: <https://crackstation.net/hashing-security.htm> – Загл. с экрана.
8. Software Licensing Cloud-Based [Электронный ресурс] // Nalpeiron. – Режим доступа: <https://www.nalpeiron.com/> – Загл. с экрана.
9. License4J [Электронный ресурс] // License4J. – Режим доступа: <https://www.license4j.com/> – Загл. с экрана.
10. IDA: About [Электронный ресурс] // Hex-Rays. – Режим доступа: <https://www.hex-rays.com/products/ida/> – Загл. с экрана.
11. OllyDbg [Электронный ресурс] // OllyDbg. – Режим доступа: <http://www.ollydbg.de/> – Загл. с экрана.
12. Tech [Электронный ресурс] // PersistentSecurity. – Режим доступа: <https://www.persistentsecurity.com/tech/> – Загл. с экрана.
13. DateCracker 2000 2.0 [Электронный ресурс] // Brothersoft. – Режим доступа: <http://www.brothersoft.com/date-cracker-2000-45650.html> – Загл. с экрана.
14. RLM [Электронный ресурс] // RepriseSoftware. – Режим доступа: <http://www.reprisesoftware.com/products/software-license-management.php> – Загл. с экрана.
15. ApacheHadoop [Электронный ресурс] // ApacheHadoop. – Режим доступа: <https://hadoop.apache.org/> – Загл. с экрана.
16. Overview of Cloudera and the Cloudera Documentation Set [Электронный ресурс] // Cloudera Documentation. – Режим доступа: <https://www.cloudera.com/documentation/enterprise/5-13-x/topics/introduction.html> – Загл. с экрана.

Кубашевский Д.В., Чернышова А.В. Исследование методов и средств защиты авторского права в области разработки программного обеспечения для распределённых систем. В статье представлено исследование недостатков существующих методов и средств защиты авторского права в области разработки программного обеспечения для распределённых систем. Предложен новый комбинированный подход к реализации распределённой системы защиты авторского права на программное обеспечение.

Ключевые слова: авторское право, защита интеллектуальной собственности, программное обеспечение, распределённые системы, безопасность данных, ApacheHadoop.

Denis Kubashevskiy, Alla Chernyshova Copyright protection methods and means research in the field of software development for distributed systems. The article deals with the research of existing

copyright protection methods and means disadvantages in the field of software development for distributed systems. The new combined approach to implementation of a software copyright protection system is proposed.

Keywords: *copyright, intellectual property protection, software, distributed systems, data security, Apache Hadoop.*

УДК 004.93; 004.5

Управление технологическими объектами посредством мобильного устройства с дополненной реальностью

Толстых В.К., Лихолетов А.В.
Донецкий национальный университет
mail@tolstykh.com, likholetovalex@gmail.com

Лихолетов А.В., Толстых В.К. Управление технологическими объектами посредством мобильного устройства с дополненной реальностью. В статье представлен краткий обзор современных технологий дополненной реальности и сферы их внедрения. Представлена концепция устройства позволяющего осуществлять управление технологическими объектами в режиме дополненной реальности. Приведен пример реализации устройства и схема взаимодействия его основных компонент.

Ключевые слова: *дополненная реальность, мобильное устройство, управление, распознавание образов.*

Введение

В настоящее время технологии дополненной реальности достигли высокого уровня развития и широко внедряются во всех сферах жизни. Образование, маркетинг, промышленные предприятия имеют множество примеров удачного использования технологий дополненной реальности, которые позволяют существенно оптимизировать существующие процессы, а также развивать их новые возможности.

Сама по себе дополненная реальность (augmented reality) представляет собой наложение виртуальных слоев, созданных посредством соответствующего программного обеспечения, на объекты реального мира в реальном времени [1].

По прогнозам ведущих мировых компаний, в сфере IT, в частности компании Apple, дополненная реальность является одной из наиболее активно развивающихся сфер IT-технологий, как в ближайшем, так и в отдаленном будущем. Также особый акцент сделан на том, что мобильные устройства будут неразрывно связаны с дополненной реальностью. По прогнозам экспертов, объем рынка дополненной и виртуальной реальности увеличится до \$150 млрд к 2020 году.

Google Glass массово поставляются на заводы многих американских компаний, включая General Electric и Boeing. Их применение повышает производительность труда, упрощает и ускоряет коммуникации. По прогнозу Forrester Research, к 2025 году почти 14,4 миллионов американских рабочих будут носить умные очки. Тем не менее, уровень развития, которого на сегодняшний день достигли технологии дополненной реальности, существенно опережает их реальное внедрение в производство, а грандиозный потенциал технологии далек от реализации [2].

Не менее актуальны и технологии удаленного управления, которые нашли применение как в промышленности, так и в повседневной жизни. Не взирая на то, что они возникли задолго до технологий дополненной реальности, именно развитие дополненной реальности открыло новые перспективы развития удаленного управления, поскольку позволило предоставить пользователю интерфейсы нового уровня, не ограниченные размерами и формой устройства.

К сожалению, на сегодняшний день комбинация технологий дополненной реальности и управления являются дорогостоящими, а потому многие прототипы так и не поступают в серийное производство.

С распространением универсальных мобильных устройств, таких как смартфоны, эти технологии становятся все более доступными как рядовому пользователю, так и промышленным предприятиям.

Как и в случае технологий дополненной реальности, потенциал технологий удаленного управления далек от реализации, а потенциал совмещения данных технологий в одном устройстве позволит серьезно расширить область их применения.

В настоящей работе рассматривается устройство позволяющее осуществлять управляющие воздействия на технологические объекты в режиме дополненной реальности.

Анализ современных разработок и сфер их внедрения

Компания Boeing – одна из крупнейших аэрокосмических компаний на планете, которая строит самолеты для 150 стран мира. В бортовых системах самолёта содержится множество компонентов, связанных между собой системой проводов. Укладка и соединение кабелей производится по специальному шаблону, после чего их скрепляют в жгуты, а на концы кабелей устанавливают разъемы. Процесс работы занимает длительное время и требует особого внимания и высокой точности. Последние 20 лет Boeing искала систему, способную сократить время на производство и устранение ошибок. В начале 2014 года компания внедрила решение дополненной реальности на платформе очков Google Glass. С помощью приложения оператор отдает голосовую команду: «Ок, Skylight. Начать создание жгута. Сканировать заказ 0447» и видит в очках дополненной реальности визуальную дорожную карту по сборке жгута № 0447.

Согласно данным из доклада компании Boeing по проекту, «использование Google Glass позволило сократить время производства на одну четверть и сократить количество ошибок в два раза» [3].

Другой пример – компания Lockheed Martin, где внедрением проекта занималась компания NGRAIN. Инженеры завода используют дополненную реальность для получения визуальной подсказки по сборке самолёта F-35. В качестве основной платформы в компании используются очки дополненной реальности Epson Moverio BT-200, оборудованные специальными датчиками движения и глубины. Когда техник монтирует деталь тормоза на шасси, в очках дополненной реальности он видит все данные о том, где и в каком порядке нужно проводить сборку и подсоединять кабели. В итоге весь процесс работы напоминает сборку конструктора LEGO, когда от рабочего требуется лишь взять подходящую деталь и поместить её в нужное место [4].

По данным компании NGRAIN, «программное обеспечение позволяет инженерам работать на 30 процентов быстрее с точностью до 96 процентов».

Модели автомобилей выпускаются в различных комплектациях и цветах. Поэтому автоконцерны часто стремятся находить баланс между массовым производством и индивидуальными предпочтениями потребителя. Конечно, во времена Генри Форда всё было намного проще. Организация такого производства часто сопровождается сложными рабочими инструкциями, которые способны спровоцировать простои, ошибки и снижение производительности.

Концерн Fiat Chrysler Automobiles (FCA) применил в своей работе проекционную систему дополненной реальности OPS Solutions. Теперь на каждом этапе сборочного процесса рабочие получают наглядную информацию о своём следующем шаге.

После внедрения инновации был организован эксперимент. Перед операторами стояла задача собрать зубчатые передачи и цепи. Весь процесс проходил в десять последовательных шагов. Сотрудники должны были выбрать правильные комплектующие, провести монтаж и убедиться, что всё сделано правильно. Часть операторов выполняла задачу с помощью бумажных инструкций, а другие использовали инструменты дополненной реальности. Результаты эксперимента:

- сокращение числа ошибок 80%;
- сокращение времени рабочего цикла 38%;
- увеличение пропускной способности 82%.

Другой автомобильной компанией, принявшей на вооружение технологии дополненной реальности, стал концерн Volkswagen. На своём заводе в Вольфсбурге Volkswagen использует очки дополненной реальности Google Glasses при комплектации заказов. Рабочие автоматически получают всю необходимую им информацию о местах хранения или номерах деталей.

Первое российское приложение дополненной реальности было выпущено компанией VR CORP для электромонтажного производства. Руководство компании «Технологии Энергосбережения Сибири» столкнулось с текучкой кадров, в результате чего много средств и усилий тратилось на обучение вновь прибывающих сотрудников. Для офисного персонала было создано приложение, с помощью которого любой сотрудник мог навести смартфон на электротехническую схему и рассмотреть в подробностях 3D-модель готового изделия, включая все её комплектующие. Особо это приложение актуально для молодых специалистов без опыта работы. Вводный курс обучения специальности сократился до простой схемы: взял смартфон, запустил приложение, навёл на электротехническую схему и рассмотрел все подробности об изделии в его готовом виде.

Похожее по функционалу приложение было разработано для рабочих сборочного цеха. Вместо электромонтажной схемы специалист получает информацию о размещении деталей и кабелей в будущем изделии, а подсветка каждой детали вовремя предупреждает, если используются неправильные комплектующие.

За свою разработку компания VR CORP получила диплом форума «Городские технологии» от Департамента промышленности, инноваций и предпринимательства мэрии города Новосибирска [5].

В результате анализа передовых технологий дополненной реальности, используемых в промышленности, не было обнаружено акцентированное внимание на проблемы и технологии управления технологическими объектами. Тем не менее, тенденции развития технологий свидетельствуют о том, что управление технологическими объектами необходимо и в ближайшее время найдет широкое применение в различных производственных процессах.

Поэтому представляется актуальным исследование и разработка технологий и устройств управления технологическими объектами, которые позволят в режиме дополненной реальности осуществлять управляющие воздействия на процессы, протекающие в них.

Концепция устройства

Анализ современного уровня развития технологий дополненной реальности, распознавания образов и удаленного управления, а также существующих устройств и концепций, базирующихся на этих технологиях, позволил разработать концепцию мобильного устройства с дополненной реальностью, позволяющего осуществлять удаленное управление технологическими объектами закрытого типа и процессами, протекающими в них.

Концепция планируемого устройства заключается в том, что, имея мобильное устройство с камерой, экраном и вычислительными модулями можно управлять технологическими объектами в режиме дополненной реальности.

При этом дополненная реальность играет крайне важную роль, поскольку демонстрирует пользователю не только базовую информацию об объекте, но и передает ему данные полученные с датчиков, находящихся внутри технологического объекта, что позволяет в реальном времени отслеживать состояние процессов, протекающих внутри объекта. Это особенно важно в тех случаях, когда речь идет о закрытых технологических объектах, чье состояние невозможно оценить невооруженным глазом.

Данная концепция подразумевает наличие возможности распознавать технологический объект, как посредством обученной нейросети, определяющей тип объекта по его форме, окрасу и иным характеристикам, так и посредством распознавания определенного маркера расположенного на поверхности объекта (например, QR-кода). Это связано с тем, что многие технологические объекты имеют схожие очертания и во избежание ошибок и потребления чрезмерного количества ресурсов устройства, для распознавания мелких деталей. Также такая возможность необходима, если устройство пользователя обладает достаточными вычислительными мощностями для работы с нейросетью. Также снижение потребляемых вычислительных мощностей позволяет существенно снизить стоимость устройства и расширяет спектр мобильных устройств совместимых в разработанным программным обеспечением для вычислительных модулей.

Реализация устройства

На основе приведенной выше концепции была разработана предполагаемая реализация устройства. Рассмотрим более подробно основные компоненты..

Прежде всего рассмотрим мобильное устройство. Оно должно включать следующие компоненты:

- камера;
- экран;
- модуль распознавания образов;
- модуль дополненной реальности;
- модуль передачи данных;
- базу данных.

Таким образом можно сделать вывод, что для реализации устройства подойдет любой смартфон, обладающий камерой и достаточными вычислительными мощностями для работы модулей распознавания образов, дополненной реальности и передачи данных. Необходимые вычислительные мощности могут варьироваться в зависимости от программного обеспечения, реализующего модули, а также количества и сложности датчиков и контролеров технологического объекта.

Роль мобильного устройства заключается в том, чтобы с помощью камеры передать изображение в реальном времени в модуль распознавания образов, который в свою очередь, должен определить, находится ли в поле зрения камеры какой-либо из объектов, загруженных в базу данных. В случае обнаружения подходящего объекта, данные передаются в модуль дополненной реальности, который выводит на экран перечень объектов, найденных в поле зрения камеры. Получив данные о найденных объектах, от пользователя требуется выбрать, с каким объектом он желает продолжать работу. После этого модуль передачи данных отправляет на сервер соответствующий запрос.

Сервер является вторым важным элементом. Он выступает посредником между мобильным устройством и технологическим объектом. Сервер получает данные от мобильного устройства о том, с каким объектом

желает взаимодействовать пользователь и подключается к нему, получая данные с внутренних датчиков устройства и передавая данные на внутренние контроллеры, отвечающие за управление процессами, протекающими внутри объекта. Для этого на сервере располагается модуль расчета управляющих воздействий, который отвечает за корректную обработку управляющих воздействий, переданных пользователем, определяет возможность подобного воздействия, рассчитывает возможные последствия, а также блокирует воздействия, которые могут привести к нежелательным последствиям вследствие ошибки пользователя.

Технологический объект включает датчики и контроллеры, которые связаны с сервером. Концепция устройства позволяет настраивать взаимодействие с любыми технологическими объектами, которые содержат датчики и контроллеры, позволяющие организовать обмен данными с сервером, а также процессы, которые подразумевают удаленное управление без непосредственных физических воздействий.

Таким образом связь этих трех компонентов позволяет осуществлять мониторинг состояния технологического объекта и удаленное управление его внутренними процессами.

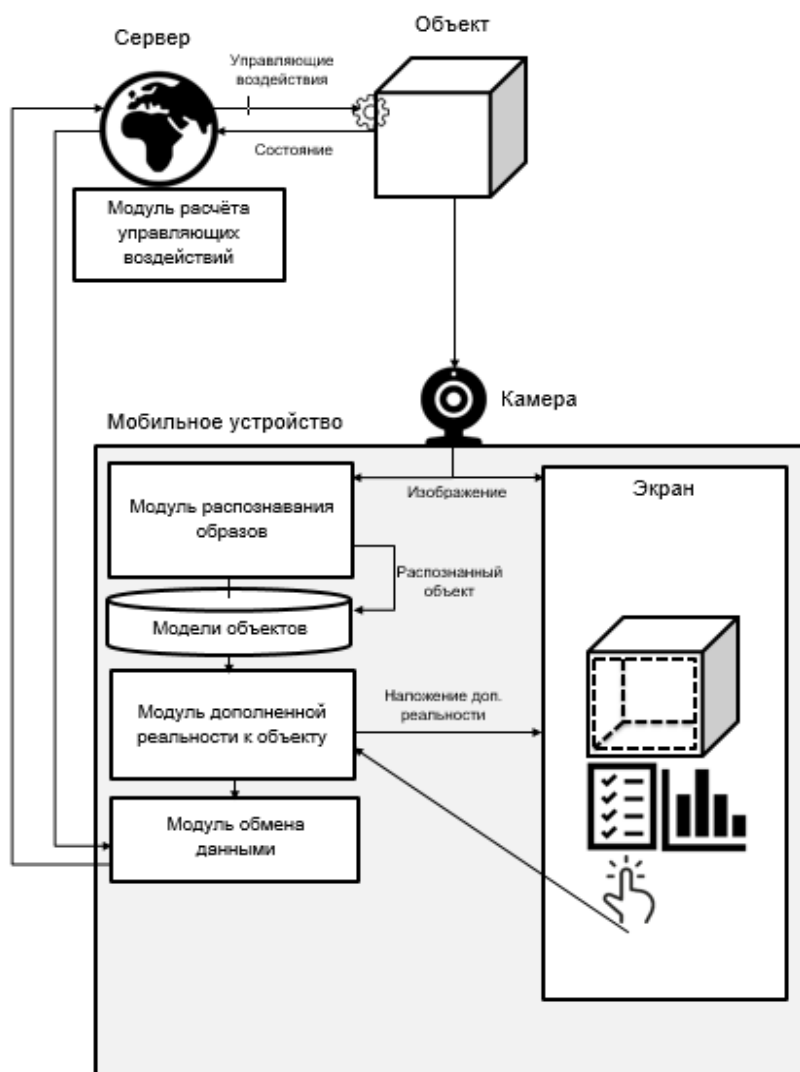


Рисунок 1 – Схема взаимодействия компонентов

Моделирование работы технологического объекта

Для управления объектами необходимо знание их достаточно точных математических моделей. Промышленные объекты обычно имеют сложные математические модели, вплоть до систем нелинейных уравнений в частных производных. Для производственных процессов они известны, известны методы их решения, но процесс их решения является достаточно ресурсоемким, что приводит к значительным временным затратам. В то же время, работа рассматриваемого мобильного устройства требует практически мгновенной реакции для контроля состояния технологических объектов и управляющих воздействий.

По этой причине необходимо отказаться от традиционного математического моделирования и

использовать упрощенные математические модели, не требующие значительных ресурсов и обеспечивающих достаточное быстродействие. Имитационное моделирование позволяет с достаточной точностью описать реальную систему и получить представление о процессах протекающих в ней, и их реакции на управляющие воздействия.

В контексте данной работы может применяться дискретно-событийное моделирование, подход предлагающий абстрагироваться от непрерывной природы событий и рассматривать только основные события моделируемой системы. Данный подход и наиболее развит на сегодняшний день и наиболее подходит для моделирования производственных процессов [6].

Также может применяться системная динамика, где для исследуемой системы строятся графические диаграммы причинных связей и глобальных влияний одних параметров на другие во времени, а затем созданная на основе этих диаграмм модель имитируется на компьютере. По сути, такой вид моделирования более всех других парадигм помогает понять суть происходящего выявления причинно-следственных связей между объектами и явлениями [7].

Вывод

В свете вышесказанного можно заключить, что описанное в статье устройство имеет большой потенциал. Оно может быть успешно внедрено на производство, как недорогостоящий и эффективный способ удаленного управления производственными процессами.

Литература

1. Дополненная реальность [Электронный ресурс] – Электрон. дан. - 2014. - Режим доступа: https://neurowareblog.blogspot.com/2014/12/blog-post_18.html - Загл. с экрана.
2. Дополненную реальность начали массово применять в американской промышленности [Электронный ресурс] – Электрон. дан. - 2017. - Режим доступа: <https://habr.com/post/402505/> - Загл. с экрана.
3. Google Glass takes flight at Boeing [Электронный ресурс] – Электрон. дан. - 2016. - Режим доступа: <https://www.cio.com/article/3095132/wearable-technology/google-glass-takes-flight-at-boeing.html> - Загл. с экрана.
4. Lockheed Is Using These Augmented Reality Glasses to Build Fighter Jets [Электронный ресурс] – Электрон. дан. - 2015. - Режим доступа: <https://www.popularmechanics.com/flight/a13967/lockheed-martin-augmented-reality-f-35/> - Загл. с экрана.
5. 2017-й год для виртуальной и дополненной реальности: как технологии придут на массовый рынок [Электронный ресурс] – Электрон. дан. - 2017. - Режим доступа: <http://www.forbes.ru/tehnologii/340601-2017-y-god-dlya-virtualnoy-i-dopolnennoy-realnosti-kak-tehnologii-pridut-na> - Загл. с экрана.
6. Хемди А. Таха, Введение в исследование операций. - 7-е изд. – М.: «Вильямс», 2007. – 737 с.
7. Строгалев В. П., Толкачева И. О., Имитационное моделирование. – МГТУ им. Баумана, 2008. – 737 с.

Толстых В.К., Лихолетов А.В. Управление технологическими объектами посредством мобильного устройства с дополненной реальностью. В статье представлен краткий обзор современных технологий дополненной реальности и сферы их внедрения. Представлена концепция устройства позволяющего осуществлять управление технологическими объектами в режиме дополненной реальности. Приведен пример реализации устройства и схема взаимодействия его основных компонент.

Ключевые слова: дополненная реальность, мобильное устройство, управление, распознавание образов.

Tolstykh V.K., Likholetov A.V. Management of technological objects through a mobile device with augmented reality. The article presents a brief overview of modern technologies of augmented reality and the scope of their implementation. The concept of the device that allows you to manage technological objects in augmented reality mode. The example of the implementation of the device and the interaction scheme of its main components is given.

Keywords: augmented reality, mobile device, control, pattern recognition.

Выбор системы управления базами данных для программного комплекса приемной комиссии ДонНТУ

Чередникова О.Ю., Щедрин С.В., Ногтев Е.А., Исаков А.Ю.
 Донецкий национальный технический университет
 – cherednikova@donntu.org, do010575ssv@gmail.com

Чередникова О.Ю., Щедрин С.В., Ногтев Е.А., Исаков А.Ю. Выбор системы управления базами данных для программного комплекса приемной комиссии ДонНТУ. В статье анализируются пять самых популярных СУБД - Oracle, MySQL, Microsoft SQL Server, PostgreSQL и MongoDB. Сравнение производилось в соответствии с критериями, выдвигаемыми к комплексу приемной комиссии ДонНТУ. Сделан вывод о наиболее предпочтительной СУБД для поставленной задачи.

Ключевые слова: СУБД, таблицы базы данных, SQL, индексы.

Введение

Рациональный выбор системы управления базами данных (СУБД), является одним из наиболее важных этапов при разработке базы данных (БД) и программного комплекса автоматизации работы приемной комиссии. Этот этап предполагает четкий и глубокий сравнительный анализ на основании ряда критериев. Выбор СУБД определяется распространенностью программного продукта, ценой, совместимостью с операционными системами, набором функций, которые СУБД предоставляет разработчику и пользователю, возможность взаимодействия с другими видами программного обеспечения, легкостью формирования запросов, отчетов, форм и особенностями конкретного проекта.

Однако есть и не сопоставимые критерии, например, оценка скорости работы двух СУБД. На первый взгляд, все просто: на различных нагрузках выполняется тестовый прогон, а затем на основе полученных результатов принимается окончательное решение [1]. Но как быть, если при одной нагрузке выигрывает одна система, а при другой — другая. Выбор осложняется, когда одна из систем работает медленнее, но зато лучше масштабируется, требуя меньше ресурсов.

Критерии популярности в интернет-источниках и популярность среди пользователей имеют преимущество перед другими, поскольку можно заметить, что они частично или полностью содержат другие предложенные критерии. Примером может служить цена СУБД в коммерческом использовании, которая, как правило, тесно коррелирует с ее популярностью. Так же следует подчеркнуть, что на популярность влияют, как положительные, так и неблагоприятные упоминания.

Рейтинг 10 наиболее популярных СУБД, согласно ресурсу DB-Engines по состоянию на октябрь 2018 года, приведен в таблице 1.

Таблица 1 –Десять самых популярных СУБД

Рейтинг	Название СУБД	Тип СУБД
1	Oracle	реляционная СУБД
2	MySQL	реляционная СУБД
3	Microsoft SQL Server	реляционная СУБД
4	PostgreSQL	постреляционная СУБД
5	MongoDB	документоориентированная СУБД
6	DB2	реляционная СУБД

7	Redis	NoSQL
8	Elasticsearch	Поиск и индексация
9	Microsoft Access	реляционная СУБД
10	Cassandra	NoSQL

Постановка задачи

Программный комплекс Приемной комиссии ДонНТУ представляет собой систему автоматизации работы операторов, членов приемной комиссии, представителей деканатов и других служб университета в период вступительной кампании. До настоящего времени в качестве СУБД в программном комплексе использовалась MS SQLServer. С каждым годом в соответствии текущими потребностями структура БД расширялась, усложнялась, появлялись новые разработчики программного обеспечения, которые вносили, как в сам программный комплекс, так и в структуру базы данных свои особенности. В итоге программный комплекс представляет собой набор разрозненных приложений, написанных с использованием различных технологий программирования. Поэтому на данном этапе перед разработчиками комплекса стоит задача проектирования нового программного обеспечения. И прежде всего необходимо оптимизировать структуру БД, в частности устранить избыточности и противоречивости существующей БД. Необходимо также определиться с СУБД, которая могла бы обеспечить большую защищенность данных и позволяла бы более оптимально работать с ними, как на стороне сервера, так и на стороне клиента.

Поэтому целью статьи является анализ современных СУБД в соответствии с критериями, предъявляемыми к организации хранения и работы с данными об абитуриентах, а также выбор наиболее оптимальной СУБД для комплекса Приемной комиссии ДонНТУ.

Анализ СУБД в соответствии с особенностями структурирования информации в программном комплексе приемной комиссии

Сравнительный анализ проведен среди наиболее популярных в настоящее время СУБД: Oracle, MS SQL Server, MySQL, PostgreSQL, MongoDB.

Рассмотрим основные критерии, которым должна соответствовать СУБД программного комплекса приемной комиссии и какие СУБД им удовлетворяют:

Возможность хранить в ячейке таблицы набор нескольких объектов.

Структура некоторых таблиц предполагает наличие столбцов, содержащих набор информации. Например, сведения о сертификатах абитуриента (ЕГЭ, ГИА, ВНО) должны содержать оценки по нескольким предметам. Логично было бы хранить их в одном столбце, который бы содержал массив записей типа «предмет – оценка».

В Oracle типом столбца может быть коллекция - структура данных, по своей функциональности сходная со списком или одномерным массивом [2]. Этот термин может иметь несколько значений: столбец таблицы, где хранятся значения типа вложенной таблицы - TABLE или массива VARRAY. Главное различие между вложенными таблицами и структурами VARRAY проявляется при использовании их в качестве типов данных столбцов. Хотя массив VARRAY, подобно вложенной таблице, позволяет сохранить в одном столбце множество значений, для него необходимо указать максимальную длину массива, который будет храниться в таблице вместе с остальными данными. Поэтому разработчики компании Oracle рекомендуют использовать столбцы типа VARRAY для «маленьких» массивов, а вложенные таблицы — для «больших». Доступ к любому элементу вложенной таблицы или varray-массива осуществляется по его индексу, который указывается в скобках после имени переменной типа коллекции.

MySQL по умолчанию не поддерживает тип данных массива [3]. Поэтому нет возможности хранить массив непосредственно в базе данных mysql.

В Transact-SQL Microsoft SQL Server нет такой структуры, как массив. T-SQL не предусматривает полностью функциональную структуру массива. Однако SQL 2000 ввел новую возможность, называемую переменной табличного типа, которая позволяет имитировать массив и/или использоваться вместо курсора в SQL Server.

PostgreSQL позволяет определять колонки таблицы как многомерные массивы переменной длины. Элементами массивов могут быть любые встроенные или определённые пользователями типы, перечисления или составные типы [4]. Массивы доменов в данный момент не поддерживаются. Столбцы в таблице в PostgreSQL могут представлять массивы, например, массивы чисел INTEGER или массивы строк VARCHAR или массивы других типов данных. Массив определяется в кавычках, как и строка, но внутри кавычек все элементы массива помещаются в фигурные скобки. Каждый отдельный элемент массива заключается в двойные кавычки. С помощью индексов можно конкретизировать, какие именно элементы массива необходимо получить:

Индексы указываются в скобках через двоеточие. Вначале идет индекс начала диапазона, а затем индекс конца диапазона. То есть в данном случае извлекаем три тега с первого по третий.

Данный критерий неприменим к MongoDB, так как это документо-ориентированная СУБД.

Возможность организации процедур, выполняемых на уровне сервера для минимизации времени выполнения запросов к базе данных.

В Oracle хранимые процедуры можно писать на языках PL/SQL и Java. Тело процедуры строится по правилам PL/SQL. Вызов процедуры состоит просто в употреблении ее имени со списком параметров.

Можно определить три типа параметров, которые могут быть объявлены:

IN — Параметр может ссылаться на процедуру или функцию. Значение параметра не может быть изменено процедурой или функцией.

OUT — параметр не может ссылаться на процедуру или функцию, но значение параметра может быть изменено процедурой или функцией.

IN OUT — Параметр может ссылаться на процедуру или функцию и значения параметра может быть изменено процедурой или функцией.

В MySQL хранимые процедуры появились в 5 версии. Они позволяют автоматизировать сложные процессы на уровне MySQL, нежели использовать для этого внешние скрипты. Процедуры можно создавать как запросы, например, через командную строку MySQL, но для удобства можно использовать MySQL GUI Tools. Данный пакет включает в себя три программы — MySQL Administrator, MySQL Query Browser и MySQL Migration Toolkit. При создании, модификации, удалении хранимых процедур сервер манипулирует с таблицей mysql.proc

Начиная с MS SQL 2005 появилась возможность написания хранимых процедур на любом языке .NET. Хранимые процедуры являются полноценными объектами базы данных, а потому каждая из них хранится в конкретной базе данных. Непосредственный вызов хранимой процедуры возможен, только если он осуществляется в контексте той базы данных, где находится процедура.

В SQL Server имеется несколько типов хранимых процедур.

Системные хранимые процедуры предназначены для выполнения различных административных действий.

Пользовательские хранимые процедуры реализуют те или иные действия.

Временные хранимые процедуры существуют лишь некоторое время, после чего автоматически уничтожаются сервером. Они делятся на локальные и глобальные. Локальные временные хранимые процедуры могут быть вызваны только из того соединения, в котором созданы. Глобальные временные хранимые процедуры доступны для любых соединений сервера, на котором имеется такая же процедура.

В PostgreSQL имеются очень мощные функции, которые в своём большинстве совместимы с тем, что в других СУБД называют хранимыми процедурами, а в некоторых случаях они могут больше. Эти функции могут быть различных типов и могут быть написаны на разных языках программирования (PgSQL, C, Perl, Java, PHP, Python, R, Ruby, Scheme, PL/sh). Основные ограничения хранимых функций в PostgreSQL по сравнению с - настоящими хранимыми процедурами это:

- неспособность возвращать несколько наборов данных,
- отсутствие поддержки автономных транзакций (BEGIN, COMMIT и ROLLBACK внутри функций),
- отсутствие поддержки стандартного SQL-синтаксиса CALL.

MongoDB позволяет сохранять функции, написанные на языке JavaScript, который используется как скриптовый язык в оболочке и как язык встроенных хранимых процедур. Это позволяет исполнять почти любой код на сервере БД. Для этого используется специальная системная коллекция system.js.

В связи с исключительной важностью хранимой информации, необходимость обеспечения максимальной защиты данных.

СУБД Oracle предлагает одни из лучших в отрасли средства защиты информации, такие, как защита ячеек/строк, защиты столбцов, фильтрация столбцов, детализированный аудит, шифрование данных, промежуточная авторизация, а также механизм контроля ролей в контексте обычных и защищенных приложения. Все это является дополнением к общепринятым средствам защиты информации, таким как аудит, проверка сложности пароля, гибкий контроль ролей, хранимых процедур и функций в базе данных.

MySQL использует безопасность, основанную на Access Control Lists (ACL) для всех соединений, запросов и других операций, которые пользователи могут попытаться выполнить. Есть также поддержка SSL-зашифрованных соединений между клиентами и серверами MySQL. Но есть и недостатки модели безопасности MySQL при её стандартной базовой настройке и установке «из коробки»:

Отсутствие встроенных возможностей по созданию и управлению группами пользователей или их ролями

Отсутствие развитой подсистемы управления паролями:

Нет возможности клонирования настроек, создания множества однотипных учетных записей, управления и мониторинга большого числа пользовательских аккаунтов.

Настройки по-умолчанию небезопасны.

Сами разработчики MySQL рекомендуют для устранения подобных «врожденных» уязвимостей использовать свой же скрипт `mysql_secure_installation` поставляемый вместе с дистрибутивом.

Среда MS SQL Server предоставляет множество различных функций для создания безопасных приложений баз данных. Платформа безопасности MS SQL Server управляет доступом к защищаемым сущностям при помощи проверки подлинности и авторизации. MS SQL Server поддерживает два режима проверки подлинности: режим проверки подлинности Windows и режим смешанной проверки подлинности.

MS SQL Server шифрует данные, используя иерархическую структуру средств шифрования и управления ключами. На каждом уровне данные низшего уровня шифруются на основе комбинации сертификатов, асимметричных ключей и симметричных ключей. Асимметричные и симметричные ключи можно хранить вне модуля расширенного управления ключами MS SQL Server.

Подсистема аудита появилась в версии MS SQL Server 2008, она позволяет проверять события ядра базы данных и настраивать параметры проверки.

Postgres обеспечивает шифрование на разных уровнях и даёт гибкость в выборе средств защиты данных в случае кражи сервера, от недобросовестных администраторов или в небезопасных сетях. Шифрование может также требоваться для защиты конфиденциальных данных, например, медицинских сведений или финансовых транзакций.

MongoDB предоставляет различные функции, такие как аутентификация, контроль доступа, шифрование, для обеспечения безопасности развертываний MongoDB. Более ранние версии MongoDB были уязвимы к автоматизированным атакам, поскольку по умолчанию для взаимодействия с БД не нужна была аутентификация. Эти уязвимости были частично устранены в релизе 3.x.

Обеспечение разграничения доступа к таблицам для различных пользователей

Доступ к данным различается для различных членов приемной комиссии, а также операторов, разработчиков. Ограничение доступа должно выполняться как на уровне доступных таблиц, так и на уровне запросов к таблицам. Т.е. деканы информацию могут только просматривать, разработчики должны иметь полный доступ к данным, а при обращении к БД с сайта университета доступ нужно давать только к определенным таблицам. Поэтому необходимо иметь возможность группировать пользователей для упрощения администрирования привилегий.

Все анализируемые СУБД имеют традиционные средства обеспечения безопасности данных с помощью прав и ролей, которые можно считать набором прав. Роли могут быть назначены пользователю или группе пользователей и определяют возможности доступа к таблицам. Однако, для Приемной комиссии актуальной является задача разграничения доступа на уровне строк таблицы. Например, представители деканатов для исключения ситуаций «переманивания» абитуриентов, подавших заявления на несколько факультетов, не должны видеть информацию других деканатов.

Такие ограничения можно осуществить с помощью политик безопасности, основанных на технологии RLS (row-level security или безопасность на уровне строк). Политики безопасности позволяют либо «закрыть» информацию полностью или частично, либо разрешить лишь определенные операции над ней.

RLS-технология была впервые представлена в Oracle 8i, но в последующих версиях ее возможности были значительно расширены. При связи объекта БД с политикой безопасности контроль доступа осуществляется через логику, занесенную в специальную PL/SQL-функцию. Для обеспечения детального управления безопасностью внутри базы данных в Oracle используются два связанных между собой механизма: контекст приложения (application context) и политику детального контроля доступа (fine-grained access control — FGAC). Для обозначения реализации политик детального контроля доступа посредством контекстов приложений в Oracle применяется термин виртуальная приватная база данных (virtual private database). Часто понятия «детальный контроль доступа», «виртуальная приватная база данных» и «безопасность на уровне строки» используют взаимозаменяемо для ссылки на возможность Oracle обеспечения безопасности на уровне отдельной строки, а не на уровне таблицы.

В MySQL существует возможность создавать только свои роли и закреплять за ними нужные привилегии, политик безопасности не реализованы.

Во всех версиях MS SQL Server используется безопасность на основе ролей. Однако, начиная с SQL Server 2016, предоставляется возможность создания политик безопасности, основанных на RLS.

PostgreSQL использует концепцию ролей (roles) для управления разрешениями на доступ к базе данных. Любая роль может использоваться в качестве пользователя, группы, и того и другого. Кроме этого, в PostgreSQL реализованы политики. Политика даёт разрешение на выборку, добавление, изменение или удаление строк, удовлетворяющих соответствующему выражению политики. Политики могут применяться для определённых команд или для определённых ролей.

В MongoDB есть только роли, которые выключены по умолчанию.

Возможность создавать код, который будет вызываться автоматически СУБД в ответ на определенные события

Функции, вызывающиеся по какому-либо событию в СУБД, называются триггерами. В программном комплексе Приемной комиссии такие функции нужны, например, чтобы как только абитуриенту будут выставлены экзаменационные оценки, тут же пересчитался бы его конкурсный балл.

Триггеры для Oracle можно писать на собственном языке программирования этой СУБД, носящим название PL/SQL (Programming Language for SQL - язык программирования для SQL), или на Java. Oracle поддерживает три вида триггеров: предваряющие (BEFORE), замещающие (INSTEAD OF) и завершающие (AFTER). Каждый из них, как и любой другой объект БД, после создания хранится в словаре данных в виде Р-кода. Ранее до версии Oracle 7.3 триггеры хранились в словаре данных, в виде исходного кода. И каждые раз при вызове компилировались, а затем исполнялись. В более старших версиях Oracle, триггеры хранятся уже в скомпилированном виде. Активация триггеров происходит при выполнении операторов DML.

Триггеры были введены в MySQL начиная с версии 5.0.2. Синтаксис триггеров несколько чужероден. MySQL использует стандарт ANSI SQL:2003 для процедур и других функций. Триггеры выполняются как процедуры при событиях UPDATE, DELETE и INSERT. Они могут быть выполнены либо до, либо после определения события. Таким образом Вы можете определить триггер, которые будет выполняться перед DELETE или после DELETE, и так далее. Это значит, что можно иметь один триггер, который выполнится до INSERT и совершенно другой, который выполнится после INSERT, что является весьма мощным инструментом.

В самом SQL Server есть два основных типа триггеров: триггеры Data Definition Language, или языка определения данных (DDL), и Data Manipulation Language, или языка управления данными (DML). До версии SQL Server 7.0, можно было создать только три триггера на одной таблице, один INSERT, один UPDATE и один DELETE. Ситуация изменилась с выходом SQL Server 7.0. Появилась возможность на одну таблицу задать большое количество (Ограничено числом объектов в базе данных) триггеров каждого типа. Порядок запуска триггеров при этом никак не определен и не гарантирован. В SQL Server 2000 появилась документированная процедура `sp_settriggerorder`, которая позволяет изменять порядок запуска триггеров. Можно указать какие триггеры AFTER срабатывают первыми или последними. Триггеры AFTER, срабатывающие между первым и последним триггерами, выполняются в неопределенном порядке.

В PostgreSQL можно создавать триггерные процедуры, которые будут вызываться при изменениях данных или событиях в базе данных. Триггерная процедура создается командой CREATE FUNCTION, при этом у функции не должно быть аргументов, а типом возвращаемого значения должен быть trigger (для триггеров, срабатывающих при изменениях данных) или event_trigger (для триггеров, срабатывающих при событиях в базе). Для триггеров автоматически определяются специальные локальные переменные с именами вида TG_переменная, описывающие условие, повлекшее вызов триггера.

В MongoDB нет поддержки триггеров.

Возможность хранить выборку из нескольких таблиц как виртуальную таблицу

Такая возможность в СУБД имеет название представления. Представления скрывают от прикладной программы сложность запросов и саму структуру таблиц БД. Ее необходимость для задач приемной комиссии обоснована тем, что именно для представлений дается доступ сайту (для обеспечения безопасности данных при взломе сайта). Также представления используются программным комплексом АСУ «Деканат» для работы с зачисленными студентами.

Представление в Oracle — результат хранимого запроса, поэтому в словаре данных сохраняется только определение представления. При экспорте базы данных Oracle можно видеть предложение “exporting views” (“экспорт представлений”), но под этим имеется в виду только определения представлений, а не физические объекты. При условии, что пользователь имеет соответствующие права доступа к лежащим в основе представления таблицам, можно запрашивать представления или даже модифицировать, удалять либо добавлять данные с использованием операторов UPDATE, DELETE и INSERT. Например, если выдать только привилегию INSERT для базовой таблицы, на которой определено представление, то этот пользователь сможет только вставлять строки в таблицу, но не сможет ни выбрать, ни вставить, ни удалить строки.

MySQL начиная с версии 5.0 поддерживает представления, включая обновляемые. Представление действует как виртуальная таблица и беспечивают гибкую настройка прав доступа к данным за счет того, что права даются не на таблицу, а на представление. Сервер MySQL довольно быстро работает с представлениями за счет того, что MySQL кэширует результаты запросов, в принципе, многие современные системы управления базами данных. Команды манипуляции данными (за исключением SELECT, который можно применять к любому представлению) можно применять к VIEW в том случае, когда строки VIEW совпадают со строками таблицы в базе данных

Ограничения представлений в MySQL:

1. Невозможно создать индекс на представлении.

2. Индекс может использоваться для представлений, обработанных, используя алгоритм слияния. Однако, представление, которое обработано алгоритмом temptable, неспособно использовать в своих интересах индекс на его основных таблицах (хотя индекс может использоваться во временных таблицах).

В SQL Server представления использовались с самого начала. Кроме основных определяемых пользователем представлений, выполняющих стандартные роли, в SQL Server предусмотрены следующие типы представлений, которые соответствуют специальным назначениям в базе данных:

1. Индексированные представления.
2. Секционированные представления.
3. Системные представления.

В PostgreSQL реализован механизм представлений: при создании представления создается правило, которое определяет, что вместо выполнения операции выборки к представлению система должна выполнять операцию выборки к базовой таблице/таблицам с учётом условий выборки, лежащих в основе определения представления. Для создания представлений, поддерживающих операции обновления, правила для операций вставки, изменения и удаления строк должны быть определены пользователем. В PostgreSQL 9.3 были добавлены материализованные представления. У материализованного представления есть правило, как и у обычного представления, а также другие физические свойства, как у таблицы. Правило используется только для наполнения таблицы, ссылки в запросах указывают на материализованные данные.

В MongoDB представления отсутствуют.

Выводы

Проанализировав пять самых популярных СУБД, можно сделать вывод, что требованиям, предъявляемым к СУБД программного комплекса Приемной комиссии, наиболее полно удовлетворяют две из них - Oracle и PostgreSQL. Однако у СУБД PostgreSQL есть одно существенное преимущество - PostgreSQL Pro сертифицирована ФСТЭК России.

Сертификат удостоверяет, что СУБД Postgres Pro может применяться для защиты информации в государственных информационных системах и автоматизированных системах управления до I класса защищенности, а также обеспечения до I уровня защищенности персональных данных в информационных системах.

Учитывая стойкие интеграционные тенденции ДонНТУ с российскими стандартами, предпочтение стоит отдать именно этой СУБД.

Дальнейшие исследования связаны с возможностями интеграции PostgreSQL с MS SQL Server, т.к. АСУ «Деканат» использует в качестве СУБД MS SQL Server и должен получать данные с сервера Приемной комиссии.

Литература

1. Андрей Николаенко. Эталонные тесты СУБД: что было, что стало, что будет // Открытые системы. СУБД. — 2017. — №2. — С. 35–39. [Электронный ресурс]. – Режим доступа: <https://www.osp.ru/os/2017/02/13052225> — Загл. с экрана.
2. Фейерштейн, С. Oracle PL/SQL для профессионалов / С. Фейерштейн, Б. Прибыл. - М.: СПб: Питер, 2005. - 941 с.
3. MySQL DEVELOPER ZONE «Data Type Overview». [Электронный ресурс]. – Режим доступа: <https://dev.mysql.com/doc/refman/8.0/en/data-type-overview.html> — Загл. с экрана.
4. Панченко И. PostgreSQL: вчера, сегодня, завтра // Открытые системы. СУБД, 2015. № 3. С. 34-37.

Чередникова О.Ю., Щедрин С.В., Ногтев Е.А., Исаков А.Ю. Выбор системы управления базами данных для программного комплекса приемной комиссии ДонНТУ. В статье анализируются пять самых популярных СУБД - Oracle, MySQL, Microsoft SQL Server, PostgreSQL и MongoDB. Сравнение производилось в соответствии с критериями, выдвигаемыми к комплексу приемной комиссии ДонНТУ. Сделан вывод о наиболее предпочтительной СУБД для поставленной задачи.

Ключевые слова: СУБД, таблицы базы данных, SQL, индексы.

Cherednikova Olga, Shchedrin Sergey, Nogtev Evgeniy, Isakov Andrey. The Choice of a database management system for admissions DonNTU. The article analyzes five most popular DBMS - Oracle, MySQL, Microsoft SQL Server, PostgreSQL and MongoDB. The comparison was made in accordance with the criteria put forward to the complex of the admissions Committee of DonNTU. The conclusion about the most preferable DBMS for the task is made.

СЕКЦИЯ 3. «ИНТЕЛЛЕКТУАЛЬНЫЙ АНАЛИЗ ДАННЫХ, РАСПОЗНАВАНИЕ И КЛАССИФИКАЦИЯ»

УДК 004.942

Прогнозирование показателей сферы инвестиций и инноваций с использованием гибридных моделей¹

Аверкин А.Н., Савинова В.М., Ярушев С.Я., Угнячев А.А.
Кафедра Информатики, ИЦЭиИТ, ФГБОУ ВО «РЭУ им. Г.В. Плеханова»
averkin2003@inbox.ru, savinovavm@gmail.com, s2017@yandex.ru

А.Н. Аверкин, В.М. Савинова, С.Я. Ярушев, А.А. Угнячев. Прогнозирование показателей сферы инвестиций и инноваций с использованием гибридных моделей. В настоящей работе рассматриваются построение гибридной модели прогнозирования показателей инвестиций и инноваций с использованием искусственных нейронных сетей (перцептронов) и множественной линейной регрессии. Модели разрабатываются с использованием системы «Горизонт»

Ключевые слова: гибридные модели, нейронные сети, линейная регрессия, верификация, прогнозирование.

Введение

Переход России от плановой к рыночной экономике создал предпосылки для развития систем прогнозирования показателей экономики. Одновременно с этим в стране проходило реформирование всей системы сбора и обработки информации в государственной системе статистической отчетности. На государственном уровне была создана Система Национальных Счетов (далее - СНС), которая позволила использовать накопленный в странах с рыночной экономикой опыт моделирования показателей финансово-экономического развития, включая показатели: макроэкономики, внешнеэкономической деятельности, финансовой системы, труда и занятости, уровня жизни населения, научных исследований и инноваций.

Как правило, экономические явления, определяются большим числом одновременно и совокупно действующих факторов. В связи с этим возникает задача исследования зависимости одной переменной от нескольких объясняющих переменных. Это задача решается с помощью множественного регрессионного анализа.

В настоящей работе с использованием системы поддержки принятия решений «Горизонт», построена гибридная модель краткосрочного прогноза показателей инноваций и инвестиций регионов Российской Федерации. Модели расчетов прогнозных показателей представляются в виде систем регрессионных уравнений и тождеств, а также дополняется интеллектуальными методами, такими, как нейронные сети. В регрессионных уравнениях исследуется поведение взаимозависимых переменных, отражающих поведение прогнозных показателей развития российской экономики в системе национальных счетов в зависимости от сценарных показателей и индикаторов, являющихся экзогенными по отношению к показателям системы моделей. По итогам математических экспериментов для каждого уравнения отбираются статистически значимые аргументы, влияние которых соответствует логике экономических процессов.

Основное преимущество использования системы «Горизонт» — это оперативная возможность изменения сценарных условий и практически незамедлительный расчет варианта прогноза. На любое сценарно-фиксируемое изменение внешней среды можно получать адекватную прогностическую реакцию.

Актуальность. Необходимость использования системы прогнозирования обуславливается

¹ Работа выполнена в рамках проекта РФФИ № 17-07-01558 «Гибридные когнитивные системы поддержки принятия решений»

необходимостью непрерывного мониторинга кризисных ситуаций экономики. Необходим инструмент, который позволит рассчитывать прогнозы различными моделями в одной среде. Такие системы являются неотъемлемой частью современных ситуационных центров, одной из задач которого является непрерывных мониторинг социально-значимых показателей регионов.

Процесс разработки модели

Данная диаграмма описывает информационно-логическую модель работы приложения при выборе модели прогнозирования. Так же разграничиваются действия пользователя-аналитика и непосредственно системы «Горизонт».

Полученные результаты работы могут использоваться в дальнейшем, как отдельный шаг расчета для других моделей, или как источник данных для внешних систем. Так же в полученную модель может быть подан произвольный набор данных для расчета на его основе результирующего набора данных, который может быть сохранен как отдельный или экспортирован во внешние системы.

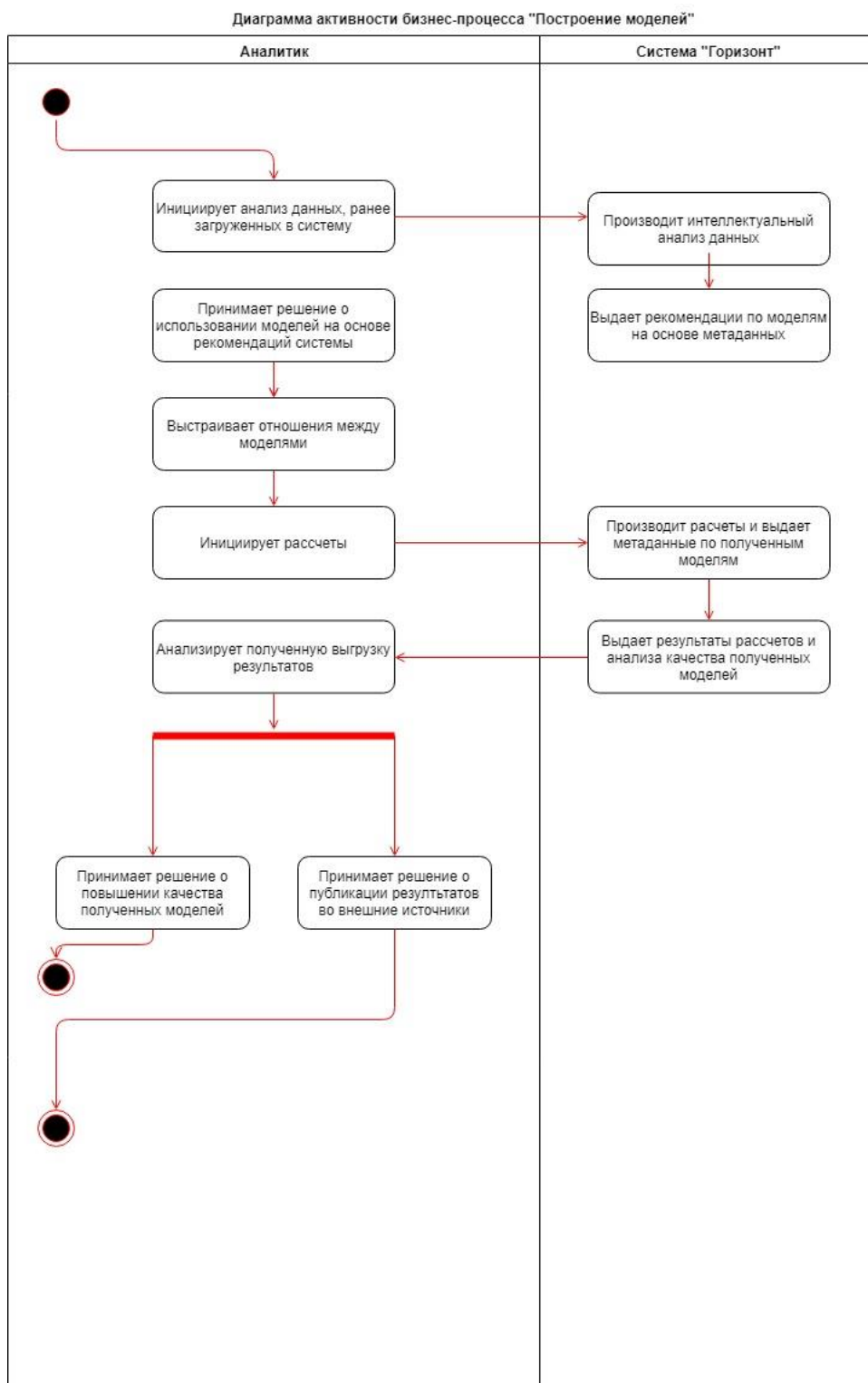


Рисунок 1 – Диаграмма активности «Построение модели»

Полученные результаты расчетов и модели могут быть экспортированы в любые поддерживаемые внешние системы. Описание этого процесса представлено на диаграмме описания бизнес-процесса публикация данных. Как видно из диаграммы, выгрузка может осуществляться во внешнюю систему или в виде отчета. На диаграмме представлен общий процесс. Конкретные шаги могут отличаться в зависимости от выбранной целевой системы выгрузки.

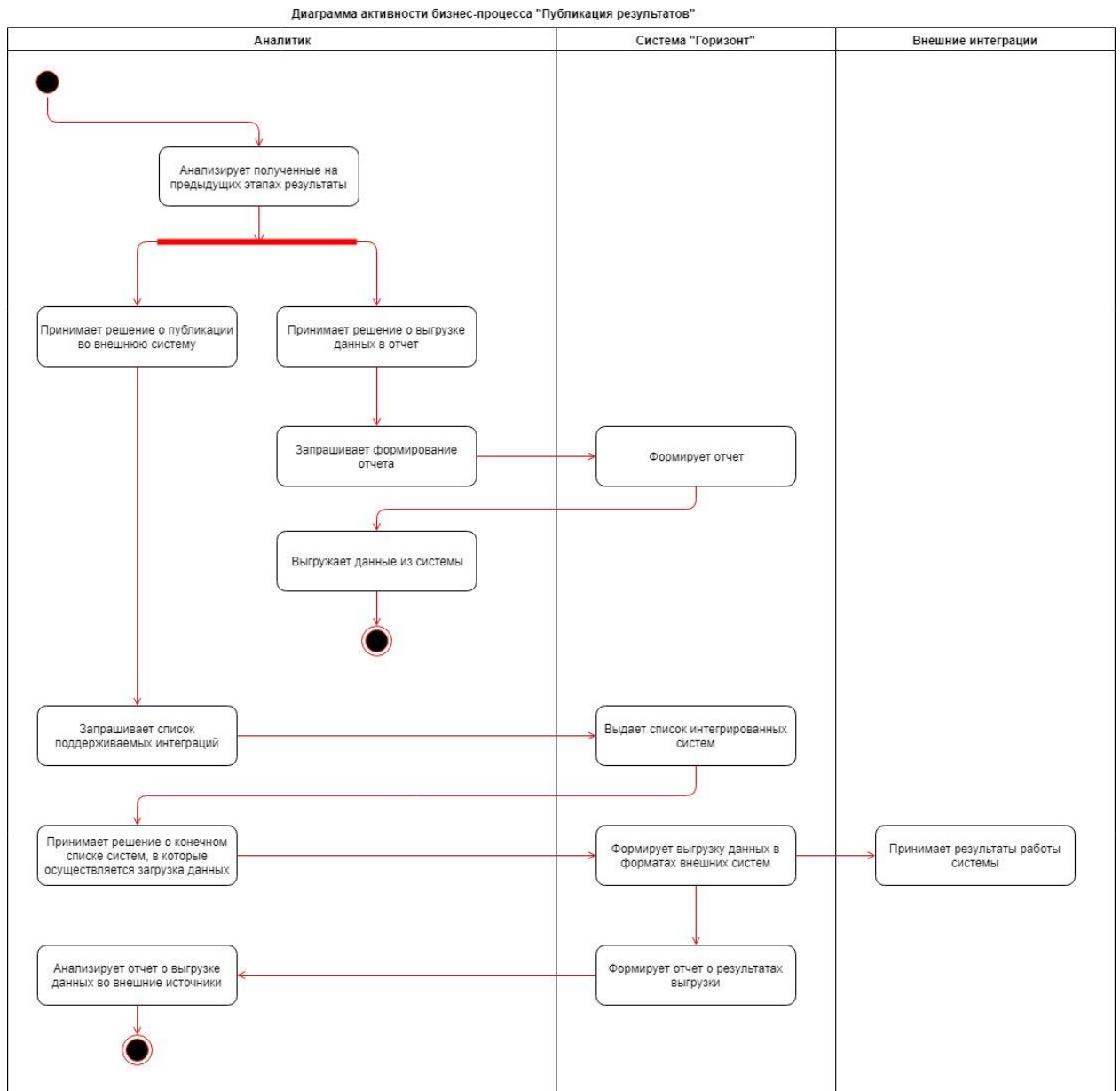


Рисунок 2 - Диаграмма активности «Публикация результатов»

Диаграмма описывает порождаемые процессы и время их жизни на этапе создания моделей на основе ранее загруженных данных.

Диаграмма последовательности взаимодействия для системы "Горизонт" процесс построения моделей

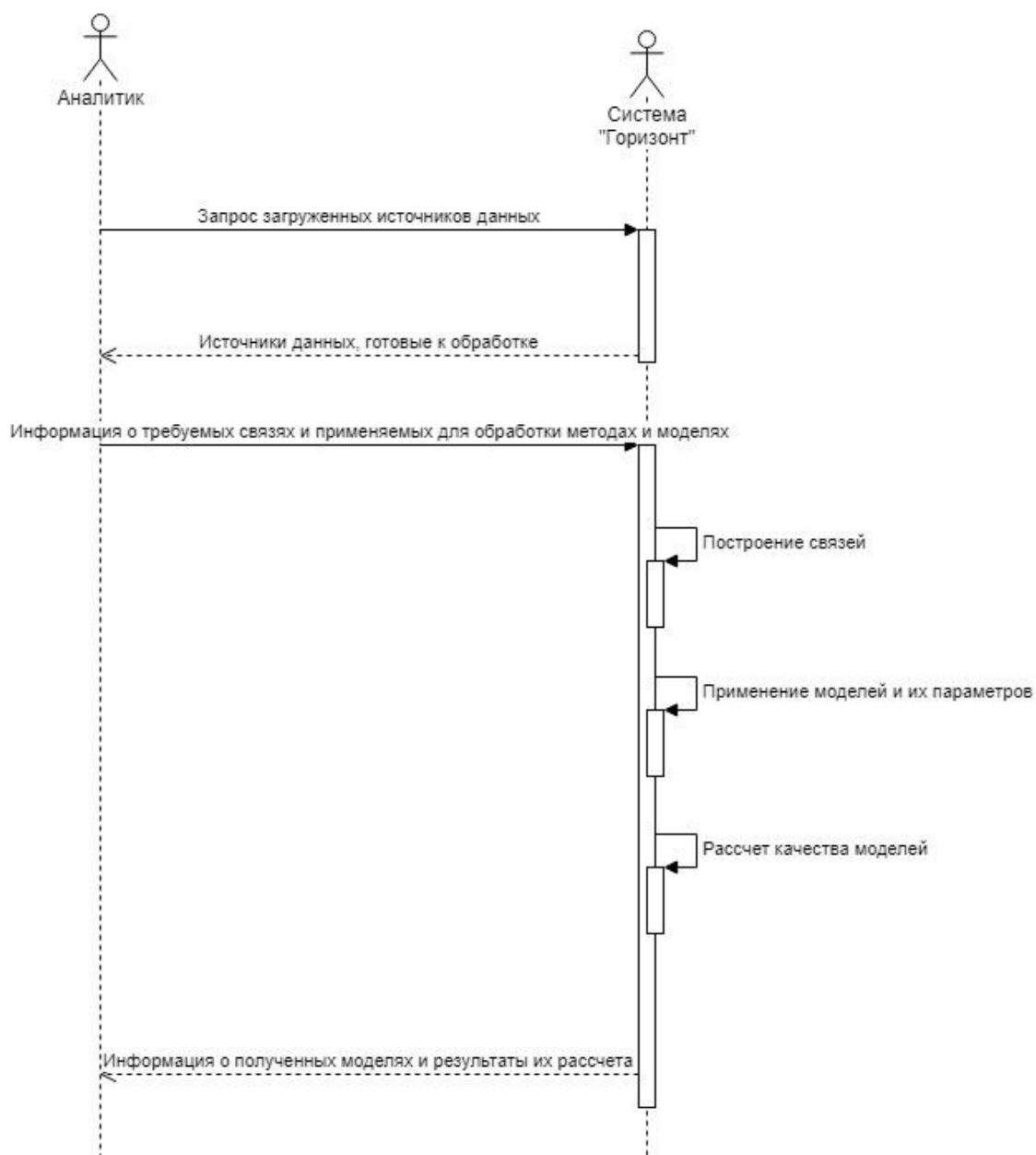


Рисунок 3 - Диаграмма последовательности. Процесс построения моделей.

Результаты моделирования показателей инвестиций привлекательности

На платформе системы «Горизонт» были построены регрессионные уравнения. Также для каждого уравнения были построены графики, иллюстрирующие отклонение расчетных значений от отчетных. Пример графика в системе «Горизонт» представлены на рисунке.



Рисунок 4 - Прогноз инвестиций в основной капитал с использованием регрессионных уравнений

Полученные регрессионные уравнения были исследованы с точки зрения их точности и качества с применением системы «Горизонт». Результат верификации представлен в таблице, которая формируется при экспорте результатов верификации из программы.

Таблица 1 – Результаты верификации регрессионных моделей

		Критерий точности		
		High	Middle	Low
Критерий качества	High	, SEI, RMO, IP1, IP3, CK1, CK2, CK3, CKD1, CKD2, CKG1, CKG2, IMC, IMCP, ITS, OVNIR, CHASP, IPT2(17)	, Ex, GI1, OPA(3)	, Nal, \$/p, Im, TIE, IPS, RIP, NIR, ZIR, PVASP, OPD, RPT, IAORG, ZTI, PVD(14)
	Low	, FS, CE3, CE1, CE2, PBT1, PBT2, INS(7)	, ip(1)	, K\$, Rmbc, DFB%, FT, TI1, TI2, GI2, SE(8)

Как видно из таблицы, прогнозы некоторых показателей не получилось рассчитать с помощью регрессионных уравнений. В связи с этим, необходимо выбрать другую модель прогнозирования. В качестве альтернативной модели выбрана нейронная сеть с архитектурой многослойного персептрона. Метод обучения в данном случае back propagation, описанный выше. Пример результатов нейронной сети для прогнозирования показателей инвестиций представлены на рисунке

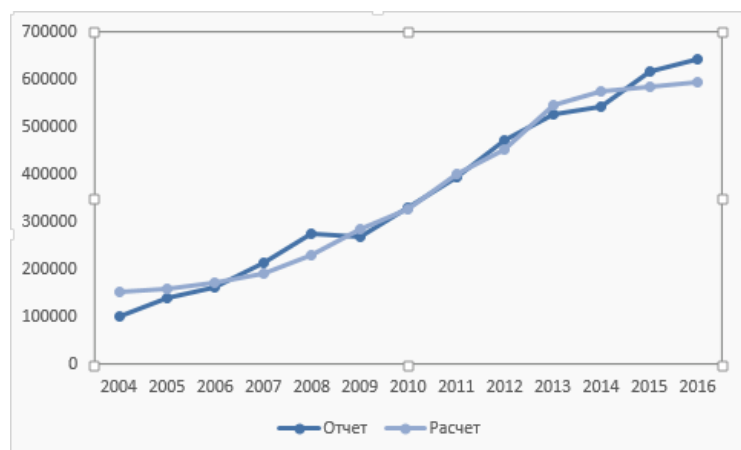


Рисунок 5 - Прогноз с использованием нейронных сетей

С использованием нейронной сети получилось улучшить качество моделей прогнозирования, полученных с помощью регрессионных моделей. Результаты верификации нейронной сети для республики Татарстан представлены в табл.

	R ²	MAPE %
IPS	0,91	16
RIP	0,95	13
NIR	0,94	13
PVASP	0,91	14
ZTI	0,94	12

Результаты верификации нейронной сети для региона Москва представлены в табл.

	R ²	MAPE %
IPS	0,99	12
RIP	0,95	14
NIR	0,96	12
PVASP	0,92	14
ZTI	0,94	12

Таим образом, использование нейронных сетей позволило улучшить качество моделей. Система поддержки принятия решений «Горизонт» является полноценным аналитическим инструментом, который позволил осуществить построение гибридной модели прогнозирования.

Литература

1. Аверкин А.Н., Прокопчина С.В. Мягкие вычисления и измерения // Интеллектуальные системы (МГУ). 1997. Т. 2. Вып. 1 – 4. С. 93 – 114
2. Sergey A. Yarushev, Alexey N. Averkin, Time Series Analysis Based on Modular Architectures of Neural Networks, Procedia Computer Science, Volume 123, 2018, Pages 562-567
3. Averkin A. N., Yarushev S. Hybrid approach for time series forecasting based on ANFIS and Fuzzy Cognitive Maps //Soft Computing and Measurements (SCM), 2017 XX IEEE International Conference on. – IEEE, 2017. – Pp. 379-381.
4. Демидова Л.А., Пылькин А.Н., Скворцов С.В., Скворцова Т.С. Гибридные модели прогнозирования коротких временных рядов. – М: Горячая линия – Телеом, 2015. – 206 с.

5. Danko T.P., Kitova O.V., Kolmakov I.B., Dyakonova L.P., Grishina O.A., Sekerin V.D. Hybrid intelligent system of forecasting of the socio-economic development of the country/ *ijaber*, Vol. 14, No.9. – 2016. - P. 5755-5766.

А.Н. Аверкин, В.М. Савинова, С.Я. Ярушев, А.А. Угнячев. Прогнозирование показателей сферы инвестиций и инноваций с использованием гибридных моделей. В настоящей работе рассматриваются построение гибридной модели прогнозирования показателей инвестиций и инноваций с использованием искусственных нейронных сетей (перцептронов) и множественной линейной регрессии. Модели разрабатываются с использованием системы «Горизонт»

Ключевые слова: гибридные модели, нейронные сети, линейная регрессия, верификация, прогнозирование.

Разработка алгоритмов и программ параметризации речевого сигнала на основе вейвлет-анализа

Жуков А.И.

Санкт-Петербургский национальный исследовательский университет информационных технологий, механики и оптики

alexzhukovwork@yandex.ru

О.И. Федяев

Донецкий национальный технический университет

fedyaev@donntu.org

Жуков А.И., Федяев О.И. Разработка алгоритмов и программ параметризации речевого сигнала на основе вейвлет-анализа. Статья посвящена проблеме автоматического распознавания речи. Речевой сигнал представлен вейвлет-спектрограммой. Распознавание вейвлет-образов речевого сигнала осуществляется нейросетевым классификатором.

Ключевые слова: автоматическое распознавание речи, анализ сигнала, вейвлет-анализ, java, нейронная сеть, python

Введение

В настоящее время устная речь всё чаще используется в человеко-машинном взаимодействии. Например, на современных смартфонах есть голосовые помощники, которые имеют возможность выполнения множества команд с помощью речевого интерфейса, например, если смартфону сказать: «Я хочу кушать», то он предложит карту с ресторанами около владельца смартфона. Также голосовые помощники оказывают различную помощь в повседневной жизни. Пользователь может голосом произносить различные поисковые запросы в Интернете: оставлять заметки, ставить будильники, звонить контактам из телефонной книги, узнавать погоду (см. рис. 1), писать сообщения контакту из телефонной книги или по указанному номеру и т.д.

Однако, применяемые в настоящее время технологии распознавания речи всё ещё обладают рядом недостатков, связанных с низкой надёжностью, скоростью распознавания, чувствительностью к помехам. Эти ограничения не позволяют использовать системы с речевыми интерфейсами во многих практических ситуациях, когда необходимо принятие ответственных решений. Это обстоятельство по-прежнему оставляет актуальными задачи разработки и исследования принципиально новых методов оценки информативных признаков речевого сигнала [1]. Поэтому в данной работе поставлена задача - разработать авторские алгоритмы и программное обеспечение, улучшающее надёжность систем распознавания речи на основе нового направления в теории обработки сигналов – вейвлет-анализа [2].



Рисунок 1 – Просмотр погоды с помощью голосового помощника «Siri»

Обзор литературы показывает, что вейвлет-анализ в отличие от Фурье-анализа является лучшим средством исследования речевых сигналов, которые характеризуются не периодичностью, не стационарностью и наличием локальных особенностей [3].

Задача автоматического распознавания речи относится к классу сложных задач классификации образов, представленных акустическими характеристиками речевого сигнала. В ней выделяются две основные подзадачи:

- предварительная обработка речевых сигналов для выделения информативных признаков, характеризующих речевые сигналы;
- классификация речевых сигналов путем сравнения с акустическими эталонами словаря.

Дискретное описание речевого сигнала

Для решения задачи автоматического распознавания речи требуется анализировать устную речь, которая вводится в компьютер с микрофона в виде амплитудно-временного звукового сигнала (рис. 2).

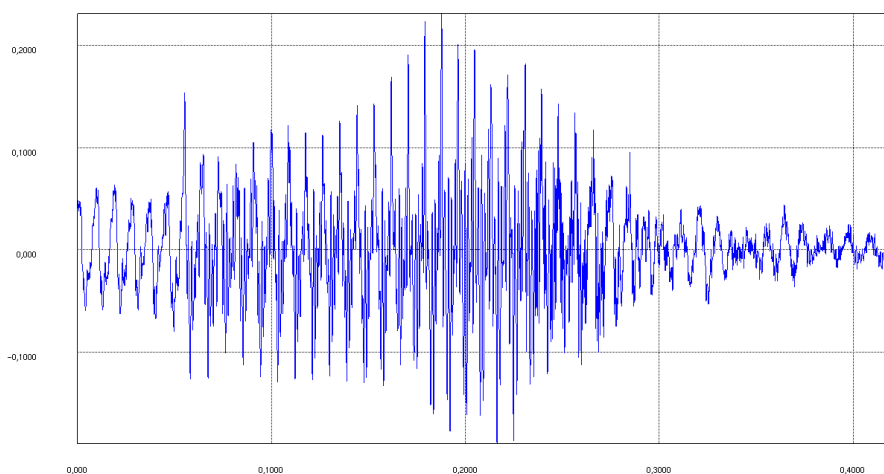


Рисунок 2 - Представление речевого сигнала для слова «Ноль»

Оцифровка речевых сигналов осуществляется в реальном времени звуковой картой, реализующей аналого-цифровое преобразование сигнала. Дискретная форма звукового сигнала представляется файлами формата WAV, которые имеют стандартную структуру (рис. 3). В работе параметры оцифровки сигнала (частота дискретизации и размер квантования сигнала) имели типовые значения, соответствующие типовым параметрам в системах передачи, хранения и обработки речевой информации.

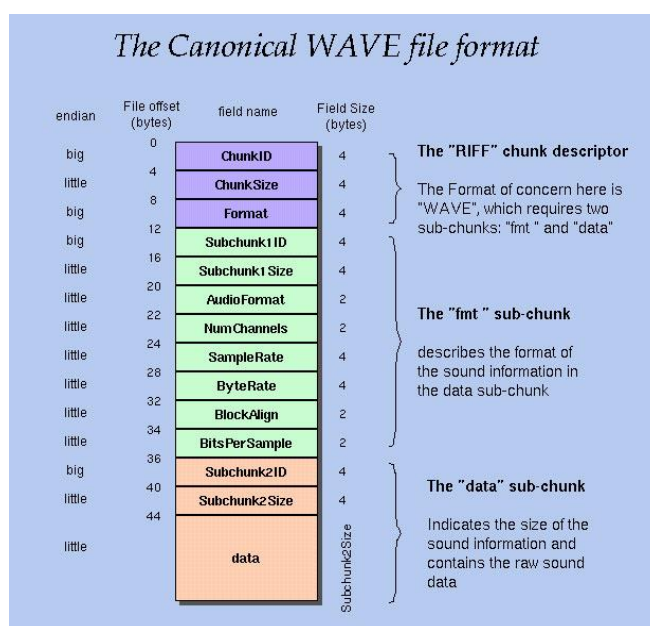


Рисунок 3 – Структура wav файла

Для распознавания речи используется нижняя часть файла под названием «Data», представляющая собой массив чисел, элементы которого являются числами типа Double. Последовательность чисел характеризует амплитуды сигнала в определенные моменты времени в зависимости от его частоты дискретизации, которая отвечает за то, сколько цифр будет приходиться на секунду сигнала. Получив эти данные, в принципе, можно приступить к следующему шагу в распознавании речи. Но распознавать речь, используя вышеприведенные данные напрямую не разумно, т.к. это гарантирует неустойчивый результат, потому что существует множество факторов, которые существенно влияют на представление сигнала в амплитудно-временном виде. К таким факторам относятся:

- тембр голоса;
- скорость произношения;
- громкость произношения;
- различные шумы на заднем плане.

При данном подходе все вышеприведенные факторы должны иметь идентичное значение в эталоне, т.е. значения поля «Data» должны полностью совпадать со значениями этого же поля эталонного сигнала, иначе распознавание даст неверный результат. Кроме того, такой подход будет иметь большую алгоритмическую сложность при большом словаре, т.к. потребуется сравнить введенный сигнал с каждым, находящимся в базе, до совпадения. При таком подходе трудно ожидать качественную классификацию образов речевых сигналов.

Поэтому для улучшения классификации сигнала, как правило, требуется предварительная его обработка с целью выделения самых существенных особенностей сигнала.

Получение вейвлет-спектрограмм речевого сигнала

Предварительной обработкой сигнала называется процесс преобразования и оценивания его информативных характеристик какой-либо технической системой. Существует несколько подходов к оцениванию этих характеристик, например, сигнал можно представить в виде суммы простых колебания (синусоид) и охарактеризовать его набором чисел, называемым спектром, определяющим долю каждого колебания в сигнале [2]. Для решения данной задачи был выбран вейвлет-анализ [1-4].

Само вейвлет-преобразование является сверткой материнской функции вейвлета и функции самого сигнала. Материнская функция – это одна из базисных функций вейвлета, ей может стать любая функция, удовлетворяющая следующим условиям:

- она должна быть локализована как по времени, так и по частоте;
- нулевой и последующие моменты равны нулю;
- функция должна быть угасающей.

В общем виде данная функция может быть записана формулой

$$\psi_{a,b}(t) = \frac{1}{\sqrt{a}} \psi\left(\frac{t-b}{a}\right) \quad (1)$$

Эта функция зависит не только от времени t , но и от двух других параметров a и b . Параметр a представляет из себя масштаб, иначе говоря, он растягивает и сжимает исходный материнский вейвлет. Чем меньше масштаб, тем детальнее можно рассмотреть более мелкие участки сигнала и наоборот. Данный подход можно сравнить с географическими картами, т.к. он позволяет проанализировать сигнал в разных масштабах [3].

Параметр b представляет из себя перемещение материнского вейвлета по сигналу. Его идея состоит в том, что материнский вейвлет должен пройти по всем участкам анализируемого сигнала. Ниже приведена зависимость функции вейвлета от параметра сдвига « b » и параметра масштабирования « a » (рис. 4).

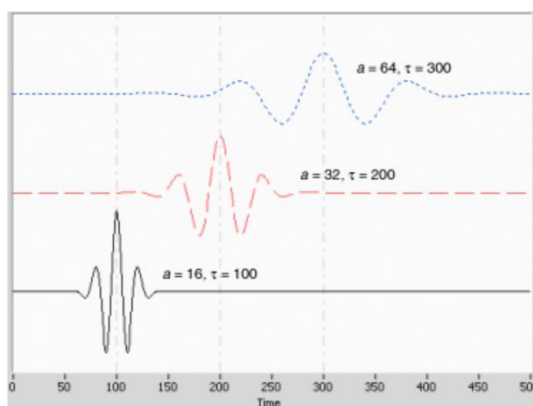


Рисунок 4 – Зависимость функции вейвлета от параметров

В роли материнского вейвлета выбран действительный вейвлет Морле, т.к. он имеет хорошую локализацию во времени [2]. Данный вейвлет представлен на рисунке 5 и в формуле 2.

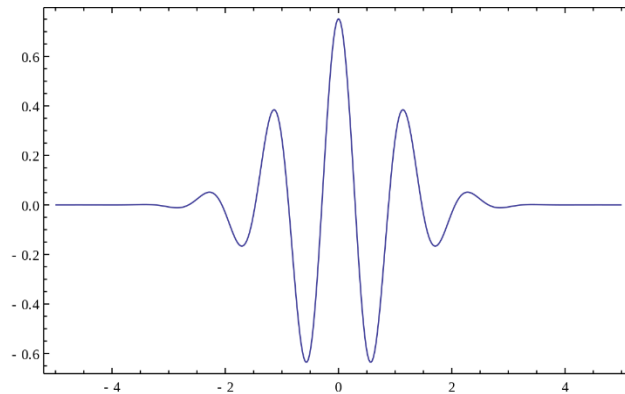


Рисунок 5 – Действительный вейвлет Морле

$$\psi(t) = -\frac{t^2}{2} * \cos(5 * t) \quad (2)$$

Примеры изображений вейвлет-спектрограмм, полученных с помощью разработанной программы, представлены на рисунках 6 и 7.

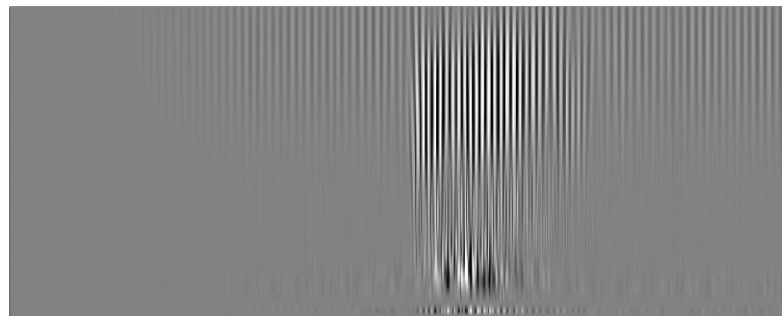


Рисунок 6 – Спектр вейвлет образа для слова «Семь»

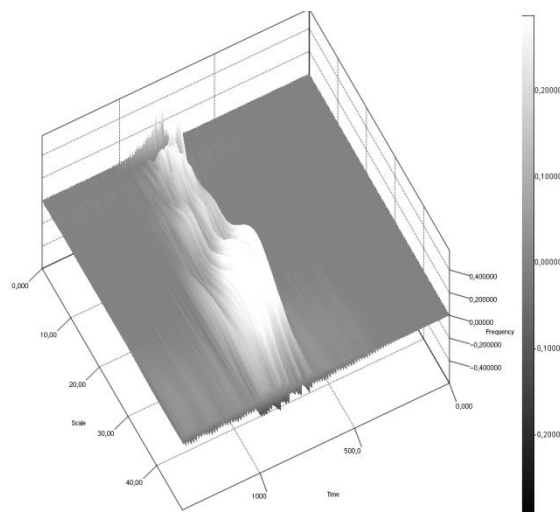


Рисунок 7 - 3D изображение вейвлет-образа для слова «Семь»

Нейросетевой распознаватель вейвлет-образов речевых сигналов

Полученные вейвлет-образы передавались на классификатор, в роли которого выступала свёрточная нейронная сеть. На входе многослойная нейросеть получала вейвлет-образ слова в формате изображения, а на выходе возвращала индекс слова из словаря. Структура сверточной нейронной сети изображена на рисунке 8.

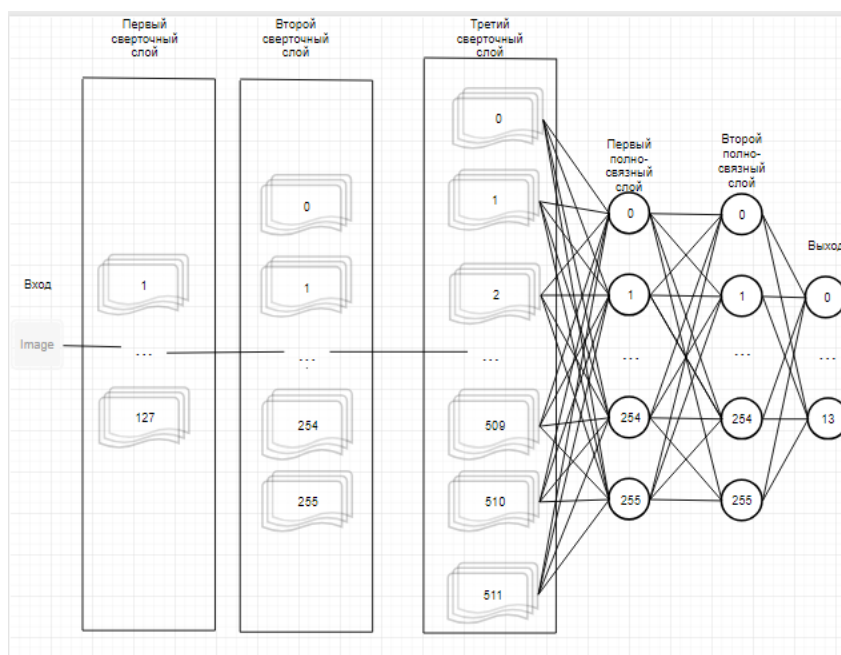


Рисунок 8 – Структура сверточной нейронной сети

Алгоритмы и программы были апробированы на распознавание набора слов, составленного из чисел от «ноль» до «девять» и основных арифметических операций «+, -, *, /» для системы «Калькулятор с речевым интерфейсом» (рис. 9). Для обучения нейронной сети использовались записанные образы одного диктора, различными тембром и с различной интонацией. Модель была протестирована на выборках различного размера. По этим данным построен график точности распознавания в зависимости от количества образов на класс изображения (см. рис. 10).

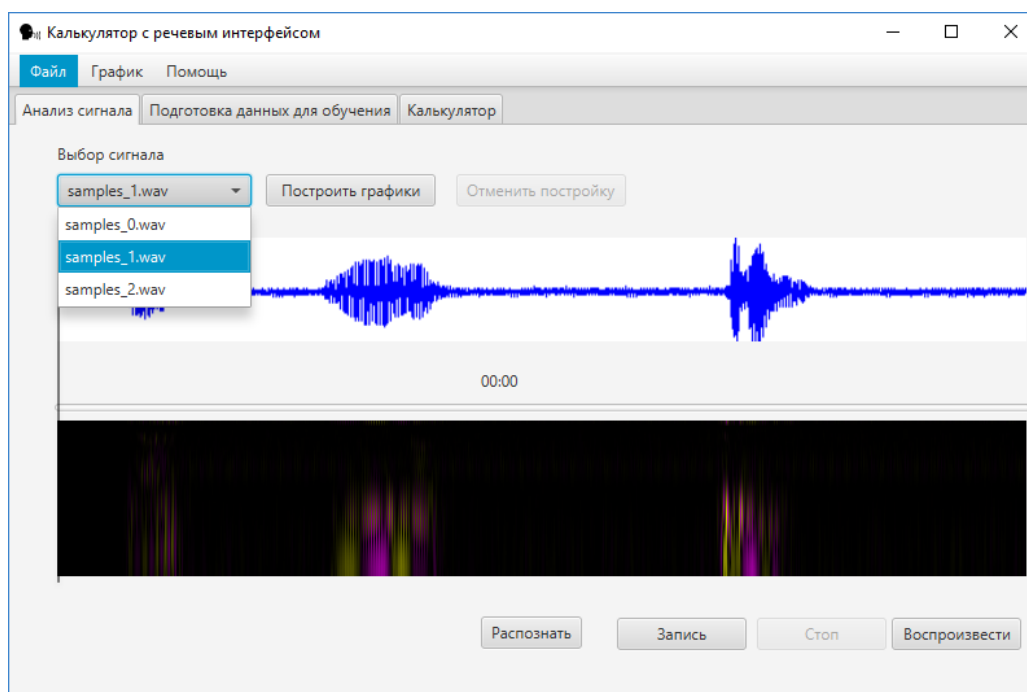


Рисунок 9 – Главное окно системы «Калькулятор с речевым интерфейсом» (режим анализа сигнала)

Вывод

В работе предложена реализация метода распознавания речи с использованием вейвлет-анализа, как средства многомасштабного представления речевого сигнала с хорошими показателями частотно-временной локализации. Исследования на разработанной программной системе «Калькулятор с речевым интерфейсом»

показали, что непрерывное вейвлет-преобразование может быть использовано как средство составления образа сигнала для дальнейшей классификации. Построен удобный интерфейс системы, предполагающий использование её в двух режимах: калькулятор (кнопка «Распознать»), подготовка и анализ графических изображений сигналов. Для ввода новой записи речевого сигнала надо нажать кнопку «Запись», произнести голосом слово и выключить запись кнопкой «Стоп». Можно просмотреть любой исходный сигнал и его вейвлет-спектрограмму из выпадающего списка кнопкой «Воспроизвести» (рис.9). В главном меню предусмотрены и другие необходимые функции для анализа сигналов и обучения нейросети.

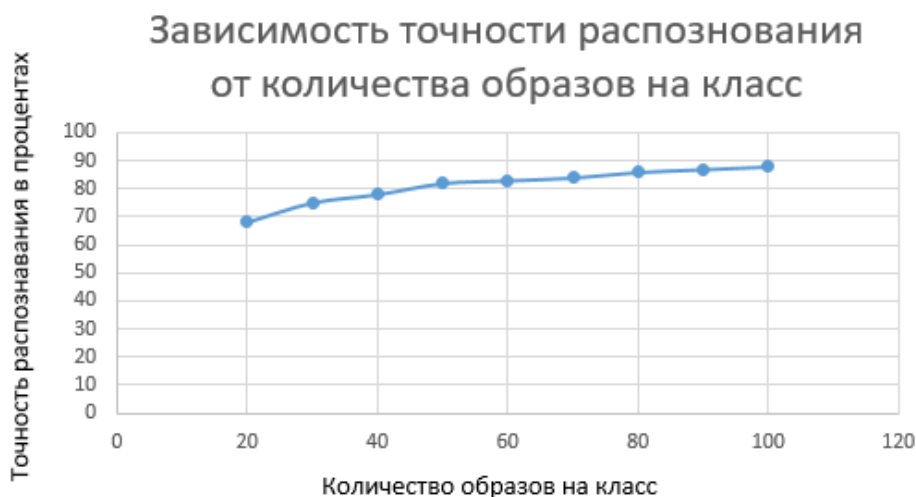


Рисунок 10 - Зависимость точности распознавания от количества образов на класс

Экспериментальное тестирование алгоритмов и программ распознавания речи показало сравнительно невысокий процент автоматического распознавания (не более 90% для раздельного произношения), что предполагает необходимость дальнейших исследований.

Литература

1. Бойков Ф.Г. Применение вейвлет-анализа в задачах автоматического распознавания речи / Автореферат канд. дисс. по специальности 05.13.18, Москва, МФТИ. - 111 с.
2. Новиков, Л. В. Основы вейвлет-анализа сигналов / Л. В. Новиков. – Санкт-Петербург; Питер, 1999. – 153 с.
3. Дьяконов В.П. Вейвлеты. От теории к практике. Изд. 2-е, перераб. и доп. – М.: СОЛОН-Пресс, 2004. - 400 с.
4. Bussow, R. An Algorithm for the Continuous Morlet Wavelet Transform / R. Bussow. – Berlin: Berlin University of Technology., 2008. – 15 p.

Жуков А.И., Федяев О.И. Разработка алгоритмов и программ параметризации речевого сигнала на основе вейвлет-анализа. Статья посвящена проблеме автоматического распознавания речи. Речевой сигнал представлен вейвлет-спектрограммой. Распознавание вейвлет-образов речевого сигнала осуществляется нейросетевым классификатором.

Ключевые слова: автоматическое распознавание речи, анализ сигнала, вейвлет-анализ, java, нейронная сеть, python

Zhukov A. I., Fedyayev O. I. Development of algorithms and programs of parametrization of the speech signal on the basis of wavelet analysis. The article is devoted to the problem of automatic speech recognition. The speech signal is represented by a wavelet spectrogram. The recognition of the wavelet images of the speech signal is carried out by the neural network classifier.

Key words: automatic speech recognition, signal analysis, wavelet analysis, java, neural network, python

Разработка программной системы распознавания лиц на основе нейросети типа неокогнитрон

Тырса С.В.

Санкт-Петербургский национальный исследовательский университет информационных технологий, механики и оптики

stas.tyrsa@gmail.com

О.И. Федяев

Донецкий национальный технический университет

fedyaev@donntu.org

Тырса С.В., Федяев О.И. Разработка программной системы распознавания лиц на основе нейросети типа неокогнитрон. В данной статье рассмотрены основные методологии распознавания лиц на фотографии. Изучены современные средства распознавания. Предложен метод распознавания лиц на изображении с помощью многослойной нейронной сети типа неокогнитрон.

Ключевые слова: нейронная сеть, неокогнитрон, машинное обучение, обучение без учителя, распознавание лиц, визуальный образ

Введение

Компьютерноераспознавание лиц – это практическое приложение теории распознавания образов, в задачу которого входит автоматическая локализация лица на фотографии и, в случае необходимости, идентификация персоны по лицу. Задача идентификации и распознавания лиц – это одна из первых важных практических задач, которая стимулировала становление и развитие теории распознавания и идентификации объектов.

В настоящее время предложено большое количество нейросетевых парадигм для решения этой задачи. Проблема распознавания лиц рассматривалась еще на ранних стадиях компьютерного зрения. Ряд компаний на протяжении более 40 лет активно разрабатывают автоматизированные, а сейчас и автоматические системы распознавания человеческих лиц: Smith & Wesson (система ASID – Automated Suspect Identification System); ImageWare (система FaceID); Imagis, Epic Solutions, Spillman, Miros (система Trueface); Vissage Technology (система Vissage Gallery); Visionics (система FaceIt) [1].

Значительные затруднения при распознавании вызывают образы, подверженные какому-либо искажению (зашумление, смещение, поворот, изменение образа в размерах). Эту проблему решают путём выбора соответствующей архитектуры и способа обучения. Ранее полученные частные решения этой проблемы не привели к созданию эталонного подхода к распознаванию. В настоящее время большие перспективы связывают с многослойными (глубокими) нейронными сетями. Поэтому данная работа посвящена построению и исследованию модели нейронной сети неокогнитрон, которая мало изучена на задачах распознавания лиц.

Объектами исследования в данной работе являются алгоритмы и программные средства моделирования многослойной нейронной сети типа неокогнитрон на языке программирования Java. Основная цель работы заключается в разработке программного эмулятора неокогнитрона с возможностью его обучения на распознавание определенного класса образов – лиц людей. Такой эмулятор позволит путём моделирования более детально изучить принципы самоорганизации многосвязной мультислойной структуры неокогнитрона.

Описываемая программная система может быть использована в различных практических задачах, в том числе, при идентификации лиц сотрудников на предприятии для обеспечения высокого уровня защиты.

Современные подходы к распознаванию и постановка задачи распознавания лиц

При постановке и решении задачи распознавания некоторые авторы стараются пользоваться математическим языком, стремясь, в отличие от теории искусственных нейронных сетей [2], где основой является получение результата путём обобщения примеров, заменить эксперимент логическими рассуждениями и математическими доказательствами [3]. К этому классу методов относятся статистические и структурные методы распознавания, которые основаны на идеях теории вероятностей и поэтому трудоемки в практической реализации [8]. Нейросетевые модели распознавания благодаря обучению легко выделяют характерные признаки распознаваемого образа, что обеспечивает качественное распознавание.

Проанализируем последние (наиболее интересные на наш взгляд) достижения в области систем

распознавания. Самый распространенный способ основан на нейросетевой парадигме [4]. Искусственная нейронная сеть – компьютерная система, вдохновлённая биологическими нейронными сетями, которые существуют в живом организме. Нейронная сеть состоит из отдельных нейронов, соединенных друг с другом. Каждое соединение между нейронами в многослойной сети называется синапсом, через который возбуждение в виде сигнала передаётся от одного нейрона к другому. термин «обучение нейросети» означает понимание входные образы.

В настоящее время разработкой нейронных сетей для распознавания различных образов занимаются значительное количество компаний, в том числе Google, которая использует нейронную сеть в онлайн переводчике Google Translate. Данная нейронная сеть получила название Google Neural Machine Translation (см. рис. 1) [5]. В основе этой нейронной сети лежит двунаправленная рекуррентная нейросеть, которая позволяет не только распознавать и переводить единичные слова, но и понимать контекст предложения, что даёт возможность перевода предложения целиком. Данная нейросеть состоит из анализирующего и синтезирующего потоков. Анализирующий поток нейросети отвечает за анализ и разбиение предложения на смысловые элементы, а синтезирующий поток – за наиболее точный вариант перевода предложений. В результате данная система анализирует с разных сторон входное текстовое сообщение с целью понимания смысла.

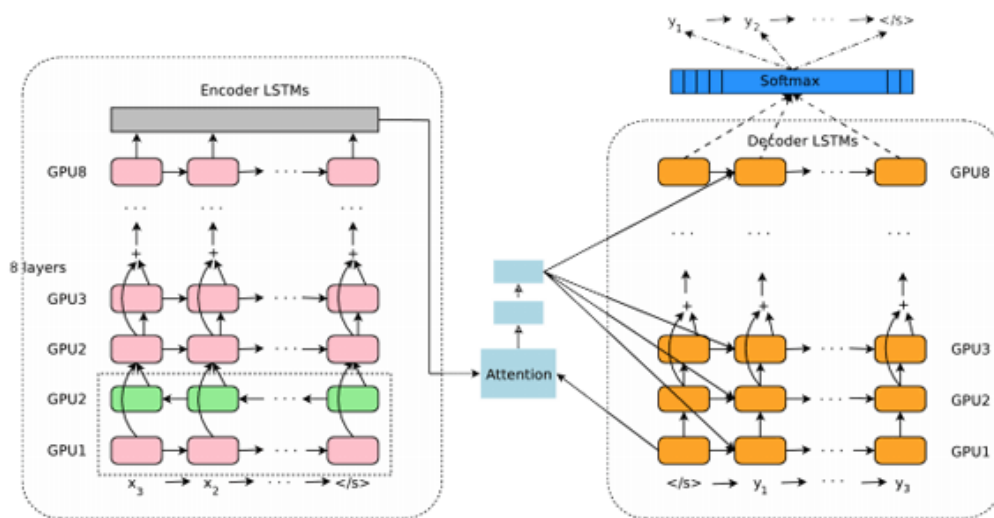


Рисунок 1 – Схема работы нейросети Google

Помимо нейросетевых подходов, существуют подходы не использующие нейронные сети. Одним из таких методов является метод главных компонент (PCA) [6]. Данный метод не искажает образ при сжатии. Берётся входной вектор X размерности N и преобразуется с помощью линейного ортогонального преобразования в вектор Y размерности M , $N > M$. При этом компоненты вектора Y являются некоррелированными, и дисперсия не изменяется. Индикаторами принадлежности выступают показатели DIFS и DFFS. DIFS показывает расстояние от образа анализируемого изображения до собственного образа. DFFS показывает расстояние от представления до проекции эталона в собственном пространстве.

В области распознавании лиц данный метод называется методом собственных лиц (см. рис.2). При распознавании матрица изображения раскладывается в набор линейных компонентов. Для каждого лица выделяются основные компоненты с применением Евклидова расстояния. Недосток данного метода заключается в трудоемкости и низком быстродействии.



Рисунок 2 – Пример изображений собственных векторов

К популярным методам можно также отнести подход на основе дискриминантного анализа и метод сравнения лиц, представленных в виде графов, вершины которого соответствуют ключевым точкам на лице человека.

Рассмотрим постановку задачи распознавания образов, которая заключается в распознавании заданного

множества объектов (монохромных плоских изображений лиц человека) $P = \{p_1, p_2, \dots, p_n\}$. Каждый объект p_i представляется некоторым графическим образом $g(x, y, f(x, y))$, где $f(x, y)$ - яркость пикселя в точке (x, y) плоского изображения лица. Для каждого объекта p_i задано множество примеров (эталонов) изображений $G_i \in G$ в виде графических образов $g(x, y, f(x, y)) \in G_i$, по которым система будет обучаться распознавать объект p_i . Множество G , состоящее из подмножеств примеров G_i , образует обучающее множество.

Задача заключается в построении такой многомерной функции $F(\cdot)$ путём обучения на обучающем множестве G , которая по графическому изображению $g(x, y, f(x, y))$ объекта p_i правильно его распознаёт

$$p_i = F(g(x, y, f(x, y))).$$

В представленной работе данная функция определяется с помощью многослойной нейронной сети, которая зарекомендовала себя как хороший аппроксиматор многомерных функций.

Разработка программной системы распознавания лиц

Неокогнитрон – это иерархическая многослойная искусственная нейронная сеть, предложенная Кунихико Фукушимой в 80-х годах [6]. С самого начала она использовалась для распознавания рукописных символов, а потом и других графических образов. Модель неокогнитрона была основана на другой модели, предложенной Дэвидом Хьюбелем и Торстеном Вайселем ещё раньше в 1959 году [7]. Нейронная сеть содержала 2 типа ячеек, называемых S-ячейками (simple – простой) и C-ячейками (complex – сложный).

Нейросеть типа Неокогнитрон реализует каскадную модель взаимодействия. Сеть состоит из множества типов ячеек, большинство из которых являются S-ячейками или C-ячейками. S-ячейки отвечают за распознавание общих черт, в то время как C-ячейки служат для определения сдвигов и деформации объекта. Общие черты распознаются постепенно и классифицируются в поздних слоях. Данная идея была обнаружена в некоторых других моделях, а именно в модели LaNet и модели SIFT.

Существует множество вариантов неокогнитрона. К примеру, некоторые виды неокогнитрона могут принимать на вход множество различных шаблонов с использованием специальных обратных сигналов для достижения избирательного внимания (свойство человека обращать внимание на одни черты и игнорировать другие).

Неокогнитрон является самоорганизующейся нейронной сетью, которая направлена на распознавание визуальных образов. Для выполнения распознавания символов, методологии должны быть способны обрабатывать изменения в сигнале. Как правило, рукописные письменные знаки имеют вариации не только для каждого пользователя, но также сильно различаются по внешнему виду между различными лицами. Эта проблема обуславливает необходимость надёжной системы распознавания символов, которая может справиться с этими изменениями. Неокогнитрон представляет собой платформу для распознавания символов, которая устойчива к изменениям вида символа и пространственного расположения. Благодаря способности к самоорганизации Неокогнитрон настраивает свою распознающую функциональность во время тренировки. С помощью иерархической системы, в которой характерные признаки извлекаются на первых слоях, более поздние слои могут обеспечить распознавание символов. Такая система может быть обучена распознавать различные типы символов.

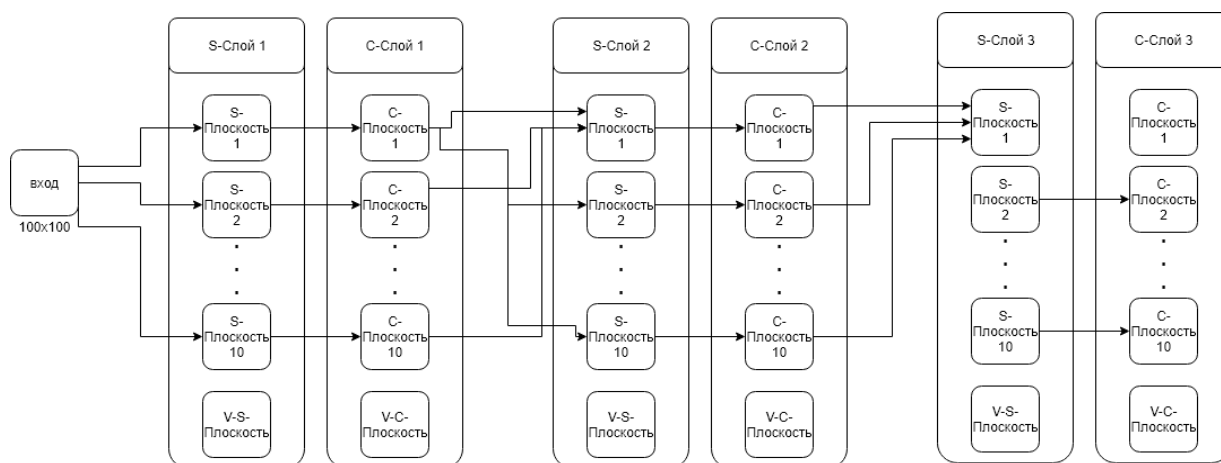


Рисунок 3 – Схема нейронной сети типа неокогнитрон со слоями и плоскостями

Входным сигналом системы распознавания лиц было квадратное изображение низкого разрешения, в нашем случае это будет 100 на 100 пикселей, с исходным визуальным образом лица человека. Сложная структура неокогнитрона позволяет извлекать черты лица на изображении, независимо от позиционных сдвигов

(см. рис.3). Каждый слой сети состоит из двух подслоев, например, слой 1 состоит из S-слоя и C-слоя. Внутри каждого слоя, есть множество S-плоскостей или C-плоскостей.

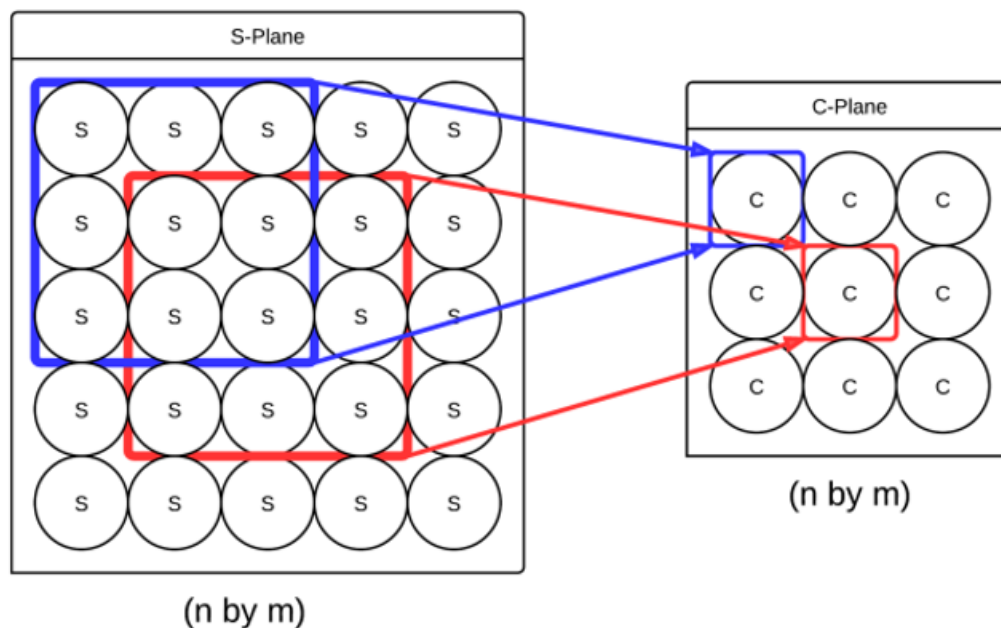


Рисунок 4 – Организация каждой плоскости

Каждая плоскость ячеек внутри каждого слоя представляет собой матрицу нейронов. Тип нейрона зависит от типа слоя, в котором он находится. Размер плоскости, как показано на рисунке 4, обычно квадрат, где n равен m . Каждая плоскость в заданном S-слое или C-слое имеет такой же размер. В многослойной архитектуре каждая C-плоскость связана только с предшествующей S-плоскостью, а каждая S-плоскость связана с каждой предшествующей C-плоскостью. Внутри каждого слоя существует также одна V-плоскость, независимо от того, сколько существует S-плоскостей и C-плоскостей. Цель каждого V-нейрона - обеспечить тормозящую реакцию на S-нейрон или C-нейрон в каждой плоскости. В связи с тем, что существует только одна V-плоскость в слое, каждая область в каждой S-плоскости и C-плоскости получает одинаковый тормозящий вход. Связи между ячейками показаны на рисунке 5.

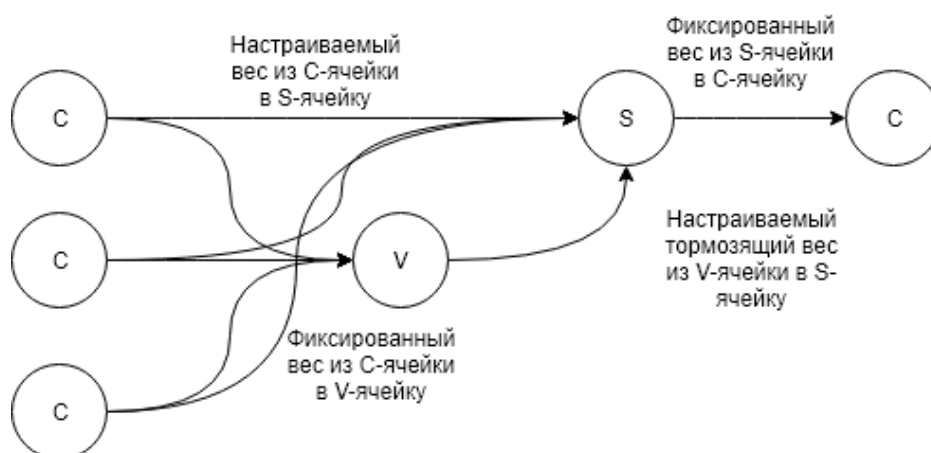


Рисунок 5 – Соединение нейронов с использованием тормозящего нейрона

Обучение неокогнитрона выполнялось по стратегии «Обучение без учителя». Для этого была разработана подсистема обучения в виде набора объектно-ориентированных классов, которая обучала нейроны сети, выполняла чтение, запись и перезапись информации о нейроне в текстовый файл, а также вычисление коэффициента ошибок. Схема алгоритма обучения нейрона представлена на рисунке 6а.

Алгоритм обучения нейронной сети

```

входы: Изображение [20]
выход: Нейрон = null
счетчик = 0
ошибка = 1
лучшая ошибка = 1
Пока (ошибка! = 0)
{
    Пока (ВерифицироватьНейрон(выход) == ложь){
        счетчик = счетчик + 1
        выход = Нейрон (Параметры_Нейрона)
        n = 0
        Пока( n < счетчик * Количество(входы){
            выход.распространить(входы[n])
            n = n + 1
        }
    }
    ошибка = РассчитатьОшибку(выход, входы)
    Если(ошибка < лучшая ошибка){
        СохранитьНейрон()
        ошибка = лучшая ошибка
    }
}

```

а)

Метод распространения нейрона

```

вход: double[][]

Если (не(квадратная_матрица(вход))){
    'ошибка входов'
}
выход = Установить выходы()
выход.УстановитьМатрицуВыходныхПлоскостей(вход)

l = 0
Пока( l < Число_S_слоев){
    выход = S-Слой.распространить(выход)
    выход = C-Слой.распространить(выход)
    l = l + 1
}

//Определить наилучший выход на последнем слое
максзнач = 0
i = 0
Пока(i < длина(последнийСлой)){
    если(последнийСлой[i] > максзнач)
        максзнач = i
}

```

б)

Рисунок 6 – Алгоритмы: а) - обучения нейрона; б) – распространения нейрона

В алгоритме на рисунке 6а присутствует метод распространения нейрона. Подробный алгоритм распространения представлен на рисунке 6б. Как видно из этого рисунка, нейрон распространяется по всем S-слоям и C-слоям с учётом базовых настроек нейрона. Далее распространяемый сигнал проходит по всем нейронам всех плоскостей (см. рис. 3) с учётом коэффициента торможения (см. рис. 5). В плоскостях сигнал попадает в каждую ячейку. Алгоритмы распространения сигнала в C-ячейке и S-ячейке представлены на рисунках 7а и 7б.

Распространения в C-ячейке

```

вход: double[]
выход: double
d: double [] //массив весов d
v: double // вес соответствующей v-ячейки
a: double // вес a
выход = УмножитьМассив(d,вход)
выход = (1+выход) / (1 + v) - 1
выход = выход / (a + выход)

```

а)

Распространения в S-ячейке

```

входы: double[][] // входное окно
vВход: double // Выход соответствующей v-ячейки
выход: double
a : double[][] // Вес
b: double // Вес
деноминатор: double

выход = 0
ck = 0
Пока (ck < Длина(входы)){
    выход += УмножитьМатрицу(a[ck], входы[ck])
}
деноминатор = 1 + 2*r/(1+r)*b*vВход
выход = (1+выход)/деноминатор - 1

Если выход < 0
    выход = 0

```

б)

Рисунок 7 – Алгоритмы распространение сигнала в: а) C-ячейке; б) S-ячейке

Информационное обеспечение для обучения и тестирования системы распознавания

Разработанное приложение использует два набора образов. Первый набор образов содержит 5 классов по 3 лица. Все образы хранятся в каталоге базы данных в форма .png. Для каждого из 5 классов существует ещё по одному лицу, из которых формируется база проб. Помимо этого, в системе используются лица из базы данных сотрудников Йельского университета. Данный каталог хранит образы в формате .grm. Он содержит 40 классов по 9 лиц. Из каталога взят 1 образ от каждого класса, который формирует базу проб. На рисунке 8 представлен набор внутренних классов для одного внешнего класса.



Рисунок 8 – Хранение одного класса изображений

Выводы

В данной работе проанализированы основные подходы к распознаванию визуальных образов. Во время исследования изучены алгоритмы работы нейронной сети типа Неокогнитрон, так как данный вид нейронной сети – один из наиболее гибких и не содержит жестких объектов для распознавания. Исследование завершено разработкой приложения по распознаванию лиц на базе неоконитрона. Структура классов приложения соответствует структуре нейронной сети. На текущий момент достигнутый уровень обучения нейронной сети выполняет распознавание с коэффициентом ошибки 40%. В дальнейшем будет продолжен подбор оптимальных значений параметров Неоконитрона для повышения качества распознавания лиц.

Литература

1. Друки, А.А. Система поиска, выделения и распознавания лиц на изображениях. // Известия Томского политехнического университета. – Томск: ТПУ, 2011. – Т. 318, №5. – С. 64-70.
2. Маткасым, Н.Н. Распознавание образов с помощью нейронных сетей. // Технологии Microsoft в теории и практике программирования: сборник трудов XIII Всероссийской научно-практической конференции студентов, аспирантов и молодых ученых. – Томск, 2011. – С. 23-25.
3. Файн, В.С. Оpoznание изображений / В.С. Файн – М.: Мир, 1970. – 153 с.
4. Журавлев, Ю.И. Об алгебраическом подходе к решению задач распознавания и классификации / Ю.Н. Журавлев // Проблемы кибернетики. – М.: Наука, 1978. – С. 5-68.
5. Выделение и распознавание лиц. [Электронный ресурс] // Техническое зрение – Режим доступа: http://wiki.technicalvision.ru/index.php/Выделение_и_распознавание_лиц. – Загл. с экрана.
6. Neocognitron: A neural network model for a mechanism of visual pattern recognition [Электронныйресурс] // IEEE Xplore Digital Library – Режимдоступа: <https://ieeexplore.ieee.org/document/6313076/>. – Загл. сэкрана.
7. Receptive fields, binocular interaction and functional architecture in the cat's visual cortex [Электронныйресурс] // NCBI – PMC – Режимдоступа: <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC1359523/>. – Загл. сэкрана.
8. Шлезингер М., Главач В. Десять лекций по статистическому и структурному распознаванию. – Киев: издательство «Наукова думка», 2004. – 545 с.

Тырса С.В., Федяев О.И. Разработка программной системы распознавания лиц на основе нейросети типа неоконитрон. В данной статье рассмотрены основные методологии распознавания лиц на фотографии. Изучены современные средства распознавания. Предложен метод распознавания лиц на изображении с помощью многослойной нейронной сети типа неоконитрон.

Ключевые слова: нейронная сеть, неоконитрон, машинное обучение, обучение без учителя, распознавание лиц, визуальный образ.

Stanislav Tyrsa, Oleg Fedyaev. Face recognition software system development based on neocognitron neural network. The article provides basic methodology of facial recognition on photos. It has been studied the most modern tools in facial recognition area. It has been proposed the method of facial recognition using neocognitron multilayer neural network.

Key words: neural network, neocognitron, machine learning, unsupervised learning, visual image

Проектирование интеллектуального речевого интерфейса для медицинской справочной системы

Бакаленко В.С., Федяев О.И.

Донецкий национальный технический университет
valeriy.bakalenko@gmail.com, fedyaev@donntu.org

Бакаленко В.С., Федяев О.И. Проектирование интеллектуального речевого интерфейса для медицинской справочной системы. Статья посвящена рассмотрению и анализу применения речевых технологий в медицинских системах. Рассматриваются проблемы разработки интеллектуального речевого интерфейса, его применение в диалоговых сценариях. Проанализированы плюсы и минусы речевых технологий в человеко-машинном взаимодействии для работников медицинской сферы.

Ключевые слова: распознавание речи, синтез речи, технологии бесконтактного взаимодействия, речевой интерфейс, человеко-машинное взаимодействие, анализ данных.

Введение

Основным преимуществом включения речевого интерфейса в приложение является то, что речь – естественный способ коммуникации для человека. Речь формируется у человека на самых ранних этапах его развития и затем поддерживается практикой всю жизнь. В речи очень важно чтобы оба взаимодействующих лица понимали друг друга для успешного сотрудничества. Следовательно, речевой интерфейс в приложении должен быть похож на естественный разговор, только этот разговор осуществляется между пользователем и компьютером. Из выше сказанного следует, что диалог пользователя и системы с речевым интерфейсом может содержать в себе различные вербальные и невербальные реакции, повороты разговора в другое направление, в целом все, что присуще естественному общению. Эффективное приложение с речевым интерфейсом должно имитировать некоторые важные аспекты разговора. Программный продукт должен не нарушать ход диалога и подсказывать пользователю, какие фразы он может сказать системе дальше для эффективного взаимодействия.

Важным фактором в успешности речевого приложения наличие явной пользы от использования речи в той или иной области человеческой деятельности. Поскольку речь является естественной средой для общения, пользователи обычно имеют очень завышенные требования к системам обладающим таким видом интерфейсного взаимодействия. Чтобы избежать такого рода ожиданий и не нагружать интерфейс, необходимо использовать речь там, где пользователь не может ввести свой запрос в систему руками (руки пользователя могут быть заняты чем-то другим; пользователь может недостаточно хорошо владеть навыками ввода или же запрос, ответ на который необходимо получить, находится где-то глубоко в системе), не может прочесть ответ, либо же получаемые данные или запросы не характерны для применения на данном электронном устройстве. Одним из наиболее полезных применений такого рода интерфейсов являются медицинские системы.

Особенности интеллектуальной медицинской системы с речевым интерфейсом

Диагностическая медицинская система должна осуществлять интеллектуальный анализ данных содержащихся в ней и на их основе осуществлять поддержку принятия решения медицинским работником. Приложение с речевым интерфейсом должно быть простым и понятным, чтобы не вызывать больших проблем, которые могут возникнуть при обучении использованию такой системы. У работников медицинской сферы часто нет времени на ввод какого-то большого количества данных в систему, поэтому приложения такого типа до сих пор и не получили широкого распространения. Одним из важных факторов, влияющих на то, будут врачи пользоваться системой или откажутся от нее, является простота и удобство использования. Из всего этого можно сделать вывод, что приложение должно минимально отвлекать работника медицинской сферы от его прямых обязанностей в работе с пациентами и предлагать самые удобные способы управления логикой приложения. Ранее уже были попытки использовать различные указатели, сенсорные экраны и прочее в таких системах. Однако, развитие технологий распознавания речи в последнее время позволяет задуматься над их применением для медицинских систем наряду с классическими способами управления приложением, которые использовались ранее.

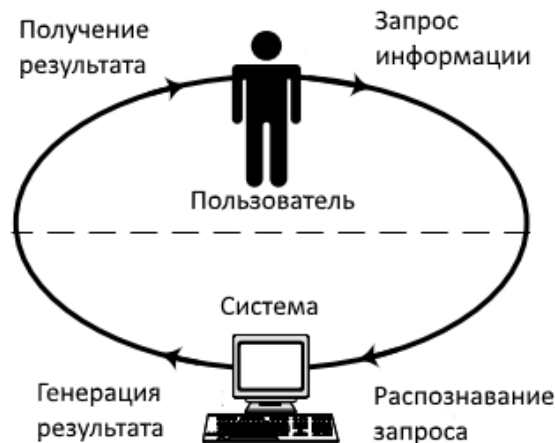


Рисунок 1 – Общая структура взаимодействия человека и системы [1]

Использование речевого интерфейса во время обследования пациента может помочь сэкономить время на ведение карточки пациента, врач может осуществлять какие-то запросы к системе, которая опираясь на прошлый опыт, с помощью интеллектуальных настроек может выдавать определенные советы. Современные речевые технологии накладывают определенные ограничения на своих пользователей, такие как ограниченный словарный запас и грамматика. До тех пор, пока будут существовать такие ограничения, речевые интерфейсы не смогут быть идеальными и поддерживать естественный ход диалога при общении с пользователем. Несмотря на то, что современные технологии распознавания речи требуют значительных доработок и улучшений, уже сейчас их можно сочетать с различными интеллектуальными настройками и использоваться в приложениях, предназначенных для медицинской сферы. Основная цель работы заключается в проектировании интерфейса, использующего речевые технологии, который был бы прост в использовании и как следствие, был бы прост в освоении врачами.

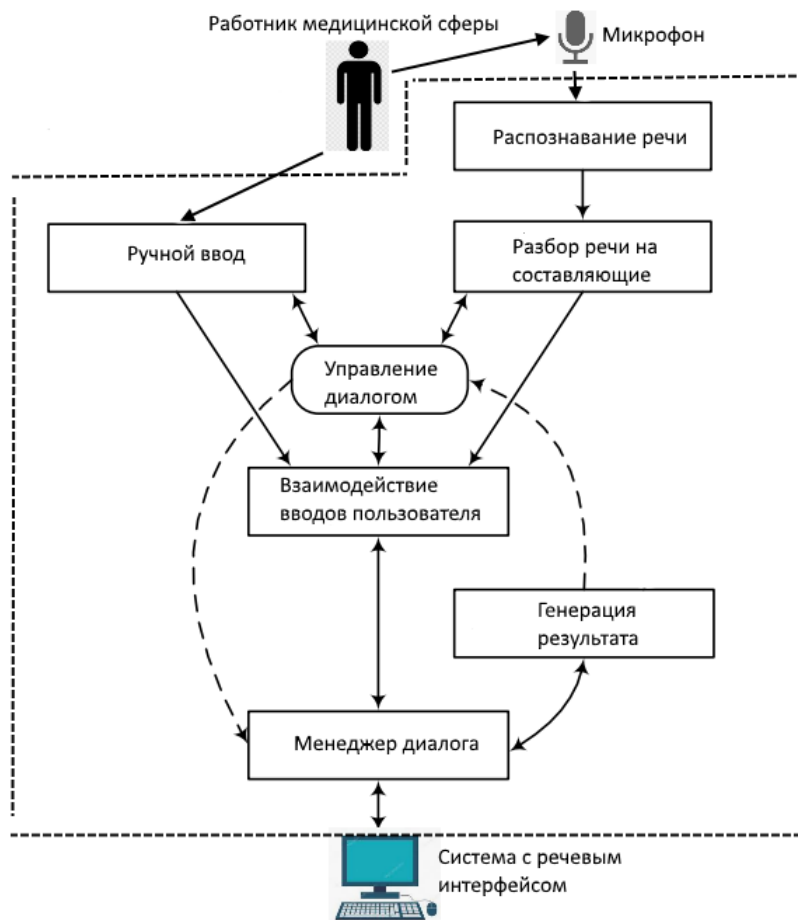


Рисунок 2 – Интерфейс системы, объединяющий в себе ручной и речевой ввод

В ходе экспериментов было выявлено, что наиболее разумным является сочетание ручного и речевого видов человеко-машинного взаимодействия для работников медицинской сферы. Речевой интерфейс удобен при вводе каких-либо данных о пациенте, а также управлении интерфейсом программы, либо же в ходе диалога, когда врач не может использовать ручное управление по каким-либо причинам (это также дает понять, что в такой системе важен и синтез речи, поскольку не всегда есть возможность стоять непосредственно рядом с компьютером, на который установлена такая система).

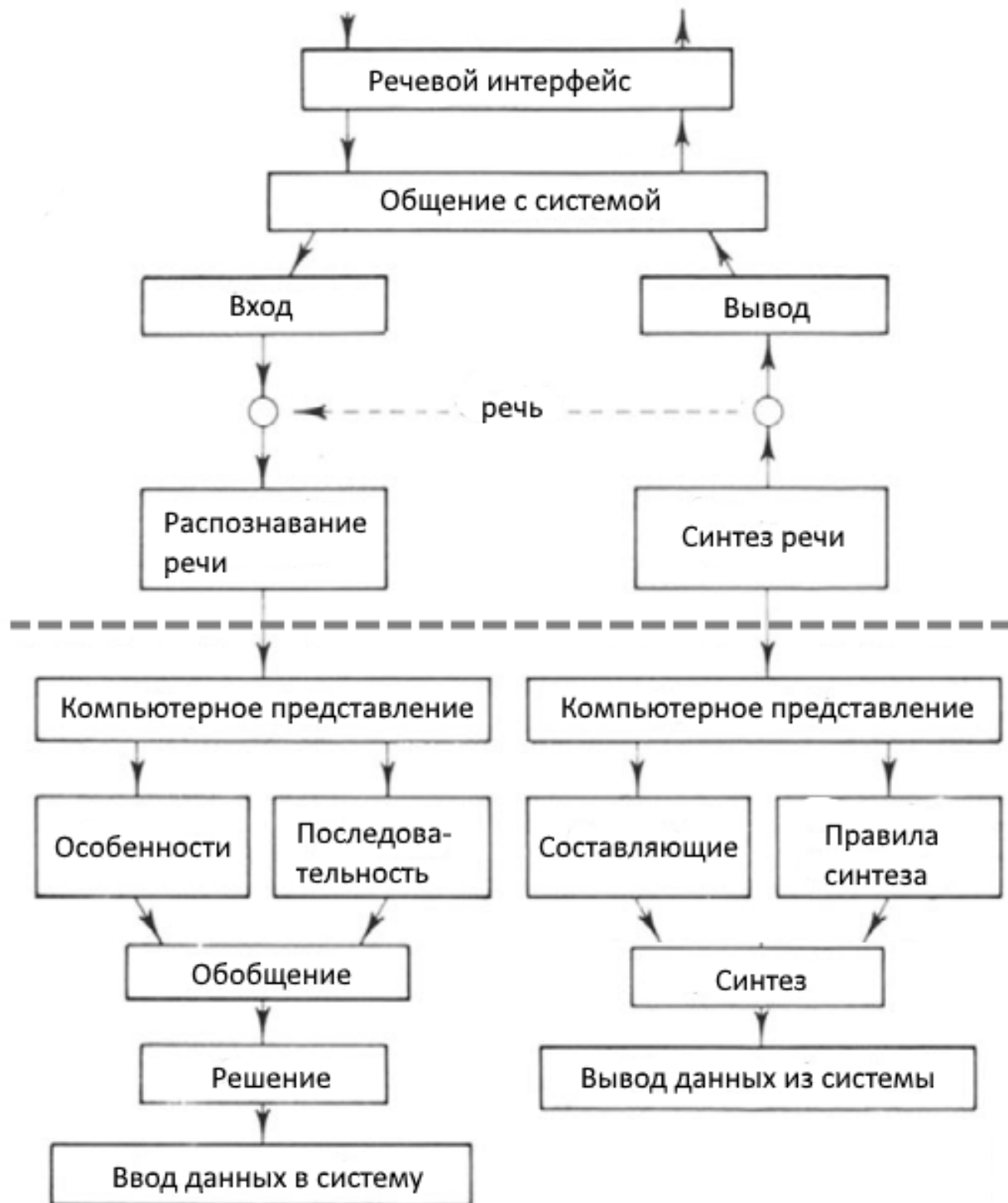


Рисунок 3 – Речевой интерфейс с вводом и выводом.

Устройство и проектирование системы

За многие годы попыток создания приложений с речевым интерфейсом была разработана базовая структура системы речевого взаимодействия между человеком и компьютером (рисунок 4). Многие составляющие этой схемы уже получили достаточно хорошую реализацию, которую можно применять в различных вариантах распознавания речи.



Рисунок 4 – Структура системы речевого человеко-машинного взаимодействия

Необходимо понимать, что в первое время интерфейс с несколькими вариантами управления может показаться менее удобным. Для того чтобы избежать такого впечатления, уже на этапе проектирования надо задуматься о том где и как будет применяться речевой интерфейс. Разрабатываемая медицинская система должна хранить в себе следующие данные о пациенте:

1. Паспортные данные
2. Данные обследований
3. Данные диагноза
4. Операционные данные
5. После операционные данные

Во всех разделах могут использоваться готовые шаблоны для ввода информации и упрощения управления приложением. В дальнейшем такое количество собираемой информации по пациенту должно помочь создать интеллектуальную надстройку, которая в ходе общения врача и системы будет выдавать свои подсказки и советы по текущему пациенту, основываясь на данных других больных. Сложность системы заключается в том, чтобы сделать простым речевой ввод сложных медицинских терминов и запросов. Эта задача является проблемной из-за построения словаря, поскольку работник медицинской сферы может описать одно и тоже заболевание различными терминами. Система должна сообщить пользователю какие именно медицинские термины в нее заложены и соответственно, какие термины она сможет понять, а также направлять пользователя во время диалога. Однако, приложение должно обладать возможностью пополнять свою базу знаний, что позволит понимать большее количество вводимых пользователем данных.

На рисунке 5 показана диаграмма типичной встроенной системы распознавания речи. Самым большим плюсом такой системы является то, что она может быть использована практически в любой сфере деятельности человека, куда можно применить распознавание речи, она может быть использована для распознавания как отдельных слов, предложений, взаимосвязанных фраз, так и непрерывной речи.

Модуль анализа особенностей речи предоставляет вектора акустических данных, которые затем используются для характеристики спектральных свойств изменяющегося во времени речевого сигнала. Модуль

акустического совпадения слов вычисляет схожесть между вектором особенностей входящего сигнала и векторами, которые хранятся в базе распознавателя для всех слов словаря, чтобы определить какое слово было сказано. Языковая модель используется в момент построения предложений, она помогает определить какое слово, вероятнее всего было сказано в этой определенной последовательности. Синтаксические и семантические правила могут быть либо заданы вручную, либо с помощью статистических моделей (N-грамм) с вероятностями.

В результате решение о правильности поиска принимается с учетом вероятных последовательностей слов и выбора наилучшего варианта.

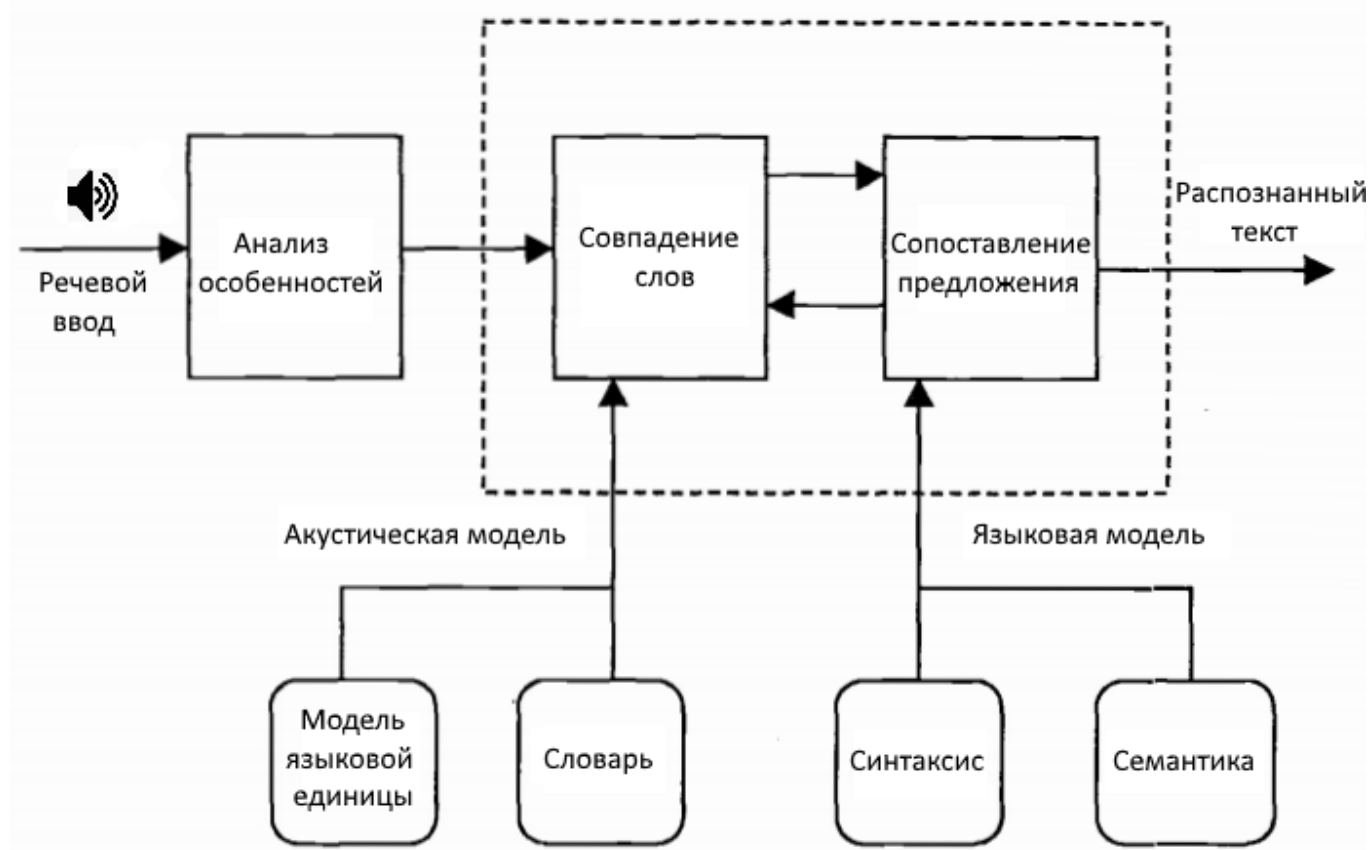


Рисунок 5 – Диаграмма классического распознавателя непрерывной речи

Практически все составляющие распознавателя с рисунка 5 были изучены и оптимизированы за многие годы. Все это помогло точно определять какие словари строить, как создавать и настраивать акустическую модель, как верно использовать языковую модель, как эффективно определять совпадение слов и выделять особенности речевого сигнала. Лучшие практики:

1. Мел-кепстральные коэффициенты
2. Скрытые Марковские модели [2]
3. Контекстно зависимые трифоны [3]
4. MMI или GPD метод обучения [4]
5. Адаптивное обучение системы для улучшения производительности системы для новых пользователей

В дополнение к хорошей технологии распознавания речи необходимо соответствовать и другим требованиям:

1. Интерфейс должен быть надежным и простым в использовании
2. Хорошие сценарии диалога должны помогать разговору двигаться вперед даже в том случае, когда возникают какие-либо замешательства в человеко-машинном взаимодействии.
3. Соответствие задачи технологиям.

Идеальный диалог позволяет обеим сторонам человеко-машинного взаимодействия инициировать запрос или выбрать один из вариантов ответа на запрос другой стороны. Очень важно, чтобы речевой интерфейс имел минимальную задержку во время ответа и работал в режиме реального времени, помогал пользователю понять на каком конкретно этапе диалога в данный момент он находится, использовал гибкую грамматику, где это возможно, подстраивал диалог под текущие ответы или запросы, полученные от работника медицинской сферы [6].

Выводы

Быстрое развитие технологий позволяет создавать все более удобные и естественные способы человеко-машинного взаимодействия применяя речевые технологии. В ходе статьи были рассмотрены очевидные плюсы использования таких интерфейсов для работников медицинской сферы, а также сложности, с которыми придется столкнуться, при их разработке. Был проведен анализ и выявлены ключевые требования к системе с речевым интерфейсом в целом, и некоторым ее составным частям. Были получены положительные отзывы о данной работе от работников медицинской сферы.

Литература

1. CMU Sphinx Open Source Toolkit For Speech Recognition Evaluation [Electronic resource] / Интернетресурс. Режим доступа <http://cmusphinx.sourceforge.net/>
2. Рабинер Л.Р. Скрытые марковские модели и их применение в избранных приложениях при распознавании речи// ТИИЭР. - 1984. - Т.72, № 2. - с. 86-120.
3. Welcome Russian Evaluation [Electronic resource] / Интернетресурс. Режимдоступа: <http://www.voxforge.org/ru>
4. Example of the Baum-Welch Algorithm [Electronic resource] / L. Moss. – Режимдоступа: <http://www.indiana.edu/~iulg/moss/hmmcalculations.pdf>. - Заголовокэкрана.
5. Руанет В.В. Информационные технологии в медицине – введение в медицинскую информатику / В.В. Руанет, А.К. Хетагурова. – М.: МАКСПресс, 2003. – 67 с.
6. Yu, D. Automatic Speech Recognition: A Deep Learning Approach [Text] / D. Yu, L. Deng. — London : Springer-Verlag, 2015

***Бакаленко В.С., Федяев О.И. Проектирование интеллектуального речевого интерфейса для медицинской справочной системы.** Статья посвящена рассмотрению и анализу применения речевых технологий в медицинских системах. Рассматриваются проблемы разработки интеллектуального речевого интерфейса, его применение в диалоговых сценариях. Проанализированы плюсы и минусы речевых технологий в человеко-машинном взаимодействии для работников медицинской сферы.*

***Ключевые слова:** распознавание речи, синтез речи, технологии бесконтактного взаимодействия, речевой интерфейс, человеко-машинное взаимодействие, анализ данных.*

***Bakalenko V.S., Fedyaev O.I. Designing an intelligent speech interface for medical reference system.** The article is devoted to the consideration and analysis of the use of speech technologies in medical systems. The problems of developing an intelligent speech interface, its application in dialog scenarios are considered. Analyzed the pros and cons of speech technology in human-machine interaction for medical workers.*

***Keywords:** speech recognition, speech synthesis, technology of contactless interaction, speech interface, human-machine interaction, data analysis.*

О методах детекции объектов из изображений в реальном времени, основанных на глубоком обучении

Пикалёв Я.С., Веретельник Д.М., Добренко Д.С.

Государственное учреждение «Институт проблем искусственного интеллекта», Донецкий национальный технический университет

pikaliov@gmail.com, dobrenko44@gmail.com, darcus.hant@mail.ru

Пикалёв Я.С., Веретельник Д.М., Добренко Д.С. О методах детекции объектов из изображений в реальном времени, основанных на глубоком обучении. В данной работе выделяются преимущества методов детекции объектов, основанных на глубоком обучении, а также даётся краткий обзор этих наиболее распространённых методов. Проводится сравнение, используя обучающий и тестовый наборы VOC-2007 и предобученную нейронную сеть VGG-16.

Ключевые слова: компьютерное зрение, глубокое обучение, свёрточные нейронные сети, детекция объектов, распознавание образов.

Введение

Детекция объектов является одной из основных проблем области компьютерного зрения. Под детекцией объектов понимается задача распознавания объекта в исходном изображении, а также определение его позиции в данном изображении. Таким образом, детекция объектов является более сложной задачей, чем классификация, в которой выполняется распознавание объектов без локализации объектов.

На сегодняшний день для идентификации объектов редко применяются традиционные методы машинного обучения (Haarfeature, HistogramofGradien, LocalBinaryPattern и т.п.). Преимуществами является: быстродействие; простота реализации и возможность отладки; неплохая точность на малом объёме данных. Однако недостатки: с ростом количества данных не увеличивается качество предсказания; сложны для распараллеливания, используя GPU. В связи с наблюдаемой тенденцией роста объёма обучающих данных, а также вычислительной мощью как персональных компьютеров, так и мобильных устройств является актуальным применение методов идентификации объектов, основанных на глубоком обучении (МИОГО). МИОГО основаны на архитектуре свёрточной нейронной сети (ConvolutionalNeuralNetwork, CNN). МИОГО классифицируют на one-stage (с одним проходом,) и two-stage (с двумя проходами: предположение о кандидате и уточнение вероятности точности детекции кандидата).

Обзор методов детекции объектов из изображений, основанных на глубоком обучении

Среди множества нейросетевых (НС) структур, применяющихся для детекции объектов были отобран ряд кандидатов, основываясь на их распространённости в научно-исследовательском сообществе (на основе анализа научно-исследовательских работ, а также платформ и сообществ).

а) **YOLO** (You Only Look Once) [1]. YOLO является одной из самых распространённых МИОГО в реальном времени. Данный подход использует одну CNN-архитектуру для задач классификации и локализации объекта, используя ограничивающие блоки (рамки, в которых происходит распознавание объекта). В отличие от простых МИОГО, где осуществляется детекция объектов, перемещая небольшое окно по изображению. И для каждого шага классификатор. На каждом шаге НС определяет какой объект находится внутри текущего окна. Использование подобной методики дает несколько сотен или тысяч прогнозов для этого изображения, но сохраняет только наиболее вероятные. Этот подход является очень медленным, так как для получения результата классификатор должен выполнить много итерационных процессов. Более эффективный подход заключается в том, чтобы сначала предсказать, какие части изображения содержат ключевую информацию – так называемые региональные предложения (RoI, regionalofproposal), а затем запустить классификатор только для этих регионов. Классификатор будет делать меньше работы, чем с раздвижными окнами, но все равно запускается много раз.

YOLO использует совершенно другой подход. YOLO выполняет только одну итерацию. YOLO делит изображение на сетку $S*S$ (S – размер стороны рамки ядра, рис. 1а). Каждая из ячеек содержит 5 величин: x , y , w , h и $P(b)$. Координаты (x, y) представляют центр поля относительно границ ячейки сетки.; w , h – ширина и высота всего изображения; c – оценка достоверности, которая определяет точность того, что ограничивающий

блок на самом деле содержит некоторый объект (чем выше показатель - тем толще окно рисуется)), причём не предсказывая, что за объект находится внутри ограничивающего блока.

$$P(box) = P(obj)IoU, \quad (1)$$

где $P(obj)$ - вероятность того, что ограничивающий блок содержит объект; IoU - расстояние пересечения над объединением.

Для каждого ограничивающего блока затем предсказывается класс ($P(class)$):

$$P(class) = P(c)IoU, \quad (2)$$

где $P(c)$ - вероятность того, что объект принадлежит классу.

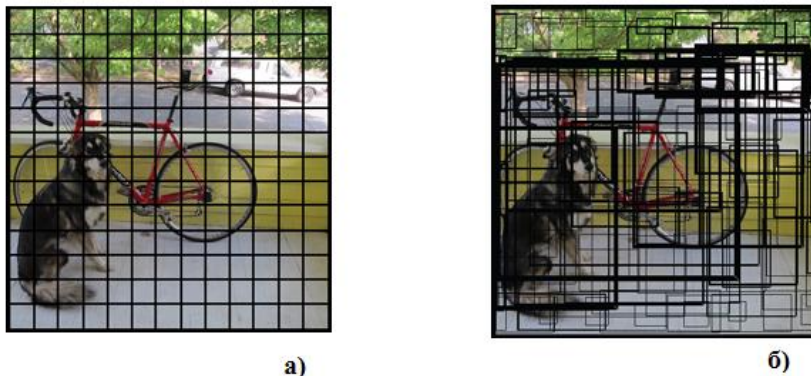


Рисунок 1 – Разделение изображения на области (а) и вычисление оценки определения достоверности локализации (б)

Оценка достоверности для ограничивающего блока и предсказанный класс затем объединяются в конечную оценку, которая отображает вероятность того, что ограничивающий блок содержит определённый класс. Например, на рис. 2а показано, что ограничивающий блок с толстой жёлтой линией с достоверностью в 85% предсказывает что в данном блоке находится объект “dog”. Т. к. изображение содержит 13×13 ячеек и каждая из них предсказывает 5 ограничивающих блоков, то всего содержится 845 ограничивающих блоков. Следовательно, большинство ячеек будут иметь незначительную оценку достоверности, поэтому учитываются только те области, чья итоговая оценка более $n\%$ (n – показатель уровня достоверности, threshold). Итог детекции объектов отображён на рис. 2б.

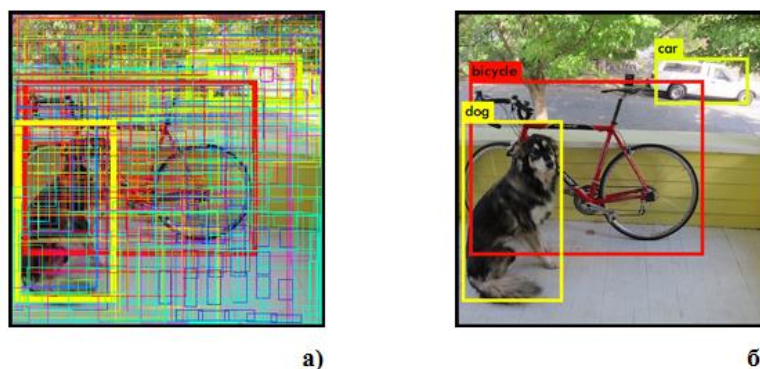


Рисунок 2 – Отображение результата детекции без threshold'a (а) и с threshold'ом (б)

Архитектура YOLO (рис. 3.) состоит из 24 сверточных слоёв, за которыми следуют 2 полностью связных слоя. Чередование 1x1 сверточных слоев уменьшает пространство возможностей из предыдущих слоев.

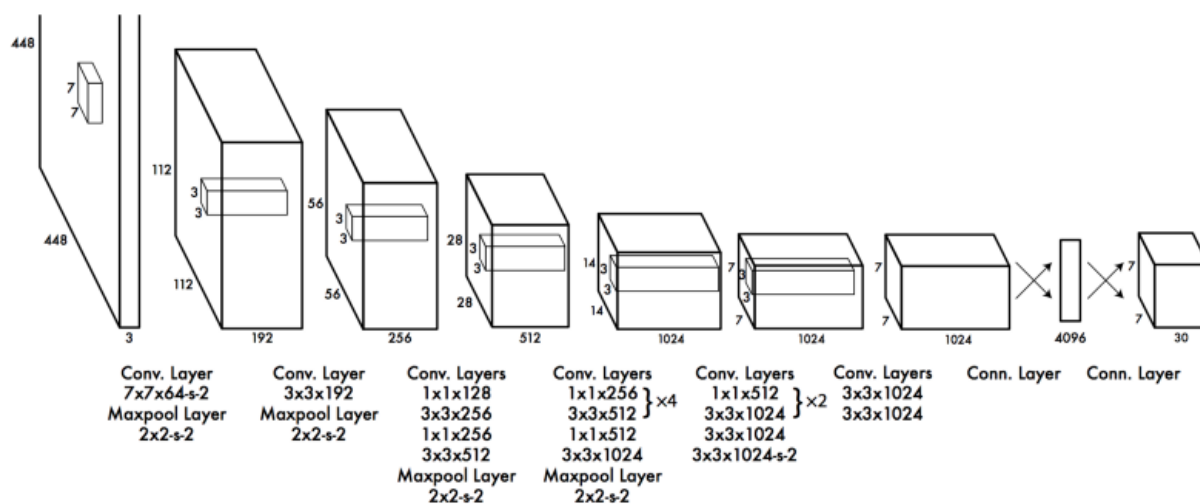


Рисунок 3 – Архитектура YOLO [1].

Таким образом, YOLO обладает следующими преимуществами: быстрое действие (можно применять для детекции объектов в реальном времени); обеспечивает пространственное разнообразие при прогнозировании. Из существенных недостатков стоит выделить тот факт, что применения подхода RoI ограничивают классификатор конкретным регионом; высокий уровень ошибок при локализации и обнаружения объектов. Помимо стандартной архитектуры YOLO существуют его модификации - YOLOv2 [2] и YOLOv3 [3]

б) YOLOv2. Основные изменения, сделанные в YOLOv2 направлены на улучшение точности предсказания детекции объекта, а также на оптимизацию быстрого действия.

Ключевыми изменениями являются:

- - пакетная нормализация для свёрточных слоёв;
- - использование классификатор высокой размерности. Процесс обучения теперь состоит из 2-х фаз: обучение НС для классификации, а затем замена в архитектуре полносвязных слоёв свёрточным и переобучение данной модели для детекции объектов
- применение якорных блоков. В отличие от стандартного подхода, в YOLOv2 вместо предсказания 5 произвольных ограничивающих блоков для одной ячейки, предсказываются смещения для каждого из вышеприведенных якорных блоков (рис.4). Т.е., если ограничить значения смещения, то можно сохранить разнообразие предсказаний и каждое предсказание будет сфокусировано на определенной форме, и более точное будет итоговая детекция.

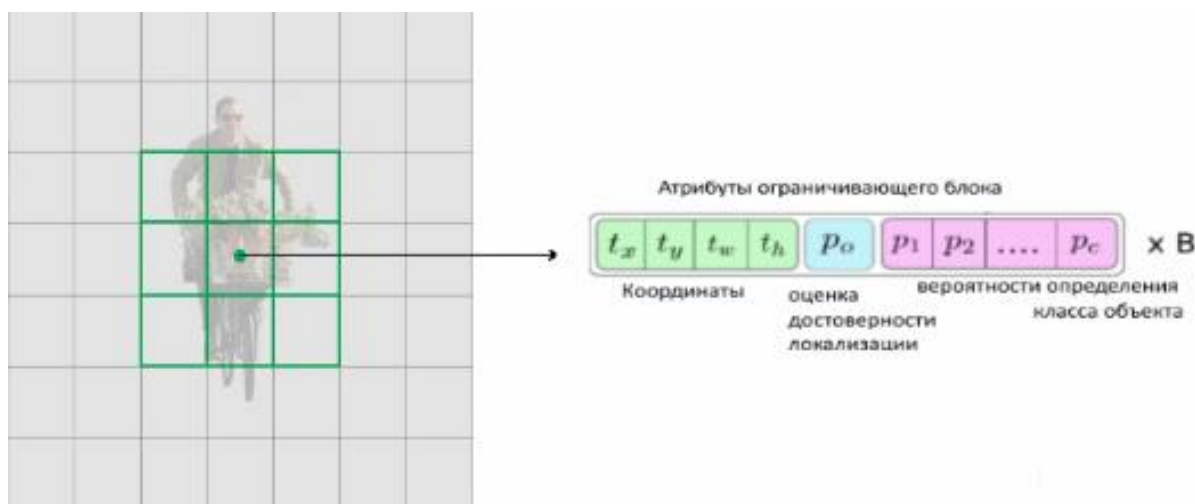


Рисунок 4 – Пример использования якорных блоков в YOLOv2.

в) YOLOv3. Главными нововведениями YOLOv3 являются:

- детекция объекта в трёх размерностях. В YOLOv3 формируется трёхмерная карта признаков (рис. 5), понижая размерность входного изображения в 32, 16 и 8 раз. YOLOv3 делает 3 предсказания для каждой ячейки:

$$K(P) = N \times N \times [3 \times (B + C)] \quad (1)$$

где B – количество ограничивающих блоков для ячейки; C – количество классов; N – размер детекционного ядра;

- возможность многоклассового предсказания. В предыдущих вариантах YOLO применялась softmax-функция для вычисления вероятности по оценкам, в YOLOv3 softmax-функция заменена независимым логистическим классификатором с применением бинарной функции потерь кросс-энтропии для предсказания каждого класса.

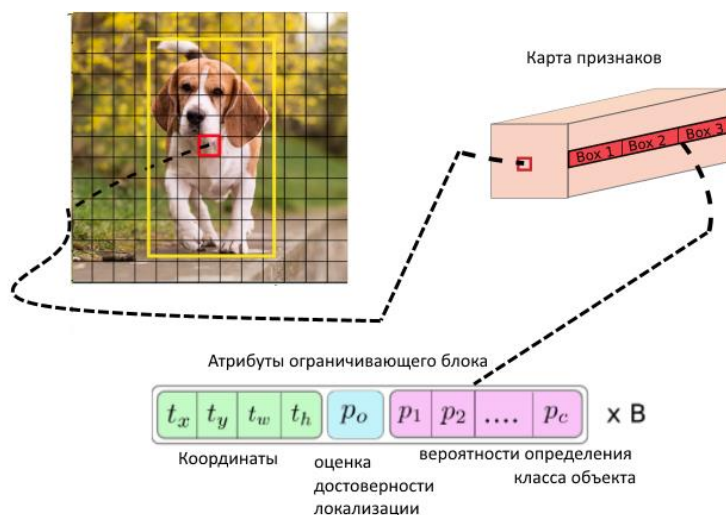


Рисунок 5 – Схема генерации признаков объекта в трёх размерностях

г) **SSD** (Single Shot MultiBox Detector) [4]. SSD является two-stage МНОГО. Подход SSD основан на сверточной сети прямого преобразования, которая генерирует множество ограничивающих блоков фиксированного размера и оценок определения объектов в этих блоках с последующим формированием итоговой оценки детекции (рис. 6, 7).

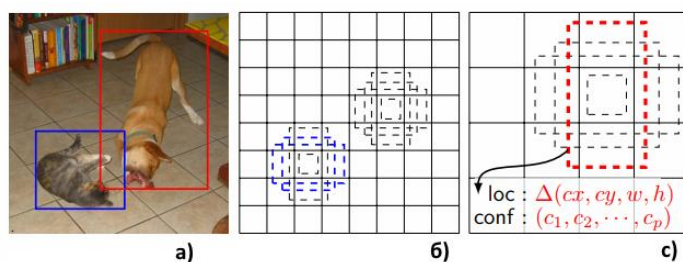


Рисунок 6 – Изображение работы SSD. (а) Результат детекции. В сверточном виде мы оцениваем небольшой набор (например, 4) полей по умолчанию с разными пропорциями в каждом местоположении в нескольких картах объектов с разными масштабами (б и с).

Главными отличиями SSD от стандартных подходов являются:

- использование многомасштабных карт признаков объектов. В SSD вместо исходных полностью подключенных слоев VGG добавлен набор вспомогательных сверточных слоев (от conv6 и далее), что позволяет извлекать объекты в нескольких масштабах и постепенно уменьшать размер ввода для каждого последующего слоя;
- использование сверточной архитектуры для локализации. Каждый добавленный векторный слой (или, возможно, существующий векторный слой из базовой сети) может создавать фиксированный набор предсказаний локализации, используя набор сверточных фильтров;
- связывание множества ограничивающих блоков с каждой ячейкой карты, для нескольких карт признаков в верхней части сети. Блоки по умолчанию разбивают карту признаков, так что положение каждого блока относительно соответствующей ячейки фиксировано. В каждой ячейке карты пространственных объектов предсказываются смещения относительно объектов

в ячейке, а также оценки по классам, которые указывают на наличие экземпляра класса в каждой из этих ячеек.

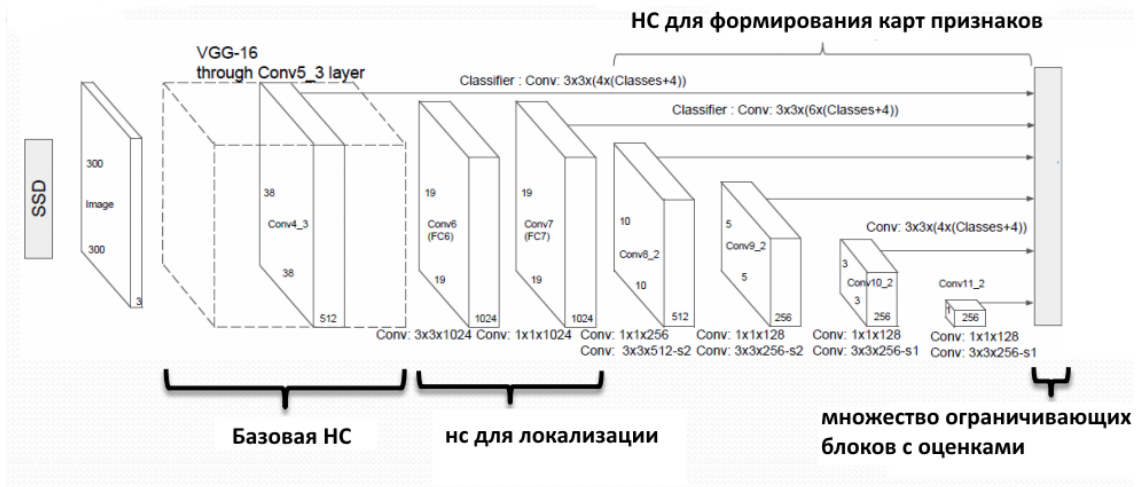


Рисунок 7 – Архитектура SSD.

д) **RetinaNet** [5]. RetinaNet - это единая унифицированная сеть, состоящая из основной НС (для извлечения признаков) и двух вспомогательных для конкретных задач. Основная НС отвечает за вычисление карты признаков объектов по всему входному изображению и является самостоятельной сетью свертки. Первая вспомогательная НС (НС локализации) выполняет классификацию на выходе основной НС; вторая вспомогательная НС (НС классификации) выполняет регрессионную свертку для ограничивающего блока. Для формирования карт пространственных признаков используется архитектура сети признаков пирамиды (FeaturePyramidNet). В качестве функции потерь применяется фокальная потеря.

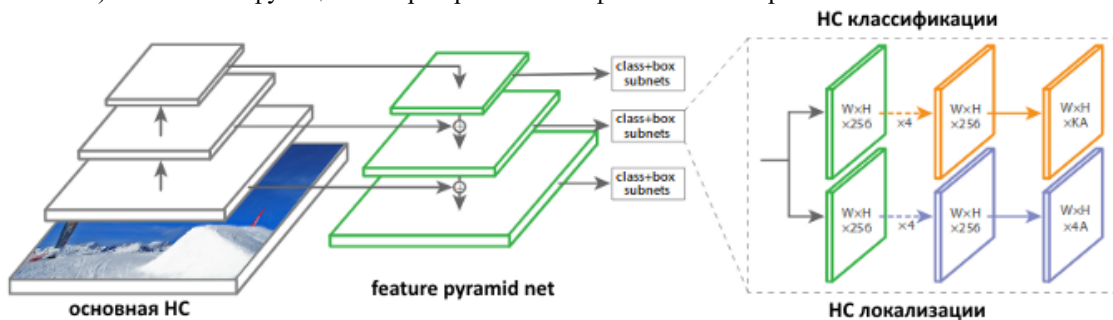


Рисунок 8 – Схемарботы RetinaNet.

е) **Faster R-CNN** (Region-based Convolutional Network, более быстрая свёрточная нейронная сеть на основе регионов) [6]. FasterR-CNN состоит из трёх основных частей частей (рис. 9):

- генерация признаков (Features), который, получив на вход изображение, выдаёт 3x-мерный тензор особенностей. При этом каждому вектору особенностей из этого тензора соответствует некоторый прямоугольник исходного изображения. В качестве такого генератора может выступать, например, сеть VGG16 или ResNet101 без полносвязных слоёв.
- Region Proposal Network (RPN). Сеть, которая получив на вход тензор особенностей, генерирует некоторое количество регионов (ROIProposal), в которых предположительно есть объекты (ROIs) при помощи метода RoIPool. Для каждого RoI на тензор особенностей накладывалась сетка с фиксированным числом ячеек и к каждой ячейке сетки применялся MaxPooling, таким образом для каждого прямоугольника претендента получался тензор особенностей, но уже фиксированных пространственных размерностей.
- полносвязные слои - это сеть, которая для каждого региона, вырезает, соответствующую этому региону, часть из тензора особенностей, и выдаёт: класс объекта (FCsoftmaxclassifier) и ограничивающий блок объекта (FCbboxregressor)

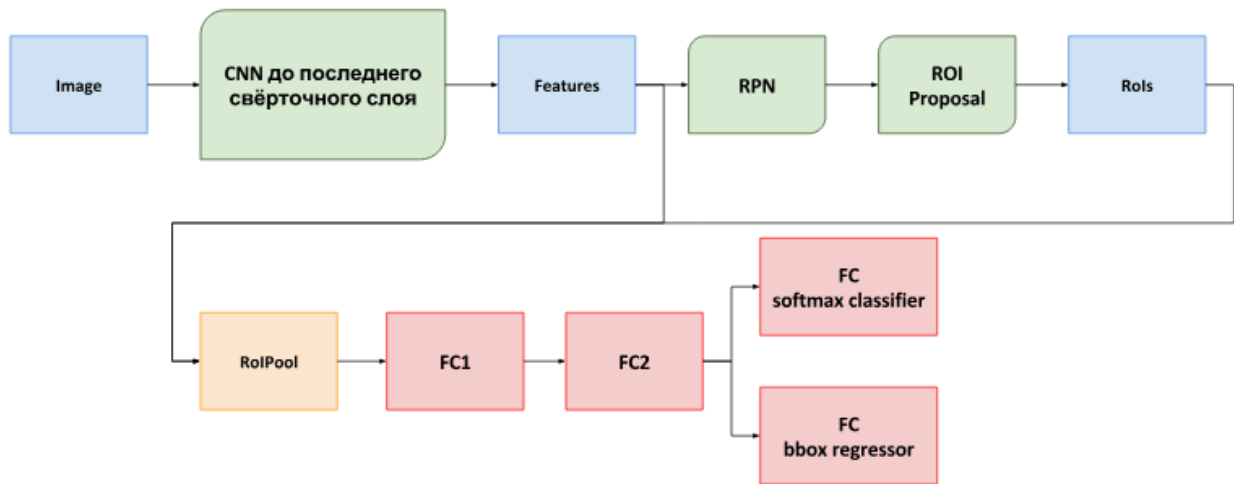


Рисунок 9 – Схема работы FasterR-CNN

ж) R-FCN [7]. R-FCN (Region-basedFullyConvolutionalNet, полностью свёрточная сеть на основе регионов). С одной стороны, при выполнении классификации объекта мы хотим изучить инвариантность местоположения в модели: независимо от того, где кошка появляется на изображении, мы хотим классифицировать ее как кошку. С другой стороны, при обнаружении объекта мы хотим узнать дисперсию местоположения: если кошка находится в верхнем левом углу, мы хотим нарисовать окно в верхнем левом углу. Для этих целей в R-FCN используется позиционно-чувствительная карта оценки, которая представляет собой одну относительную позицию одного класса объектов. Например, одна карта баллов может активироваться везде, где она обнаруживает верхнюю правую часть кошки. Другая карта баллов может активироваться там, где она видит нижнюю левую часть автомобиля. По сути, эти карты баллов представляют собой свёрточные карты признаков, которые были обучены распознаванию определенных частей каждого объекта.

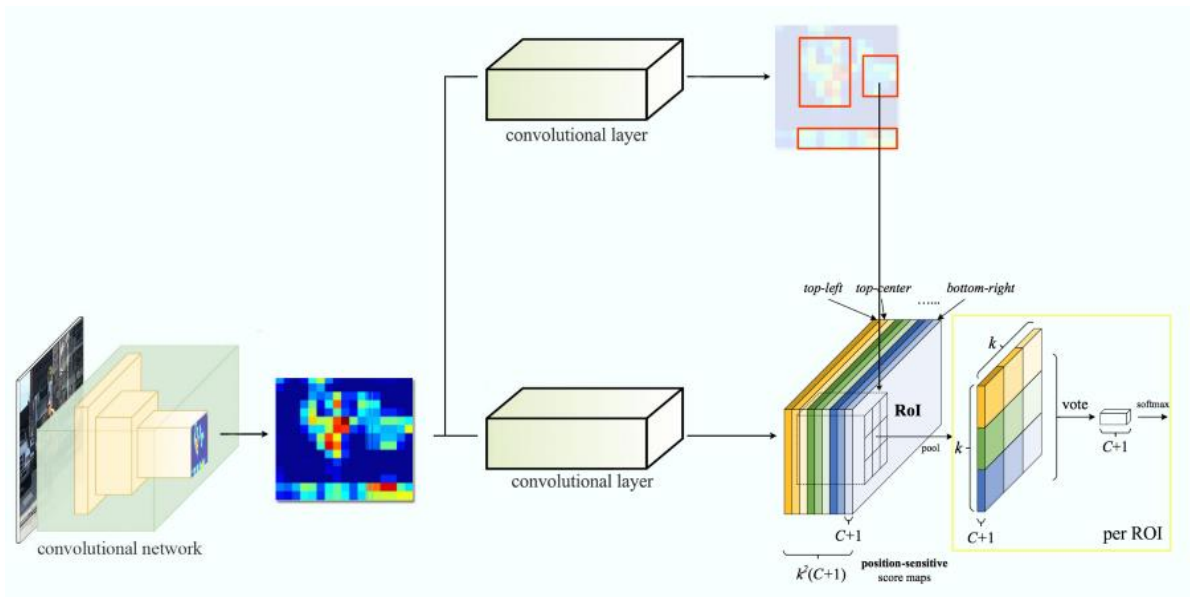


Рисунок 10 –Схема работы RFCN [7]

з) Mask R-CNN [8]. MaskR-CNN является two-stage МИОГО, основанной на Faster R-CNN с дополнительной операцией генерации бинарной маски объекта в дополнение к ограничивающему прямоугольнику и классу. Таким образом у сети появляется дополнительный выход, для которого из каждого класса снимается бинарная маска.

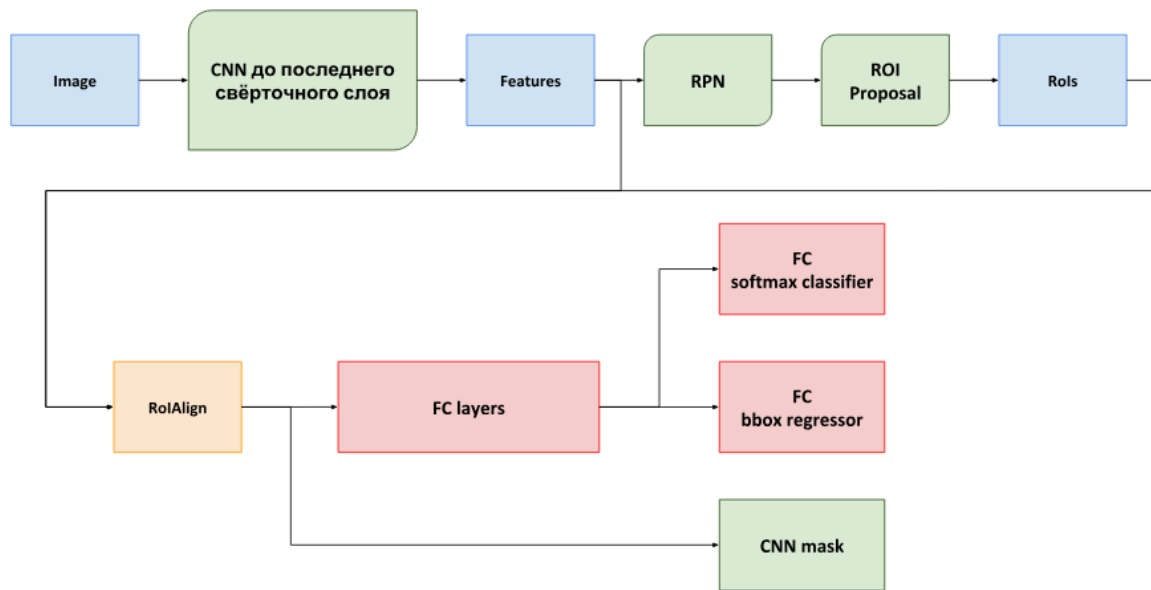


Рисунок 11 – Схема работы MaskR-CNN

Т.к. для каждого класса делается своя маска, поэтому выбор правильной маски осуществляется в зависимости от класса, который выдала сеть. Т.е. в данном подходе разделяется предсказание класса и предсказание маски. Это существенное отличие от случая простой сегментации изображения, где на выходе для каждого пикселя изображения получают вектор с вероятностями того, что этот пиксель принадлежит соответствующему классу. В связи с этим для MaskR-CNN в качестве функции ошибки применяется *sigmoid* с бинарным штрафом. Стоит отметить, что НС генерирует маску фиксированного (не зависящего от размера ROI).

Большая часть архитектуры MaskR-CNN подобна Faster R-CNN (генератор особенностей, RPN, полносвязные слои для получения класса объекта и уточнения прямоугольника). Класс и прямоугольник объект сеть выдаёт в виде векторов небольшой размерности (для класса это вектор вероятностей, для прямоугольника вектор сдвигов координат), и их логично генерировать, используя полносвязные слои. Однако, маска, являющаяся бинарным изображением, генерируется с сохранением пространственной структуры. Т. е. вместо полносвязных слоёв используются свёрточные, что уменьшает число параметров.

Еще одно изменение в структуре сети, связанное с желанием получить маску объекта, это замена RoIPool слоя на RoIAlign, особенность которого состоит в том, что не округляя координаты и не выравнивая, в каждой ячейке заводится дополнительную сетку (зеленые круги на рис. 12). В каждой точке дополнительной сетки применяется билинейная интерполяция, чтобы получить значения признаков, а затем используется MaxPool для генерации выхода. Главным недостатком MaskRCNN из-за подобного подхода является сложность детекции больших протяженных объектов (например, дороги).

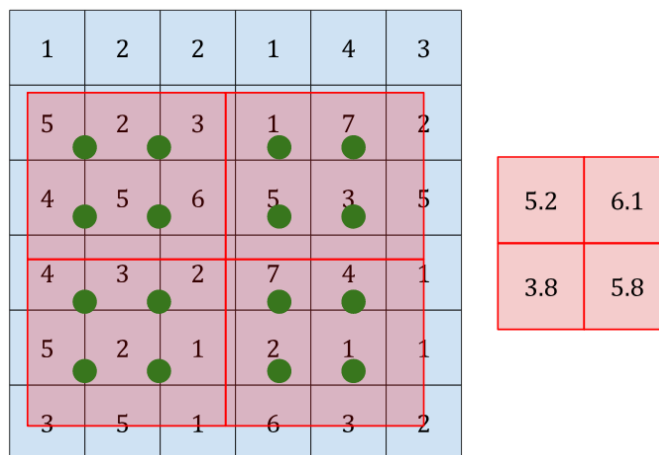


Рисунок 12 – Изображение процесса RoIAlign

Тестирование и сравнение методов детекции объектов из изображений, основанных на глубоком обучении

Тестирование проводилось на датасете VOC-2007 [9]. В качестве базовой НС использовалась VGG-16 [10]. Обучение производилось на облачной платформе глубокого обучения GoogleColab, используя видеокарту TeslaK80. Результаты тестирования отображены в табл. 1.

Таблица 1 – Результаты тестирования.

№	МИОГО	Точность детекции,	Средняя скорость работы, сек	Ресурсоёмкость (видеопамять), Гб
1	YOLOv1	78,3	0,14	1,1
2	YOLOv2	81	0,11	1,4
3	YOLOv3	83,4	0,05	1,5
4	SSD	79,8	0,15	0,7
5	Faster R-CNN	73,2	0,2	2,1
6	MaskR-CNN	77	0,17	1,6
7	R-FCN	79,5	0,170	3,1

Для сравнения МИОГО была составлена оценочная таблица, с их последующей оценкой, к которой был применён статистический метод оценивания - метод Борда [12]. В основе метода Борда лежит упорядочивание альтернатив на основе сумм рангов, назначенных альтернативам экспертами. Метод состоит из двух этапов: на первом этапе для каждого объекта с номером k определяется величина S_k , равная сумме рангов, присвоенных объекту всеми экспертами; на втором этапе определяется ранг объекта – чем больше величина S_k , тем выше место альтернативы в искомом ранжировании. Т.е. по каждому локальному критерию проводится сравнение вариантов. В случае если вариант a_{ij}^j лучший, то он занимает первое место и ему даётся 1 балл, второму – 2 балла; затем для каждого варианта суммируется количество его набранных баллов; наилучшим вариантом считается тот, чья сумма набранных баллов является наименьшей:

$$S_k = \min_j \sum_i a_{ij}^j, i = 1 \dots n; j = 1 \dots m \quad (2)$$

Таблица 2 – Оценочная таблица для МИОГО.

Баллы	Точность	Скорость	Память
5	>80	<0.1	<1000
4	75...80	0.15...0.1	1000...1500
3	70...75	0.2...0.15	1500...1700
2	65...70	0.25...0.2	1700...2000
1	<65	>0.25	>2000

Используя табл. 1 и табл.2, рассчитываем величину Борда и получаем следующее ранжирование МИОГО: $V3 > V2 > V1, V4 > V6 > V5$.

Выводы

Таким образом, наилучшим МИОГО среди рассматриваемых в данной работе является YOLOv3. Преимуществами данного метода является: использование трёхмерной карты признаков; возможность многоклассовой детекции; быстродействие. Главным недостатком данного подхода является довольно высокая ресурсоёмкость. Данную проблему можно решить путём уменьшения размера входного изображения, а также уменьшения гиперпараметров, но за счёт этого уменьшится точность.

Стоит также отметить МИОГО, который не требователен к ресурсам – SSD. Данный подход обеспечивает неплохую точность детекции за счёт использования многомасштабных карт признаков объектов и свёрточной архитектуры для локализации.

В дальнейшем планируется рассмотреть НС-структуры для извлечения признаков, а также применить полученные знания и навыки на практике.

Литература

1. Redmon J. et al. You only look once: Unified, real-time object detection //Proceedings of the IEEE conference on computer vision and pattern recognition. – 2016. – С. 779-788.
2. Redmon J., Farhadi A. YOLO9000: better, faster, stronger //arXiv preprint. – 2017.
3. Redmon J., Farhadi A. Yolov3: An incremental improvement //arXiv preprint arXiv:1804.02767. – 2018.
4. Liu W. et al. Ssd: Single shot multibox detector //European conference on computer vision. – Springer, Cham, 2016. – С. 21-37.
5. Ren S. et al. Faster r-cnn: Towards real-time object detection with region proposal networks //Advances in neural information processing systems. – 2015. – С. 91-99.
6. Dai J. et al. R-fcn: Object detection via region-based fully convolutional networks //Advances in neural information processing systems. – 2016. – С. 379-387.
7. He K. et al. Mask r-cnn //Computer Vision (ICCV), 2017 IEEE International Conference on. – IEEE, 2017. – С. 2980-2988.
8. Everingham M. et al. The pascal visual object classes challenge 2007 (voc 2007) results (2007). – 2008.
9. Simonyan K., Vedaldi A., Zisserman A. Deep inside convolutional networks: Visualising image classification models and saliency maps //arXiv preprint arXiv:1312.6034. – 2013.
10. Кичмаренко О. Д., Огуленко А. П. Теория принятия решений //Одесса: ОНУ им. ИИ Мечникова. – 2012. – Т. 52.

Пикалёв Я.С., Веретельник Д.М., Добренко Д.С. О методах детекции объектов из изображений в реальном времени, основанных на глубоком обучении. В данной работе выделяются преимущества методов детекции объектов, основанных на глубоком обучении, а также даётся краткий обзор этих наиболее распространённых методов. Проводится сравнение, используя обучающий и тестовый наборы VOC-2007 и предобученную нейронную сеть VGG-16.

Ключевые слова: компьютерное зрение, глубокое обучение, свёрточные нейронные сети, детекция объектов, распознавание образов.

Pikaliov Ya.S., Veretelnik D.M., Dobrenko D.S. About methods of detection of objects from real-time images based on deep learning. This paper highlights the advantages of object detection methods based on deep learning, and provides a brief overview of these most common methods. A comparison is made using the training and test datasets VOC-2007 and pre-trained neural network VGG-16.

Key words: computer vision, deep learning, convolutional neural networks, detection of objects, image recognition.

Реализация прототипа анализатора жанровой принадлежности произведений на основании векторно – пространственной модели представления документов

Сторожук Н.О., Коломойцева И.А.
Донецкий национальный технический университет
storozhuk2909@mail.ru, bolatiger@mail.ru

Сторожук Н.О., Коломойцева И.А.. Реализация прототипа анализатора жанровой принадлежности произведений на основании векторно – пространственной модели представления документов. В работе представлено описание прототипа программной системы, реализующей анализ жанровой принадлежности документа с использованием векторно–пространственной модели. Приведен алгоритм работы программы, выполнен анализ результатов, на основании которого сделаны выводы о недостатках построенной модели, определены последующие направления работы.

Ключевые слова: документ, жанр, сходство, терм, информационный поиск векторно-пространственная модель, прототип, анализ.

Постановка проблемы

Поиск семантического сходства между текстами является серьезной проблемой для автоматической обработки текста. Необходимость поиска расстояния между документами возникает в различных задачах, таких как обнаружение плагиата, определение авторства документа, поиск информации, машинный перевод, формирование тестов и задач, автоматическое построение рефератов и пр. Близка к классификации задача рубрицирования текста – его отнесение к одной из заранее известных тематических рубрик или жанру.

Целью данной работы является обоснование выбора модели представления документов в системе, описание программной реализации прототипа программы, определяющей жанровую принадлежность документа, анализ результатов и формирование направлений дальнейшей работы.

Анализ основных моделей информационного поиска

Полнотекстовый поиск документов в больших массивах документов, проводится обычно на основе их поисковых образов, под которыми понимается набор ключевых слов – слов, отражающих основную тему документа. В качестве ключевых слов можно рассматривать отдельные слова, словосочетания, однако в силу особенностей русского языка (наличия склонений и различных словоформ) такие методы не являются приемлемыми. Запрос на поиск можно представить в виде набора слов, тогда подходящие документы можно определить на основе схожести запроса и поискового образа документа. Создание поискового образа документа предполагает индексирование его текста, т.е. выделение в нем ключевых слов.

В основу традиционных методов поиска положены три главных подхода, первый из которых базируется на теории множеств (булева модель), второй — на векторной алгебре (векторно-пространственная модель), а третий — на теории вероятностей (вероятностная модель). Эти подходы могут применяться на практике и в каноническом виде, однако у них есть общий недостаток, обусловленный предположением, что содержание документа определяется множеством слов и устойчивых словосочетаний — термов, которые входят в него без учета взаимосвязей, как «мешок слов», и, более того, считаются независимыми. Такое предположение ведет к потере содержательных оттенков, тем не менее, оно позволяет реализовать поиск и группирование документов по формальным признакам.

Классическая булева модель базируется на теории множеств и математической логике. Популярность этой модели связана, прежде всего, с простотой ее реализации, которая позволяет индексировать и выполнять поиск в больших документальных массивах [1, с. 34].

В рамках булевой модели документы и запросы представляются в виде множества термов — ключевых слов и устойчивых словосочетаний. Каждый терм представлен как булева переменная: 0, если данный терм из запроса отсутствует в документе, или 1 в противном случае. При этом весовое значение термина в документе принимает лишь два этих значения и число повторов термина в документе и их местоположение не учитывается.

В булевой модели запрос пользователя представляет собой логическое выражение, в котором термины связываются логическими операторами конъюнкции, дизъюнкции и отрицания. Так как любое логическое

выражение можно представить дизъюнкцией некоторых выражений, соединенных между собой операцией конъюнкции (дизъюнктивной нормальной формой), то для любого запроса можно выполнить ряд операций поиска по каждому слову запроса и выполнить необходимые логические операции, представленные выше.

На основании этой информации можно перечислить следующие недостатки булевой модели:

- отсутствие весовых и частотных различий использованных термов, что не предоставляет возможности в дальнейшем классифицировать и ранжировать ряд документов по их уровню соответствия булевой модели запроса;

- отсутствие контекстных операторов, невысокая эффективность поиска;

- алгоритмическая необходимость хранения исходных документов, таблиц с булевыми значениями для термов и документов, инверсной таблицы и прочей вспомогательной информации большого объема.

Усовершенствованная булева модель устраняет первые два недостатка введением показателя частоты встречаемости документа, которая нормируется с помощью деления частоты текущего терма в документе на максимальную частоту. Тогда близость между двумя документами можно определить как нормированное расстояние между частотами соответствующих термов. Если предположить, что в запросе каждый терм также имеет какой-то вес, то модель усложняется представлением запроса в рамках булевой модели и усложнением вычислительной сложности получения результата. Кроме того, такая модель усугубляет объем хранимой информации.

В вероятностной модели поиска вероятность того, что документ соответствует запросу, основывается на предположении, что термы запроса по-разному распределены среди релевантных и нерелевантных документов. При этом используются формулы расчета вероятности, базирующиеся на теореме Байеса.

Основной вопрос, который решается с помощью модели: как велика вероятность того, что документ d релевантен запросу q ? Релевантность в данном случае рассматривается как вероятность того, что данный документ может оказаться подходящим для пользователя. Функционирование модели основывается как на экспертных оценках, получаемых в результате обучения модели, которые признают документы из учебной коллекции релевантными/нерелевантными, так и на последующих оценках вероятности того, что документ является релевантным запросу, исходя из состава его термов.

Вероятностная модель поиска предполагает вычисление вероятности соответствия запросу документов из множества, сортировку результатов по релевантности. Следует отметить, что в вероятностной модели поиска предполагается наличие учебных наборов релевантных и нерелевантных документов, показатели которых определяются экспертами, или полученных автоматически при каком-то начальном предположении. Вероятность того, что поступивший документ является релевантным, рассчитывается на основании соотношения появления термов в релевантном и нерелевантном массиве документов.

В случае применения экспертных оценок процесс поиска является итерационным: на каждом шаге итерации, благодаря режиму обратной связи, определяется множество документов, отмеченных пользователем как удовлетворяющих его информационным потребностям.

Недостатками векторной модели являются:

- низкая вычислительная масштабируемость (существенное снижение эффективности при росте объемов данных);

- высокая вычислительная сложность;

- необходимость постоянного задействования узкоспециализированных экспертов для подготовки обучающего множества.

Другой подход к анализу текстов – использование векторно–пространственной модели – предполагает выявление схожести документов на основании пропорций вхождения слов в каждый из документов. Для подготовки данных для использования этого подхода используется метод под названием «мешок слов». Его суть состоит в том, что для дальнейшего анализа важно число вхождения конкретных слов, а не морфологические формы слов и их порядок в документе.

В данной модели документ описывается вектором в евклидовом пространстве, в котором каждому терму, используемому в документе, ставится в соответствие его весовое значение, определяемое на основе статистической информации о его появлении как в отдельном документе, так и во всем документальном массиве. Описание запроса, соответствующего необходимой пользователю тематике, также представляет собой вектор в том же евклидовом пространстве термов.

Предположим, что каждый документ можно охарактеризовать определенным набором слов и частотой их появления. Тогда при условии, что если в документе конкретный набор слов употребляется с определенными частотами, то этот документ отвечает требованиям этого запроса. На основании этой информации строится таблица «слово–документ», где строки соответствуют терминам, а столбцы – исследуемым документам. В каждой ячейке может храниться булево значение, которое показывает наличие хотя бы одного вхождения термина в документ, частота слова в документе или вес термина [2, с. 143]. Теперь, для того чтобы сравнить документ d и запрос q , нужно определить меру схожести двух столбцов таблицы.

В рамках этой модели каждому терму t_i документе d^j соответствует некоторый неотрицательный вес w_i^j . Каждому запросу q , который представляет собой множество термов, не соединенных между собой никакими

логическими операторами, также соответствует вектор весовых значений w_i^j .

При этом вес отдельных термов можно вычислять разными способами. Один из возможных простейших подходов – использование нормализованной частоты появления терма в документе (см. формулу 1):

$$w_i^j = \frac{\text{freq}_i^j}{\max_{1 \leq k \leq n} \text{freq}_k^j}, \quad (1)$$

где freq_i^j – число употребления терма в документе.

Такой показатель веса терма в документе обозначают аббревиатурой tf^j (от английского *termfrequency* – частота термина). Однако этот подход не учитывает, насколько часто рассматриваемый терм используется во всем массиве документов, иными словами – дискриминационную силу терма. Поэтому в случае, когда доступна статистика использования термов во всем документальном массиве, более эффективно следующее правило вычисления веса (см. формулу 2):

$$w_i^j = tf_i^j \times \log \frac{N}{n_i}, \quad (2)$$

где n_i – количество документов, в которых используется терм t_i ,

N – общее количество документов в массиве.

Следует отметить, что основание логарифма может быть любым, для простоты чаще всего используют 2 или 10. Приведенная выше формула многократно уточнялась с целью наиболее точного соответствия выдаваемых документов запросам пользователей. В 1988 году Солтоном был предложен такой вариант для вычисления веса терма t_i из запроса в документе, представленный в формуле 3 [3, с.41]:

$$w_i^q = \left(0.5 + \frac{0.5 * \text{freq}_i^q}{\max_{1 \leq l \leq n} \text{freq}_l^q}\right) \times \log \frac{N}{n_i}, \quad (3)$$

где freq_i^q – частота терма t_i из запроса в тексте этого документа.

Использование такого показателя позволяет усилить отличие термина, если он встречается часто в небольшом количестве документов и уменьшить значение термина и релевантность документа, если термин встречается во многих документах или используется всего несколько раз.

Векторно-пространственная модель представления данных обеспечивает системам, построенным на ее основе, следующие возможности:

- обработку запросов без ограничений их длины;
- простоту реализации режима поиска подобных документов (каждый документ может рассматриваться как запрос);
- сохранение результатов поиска с возможностью выполнения уточняющего поиска;
- возможность установки дополнительных весовых коэффициентов для усовершенствования более узконаправленного поиска и анализа сходства документов [4].

Вместе с тем в векторно-пространственной модели не предусмотрено использование логических операций в запросах, что определенным образом ограничивает ее применимость.

Описание прототипа программной системы

Рассмотрим этапы работы программной модели анализатора документов (см. рис.1). Для разработки прототипа системы и тестирования её работы, цель которой состоит в определении принадлежности документа к некоторому литературному жанру, были выбраны 2 жанра – детектив и любовный роман, подобрано по 5 книг каждого жанра. По нажатию на кнопку в систему загружаются 6 файлов – первый файл является запросом, ещё 5 – предварительно отобранные файлы одного жанра. Программный модуль, отвечающий за лексический анализ, запускается в первую очередь и готовит данные для последующих этапов анализа. Основная задача лексического анализа – распознать лексические единицы текста, удалить стоп-слова и привести их к единому представлению с помощью стемминга. Стоп-словами называются слова, которые являются вспомогательными и несут мало информации о содержании документа. Наборы стоп-слов могут быть разными и зависеть от целей и задач исследования. В данном случае используется список стоп-слов, состоящий из частиц, предлогов и некоторых местоимений. По результатам собственных исследований, удаление стоп-слов отсекает в среднем 30% слов (см. рис.2), что ускоряет дальнейшую обработку данных и повышает точность результатов.



Рисунок 1 – Диаграмма состояний программной модели

В процессе доработки программной модели и дальнейшего анализа полученных токенов метод первичной обработки текста будет усовершенствован на основании экспериментальных данных.

Далее все 6 файлов преобразуются в векторно–пространственную модель согласно описанию выше, т.е. в список весов термов в каждом документе. На данном этапе разработки вычисление весов происходит по формуле Солтона (см. формулу 3), которая предполагает расчет весов одинаково для всех термов. Однако для последующего определения жанровой принадлежности следует выделить термы, которые имеют наибольший вес в интересующем жанре и присутствуют в произведениях другого в минимальном количестве, либо отсутствуют вообще. С этой целью программа сохраняет промежуточные результаты работы для дальнейшего анализа. Так как готовых словарей или прочих вспомогательных материалов инженерного либо литературного характера не найдено, возникла необходимость анализировать данные самостоятельно, т.е. проводить экспертное оценивание. В дальнейшем будет разработан вспомогательный инструмент, анализирующий веса и термы на базовом, «грубом» уровне, на основании полученных им данных будет произведены элементы ручной обработки списка.

Для определения жанровой принадлежности текста в реализовываемом прототипе программной системы был разработан анализатор близости загружаемого документа к документам, жанровая принадлежность которых уже определена. С этой целью была использована векторно–пространственная модель, как наиболее подходящая для решения поставленной задачи. Для определения близости документов были использованы 4 наиболее популярных показателя:

- скалярное произведение векторов;
- евклидово расстояние;
- манхэттенское расстояние;
- метод косинусной меры сходства.

Анализируя эти методы, следует обратить внимание, что значения схожести по евклидовому и манхэттенскому методам на самом деле являются расстояниями между векторами. Таким образом, чем расстояние меньше, тем более документ схож с запросом. Методы, использующие скалярное произведение и косинусные показатели, в результате возвращают именно сходство, то есть наилучший документ тот, который имеет наибольший показатель сходства. Весомым недостатком методов манхэттенского расстояния и евклидова расстояния является полный перебор векторов и вычисление расстояний даже для нулевых точек, что несет за собой увеличение требований к производительности.

Одна из причин популярности косинусного сходства состоит в том, что оно эффективно в качестве оценочной меры, особенно для разреженных векторов, так как необходимо учитывать только ненулевые

измерения [5]. На основании проделанных расчетов и анализа описанных методов, основанных на векторной модели, лучшим методом для определения близости документов был определен метод косинусного сходства (см. рис. 2).

	Исходный текст	Необходимый жанр	Другой жанр
Жанр	Детектив	Детектив	Любовный роман
Число лексем, шт.	53943	434754	409466
Число лексем после стоп - слов, шт.	36452	298838	280789
Длительность обработки, миллисек.	372	3065	2897
Скалярное произведение		10927.2944424631	13648.484940235567
Манхэттенское расстояние		12.863416751657816	14.854492077544691
Косинусная мера сходства		0.008098690186879687	0.00730053182503157
Евклидово расстояние		0.6889242973072531	0.7163377084136358
Длительность вычислений, миллисек.		106401	133637

Рисунок 2 – Результаты работы программы

Так как показатели расстояний между запросом и документами численно меньше у выборки детективного жанра, а значения скалярного произведения и косинусной меры сходства наоборот, больше, можно сделать вывод, что исследуемый документ детективного жанра. Однако разительной разницы заметить не удастся – выбранная для дальнейшего исследования косинусная мера сходства в общем случае принадлежит промежутку $[0, 1]$, т.е., чем ближе значение к единице, тем более вероятно, что жанр определен верно

Следовательно, показатели в 0,0081 и 0,0073 являются недостаточно весомыми, чтобы с уверенностью классифицировать выбранный текст. Для того, чтобы считать документ принадлежащим к определенному жанру, следует максимально приблизить показатель косинусной меры к единице, следовательно, необходимо доработать систему.

Выводы

В статье представлен обзор моделей представления документов для последующего анализа, описание прототипа программной системы, установлены основные направления работы в рамках дальнейшей работы. На данном этапе разработки были выявлены следующие недостатки программной системы:

- высокая продолжительность первичной обработки текстов;
- низкие показатели схожести документов, релевантность которых определена на основе экспертных оценок;
- методы манхэттенского и евклидова расстояния требуют большей временной и вычислительной сложности;
- отсутствует возможность дообучения, то есть дополнения обучающей выборки документами, которые в результате программного анализа были определены как принадлежащие к выбранному жанру;
- требуется выполнять загрузку обучающей выборки каждый раз в начале работы программы.

В связи с этим следует исследовать следующие направления работы:

- доработать первичную обработку документов посредством удаления личных имен, названий населенных пунктов, других малозначащих конструкций;
- использовать распараллеливание в процессе обработки документов;
- использовать для определения схожести документов метод косинусного сходства;
- считать документом, принадлежащим к выбранному жанру такой, косинусная мера сходства которого будет более чем 0.7;
- разработать методологию сохранения результатов анализа для последующего обучения модели;
- увеличить число обучающей выборки минимум до 100 книг выбранного жанра;

– с помощью экспертов выполнить анализ весов часто употребляемых термов в документах определенного жанра, повысить степень их влияния на конечное вычисление схожести документов.

Литература

1. Ландэ, Д. В. Интернетика: Навигация в сложных сетях: модели и алгоритмы / Д. В. Ландэ, А. А. Снарский, И. В. Безсуднов. – Москва : М.: Книжный дом «ЛИБРОКОМ», 2009. – 264 с.
2. Маннинг, К.Д. Введение в информационный поиск / К. Д. Маннинг, П. Рагхаван, Х. Шютце. – Москва : М.: ООО «И.Д. Вильямс», 2011. – 528 с.
3. Salton G., Fox E., Wu H. Extended Boolean information retrieval. Communications of the ACM. 2001./ G.Salton, E.Fox, H.Wu.- Vol. 26. № 4. P.35–43.
4. Векторно-пространственная модель поиска [Электронный ресурс] / Д.В.Ландэ [и др.] // Информатика. – Электрон. дан. 2015. Режим доступа: https://pidruchniki.com/71826/informatika/zadachi_poiska_setyah#42. - Загл. с экрана.
5. Сторожук, Н. О. Анализ методов определения текстовой близости документов / Н. О. Сторожук, И. А. Коломойцева // Материалы студенческой секции IX Международной научно-технической конференции «Информатика, управляющие системы, математическое и компьютерное моделирование» (ИУСМКМ-2018). – Донецк: ДонНТУ, 2018. – С. 43-47.

Сторожук Н.О., Коломойцева И.А.. Реализация прототипа анализатора жанровой принадлежности произведений на основании векторно – пространственной модели представления документов. В работе представлено описание прототипа программной системы, реализующей анализ жанровой принадлежности документа с использованием векторно–пространственной модели. Приведен алгоритм работы программы, выполнен анализ результатов, на основании которого сделаны выводы о недостатках построенной модели, определены последующие направления работы.

Ключевые слова: документ, жанр, сходство, терм, информационный поиск векторно-пространственная модель, прототип, анализ.

Storozhuk N.O., Kolomoitseva I.A.. The implementation of the prototype of the analyzer of the genre affiliation of works based on the vector - spatial model of the presentation of documents. The paper presents a description of the prototype of a software system that implements the analysis of the genre affiliation of a document using a vector-spatial model. An algorithm of the program is presented, an analysis of the results is carried out, on the basis of which conclusions are drawn about the shortcomings of the constructed model, and subsequent work directions are defined.

Keywords: document, genre, similarity, term, information search vector-spatial model, prototype, analysis.

Обзор и анализ алгоритмов для осуществления бинарной классификации информации о внешнеторговой деятельности государств

Титаренко М.Г., Коломойцева И.А.
Донецкий национальный технический университет, г. Донецк
кафедра программной инженерии
timigen2010@yandex.ua
bolatiger@mail.ru

М.Г. Титаренко, И.А. Коломойцева Обзор и анализ алгоритмов для осуществления бинарной классификации информации о внешнеторговой деятельности государств. Представлен анализ существующих алгоритмов классификации, выполнен отбор признаков и тестовых данных, произведено тестирование классификаторов, их оценка и сравнение результатов по бинарной классификации внешнеторговой информации.

Ключевые слова: SVM, KNearestNeighbours, Гауссов классификатор, деревья решений, RandomForest, наивный Байесовский классификатор.

Актуальность и цель работы

В современном мире в интернете ежедневно появляется огромное количество новостных заголовков о внешнеторговой деятельности государств, однако часто эти статьи, заметки и обзоры представляются пользователю общим списком, который, обычно отсортирован по времени добавления и не позволяет оценить полезность информации, действительно ли она соответствует искомой категории. В связи с этим, возникает необходимость автоматической классификации внешнеторговой информации. Данная работа является актуальной для информационно-поисковых систем, направленных на поиск и обработку информации по международной торговле.

В статье приведен обзор алгоритмов классификации информации и их сравнение при работе с данными по внешнеторговой экономической деятельности государств.

Отбор терминов для классификации

Любая классификация производится на основе каких-либо признаков. Для того чтобы классифицировать текст прежде всего необходимо определить значения выбранных признаков для этого текста. На сегодняшний день одной из наиболее эффективных для автоматического определения необходимых признаков является TF-IDF мера [1]. TF-IDF - статистическая мера, которая используется для оценки значимости слова в документе, который является частью набора документов. Вес слова пропорционален частоте его употребления в документе и обратно пропорционален частоте его употребления во всем наборе (документах). Количество признаков было выбрано 10 произвольно, однако при его выборе учитывалось изменение f1 метрики на более высоких показателях данного параметра.

Отбор материала для классификации

Для тестирования алгоритмов классификации решено использовать набор классифицированных статей от reuters в количестве 10788 штук, из которых 7769 приходятся на обучающую выборку и 3019 – на тестируемую. Статьи классифицированы на 90 категорий. В исследовании реализован бинарный классификатор статей по внешнеторговым признакам, поэтому остальные 89 категорий были помечены, как «other».

В качестве классификаторов в исследовании выбраны следующие алгоритмы: SVM (support vector machine), KNearestNeighbours, Гауссов классификатор, Деревья решений, классификатор RandomForest и Наивный Байесовский классификатор.

SVM

SVM (support vector machine) - набор алгоритмов классификации, которые переводят полученные исходные векторы в пространство большей размерности и находят разделяющую гиперплоскость, которая разделяет представленные классы. [2].

В исследовании проводилось тестирование данного классификатора на разных показателях вводимого ядра, гаммы и параметр штрафа. При этом были рассчитаны метрики точности, полноты и f1 метрика. Результаты приведены в таблице 1

Таблица 1 – Метрики для классификатора SVC

Параметры	Precision	Recall	F1
kernel="linear", C=0.025	0.924	0.9612	0.9423
gamma=2, C=1	0.9578	0.9626	0.946
gamma=3, C=1	0.9522	0.9626	0.9477

Согласно полученным данным третье значение является оптимальным согласно F1 - мере. В дальнейшем сравнении будут использованы данные по этим параметрам.

KNearestNeighbours

В основе алгоритма ближайших соседей (kNN) лежит правило, что тестируемый объект со своим набором признаков принадлежит классу, которому принадлежат большинство из k его ближайших соседей [3].

В исследовании проводилось тестирование данного классификатора на разных показателях вводимого k, а именно на 3-х, 5-ти и 10-ти соседях. При этом были рассчитаны метрики точности, полноты и f1 метрика. Результаты приведены в таблице 2.

Таблица 2 – Метрики для классификатора kNN

k	Precision	Recall	F1
3	0.946	0.9559	0.95
5	0.9494	0.9603	0.9527
10	0.9528	0.9566	0.9498

Согласно полученным данным значение кол-ва соседей 5 является оптимальным согласно F1 - мере. В дальнейшем сравнении будут использованы данные по этому параметру.

Гауссов классификатор

Основная идея гауссовского классификатора заключается в предположении того, что функция правдоподобия (тренировочный набор) известна для каждого класса и равна плотности гауссовского нормального распределения [4].

В исследовании проводилось тестирование данного классификатора на разных показателях вводимого аргумента радиально-базисной функции. При этом были рассчитаны метрики точности, полноты и f1 метрика. Результаты приведены в таблице 3.

Таблица 3 – Метрики для Гауссова классификатора

RBF(x)	Precision	Recall	F1
1.0	0.924	0.9612	0.9423
0.5	0.924	0.9612	0.9423
1.5	0.924	0.9612	0.9423

Согласно полученным данным значение RBF слабо влияет на показатели согласно F1 - мере.

Деревья решений

Дерево решений – классификатор, который на тренировочных данных выстраивает структуру, узлам которого являются атрибуты различий, в листьях записаны атрибуты целевой функции, а на ребрах – необходимое множество атрибутов. Задача дерева решений – создать модель, которая предсказывает значение целевой функции на основе нескольких входов [4].

В исследовании проводилось тестирование данного классификатора на разных показателях вводимого аргумента максимальной глубины дерева. При этом были рассчитаны метрики точности, полноты и f1 метрика. Результаты приведены в таблице 4.

Таблица 4 – Метрики для дерева решений

max	Precision	Recall	F1
5	0.9458	0.9573	0.9501
10	0.9421	0.9523	0.9465
15	0.943	0.95	0.9462

Согласно полученным данным глубина дерева 5 является оптимальным согласно F1 - мере. В дальнейшем сравнении будут использованы данные по этому параметру.

Классификатор RandomForest

RandomForest – это алгоритм машинного обучения, который заключается в использовании гомогенного ансамбля деревьев решений. Основная идея состоит в использовании большого ансамбля деревьев решений, который за счет их большого количества улучшает результат классификации [5].

В исследовании проводилось тестирование данного классификатора на разных показателях вводимого аргумента максимальной глубины дерева. При этом были рассчитаны метрики точности, полноты и f1 метрика. Результаты приведены в таблице 5.

Таблица 5 – Метрики для RandomForest

max	Precision	Recall	F1
5	0.924	0.9606	0.9419
10	0.9502	0.9626	0.953
15	0.9487	0.9612	0.9527

Согласно полученным данным глубина дерева 10 является оптимальным согласно F1 - мере. В дальнейшем сравнении будут использованы данные по этому параметру.

Наивный Байесовский классификатор

В основе наивного Байесовского классификатора лежит теорема Байеса. Данный классификатор стал одним из стандартных универсальных методов классификации. Достоинством данного классификатора является относительно небольшое количество данных, необходимых для обучения [6].

В исследовании проведено тестирование данного классификатора. При этом были рассчитаны метрики точности, полноты и f1 метрика. Результаты приведены в таблице 6.

Таблица 6 – Метрики для NaiveBayesian

Precision	Recall	F1
0.9551	0.6568	0.7602

Сравнение классификаторов

После проведения тестирования выполнен сравнительный анализ классификаторов по точности, полноте и f1 мере. В связи с тем, что была использована достаточно большая коллекция тренировочных документов, а также благодаря использованию алгоритма отбора признаков TF-IDF, полученные результаты отличаются достаточно незначительно и все имеют хорошие показатели распознавания текстов с внешнеторговой международной информацией. Исключение составляет лишь наивный Байесовский алгоритм, который показал уровень F1 в 0.7602, что не является удовлетворительным результатом для бинарной классификации. Результаты сравнения представлены на рисунке 1. По взвешенной оценке, наилучшим образом себя показал гомогенный ансамбль RandomForest на глубине дерева 10.

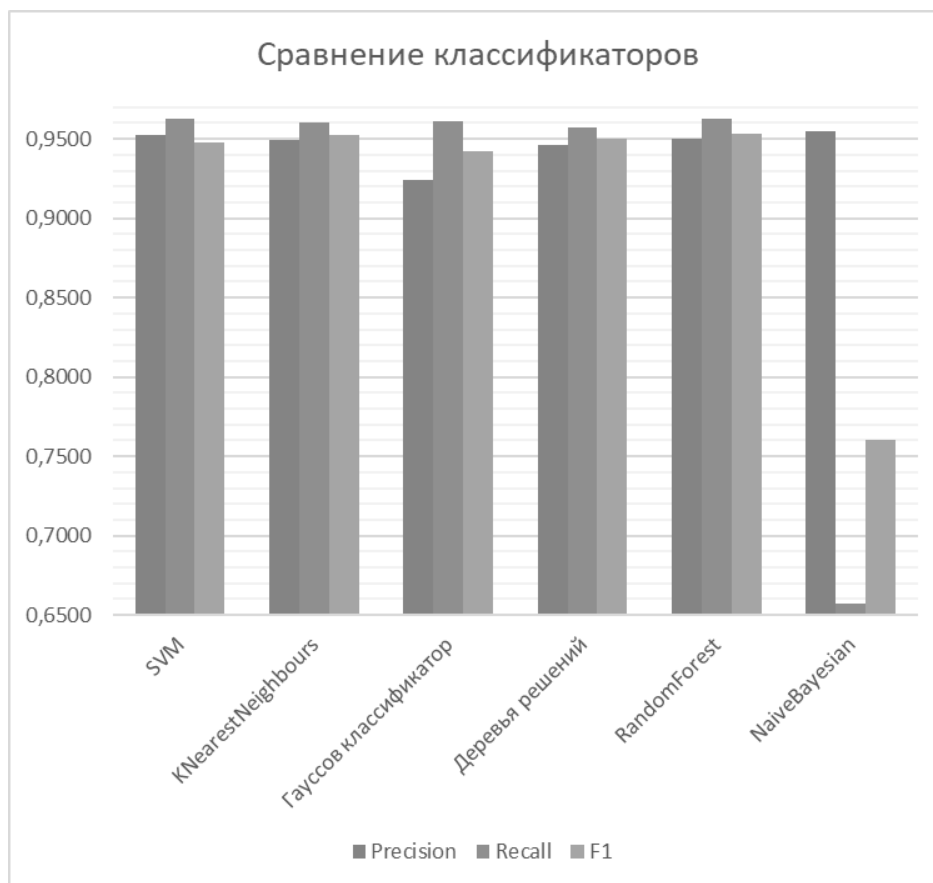


Рисунок 1 – Сравнение классификаторов

Выводы

Проведен анализ алгоритмов классификации, таких как SVM (support vector machine), KNearestNeighbours, Гауссов классификатор, Деревья решений, классификатор RandomForest и Наивный Байесовский классификатор. Реализован алгоритм отбора и отобранные признаки классификации по TF-IDF. Проведено тестирование алгоритмов с различными параметрами по тренировочным и тестовым данным, определены оптимальные параметры для каждого алгоритма на основании оценки F1 меры. Проведено сравнение алгоритмов сравнением оптимального значения F1 меры, полноты и точности для каждого из них. Гомогенный ансамбль RandomForest установлен, как оптимальный классификатор для бинарной классификации внешнеторговой информации. Установлены неудовлетворительные результаты классификации наивным Байесовским классификатором.

Литература

1. Salton, G. and Buckley, C. Term-weighting approaches in automatic text retrieval. Information Processing & Management, 1988
2. Nello Cristianini, John Shawe-Taylor An Introduction to Support Vector Machines and Other Kernel-based Learning Methods. — Cambridge University Press, 2000
3. Brett Lantz Machine Learning with R. Pack Publishing. Birmongham-Mumbai, 2013
4. Breiman, Leo; Friedman, J. H., Olshen, R. A., & Stone, C. J. Classification and regression trees. Monterey, CA: Wadsworth & Brooks/Cole Advanced Books & Software, 1984
5. Hastie, T., Tibshirani R., Friedman J. Chapter 15. Random Forests // The Elements of Statistical Learning: Data Mining, Inference, and Prediction. — 2nd ed. — Springer-Verlag, 2009. — 746 с.
6. Hand, DJ, & Yu, K. «Idiot's Bayes — not so stupid after all?» International Statistical Review, 2001. - с 385—399.
7. Е.И. Большакова Автоматическая обработка текстов на естественном языке и компьютерная лингвистика: учеб. пособие / Большакова Е.И., Клышинский Э.С., Ландэ Д.В., Носков А.А., Пескова О.В., Ягунова Е.В. — М.: МИЭМ, 2011. — 272 с.

М.Г. Титаренко, И.А. Коломойцева Обзор и анализ алгоритмов для осуществления бинарной классификации информации о внешнеэкономической деятельности государств. Представлен анализ существующих алгоритмов классификации, выполнен отбор признаков и тестовых данных, произведено тестирование классификаторов, их оценка и сравнение результатов по бинарной классификации внешнеэкономической информации.

Ключевые слова: SVM, KNearestNeighbours, Гауссов классификатор, деревья решений, RandomForest, наивный Байесовский классификатор.

M.G. Titarenko, I.A. Kolomoitseva Review and analysis of algorithms for the implementation of the binary classification of information on international trade activities of states. The analysis of the existing classification algorithms is presented, the selection of signs and test data is carried out, the classifiers are tested, they are evaluated and the results of the binary classification of international trade information are compared.

Keywords: SVM, KNearestNeighbours, Gauss Classifier, Decision Trees, RandomForest, Naive Bayes Classifier.

Мультиагентное имитационное моделирование системы трудоустройства студентов

Федяев О.И., Медгаус С.В.

Донецкий национальный технический университет
3 fedyayev@donntu.org, medgaus-sergey@yandex.ru

Федяев О.И., Медгаус С.В. Мультиагентное имитационное моделирование системы трудоустройства студентов. Архитектура модели представлена в виде мультиагентной системы. Задача мультиагентного моделирования полностью формализована. Она представлена как задача коллективного выбора с учётом взаимных требований сторон. Эксперимент по анализу трудоустройства проведен в режиме координированного взаимодействия агентов.

Ключевые слова: агенты, мультиагентная система, имитационная модель, сотрудничество агентов, трудоустройство студентов

Введение

Вопросы, касающиеся прогнозирования трудоустройства студентов, являются очень актуальными в наше время. Данная работа посвящена созданию программной системы, которая позволит, исходя из полученных знаний студентов, прогнозировать успешность их трудоустройства с учётом профессиональных требований фирмы. Эта система также позволит выявлять возможные пробелы в подготовке студентов на выпускающей кафедре и оценивать качество подготовки.

Оценка эффективности процессов подготовки специалистов и их трудоустройство очень важна для анализа и управления сложной и инерционной системой образования, в рамках которой решаются задачи составления правильных государственных стандартов, организации эффективного учебного процесса, распределения выпускников на предприятия в соответствии с полученной квалификацией и требованиями заказчиков [1]. Перечисленные задачи являются трудно формализуемыми и поэтому не могут быть решены традиционными математическими методами [2]. Кроме того, участники рассматриваемых процессов взаимосвязаны и образуют распределённую, неоднородную и интеллектуальную систему.

Чтобы учесть особенности анализируемых процессов и построить адекватную модель, в данной работе имитационная система строится методами агентно-ориентированного моделирования, которые сейчас успешно применяются для описания поведения неоднородных систем с распределённым интеллектом [3].

Структура системы моделирования подготовки и трудоустройства студентов

Общая структура системы подготовки специалистов, частью которой является подсистема трудоустройства выпускников, представлена на рисунке 1.

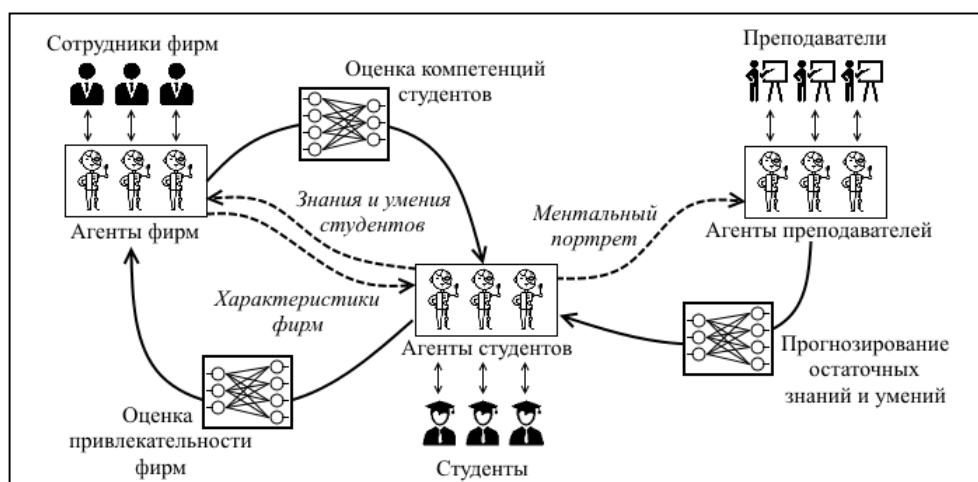


Рисунок 1 - Многоагентная система имитационного моделирования процесса подготовки и распределения студентов на фирмы

Из рисунка видно, что существуют основные три группы людей, между которыми происходит взаимодействие. Каждая из групп передаёт свои цели и способы их достижения своим программным агентам, которые в свою очередь моделируют поведение реальных людей.

Самой большой группой агентов являются агенты студентов. Каждый студент обладает некоторым ментальным портретом (темперамент, уровень IQ, тип личности и т.д.), уровнем знаний и умений по каждой учебной дисциплине, предусмотренной в учебном плане. Для того, чтобы агенты преподавателей смогли спрогнозировать уровень остаточных знаний и умений, студент должен предоставить им свой ментальный портрет. Далее, после того, как остаточные знания определились, агенты-студенты начинают поиск привлекательных фирм, для этого они «обходят» все фирмы и узнают у них условия работы (зарплата, количество рабочих часов, и другие важные социальные факторы), после этого студент формирует для каждой фирмы оценку привлекательности, на основе которой он предлагает свою кандидатуру фирмам. В свою очередь фирмы, при выявлении желания студента работать у них, запрашивают уровни остаточных знаний и умений. После этого, каждая фирма оценивает их, пробуя применить для решения задач каждой конкретной фирмы, и успешно прошедшим этап отбора задачами фирмы предлагают трудоустроиться.

Формальная постановка задачи моделирования процессов трудоустройства

Модель динамического процесса трудоустройства студентов на фирмы (предприятия) представлена двумя группами взаимодействующих искусственных агентов:

$X = \{x_1, x_2, \dots, x_N\}$ - множество агентов, представляющих студентов, которые хотят трудоустроиться по специальности (N - количество студентов);

$F = \{f_1, f_2, \dots, f_M\}$ - множество агентов, представляющих фирмы разного профиля, которые предлагают вакантные места для работы (M – количество фирм).

Процесс трудоустройства начинается с того, что каждый студент по своим критериям оценивает для себя привлекательность каждой фирмы

$$t_{n,m} = \varphi_n(c_m) ,$$

где $t_{n,m}$ – оценка привлекательности m -ой фирмы для n -го студента; φ_n - многомерная функция субъективной оценки студентом x_n привлекательности фирмы f_m ; c_m – вектор значений социально-экономических характеристик, т.е. социальный пакет, который предлагается студенту на фирме f_m . Компоненты вектора c_m определяют размер зарплаты, продолжительность рабочего дня, форму собственности, обеспеченность жильём, возможность удалённой работы и другие показатели, значения которых характеризуют условия работы на m -ой фирме.

В свою очередь каждая фирма оценивает уровень знаний и умений претендентов, предлагая им тестовые задания по профилю деятельности фирмы

$$g_{n,m} = \sum_{j=1}^{J_m} \mu_{m,j}(z_n) ,$$

где z_n – вектор значений уровней знаний и умений, которыми обладает студент x_n (в частном случае – по одной дисциплине); $\mu_{m,j}$ - многомерная функция субъективного оценивания m -ой фирмой способность студента x_n решать j -ое задание, предлагаемое фирмой; $g_{n,m}$ - оценка профессиональных компетенций n -го студента, выставленная m -ой фирмой; J_m - количество тестовых заданий у m -ой фирмы ($1 \leq j \leq J_m$).

Поставленная в статье цель, ориентированная на оптимальное распределение студентов на фирмы, может быть достигнута путём решения следующих двух задач [4].

Задача 1. Она относится к классу обратных задач и заключается в нахождении функций $\varphi_n(\cdot)$ и $\mu_{m,j}(\cdot)$. Функция $\varphi_n(\cdot)$ отражает мнение конкретного студента или группы студентов и описывает зависимость привлекательности фирмы от предлагаемого фирмой социального пакета. Вторая функция $\mu_{m,j}(\cdot)$ имитирует поведение работников фирмы, занимающихся подбором кадров, при тестировании уровня компетенции студентов в зависимости от их знаний и умений по профилю деятельности фирмы. Поскольку обе функции связывают качественные данные, то для их конструирования применена нейросетевая методология как универсальное средство функциональной аппроксимации. Для обучения нейросетевых моделей функций использовались данные реально проводимых опросов нескольких десятков респондентов. Решение этой задачи не рассматривается в данной статье.

Задача 2. Эта задача относится к проблеме коллективного выбора с учётом взаимных требований сторон, которая типична для задач многокритериального принятия решений. В содержательной форме эта задача состоит в распределении студентов на фирмы так, чтобы отклонения от планов приёма отобранных лучших студентов были минимальными, а желания студентов были максимально учтены.

Рассмотрим математическую постановку данной задачи. Каждый студент x_n с учётом своих желаний создаёт для себя список фирм $Q_n = (q_{n,1}, q_{n,2}, \dots, q_{n,M}), q_{n,m} \in F$, который показывает очерёдность посещения им фирм при поиске работы, т.е. в каком порядке студент x_n будет обходить фирмы. Для этого списки Q_n , входящие во множество $Q = \{Q_1, Q_2, \dots, Q_N\}$, упорядочены по убыванию оценок привлекательности фирм, полученных с помощью функции $\varphi_n(\cdot)$:

$$t_{n,k}(q_{n,i}) \geq t_{n,l}(q_{n,i+1}) \text{ если } q_{n,i} = f_k, q_{n,i+1} = f_l.$$

Поскольку один студент не может одновременно посетить все фирмы, то трудоустройство всех студентов происходит за несколько временных циклов посещения ими выбранных фирм. На каждом цикле для каждой фирмы f_m формируется очередной поток (множество) $R_m = \{r_{m,i} | r_{m,i} \in X, 1 \leq i \leq I\}$ из студентов x_n , которые хотели бы устроиться на данную фирму, где I – размер группы претендентов в данном потоке; $\bigcap_{m=1}^M R_m = \emptyset$; \emptyset – пустое множество.

Перед каждым новым циклом посещения все потоки во множестве $R = \{R_1, R_2, \dots, R_M\}$ обновляются. Например, в очередной k -й поток соискателей на фирму R_m попадают те студенты x_n , у которых первыми в очереди Q_n стоит фирма f_m , т.е. $q_{n,1} = f_m (1 \leq n \leq N)$. Целое число k также можно трактовать как порядковый номер цикла посещения студентами фирм.

Перед каждым новым циклом посещения все потоки во множестве $R = \{R_1, R_2, \dots, R_M\}$ обновляются. Например, в очередной k -й поток соискателей на фирму R_m попадают те студенты x_n , у которых первыми в очереди Q_n стоит фирма f_m , т.е. $q_{n,1} = f_m (1 \leq n \leq N)$. Целое число k также можно трактовать как порядковый номер цикла посещения студентами фирм.

Пусть планы приёма студентов на фирмы задаются множеством $L = \{l_1, l_2, \dots, l_M\}$, где l_m – количество вакансий на m -ой фирме ($0 \leq l_m \leq N$). Введём множество последовательностей $Y = \{Y_1, Y_2, \dots, Y_M\}$, в котором каждая Y_m определяет список кандидатов, отобранных в ходе тестирования студентов на фирме f_m . Каждая последовательность Y_m представляет собой множество студентов, упорядоченное по убыванию полученной ими оценки по тестированию. В каждой последовательности (упорядоченном множестве) $Y_m = \{y_{m,1}, y_{m,2}, \dots, y_{m,N}\}$ элементы $y_{m,n}$ состоят из студентов, входящих во множество X , т.е. $y_{m,n} \in X$. Из этого следует, что $g_{k,m}(y_{m,i}) \geq g_{l,m}(y_{m,i+1})$, если $y_{m,i} = x_k, y_{m,i+1} = x_l, 1 \leq i \leq N - 1$. В целом Y – это мультимножество, т.е. каждый студент x_n может успешно пройти тестирование на нескольких фирмах и таким образом принадлежать нескольким последовательностям Y_n .

Кроме того, студенты и фирмы могут устанавливать для себя пороговые значения соответственно по уровню привлекательности фирмы и уровню компетентности студента в виде множеств $P = \{p_1, p_2, \dots, p_N\}$ и $D = \{d_1, d_2, \dots, d_M\}$, где p_n – минимально-допустимое значение привлекательности фирмы у n -го студента; d_m – минимально допустимый уровень компетентности для соискателей на m -ой фирме.

На каждом k -м цикле отбора студентов каждая фирма f_m принимает на входе соответствующий поток студентов $R_m = \{r_{m,i} | r_{m,i} \in X\}$, которые желают работать на фирме. Фирмы в своих списках кандидатов Y_m всем элементам $y_{m,n} \in X$ из пересечения множеств $R_m \cap Y_m$ приписывают приоритет, равный номеру цикла k . Условимся считать, что чем меньше значение k , тем выше приоритет студента x_n в последовательности Y_m . Обозначим через $p(y_{m,n})$ значение приоритета студента x_n , находящегося на месте $y_{m,n}$ в последовательности Y_m . В ходе циклического процесса все студенты-кандидаты в мультимножестве Y получают свои значения приоритетов.

Таким образом, задача квазиоптимального распределения студентов на фирмы сводится к такой расстановке элементов $y_{m,n} \in X$ в последовательностях Y_1, Y_2, \dots, Y_M , где

$$Y_m = \{x_n | (x_n \in X) \& (g_{n,m}(x_n) \geq d_m) \& (t_{n,m}(f_m) \geq p_n)\},$$

при которой минимизируются, во-первых, невыполнение заявок фирм на молодых специалистов с учётом сохранения ранжирования отобранных кандидатов по уровню их квалификации $g_{n,m}(\cdot)$ и, во-вторых, сумма всех значений приоритетов $p(y_{m,n})$ во множестве Y (оно должно быть уже не мультимножеством), гарантирующая рациональную расстановку студентов на основе введенных приоритетов $p(y_{m,n})$, выражающих желания студентов посредством функции $t_{n,m}(\cdot)$.

Анализ процессов трудоустройства с помощью имитационной модели

В виде программы многоагентная система моделирования была реализована с помощью инструментальной среды Jason, которая ориентирована на создание интеллектуальных программных агентов с открытой BDI-архитектурой, позволяющей описывать на языке Java поведение агентов с учётом специфики решаемой задачи [5]. На рисунке 2 представлен интерфейс программного обеспечения системы, реализующей взаимодействие с системой посредством удобных окон. Окна позволяют настраивать агентов-студентов и агентов-фирм, задавая их характеристики, а также редактировать список доступных компетенций и выбирать различные режимы моделирования процесса трудоустройства.

Окно диалога для задания социально-экономических характеристик фирмы представлено на рисунке 3.

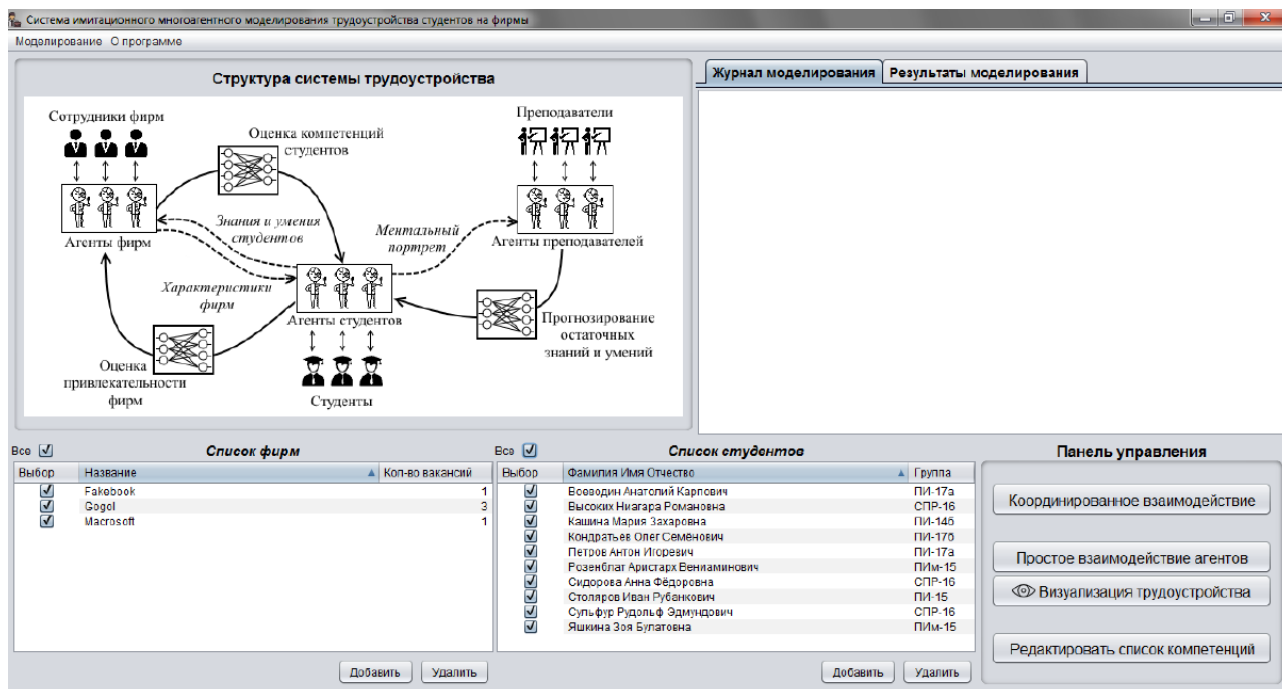


Рисунок 2 - Пользовательский интерфейс системы имитационного моделирования

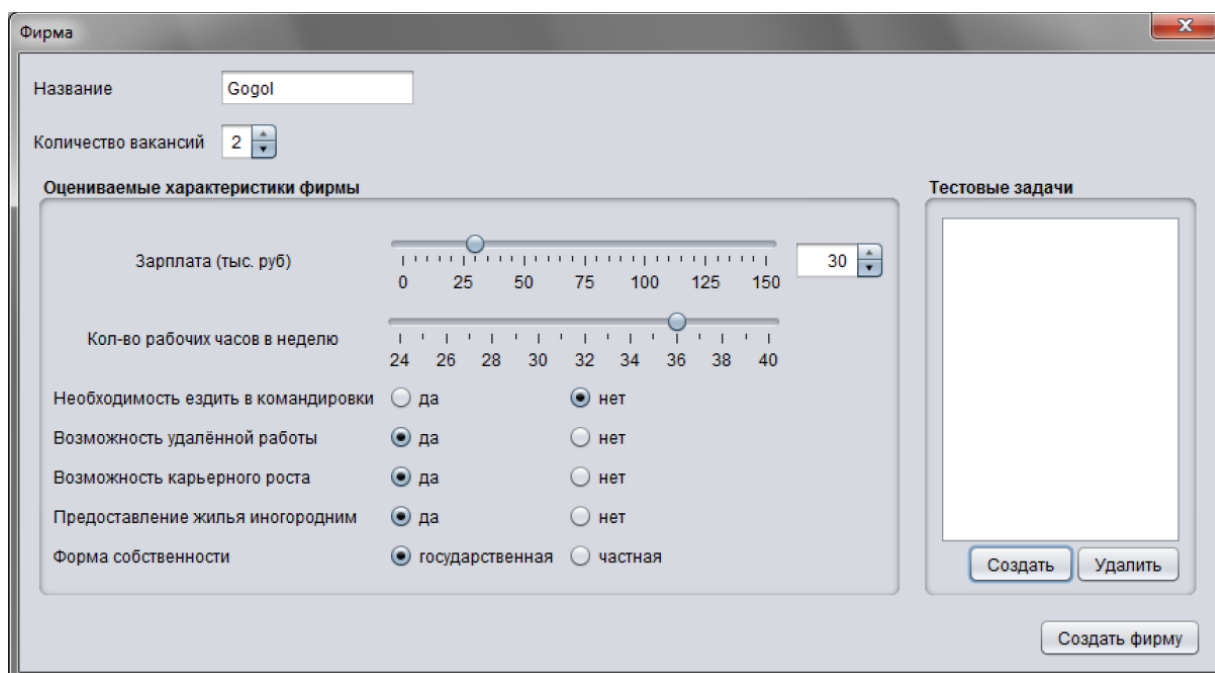


Рисунок 3 – Диалоговое окно для задания условий работы на фирме

Первым этапом моделирования процесса трудоустройства является «собеседование», которое происходит между агентами студентов и работниками фирм. На рисунке 4 представлена диаграмма последовательности на языке UML, показывающая динамику взаимодействия агентов при наличии Координатора (арбитра), роль которого может выполнять, например, ответственный за распределение студентов на кафедре или заведующий кафедрой.

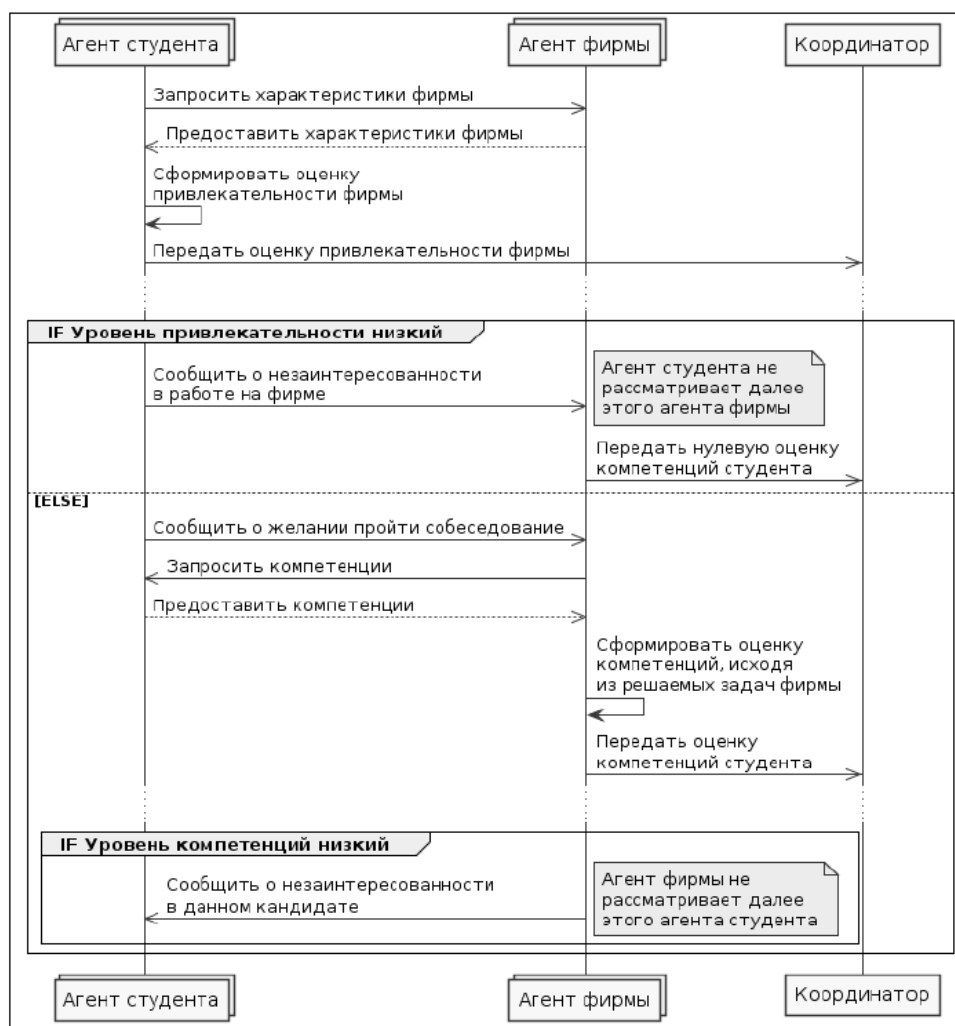


Рисунок 4 – Диаграмма последовательности взаимодействия агентов-студентов, агентов-фирм и Координатора

При координированном взаимодействии агентов для получения плана распределения агент Координатор, получив на этапе знакомства списки предпочтения агентов, может сразу применять оптимизационный метод и выдать оптимальное решение по трудоустройству. При таком виде трудоустройства агенты взаимодействуют только на этапе знакомства.

Функция вычисления оценки привлекательности фирмы, которая используется в самом начале взаимодействия агентов, реализована в нейросетевом базисе, так как она является трудно формализуемой. Поэтому для её построения использовалась модель нейронной сети - многослойный персептрон, а проверка адекватности модели нейросети выполнялась на реальных данных (см. рис. 3), полученных от нескольких десятков студентов-респондентов. После настройки нейросетевой функции по стратегии «обучение с учителем» были проведены эксперименты для получения субъективной оценки привлекательности фирмы со стороны конкретного студента (см. рис.5).

Следующая важная задача связана с получением оценки уровня владения компетенциями (знаниями и умениями) в виде коэффициентов уверенности из диапазона [0,1], которые отражают наличие остаточных знаний по той или иной дисциплине. В свою очередь работодатель формирует список тестовых задач, которые являются профильными для его предприятия, попутно выбирая те знания и умения, которые, по его мнению, действительно важны в решении соответствующей задачи и ставит каждой компетенции свой коэффициент важности по фиксированной четырёхбалльной шкале (1 – высокий, 0.8 – выше среднего, 0.5 – средний, 0.3 – малый). Для удобства коэффициенты важности (P_i) целесообразно нормализовать в диапазон от 0 до 1, так

чтобы $\sum P_i = 1$. Поскольку все входные сигналы изменяются в одном диапазоне от 0 до 1 и оценки важности, установленные экспертом, также нормализованы, то оценку успешности решения одной задачи можно реализовать одним нейроном с линейной функцией активации без обучения (весовыми коэффициентами выступают нормализованные коэффициенты важности P_i). Также необходимо отметить, что итоговым результатом оценки уровня компетенций студента является обобщающий результат оценок по всем задачам фирмы, поэтому итоговая оценка высчитывается по формуле

$$\text{Итоговая оценка} = \frac{1}{k} \sum_{j=1}^k \sum_{i=1}^{N+M} x_i * P_i ,$$

где k - количество тестовых задач; N, M – количество соответственно знаний и умений по дисциплине; x_i - степень владения компетенцией студентом; P_i - нормализованный коэффициент важности.

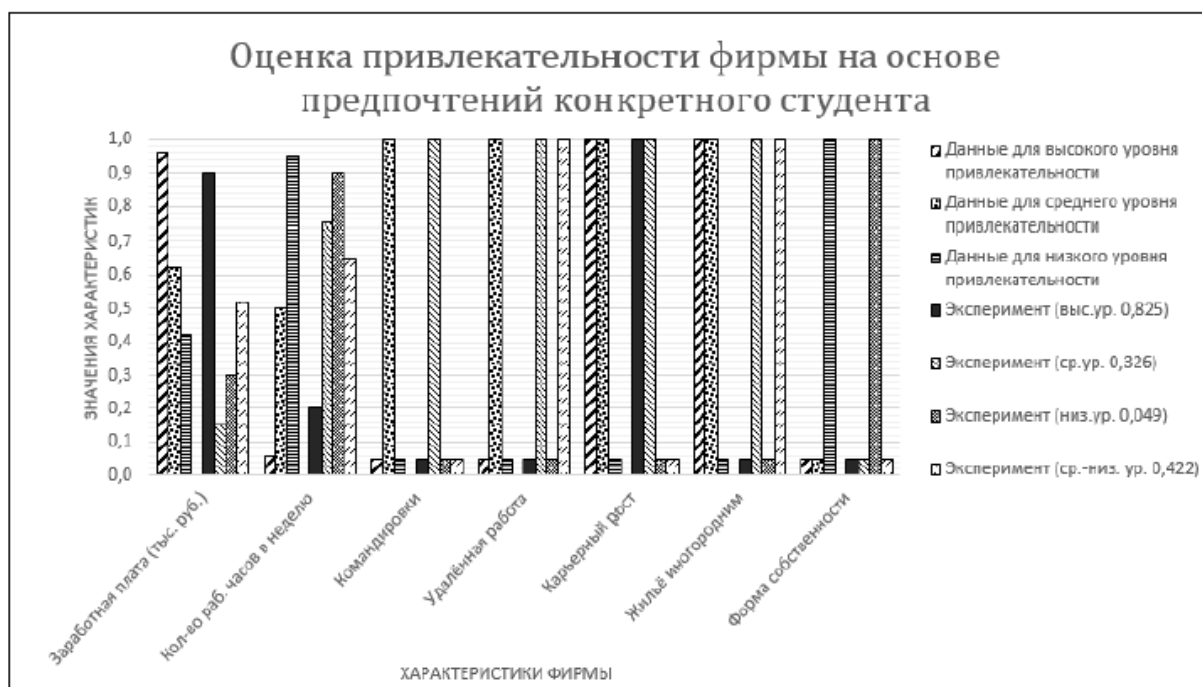


Рисунок 5 – Эксперименты с нейросетью, обученной на основе предпочтений конкретного студента

Ниже приведены результаты моделирования процесса трудоустройства, в котором участвовали 3 фирмы (у всех по 4 вакантных места) и 10 студентов (у всех принята минимальная привлекательность и минимальный уровень знаний, равные 0). При запуске процесса моделирования все агенты студентов проходят «собеседование» с агентами фирм в соответствии с алгоритмом на рисунке 4 и по ранее рассмотренным методикам агенты с нейросетевой архитектурой формируют значения привлекательности и компетенции, которые передаются агенту Координатору (см. табл. 1).

Таблица 1 – Предпочтения агентов

Студенты	f_1 – кол-во вакансий - 4		f_2 – кол-во вакансий - 4		f_3 – кол-во вакансий - 4	
	Привлекательность	Компетенция	Привлекательность	Компетенция	Привлекательность	Компетенция
x_1	0,44	0,89	0,38	0,76	0,5	0,25
x_2	0,62	0,25	0,41	0,23	0,23	0,84
x_3	0,63	0,79	0,5	0,73	0,43	0,37
x_4	0,46	0,72	0,57	0,44	0,35	0
x_5	0,79	0,04	0,77	0,85	0,76	0,86
x_6	0,4	0,27	0,74	0,76	0,3	0,26
x_7	0,51	0,24	0,08	0,91	0,43	0,84
x_8	0,06	0,48	0,51	0,1	0,92	0,77
x_9	0,82	0,22	0,99	0,42	0,5	0,24
x_{10}	0,01	0,04	0,31	1	0,47	0,8

На рисунке 6 представлен журнал моделирования, в котором записываются все действия, совершённые агентами в ходе моделирования (строки без сдвига – действия агентов-фирм, строки со сдвигом – действия агентов-студентов).

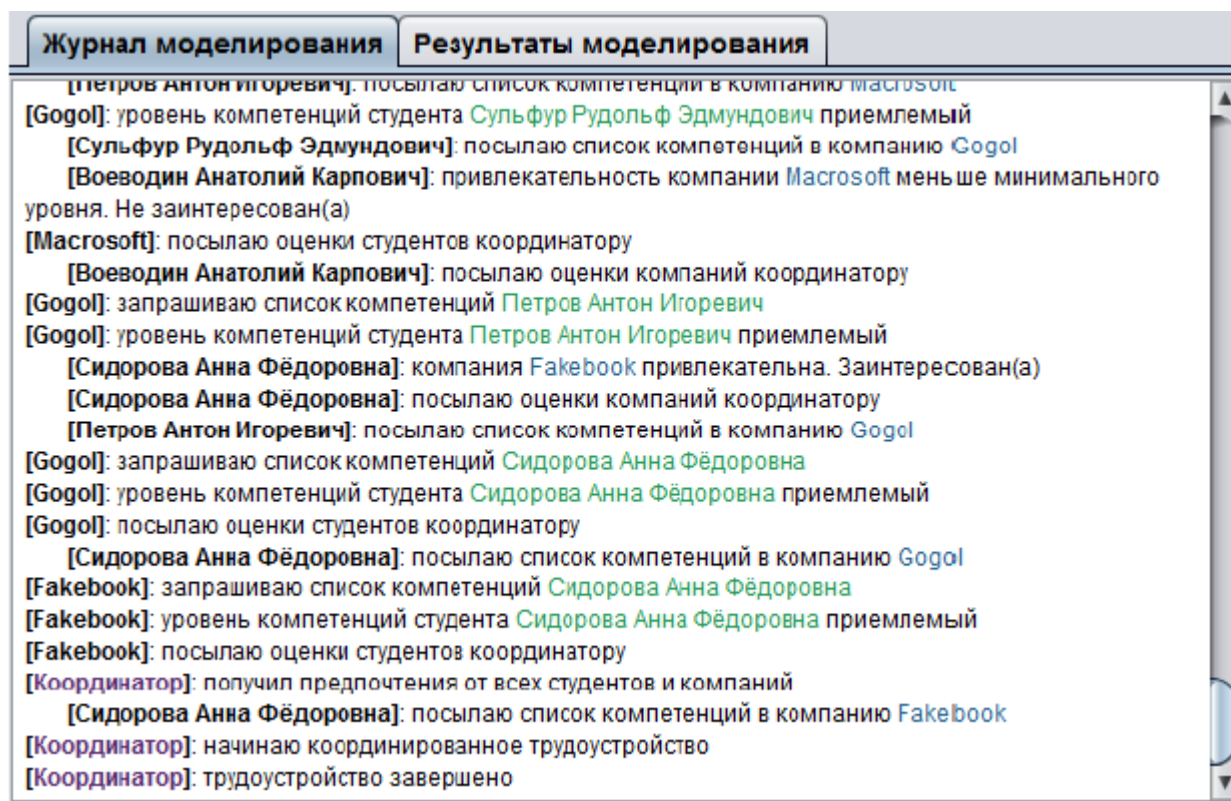


Рисунок 6 – Журнал моделирования

На рисунке 7 представлены результаты моделирования трудоустройства с помощью координированного взаимодействия, используя венгерский метод оптимизации [6].

Студенты	Fakebook (4)	Gogol (4)	Microsoft (4)
Воеводин Анатолий Карпович	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Высоких Ниагара Романовна	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Кашина Мария Захаровна	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Кондратьев Олег Семёнович	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Петров Антон Игоревич	<input type="checkbox"/>	<input type="checkbox"/>	<input checked="" type="checkbox"/>
Розенблат Аристарх Вениаминович	<input type="checkbox"/>	<input type="checkbox"/>	<input checked="" type="checkbox"/>
Сидорова Анна Фёдоровна	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>
Столяров Иван Рубанкович	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>
Сульфур Рудольф Эдмундович	<input type="checkbox"/>	<input type="checkbox"/>	<input checked="" type="checkbox"/>
Яшкина Зоя Булатовна	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>

Рисунок 7 – Результаты трудоустройства с помощью координированного взаимодействия

В итоге все студенты нашли себе рабочие места. Однако, необходимо отметить, что не было элемента соревнования, то есть везде были места и студенты беспрепятственно занимали их.

Заключение

Имитационная модель представлена в виде многоагентной системы, учитывающей распределённость, неоднородность и интеллектуальность объекта исследования. Кроме того, в статье рассмотрена постановка

задачи моделирования процесса трудоустройства студентов. Динамика процесса распределения студентов на фирмы реализована совокупностью взаимодействующих программных агентов. Для каждого типа агентов формализованы его роли и протоколы взаимодействия с другими агентами. Предложен подход к реализации сотрудничества агентов при поиске компромиссного решения, в основе которого лежит принцип гомеостатического (равновесного) управления динамикой многоагентной системы. С помощью разработанной имитационной модели проведен эксперимент, посвященный анализу процесса трудоустройства студентов в режиме координированного взаимодействия агентов.

Литература

1. Тельнов Ю.Ф., Данилов А.В., Казаков В.А. Применение многоагентной технологии для решения образовательных задач в информационно-образовательном пространстве // Инжиниринг предприятий и управление знаниями. Сб. науч. тр. 18-й научно-практической конференции (ИПиУЗ-2015, 21-24 апреля 2015 г., Москва, МЭСИ). – М.: МЭСИ, 2015.
2. Андрейчиков А.В., Андрейчикова О.Н. Интеллектуальные информационные системы: Учебник. – М.: Финансы и статистика, 2006.
3. Тарасов В.Б. От многоагентных систем к интеллектуальным организациям: философия, психология, информатика. – М.: Эдиториал УРСС, 2002. – 352 с.
4. Федяев О.И. Модель системы подготовки и трудоустройства специалистов на основе программных агентов с нейросетевой архитектурой // Пятнадцатая национальная конференция по искусственному интеллекту с международным участием КИИ-2016, Труды конференции. В 3-х томах. Т.2. – Смоленск: Универсум, 2016.
5. Programming multi-agent systems in AgentSpeak using Jason – <http://jason.sourceforge.net/wp/>
6. Железова Е., Измалков С., Сонин К., Хованская И. Теория и практика двусторонних рынков // Вопросы экономики, № 1, 2013.

Федяев О.И., Медгаус С.В. Мультиагентное имитационное моделирование системы трудоустройства студентов. Архитектура модели представлена в виде мультиагентной системы. Задача мультиагентного моделирования полностью формализована. Она представлена как задача коллективного выбора с учётом взаимных требований сторон. Эксперимент по анализу трудоустройства проведен в режиме координированного взаимодействия агентов.

Ключевые слова: агенты, мультиагентная система, имитационная модель, сотрудничество агентов, трудоустройство студентов

Fedyayev O.I., Medgaus S.V. Multi-agent simulation modeling of the student employment system. The architecture of the model is presented in the form of a multi-agent system. The task of multi-agent modeling is fully formalized. It is presented as a task of collective choice, taking into account the mutual demands of the parties. The employment analysis experiment was conducted in the mode of coordinated interaction of agents.

Key words: agents, multi-agent system, simulation model, cooperation of agents, employment of students

Современное решение проблемы распознавания лиц на основе нейронных сетей

Чернышов Б.С., Фоминых И.Д., Рудак Л.В., Федяев О.И.

Донецкий национальный технический университет

science@art9.studio, fedyaev@donntu.org

Чернышов Б.С., Фоминых И.Д., Рудак Л.В., Федяев О.И. Современное решение проблемы распознавания лиц на основе нейронных сетей. В статье описаны современные способы решения следующих проблем задачи распознавания лиц: выделение и выравнивание лица, генерация из полученного изображения векторного идентификатора лица и определение человека по идентификатору.

Ключевые слова: распознавание лиц, нейросеть, обучение, классификатор, МТСNN, датасет.

Введение

В современном мире системы распознавания лиц пользуются огромной популярностью и востребованы везде, где задействован человек. Их можно внедрять, например, в охранные системы для того, чтобы определять нежелательные лица, которые проникают на охраняемую территорию. Требования к подобной системе достаточно серьезные: нужно, чтобы всё быстро работало и при этом показывало высокую точность распознавания [1]. Именно эти показатели являются ключевыми в разработке такой системы. Кроме этого, существует и ряд проблем, с которыми сталкиваются разработчики, такие как освещенность объекта, разрешение распознаваемого изображения, ракурс и другие.

Архитектура решения

1. Каскадная нейронная сеть – для поиска лиц на фото.
2. Свёрточная НС – для генерации векторного идентификатора лица.
3. Классификатор на основе метода опорных векторов – определение человека по идентификатору лица.

Каскадная нейронная сеть

Первая проблема заключается в выделении лица человека и его последующем выравнивании. На сегодняшний день сложность её решения заключается в возможности присутствия на изображении различных ракурсов, источников освещения и перекрывающего взаимного расположения лиц.

Рассмотрев существующие программные реализации, позволяющие решить задачи выделения и выравнивания лица, мы выделили 3 подхода: Виола-Джонс в реализации OpenCV, HOG в реализации dlib и R-FCN (свёрточная каскадная нейронная сеть). На датасете Fddb выбранные алгоритмы продемонстрировали следующие результаты (рис. 1).

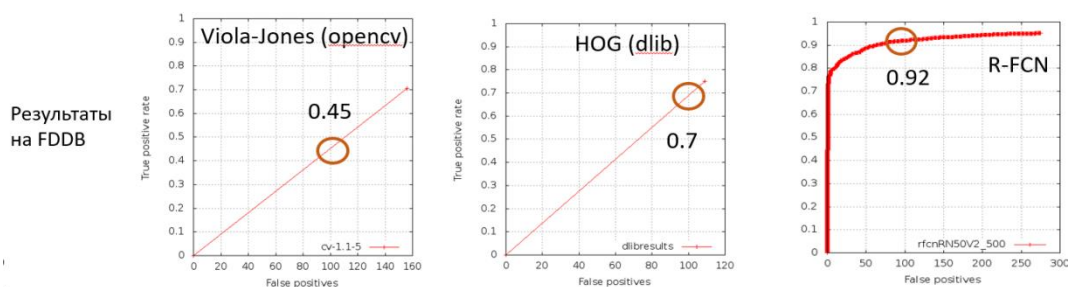


Рисунок 1 – Результаты алгоритмов на Fddb

Наилучший результат в проведённых тестах продемонстрировала R-FCN. Однако время работы составило 120 мс на GPU, что для системы, работающей в реальном времени, является низким показателем. Поэтому было принято решение найти более быструю альтернативу с похожей точностью. В результате было

обнаружено решение с использованием MTCNN, показывающая 51мс на Fddb, но слегка уступающая в точности – 90%.

Данная структура использует каскадную архитектуру с тремя этапами тщательно разработанных глубоких свёрточных нейронных сетей для прогнозирования местоположения лица и его ключевых точек. На лице выделяется 5 точек: по две – на глазах и углах губ и одна на кончике носа (см.рис. 2). Точки помогают выровнять лицо до положения анфас, что в будущем облегчит задачу распознавания лица.

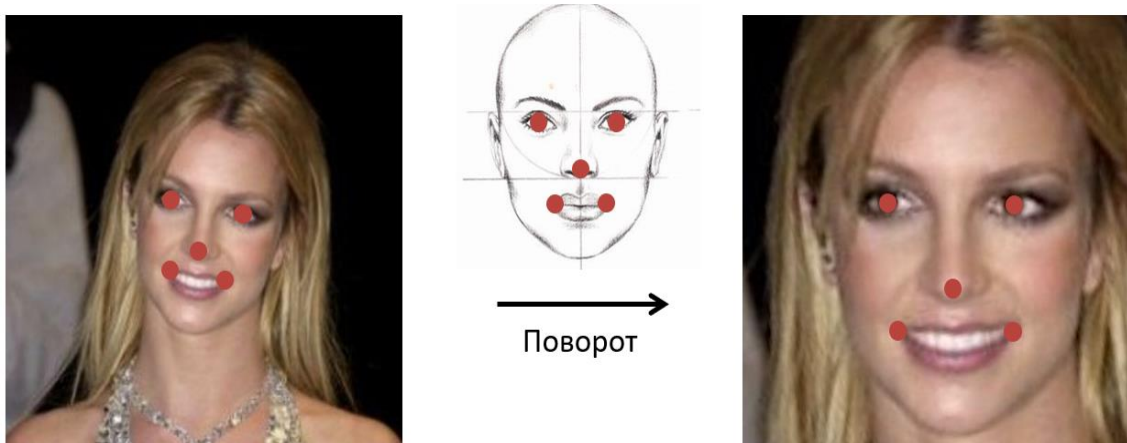


Рисунок 2 – Процесс выравнивания лица

Свёрточная нейронная сеть

Следующая проблема состоит в преобразовании выделенного лица к виду векторного идентификатора. Наиболее оптимальным средством решения проблемы являются свёрточные нейронные сети.

Свёрточная нейронная сеть — специальная архитектура искусственных нейронных сетей, предложенная Яном Лекуном в 1988 году нацеленная на эффективное распознавание изображений, входит в состав технологий глубокого обучения (англ. deep learning). Она использует некоторые особенности зрительной системы человека, в которой были открыты так называемые простые клетки, реагирующие на прямые линии под разными углами, и сложные клетки, реакция которых связана с активацией определённого набора простых клеток. Таким образом, идея свёрточных нейронных сетей заключается в чередовании свёрточных слоёв (англ. convolution layers) и субдискретизирующих слоёв (англ. subsampling layers или англ. pooling layers, слоёв подвыборки). Структура сети — однонаправленная (без обратных связей), принципиально многослойная. Для обучения используются стандартные методы, чаще всего метод обратного распространения ошибки. Функция активации нейронов (передаточная функция) — любая, по выбору исследователя.

Название архитектура сети получила из-за наличия операции свёртки, суть которой в том, что каждый фрагмент изображения умножается на матрицу (ядро) свёртки поэлементно, а результат суммируется и записывается в аналогичную позицию выходного изображения.

В обычном перцептроне, который представляет собой полносвязную нейронную сеть, каждый нейрон связан со всеми нейронами предыдущего слоя, причём каждая связь имеет свой персональный весовой коэффициент. В свёрточной нейронной сети в операции свёртки используется лишь ограниченная матрица весов небольшого размера, которую «двигают» по всему обрабатываемому слою (в самом начале — непосредственно по входному изображению), формируя после каждого сдвига сигнал активации для нейрона следующего слоя с аналогичной позицией. То есть для различных нейронов выходного слоя используются одна и та же матрица весов, которую также называют ядром свёртки. Её интерпретируют как графическое кодирование какого-либо признака, например, наличие наклонной линии под определённым углом. Тогда следующий слой, получившийся в результате операции свёртки такой матрицей весов, показывает наличие данного признака в обрабатываемом слое и её координаты, формируя так называемую карту признаков (англ. feature map). Естественно, в свёрточной нейронной сети набор весов не один, а целая гамма, кодирующая элементы изображения (например линии и дуги под разными углами). При этом такие ядра свёртки не закладываются исследователем заранее, а формируются самостоятельно путём обучения сети классическим методом обратного распространения ошибки. Проход каждым набором весов формирует свой собственный экземпляр карты признаков, делая нейронную сеть многоканальной (много независимых карт признаков на одном слое). Также следует отметить, что при переборе слоя матрицей весов её передвигают обычно не на полный шаг (размер этой матрицы), а на небольшое расстояние. Так, например, при размерности матрицы весов

5×5 её сдвигают на один или два нейрона (пикселя) вместо пяти, чтобы не «перешагнуть» искомый признак.

Операция субдискретизации (англ. subsampling, англ. pooling, также переводимая как «операция подвыборки» или операция объединения), выполняет уменьшение размерности сформированных карт признаков. В данной архитектуре сети считается, что информация о факте наличия искомого признака важнее точного знания его координат, поэтому из нескольких соседних нейронов карты признаков выбирается максимальный и принимается за один нейрон уплотнённой карты признаков меньшей размерности. За счёт данной операции, помимо ускорения дальнейших вычислений, сеть становится более инвариантной к масштабу входного изображения.

Рассмотрим типовую структуру свёрточной нейронной сети более подробно. Сеть состоит из большого количества слоёв. После начального слоя (входного изображения) сигнал проходит серию свёрточных слоёв, в которых чередуется собственно свёртка и субдискретизация (пулинг). Чередование слоёв позволяет составлять «карты признаков» из карт признаков, на каждом следующем слое карта уменьшается в размере, но увеличивается количество каналов. На практике это означает распознавания сложных иерархий признаков. Обычно после прохождения нескольких слоёв карта признаков вырождается в вектор или даже скаляр, но таких карт признаков становится сотни. На выходе свёрточных слоёв сети дополнительно устанавливают несколько слоёв полносвязной нейронной сети (перцептрон), на вход которому подаются оконечные карты признаков (см.рис.4).

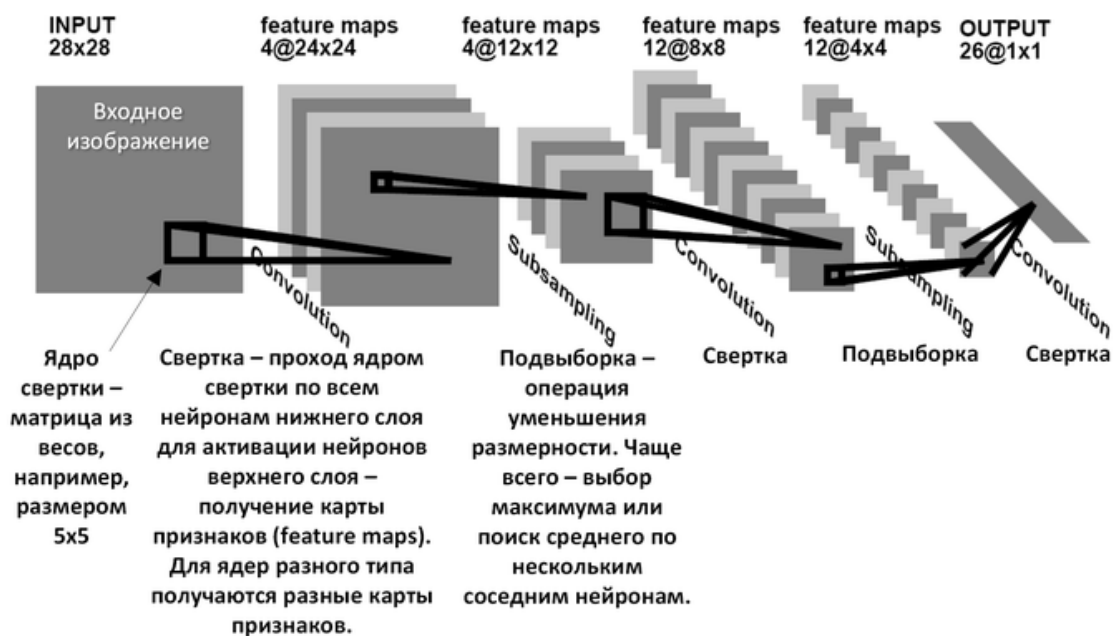


Рисунок 4 - Архитектура свёрточной нейронной сети

В итоговой реализации была использована архитектура ResNet-152 с модификациями, разработанная компанией Google. Она содержит 152 слоя, вычисляется за 11.3 млрд операций с плавающей запятой, что намного меньше, чем у конкурентов, при превосходящем количестве слоев (см.рис. 5).

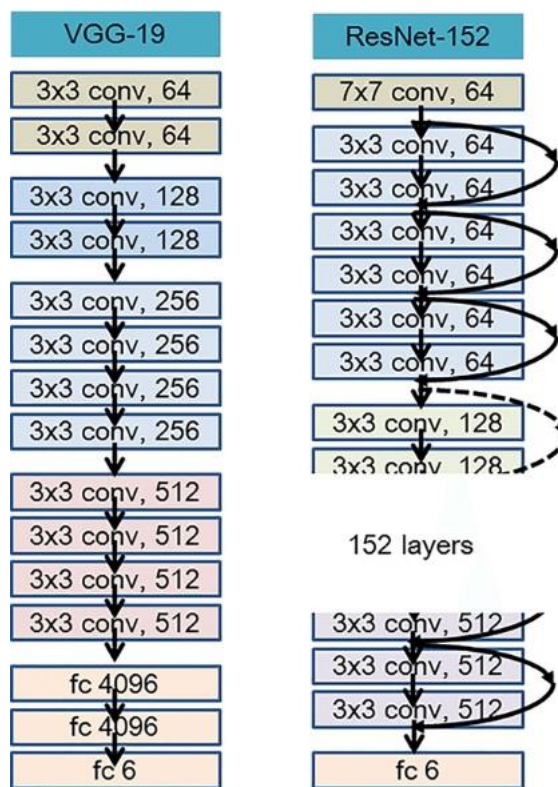


Рисунок 5 – Сравнение архитектур VGG-19 и ResNet-152

Классификация векторов

Следующей проблемой в ходе решения задачи распознавания лиц является классификация полученных в результате работы свёрточной нейронной сети векторных идентификаторов лиц. Для решения этой проблемы необходимо воспользоваться классификатором.

Классификация — один из разделов машинного обучения, посвященный решению следующей задачи. Имеется множество объектов (ситуаций), разделённых некоторым образом на классы. Задано конечное множество объектов, для которых известно, к каким классам они относятся. Это множество называется обучающей выборкой. Классовая принадлежность остальных объектов не известна. Требуется построить алгоритм, способный классифицировать произвольный объект из исходного множества.

Классификатор – алгоритм в машинном обучении, позволяющий решить задачу классификации объекта.

Среди имеющихся реализаций классификаторов наиболее подходящими для решения нашей проблемы являются: классификатор на основе метода опорных векторов (SVC) и логистическая регрессия.

Метод опорных векторов (англ. SVM, support vector machine) — набор схожих алгоритмов обучения с учителем, использующихся для задач классификации и регрессионного анализа. Принадлежит семейству линейных классификаторов и может также рассматриваться как специальный случай регуляризации по Тихонову. Особым свойством метода опорных векторов является непрерывное уменьшение эмпирической ошибки классификации и увеличение зазора, поэтому метод также известен как метод классификатора с максимальным зазором.

Основная идея метода — перевод исходных векторов в пространство более высокой размерности и поиск разделяющей гиперплоскости с максимальным зазором в этом пространстве. Две параллельных гиперплоскости строятся по обеим сторонам гиперплоскости, разделяющей классы. Разделяющей гиперплоскостью будет гиперплоскость, максимизирующая расстояние до двух параллельных гиперплоскостей. Алгоритм работает в предположении, что чем больше разница или расстояние между этими параллельными гиперплоскостями, тем меньше будет средняя ошибка классификатора.

Логистическая регрессия или логит-регрессия (англ. logit model) — это статистическая модель, используемая для предсказания вероятности возникновения некоторого события путём подгонки данных к логистической кривой (см.рис. 6).

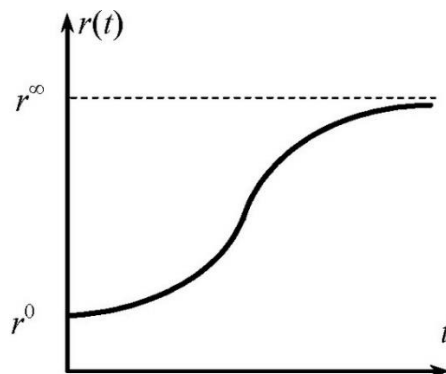


Рисунок 6 – Логистическая кривая

Среди этих двух классификаторов по показателям точности и времени работы был выбран SVC.

Вывод

В данной статье были рассмотрены наиболее современные подходы к реализации решения проблемы распознавания лиц. Также был выбран и реализован наиболее оптимальный вариант программы, использующей технологии, описанные в статье.

Литература

1. Кулябичев Ю.П. Обзор методов идентификации людей на основе изображений лиц / Пивторацкая С.В // Московский инженерно-физический институт
2. WenYi Zhao Imagebased Face Recognition Issues and Methods / Rama Chellappa. // Sarno Corporation Center for Automation Research, Washington Road University of Maryl
3. Samaria F. Face Recognition Using Hidden Markov Models // PhD thesis, Engineering Department, Cambridge University, 1994.
4. Форсайт Д.А. Компьютерное зрение: современный подход / Понс Ж.// пер. с англ. – М.: Издательский дом «Вильямс», 2004.
5. Уоссермен Ф. Нейрокомпьютерная техника: теория и практика / Ф. Уоссермен. М. // Перевод на русский язык, Ю. А. Зуев, В. А. Точенов, 1992.
6. Moghaddam B. IEEE Transactions on Pattern Analysis and Machine Intelligence. / Pentland A // 1997. V. 19. P. 696.
7. Hurieh Khalajzadeh FaceRecognition using Convolutional Neural Network and Simple Logistic Classifier. // Intelligent Systems Laboratory (ISLAB) Faculty of Electrical & Computer Engineering K.N. Toosi University of Technology, Tehran, Iran

Чернышов Б.С., Фоминых И.Д., Рудак Л.В., Федяев О.И. Современное решение проблемы распознавания лиц на основе нейронных сетей. В статье описаны современные способы решения следующих проблем задачи распознавания лиц: выделение и выравнивание лица, генерация из полученного изображения векторного идентификатора лица и определение человека по идентификатору.

Ключевые слова: распознавание лиц, нейросеть, обучение, классификатор, MTCNN, датасет.

B.S. Chernyshov, I.D. Fominykh, L.V. Rudak, O.I. Fedyaev. A modern solution of the face recognition problem based on neural networks. The article describes the modern ways of solving the following face recognition problem: selecting the face and aligning it, generating a vector identifier of the face from the resulting image, and identifying a person by the identifier.

Keywords: face recognition, neural network, training, classifier, MTCNN, dataset.

Системы распознавания изображений как средство автоматизации процессов документооборота

Шевцов Д.В.

Донецкий национальный университет
shedmival@mail.ru

Шевцов Д.В. Системы распознавания изображений как средство автоматизации процессов документооборота. В работе проанализированы современные тенденции и обоснована актуальность автоматизации процессов поиска и аналитической обработки электронных информационных ресурсов библиотек, музеев, архивных учреждений, промышленных предприятий, использующих системы электронного документооборота, конструкторских и проектных бюро. Исследованы современные средства автоматического распознавания документов электронных информационных ресурсов рассмотренных отраслей и осуществлены анализ и оценка результатов их использования для приведения сканированных документов к редактируемому формату. Изучены недостатки современных OCR-систем, снижающие эффективность автоматизации поиска и аналитической обработки электронных информационных ресурсов. Определены и обоснованы перспективные направления разработки методов и информационных технологий аналитической обработки документов электронных информационных ресурсов.

Ключевые слова: распознавание образов, электронный документооборот, электронные информационные ресурсы, цифровые изображения, знаки открытых алфавитов.

Введение

Анализ современных тенденций развития информационных технологий обосновывает актуальность автоматизации процессов поиска и аналитической обработки документов электронных информационных ресурсов. В Статье 8 Закона об информации и информационных технологиях, принятого Народным Советом Донецкой Народной Республики 07 августа 2015 года (Постановление № I-278П-НС), отмечена необходимость обеспечения права граждан на доступ к информации, накапливаемой в открытых фондах библиотек, музеев и архивов, а также в государственных информационных системах, информационных системах органов местного самоуправления и иных информационных системах, созданных или предназначенных для обеспечения физических и юридических лиц такой информацией.

Использование современных технологий распознавания зрительных образов предоставляет пользователям возможности поиска информации по содержимому файлов в форматах PDF, Jpeg, TIFF, bmp, gif, png, dib в отсканированных страницах книг, рекламных буклетов, электронных копиях подписанных договоров, презентациях, анкетным формам и других документах; автоматической индексации электронных копий документов для поиска по содержимому на Internet-серверах Google; обучения системы распознавания знакам алфавитов, изначально не присутствовавшим в базовом наборе эталонов, с целью анализа текстов, содержащих нестандартные символы.

Однако, не смотря на то, что системы приведения сканированных копий документов успешно внедрены более чем в 30 странах мира, что позволило существенно повысить эффективность работы с электронными информационными ресурсами за счет использования технологий распознавания, результаты автоматической обработки широкого класса документов (сканированных книг, содержащих математические, химические символы, знаки нестандартных алфавитов, анкет, бланков, заявок, удостоверений личности, карточек складского учета, приказов, форм, таблиц, коммерческих предложений, чертежной и проектной документации) в общем случае не соответствуют пользовательским потребностям. Данный факт обусловлен, в частности, значительным количеством ошибок, допускаемых при распознавании слабо структурированных либо неструктурированных текстов; сложностью автоматической интерпретации нестандартных, специальных и рукописных символов; низким качеством электронных копий документов, полученных в результате сканирования и, как следствие, содержащих помехи (шумы), что во многих случаях приводит либо к неверной интерпретации символов, либо делает распознавание невозможным; отсутствием единых правил анализа полнотекстовых документов; ограниченностью наборов словарей, используемых при распознавании.

Избежать указанных недостатков возможно путем разработки методов и информационных технологий обработки сканированных копий документов электронных информационных ресурсов, содержащих знаки

различных алфавитов, в том числе открытых. При этом их реализация должна осуществляться в автоматическом режиме безотносительно к априорному выбору значащих свойств, множеств имен классов и изображений знаков – объектов обучающей выборки, которые существенно ограничивают функциональность систем распознавания изображений.

1 Тенденции автоматизации поиска и аналитической обработки электронных информационных ресурсов

В настоящее время необходимость учета, обработки, анализа и хранения больших объемов информации, содержащейся в различного рода документах, является неотъемлемой частью работы практически любого учреждения [1]. В частности, согласно [2, 3], библиотеки, архивные организации и музеи, выполняющие общественно-коммуникативную функцию как базовые элементы культурной, образовательной и информационной инфраструктуры, должны эффективно обеспечивать хранение, накопление и доступ к документам и содержащейся в них информации с целью реализации конституционных прав граждан, связанных с удовлетворением информационных потребностей, свободы литературного, художественного, научного и технического творчества [3], защитой интересов субъектов авторского права и смежных прав [20].

Функционирование органов государственного управления, предприятий крупного, среднего и малого бизнеса различных форм собственности, страховых обществ, банковских учреждений, инвестиционных компаний и прочих невозможно без делопроизводства [1-3] – совокупности процессов, которые обеспечивают документирование управленческой информации и организацию работы со служебными документами. Основными задачами служб делопроизводства являются: установление единого порядка работы с документами; документационное обеспечение на основе использования современной техники; автоматизация работы с документами и сокращение количества документов; организация документооборота – движения служебных документов от даты их создания или получения до даты завершения выполнения или отправления [2].

Подготовка, формирование и сопровождение рабочей конструкторской документации, необходимой для производства разнообразного оборудования, станков, изделий, промышленных товаров; изготовление чертежей, проектной, конструкторско-технологической и эксплуатационной документации; разработка систем учета и контроля ресурсов предприятий и другие виды работ [3], выполняемых конструкторскими и проектными бюро (КПБ), предполагают необходимость создания, накопления, учета, анализа, систематизации и хранения больших объемов различных документов, таких как чертежи, схемы, графики, таблицы, отчеты и прочие [2, 3]. Решение перечисленных задач указанными бюро позволяет, в частности, промышленным предприятиям: автоматизировать технологические процессы; усовершенствовать производство; модернизировать оборудование; повышать качество, расширять ассортимент и снижать себестоимость готовой продукции; увеличивать надежность и безопасность производственных линий [4].

Эффективная организация труда сотрудников библиотек, архивных и музейных учреждений, отделов делопроизводства промышленных предприятий, КПБ на сегодняшний день предполагает внедрение и эксплуатацию современных информационных технологий [5]. Применение информационных технологий (ИТ) позволяет ставить и решать такие актуальные задачи, как:

- создание сводного каталога, единого национального депозитария, а также повышение эффективности использования и обеспечения доступности документов, хранящихся в библиотечных, архивных и музейных фондах;

- разработка информационно-поисковых систем, средств информационного обеспечения организаций и учреждений в рамках интеграции с отечественными и зарубежными научными и образовательными центрами;

- ведение, учет и систематизация документов, а также снижение временных затрат на поиск, сбор и копирование документов сотрудниками предприятий, КПБ;

- обеспечение конфиденциальности и сохранности документов и содержащейся в них информации;

- проверка подлинности документов, соблюдение авторских и смежных прав [2], и другие [3, 4].

Объективная необходимость автоматизации библиотечного, архивного и музейного дела, делопроизводства промышленных предприятий, КПБ с использованием ИТ, обусловленная многообразием задач по аналитической обработке и поиску информации [1], а также значительными затратами временных и трудовых ресурсов на их решение [2, 5], предопределила появление и развитие автоматизированных библиотечно-информационных систем (АБИС), электронных архивов документации (ЭАД), систем электронного документооборота (СЭД), систем автоматического проектирования (САПР).

Рассмотрим основные задачи и функции систем перечисленных типов.

АБИС – автоматизация и систематизация процесса комплектования и обработки фонда библиотеки, создание баз данных электронных каталогов и ресурсов, обеспечение поиска информации, а именно: каталогизация, систематизация и перевод в электронную форму библиотечных ресурсов, осуществление в автоматизированном режиме библиотечно-информационных процессов, обслуживание пользователей и обеспечение доступа к информационным ресурсам библиотеки, отслеживание библиотечных фондов.

ЭАД – упорядоченное хранение электронных копий документов, предполагающее: формирование, ведение и учет документального архивного фонда организации; безопасное хранение, оперативный доступ,

эффективный поиск и отображение архивных документов, поточное сканирование и распознавание документов, классификация дел и рубрикация документов с учетом различных форм, хранение не только карточек документов, но и соответствующих электронных образов документов (текстов/отсканированных изображений), обеспечение удаленного доступа пользователей к электронным образам дел и документов с гибкой иерархией разграничения прав.

СЭД – обеспечение процессов создания, управления доступом и распространения электронных документов, контроля над потоками документов в организации, а также: контроль отправления, передачи, получения, хранения, использования и уничтожения документов (в том числе электронных), которые выполняются с применением проверки целостности и, в случае необходимости, с подтверждением факта получения таких документов; сканирование и распознавание документов; резервное копирование, осуществление систематизации архива.

САПР – автоматизация проектной и конструкторской деятельности, направленная, в частности, на эффективное использование бумажных чертежей, планов, схем при создании новых проектов; сокращение трудоемкости проектирования и планирования; сокращение сроков, а также себестоимости проектирования и изготовления продукции; уменьшение затрат на эксплуатацию; повышение качества и технико-экономического уровня результатов проектирования; сокращение затрат на натурное моделирование и испытания.

Системы каждого из перечисленных типов получили широкое распространение в Российской Федерации и за ее пределами. Одними из наиболее известных систем, используемых для автоматизации процессов соответствующих отраслей, являются: АБИС «ИРБИС» (разработчик – Международная Ассоциация пользователей и разработчиков электронных библиотек и новых информационных технологий ЭБНИТ) и «Фолиант» (разработчик – РЦ НИТ совместно с НБ ПетрГУ); ЭАД «АРХИВНОЕ ДЕЛО» (разработчик – «ЭОС – системы») и «СПДС-АРХИВ» (разработчик – «АрхиСОФТ»); СЭД «АСКОД» (разработчик – ЦКТ «ИнфоПлюс») и ОРТИМА WorkFlow (разработчик – группа компаний ОРТИМА и ЕС/RG Data); САПР «КОМПАС» (разработчик – АО АСКОН) и «Schemagee» (разработчик – ООО «Евроинтех»).

Современные тенденции к компьютеризации информационных структур в рамках используемых средств автоматизации предопределяют создание и применение электронных информационных ресурсов (ЭИР), которые постепенно вытесняют традиционные документы на бумажных носителях [5]. Отказ от бумажных документов и переход к электронным аналогам определен множеством факторов, среди которых: физический износ и моральное устаревание бумажных носителей информации; значительные затраты времени высококвалифицированных специалистов на поиск, сбор и копирование документов; отсутствие доступа к оригиналам документов сотрудников из различных территориальных подразделений учреждений; потеря бумажных документов; затраты, связанные с переоформлением найденных договоров; трудности и расходы, обусловленные необходимостью хранения и поддержания бумажных носителей в актуальном состоянии; организация иерархии доступа к информации, содержащейся в документах; контроль за соблюдением авторских прав и другие [5-8].

Кроме того, использование бумажных носителей информации делает затруднительным решение таких актуальных задач, как

- создание и усовершенствование интеллектуальных технологий поиска и аналитической обработки информации, а также внедрение и использование программного обеспечения для информатизации библиотек, архивов и музеев;
- создание информационно-телекоммуникационных систем с высокой пропускной способностью для обеспечения доступа пользователей и постоянного копирования информации;
- автоматизированное и автоматическое ведение учета и систематизации документов, обеспечение полноты и не избыточности хранимой и обрабатываемой информации;
- осуществление автоматизированного импорта и конвертации документов.

Однако, объемы бумажных документов, участвующих в документообороте, а также специфика и формы представления информации, содержащейся в них, являются факторами, существенно снижающими уровень использования ЭИР в АБИС, СЭД, ЭАД, САПР [1-7]. Структура информационных систем библиотек, архивов, промышленных предприятий, КИБ не позволяет на сегодняшний день полностью отказаться от морально (а зачастую и физически) устаревших бумажных носителей [5], вследствие чего одной из компонент, традиционно интегрируемых в контур соответствующих средств автоматизации, являются системы сканирования и распознавания бумажных документов [7].

При этом возникают задачи, связанные с потоковым сканированием, вводом и распознаванием полнотекстовых документов; сканированием и распознаванием документов с занесением данных в систему для ускорения анализа и увеличения достоверности информации; преобразованием чертежей в векторную форму представления; конвертации форматов документов; приведения документов к редактируемому формату; индексации сканированных копий документов; эффективной организацией процессов хранения электронных документов и поиска, и иные, предполагающие аналитическую обработку ЭИР.

Решение указанных задач при разработке средств автоматизации поиска и аналитической обработки ЭИР осуществляется на основе применения современных подходов, методов, способов и соответствующих

алгоритмов распознавания зрительных образов [8, 9] к электронным копиям документов. Анализ средств автоматизации информационных процессов, имеющих место в рассмотренных организациях и учреждениях, а также литературных источников [1-7], позволил представить обобщенную функциональную схему формирования, обработки и хранения документов ЭИР, приведенную на рисунке 1.

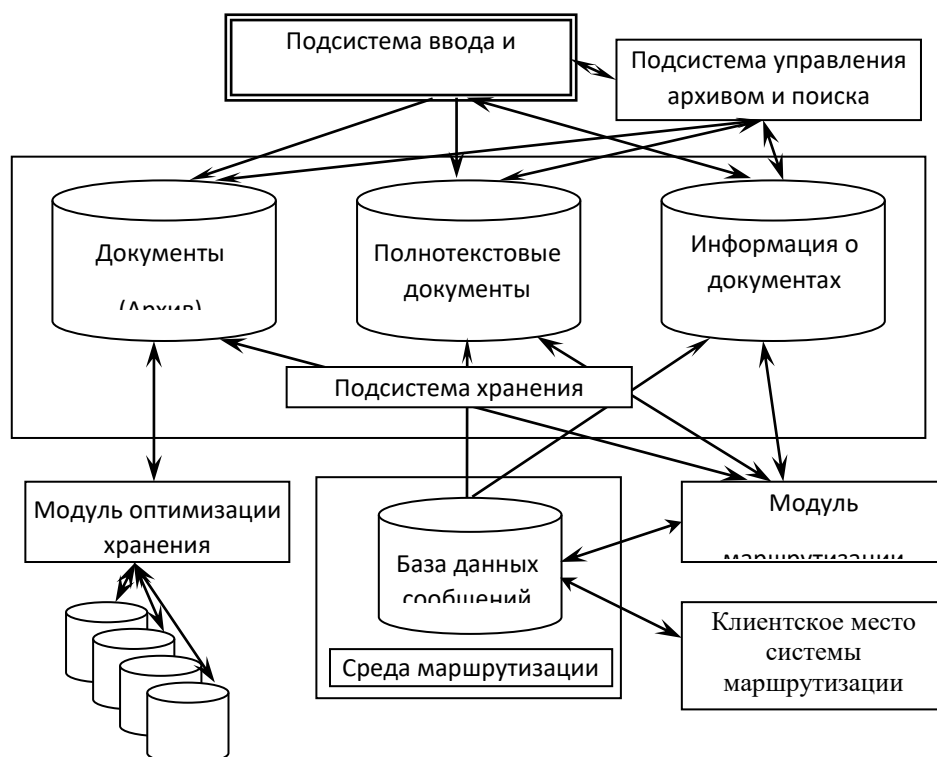


Рисунок 1 – Обобщенная функциональная схема формирования, обработки и хранения ЭИР

Одними из функций, выполняемых подсистемой ввода и распознавания документов, как показано на рисунке 1, является формирование ЭИР и преобразование сканированных документов к редактируемому формату, что позволяет в последующем вносить изменения в электронный документ; использовать содержащуюся в нем информацию для анализа, передачи, формирования отчетов; реализовывать полнотекстовый поиск; классифицировать, хранить и систематизировать данные; оптимизировать хранение информации и т.д.

Указанное преобразование является одной из наиболее сложных и трудно формализуемых задач по обработке ЭИР, поскольку многообразие форм представления данных на бумажных носителях не предполагает создания единых эффективных и унифицированных средств распознавания для произвольных типов документов [1-4].

В соответствии с этим, автоматизация процессов поиска и аналитической обработки ЭИР предполагает качественное решение задач распознавания сканированных документов различных предметных областей, присущих таким учреждениям, как библиотеки, архивы, музеи, промышленные предприятия, КПБ. Согласно [5, 7], распознавание и приведение к редактируемому формату указанных документов целесообразно осуществлять в автоматическом режиме, что позволит, в частности, без существенных финансовых затрат эффективно обеспечить систематизацию библиографических и музейных фондов, сформировать единый центр каталогизации с соответствующими справочно-информационными средствами, организовать гибкие системы поиска документов в СЭД и ЭАД, сократить ресурсы, необходимые для хранения и обслуживания архивов, а также улучшить качество и условия работы сотрудников КПБ с документацией, выполненной на бумажных носителях.

Задача распознавания и приведения к редактируемому формату сканированных документов в настоящее время решается разработчиками систем оптического (OCR – opticalcharacterrecognition) и интеллектуального (ICR – intelligentcharacterrecognition) распознавания текста. Признанными лидерами на рынке программных продуктов, реализующих ввод, потоковый ввод и распознавание сканированных документов, являются компании ABBYY и Cognitive Technologies. Такие OCR-системы, как FineReader (разработчик – компания ABBYY) и CuneiForm (разработчик – компания Cognitive Technologies), лицензированы и успешно

используются в качестве подсистем практически во всех системах автоматизации делопроизводства, в которых осуществляется преобразование ЭИР из графического в текстовый формат.

Использование OCR и ICR-технологий предоставляет пользователям при работе с различными ЭИР возможности: поиска информации по содержимому файлов в форматах PDF, Jpeg, TIFF, bmp, gif, png, dib по отсканированным страницам книг, рекламных буклетов, электронным копиям подписанных договоров, презентациям, анкетным формам и другим документам [1-4]; автоматической индексации электронных копий документов для поиска по содержимому на Internet-серверах (например, Google); обучения системы распознавания знакам алфавитов, изначально не присутствовавшим в базовом наборе эталонов, с целью анализа текстов, содержащих нестандартные символы.

Как следствие, увеличивается эффективность работы пользователей, поскольку, используя универсальный поиск, имеется возможность найти нужные для работы документы гораздо быстрее, вследствие чего больше времени уделяется решению непосредственных задач делопроизводства [2]. Кроме того, сокращается дублирование работ за счет обеспечения доступа к полному архиву документов и возможности использования подготовленных ранее документов; ускоряется необходимый доступ к полной актуализированной информации, что позволяет принимать обоснованные решения в более короткие сроки [3, 4]; сокращаются затраты на обучение сотрудников; обеспечивается доступ сотрудников из различных территориальных подразделений к оригиналам и копиям необходимых документов; снижаются издержки, связанные с обслуживанием бумажного документооборота.

Однако, не смотря на то, что системы FineReader и CuneiForm успешно внедрены более чем в 30 странах мира, что позволило существенно повысить эффективность работы с ЭИР за счет использования технологий распознавания, результаты обработки OCR и ICR-системами широкого класса документов (сканированных книг, содержащих математические, химические символы, знаки нестандартных алфавитов, анкет, бланков, заявок, удостоверений личности, карточек складского учета, приказов, форм, таблиц, коммерческих предложений, чертежной и проектной документации) в общем случае не всегда соответствуют пользовательским потребностям.

Данный факт обусловлен, в частности, значительным количеством ошибок, допускаемых при распознавании слабо структурированных либо неструктурированных текстов; сложностью автоматической интерпретации нестандартных, специальных и рукописных символов; низким качеством документов ЭИР, полученных в результате сканирования и, как следствие, содержащих помехи (шумы), что во многих случаях приводит либо к неверной интерпретации символов, либо делает распознавание невозможным; отсутствием единых правил анализа полнотекстовых документов; ограниченностью наборов словарей, используемых OCR и ICR-технологии при распознавании [8, 9].

2. Перспективные направления в разработке методов и информационных технологий автоматизации аналитической обработки документов электронных информационных ресурсов

Как следует из [3-5, 7], средствами автоматизации процессов поиска и аналитической обработки документов ЭИР библиотек, архивных и музейных организаций, предприятий и КПБ осуществляется, в частности, автоматическое моделирование и распознавание цифровых изображений, представленных отмеченной на рисунке 1 подсистемой ввода и распознавания в виде файлов одного из графических форматов.

В соответствии с этим, при проведении последующих рассуждений, не нарушая их общности, будем полагать, что указанными средствами автоматизации анализируются растровые цифровые изображения (РЦИ), сформированные, хранимые и отображаемые видеосистемой ЦЭВТ [9]. Под РЦИ понимается множество значений функции яркости $f(x,y)$, областью определения которой является множество $D(f)=X \times Y$, множеством значений – $R(f)=\{0, 1, \dots, \tilde{L}-1\}$, где $X=\{0, 1, \dots, \tilde{M}-1\}$, $Y=\{0, 1, \dots, \tilde{N}-1\}$, \tilde{M} , \tilde{N} – размеры цифрового изображения в пикселях, \tilde{L} – число уровней (градаций) яркости пикселей, определенные используемой видеосистемой [8, 9].

Как правило, РЦИ представляют в виде матрицы размера $\tilde{M} \times \tilde{N}$ следующего вида:

$$f(x,y)= \begin{pmatrix} f(0,0) & f(0,1) & \dots & f(0,\tilde{N}-1) \\ f(1,0) & f(1,1) & \dots & f(1,\tilde{N}-1) \\ \vdots & \vdots & & \vdots \\ f(\tilde{M}-1,0) & f(\tilde{M}-1,1) & \dots & f(\tilde{M}-1,\tilde{N}-1) \end{pmatrix} [92]. \quad (1)$$

При этом авторы [9] отмечают, что в ряде случаев для обозначения цифрового изображения и его элементов бывает полезно использовать более традиционную матричную запись:

$$\tilde{A} = \begin{pmatrix} a_{0,0} & a_{0,1} & \dots & a_{0,\tilde{N}-1} \\ a_{1,0} & a_{1,1} & \dots & a_{1,\tilde{N}-1} \\ \vdots & \vdots & & \vdots \\ a_{\tilde{M}-1,0} & a_{\tilde{M}-1,1} & \dots & a_{\tilde{M}-1,\tilde{N}-1} \end{pmatrix}, \quad (2)$$

где $a_{i,j} = f(x=i, y=j) = f(i, j)$, $i = \overline{0, \tilde{M}-1}$, $j = \overline{0, \tilde{N}-1}$,

С учетом введенных обозначений полагается, что формы представления (1) и (2) считаются идентичными.

В зависимости от используемого динамического диапазона функции яркости различают, в частности, полноцветные, полутоновые и бинарные РЦИ. При автоматическом моделировании и распознавании знаков большинства алфавитов анализ полноцветных и полутоновых цифровых изображений зачастую затруднителен по причине избыточности информации, представленной значениями функции яркости в формуле (1), принимаемыми в большом диапазоне.

Кроме того, использование бинарных РЦИ предполагает существенное сокращение временных и аппаратных затрат на их обработку и хранение по сравнению с иными формами представления растровых изображений [10]. Как следствие, в системах автоматического анализа РЦИ документов ЭИР осуществляется бинаризация [9] полноцветных и полутоновых цифровых изображений с использованием известных методов и алгоритмов.

Методы глобальной и локальной бинаризации, отличающиеся по качеству, скорости обработки РЦИ и прочим параметрам, характеризуются также различными недостатками [8, 9], вследствие чего априорно обосновать выбор того или иного алгоритма проблематично.

В соответствии с этим, эмпирически выбирается один из методов бинаризации, соответствующий алгоритм которого продемонстрировал наилучший, с точки зрения исследователя, результат для РЦИ знаков алфавитов, представленных в документах ЭИР, обрабатываемых системой автоматизации аналитической обработки.

Поскольку выбор метода и алгоритма бинаризации не имеет непосредственного отношения к обоснованию в данном разделе перспективных направлений разработки методов и информационных технологий автоматизации аналитической обработки документов ЭИР, при проведении дальнейших рассуждений будем полагать, что указанные документы вследствие применения к ним алгоритмов бинаризации представлены в виде бинарных РЦИ, то есть множеством значений функции яркости из формулы (1) является $R(f) = \{0, 1\}$. Соответственно, для матричной формы записи (2) цифрового изображения выполнено условие:

$a_{i,j} \in \{0, 1\}$, $i = \overline{0, \tilde{M}-1}$, $j = \overline{0, \tilde{N}-1}$.

Согласно [9], каждому из бинарных изображений знаков алфавитов, присутствующих в документах ЭИР, соответствует некоторое подмножество элементов матрицы $\tilde{A}_{\tilde{M} \times \tilde{N}}$ (2), состоящей из нулей и единиц.

На основании проведенных рассуждений, про анализе документов ЭИР необходимо осуществлять, в частности, автоматическое именование указанных подмножеств элементов матриц в формулах (1), (2) именами знаков используемого алфавита [11].

Одними из актуальных задач автоматизации аналитической обработки документов ЭИР являются моделирование и распознавание знаков открытых алфавитов [11]. В рамках последующих рассуждений будем полагать, что рассматриваемый алфавит, как и множество имен содержащихся в нем знаков, изначально представляют собой пустые множества. Данная посылка, в силу своей общности, очевидно, не противоречит выбору открытого алфавита для проведения исследований [8].

Решение указанных задач предполагает формирование (дополнение) алфавита, которое осуществляется вследствие автоматической генерации модели образа произвольного знака [11] как подмножества элементов бинарной матрицы $\tilde{A}_{\tilde{M} \times \tilde{N}}$ в формуле (2) и последующего назначения пользователем соответствующего имени.

Поскольку множество имен и алфавит, в силу сделанных предположений, пусты, априорный выбор элементов обучающего множества, элементарных составляющих математических моделей (ММ) и методов моделирования знаков не может быть осуществлен.

Рассмотрим имеющиеся предпосылки к разработке искомых методов и информационных технологий аналитической обработки документов ЭИР, предполагающих автоматическое моделирование и именование знаков открытых алфавитов без реализации стандартных этапов разработки OCR-систем, что обосновано в работах [11, 12].

Учитывая, согласно формулам (1), (2), дискретность представления видеосистемой бинарных РЦИ знаков алфавитов, формирование, обработка и анализ которых осуществляется средствами современной ЦЭВТ, а также приведенные в [11-13] рассуждения, автоматическое моделирование и именование соответствующих подмножеств элементов бинарной матрицы $\tilde{A}_{\tilde{M} \times \tilde{N}}$ целесообразно осуществлять в терминах дискретного множества A атомарных элементов (АЭ) [12].

Следовательно, согласно изложенным в работе [12] основным теоретическим положениям моделирования знаков на дискретных множествах, в рамках данной работы будем полагать, что документу ЭИР, подлежащему аналитической обработке, в частности, с целью автоматического моделирования и именования содержащихся в нем знаков, соответствует множество АЭ $A = \hat{A} \cup \check{A}$, где \hat{A} – множество активных атомарных элементов (ААЭ), \check{A} – множество пассивных АЭ.

В соответствии с гипотезой о способе формирования изображений знаков, приведенной в [14], будем полагать, что автоматическому анализу подлежат максимальные по мощности связанные подмножества ААЭ, каждое из которых соответствует отдельному знаку алфавита документов ЭИР [11] как подмножеству элементов бинарной матрицы $\tilde{A}_{\tilde{M} \times \tilde{N}}$.

При проведении дальнейших рассуждений воспользуемся определением D -знака Z , предложенным в [13]: D -знаком образа изображения на множестве A называется каждое максимальное по мощности связанное подмножество множества \hat{A} .

Согласно [9, 11, 14], D -знаки характеризуются, в частности, различной толщиной. Как отмечено в [11], при моделировании и именовании D -знаков актуальной задачей является формирование скелета [9] или остова [14] с целью генерации изображения, к которому эффективно применимы как методы векторизации [8], так и известные способы определения количественных и качественных характеристик [9].

В рамках принятой гипотезы [14] о способе формирования изображений знаков, в работах [12, 13] обосновано, что на множестве АЭ путь является образом порождающей траектории [15], вследствие чего разработано конструктивное определение ГТ-объекта как, в общем случае, множества кратчайших путей (КП), результирующих скелетизацию, и предложен способ автоматической генерации ГТ-объектов произвольных D -знаков.

Поскольку в работах [8, 9] формальное определение скелета или остова не приводится, а под соответствующим объектом понимается, как правило, результат применения того или иного алгоритма скелетизации к исходному РЦИ, согласно [14] будем полагать, что каждому D -знаку Z ставится в соответствие ГТ-объект как связанное множество КП [15]. Генерация ГТ-объектов образов произвольных бинарных РЦИ знаков осуществляется безотносительно к априорному введению морфологических операций и примитивов [11], являясь константно-независимой и позволяет осуществлять предварительную обработку исходных D -знаков с целью формирования множества конструктивно определенных элементарных составляющих модели – кратчайших путей или D -отрезков [12].

Так как ГТ-объект для произвольного D -знака Z определен конструктивно, не ограничивая общности проводимых рассуждений будем полагать, что автоматическому моделированию и анализу подлежат множества КП ГТ-объектов, результирующие скелетизацию исходных бинарных РЦИ знаков.

Поскольку в общем случае произвольный D -знак Z представим множеством ГТ-объектов [14], а также имеет место известная неоднозначность в изображении одноименных знаков, разрабатываемые методы автоматического моделирования и именования должны учитывать допустимую вариативность представления знаков алфавитов, которые, с точки зрения пользователя, соответствуют одному имени.

Для решения указанной проблемы в работах [11, 16] предложено использовать покрытие знака элементами представления для автоматического формирования ММ и последующего именования бинарных РЦИ знаков, сгенерированных отрезками прямых [12]. В рамках разработанного способа моделирования определены основные характеристики отрезков, формирующих исходный объект, а именно: количество, значения длин и углов наклонов к горизонтали отрезков, а также количественные параметры, описывающие взаимное расположение отрезков [8].

Однако, проведенные компьютерные эксперименты по практическому применению предложенного способа автоматического моделирования показали, что использование современной ЦЭВТ в качестве элементной базы для его реализации предопределяет целесообразность формирования множества ММ, сгенерированных при всевозможных значениях элемента представления и его секторов, причем размещение секторов на границе указанных элементов не должно быть фиксировано [16]. Кроме того, в некоторых случаях было установлено, что изображение, сгенерированное по автоматически сформированной модели, этой же модели не удовлетворяет, что обосновывает необходимость модификации предложенного метода с целью устранения выявленных недостатков.

В соответствии с [12], дискретный аналог отрезка, заданного аналитически в терминах свойств всюду плотных множеств, является частным случаем КП, участвующих в формировании ГТ-объектов D -знаков на множестве АЭ. Следовательно, предложенный способ автоматического моделирования [11, 16] может быть использован для формирования ММ ГТ-объектов D -знаков различных алфавитов.

Таким образом, одним из перспективных направлений исследований в области автоматического моделирования и именования образов бинарных РЦИ является разработка, на основе описанного способа, метода автоматического формирования ММ *D*-знаков, заданных в терминах свойств множества АЭ, предполагающая исследование свойств элементов представления применительно к дискретным образам анализируемых объектов как GT-объектов, результирующих их скелетизацию [15].

При этом для каждой из измеримых характеристик элементарных составляющих ММ целесообразно задавать, в общем случае, различные уровни детализации, что позволит более точно и эффективно учитывать особенности моделируемых и именуемых бинарных РЦИ знаков открытых алфавитов [11].

Согласно проведенным рассуждениям, указанными элементарными составляющими автоматически генерируемых ММ *D*-знаков являются КП или *D*-отрезки, формирующие соответствующие GT-объекты [14].

Таким образом, перспективные направления исследований по разработке методов и информационных технологий автоматизации процессов аналитической обработки документов ЭИР библиотек, музейных и архивных организаций, предприятий и КПБ предполагают, в частности, следующее:

– выбор метода бинаризации полноцветных и полутоновых растровых цифровых изображений знаков открытых алфавитов, характеризуемого наилучшими значениями показателей качества по сравнению с остальными;

– определение в терминах свойств дискретного множества АЭ измеримых характеристик и способов установления их значений для кратчайших путей и *D*-отрезков как элементарных составляющих автоматически генерируемых ММ;

– обоснование основных теоретических положений автоматического моделирования бинарных РЦИ произвольных знаков открытых алфавитов, основанного на определениях и свойствах элементов покрытия, секторов и технологии полного покрытия, сформулированных в терминах дискретного множества АЭ для КП и *D*-отрезков, являющихся элементарными составляющими формируемых ММ;

– разработку метода автоматического распознавания знаков открытых алфавитов, основанного на автоматическом моделировании и последующем именовании бинарных РЦИ, заданных в терминах свойств дискретных множеств.

Выводы

В статье исследованы современные тенденции в разработке средств автоматизации поиска и аналитической обработки документов электронных информационных ресурсов библиотек, музеев, архивных учреждений, промышленных предприятий, конструкторских, проектных бюро и прочих. Проанализированы принципы построения систем автоматизации приведения электронных копий сканированных документов к редактируемому формату. Обоснована актуальность применения для автоматизации процессов обработки электронных документов методов распознавания зрительных образов. Изучены методы и способы решения задач распознавания образов, а также причины, обуславливающие снижение значений точности распознавания. Определены перспективные направления разработки методов и информационных технологий аналитической обработки электронных сканированных документов, и изложены основные предпосылки к их реализации.

Литература

1. Кузнецов И.Н. Документационное обеспечение управления. Документооборот и делопроизводство: Учебник и практикум / И.Н. Кузнецов. - Люберцы: Юрайт, 2016. - 477 с.
2. Басаков, М.И. Делопроизводство (документационное обеспечение управления на основе ГОСТ Р 6.30-2003): Учебное пособие для студентов образовательных учреждений среднего профессионального образования / М.И. Басаков. - М.: Дашков и К, 2012. - 348 с.
3. Басовская, Е.Н. Делопроизводство: Учебное пособие / Е.Н. Басовская, Т.А. Быкова, Л.М. Вялова, Е.М. Емышева. - М.: Форум, 2012. - 256 с.
4. Грувер, М. САПР и автоматизация производства / М. Грувер, Э. Зиммерс. - Москва: Наука, 1987. - 528 с.
5. Бобылева М.П. Эффективный документооборот: от традиционного к электронному / М.П. Бобылева. - М.: Изд-во МЭИ, 2004. - 172 с.
6. Гугуева, Т.А. Конфиденциальное делопроизводство: Учебное пособие / Т.А. Гугуева. - М.: Альфа-М, НИЦ ИНФРА-М, 2012. - 192 с.
7. Чернов В.Н. Системы электронного документооборота. - М.: РАГС, 2009. - 84 с.
8. Журавлев Ю.И. Распознавание. Математические методы. Программная система. Практические применения / Ю.И. Журавлев, В.В.Рязанов, О.В. Сенько. - М.: Фазис, 2006. - 176 с.
9. Гонсалес Р. Цифровая обработка изображений. Третье издание, исправленное и дополненное: пер. с англ. / Р. Гонсалес, Р. Вудс; под общ. ред. П.А. Чочиа. - М.: Техносфера, 2012. - 1105 с.
10. Попов С.Н. Аппаратные средства мультимедиа. Видеосистема РС. / Н.С. Попов. - СПб.: ВHV-Санкт-

Петербург, 2000 - 400 с.: ил.

11. Шевцов Д.В. Обоснование перспективных направлений при проектировании систем автоматизированной обработки видеoinформации / Д.В. Шевцов // Вестник ХНТУ. – Херсон: ХНТУ, 2009 г. – № 1 (34). – С. 231-240.

12. Шевцов Д.В. Основные теоретические положения моделирования знаков изображений, подлежащих распознаванию в интеллектуальных робототехнических системах: Учебное пособие / С.В. Мышко, Д.В. Шевцов. – Донецк: ДонНУ, 2006. – 75 с.

13. Шевцов Д.В. Подпути кратчайших путей в системах автоматического опознавания знаков / Н.В. Вайсруб, С.В. Мышко, Д.В. Шевцов // Вестник Донецкого национального университета. Серия А: Естественные науки. – Донецк: ДонНУ, 2005. – № 2. – С. 403-411.

14. Шевцов Д.В. Моделируемость GT-объекта на дискретном множестве атомарных элементов / А.-В.В. Мельник, С.В. Мышко, Д.В. Шевцов // Вестник ХНТУ. – Херсон: ХНТУ, 2008 г. – № 3. – С. 223-300.

15. Мельник А.-В.В. Остов образа изображения, заданного на множестве атомарных элементов, как множество кратчайших путей / А.-В.В. Мельник, С.В. Мышко // Труды 9 международной научно-практической конференции «Современные информационные и электронные технологии» в 2 т.– Одесса, ОНПУ.– 2008. – Т.1. – С.29.

16. Шевцов Д.В. Критерии сравнения моделей GT-объектов знаков цифровых бинарных изображений / Д.В. Шевцов, С.А. Сопин // Wschodnioeuropejskie Czasopismo Naukowe (Warszawa), Том 4, № 2, 2015. – С. 87-94.

***Шевцов Д.В. Системы распознавания изображений как средство автоматизации процессов документооборота.** В работе проанализированы современные тенденции и обоснована актуальность автоматизации процессов поиска и аналитической обработки электронных информационных ресурсов библиотек, музеев, архивных учреждений, промышленных предприятий, использующих системы электронного документооборота, конструкторских и проектных бюро. Исследованы современные средства автоматического распознавания документов электронных информационных ресурсов рассмотренных отраслей и осуществлены анализ и оценка результатов их использования для приведения сканированных документов к редактируемому формату. Изучены недостатки современных OCR-систем, снижающие эффективность автоматизации поиска и аналитической обработки электронных информационных ресурсов. Определены и обоснованы перспективные направления разработки методов и информационных технологий аналитической обработки документов электронных информационных ресурсов.*

***Ключевые слова:**распознавание образов, электронный документооборот, электронные информационные ресурсы, цифровые изображения, знаки открытых алфавитов.*

***Shevtsov Dmitry. Image recognition systems as a means of automation of document management processes.** The article analyzes current trends and substantiates the relevance of automation of search and analytical processing of electronic information resources of libraries, museums, archival institutions, industrial enterprises using electronic document management systems, design and design bureaus. Modern means of automatic recognition of documents of electronic information resources of the considered branches are investigated and the analysis and an assessment of results of their use for reduction of the scanned documents to the edited format are carried out. The shortcomings of modern OCR systems that reduce the efficiency of search automation and analytical processing of electronic information resources are studied. The perspective directions of development of methods and information technologies of analytical processing of documents of electronic information resources are defined and proved.*

***Key words:** pattern recognition, electronic document management, electronic information resources, digital images, open alphabet characters.*

СЕКЦИЯ 4. «ПРИКЛАДНЫЕ ИНТЕЛЛЕКТУАЛЬНЫЕ СИСТЕМЫ УПРАВЛЕНИЯ, ПРОЕКТИРОВАНИЯ И ПРИНЯТИЯ РЕШЕНИЙ»

УДК 004.514

Обзор проблем реализации человеко-машинного интерфейса в системах SCADA

Савкова Д.Г.
ИнСАТ, Москва
das.savkova@gmail.com

Савкова Д.Г. Обзор проблем реализации человеко-машинного интерфейса в системах SCADA. В данной статье описана базовая схема работы автоматизированного объекта, рассмотрены особенности человеко-машинного интерфейса на разных устройствах, перечислены проблемы, с которыми может столкнуться инженер при автоматизации объекта и оператор при использовании уже налаженной системы

Ключевые слова: человеко-машинный интерфейс, SCADA, автоматизация, браузер, АРМ, мобильное устройство.

Введение

Scada-система – это система сбора данных, контроля и управления процессами, предназначенная для комплексной автоматизации процессов. Система позволяет создать удобный операторский интерфейс и исполняет функции контроля и управления процессами в реальном времени.

Современный человеко-машинный интерфейс – это совокупность экранов, графических дисплеев, клавиатур, переключателей, сенсорных устройств и т.п., используемых оператором для взаимодействия со SCADA-системой. Плохо спроектированный человеко-машинный интерфейс может привести к снижению производительности оператора, увеличению числа ошибок, пропуску критической информации и большим убыткам в производстве.

С развитием интернета серверная часть программного обеспечения начала переходить на облачные технологии для повышения доступности и надёжности, а визуализация клиентской части – на html 5 [1]. Даже без интернета Web-визуализация имеет множество преимуществ, основным из которых, является кроссплатформенность. Для просмотра мнемосхемы и управления технологическими процессами достаточно иметь установленный браузер на устройстве. Существуют и текстовые браузеры (Lynx), позволяющие использовать Web-технологии даже на устройствах с очень скромными техническими характеристиками и текстовым дисплеем.

Операторы могут управлять технологическими процессами не только со своих автоматизированных рабочих мест, но и с планшетов, смартфонов, сенсорных панелей. Такое разнообразие устройств усложняет требования к человеко-машинному интерфейсу и к системе в целом.

Общая схема автоматизации объекта

Рассмотрим общую схему программно-аппаратного комплекса в автоматизации (рис. 1). Проект создается в инструментальной среде инженерами по автоматизации. Инструментальная среда в себя включает:

9. драйверы для оборудования;
10. сервер с БД;
11. инструменты для организации связи между узлами;
12. конструктор графического интерфейса;
13. поддержку языков программирования для ПЛК [2] (FBD, ST, LD и т. п.);
14. средства отладки.

Конфигурация проекта загружается на микроконтроллер, запускается сервер и веб-сервер. К входам и выходам микроконтроллера подключаются внешние устройства: индикаторы, кнопки, датчики, измерительные приборы, дополнительная память. Сервер заносит показания датчиков в базу данных. Конечные пользователи подключаются к веб-серверу с помощью браузера со своих ПК и мобильных устройств. Сервер следит за актуальностью данных как для веб-представления, так и для внешних подключенных устройств. Оператор

может не только следить за данными, но и осуществлять управление. От удобства интерфейса зависит как быстро он заметит критическую ситуацию и своими действиями ее нормализует.

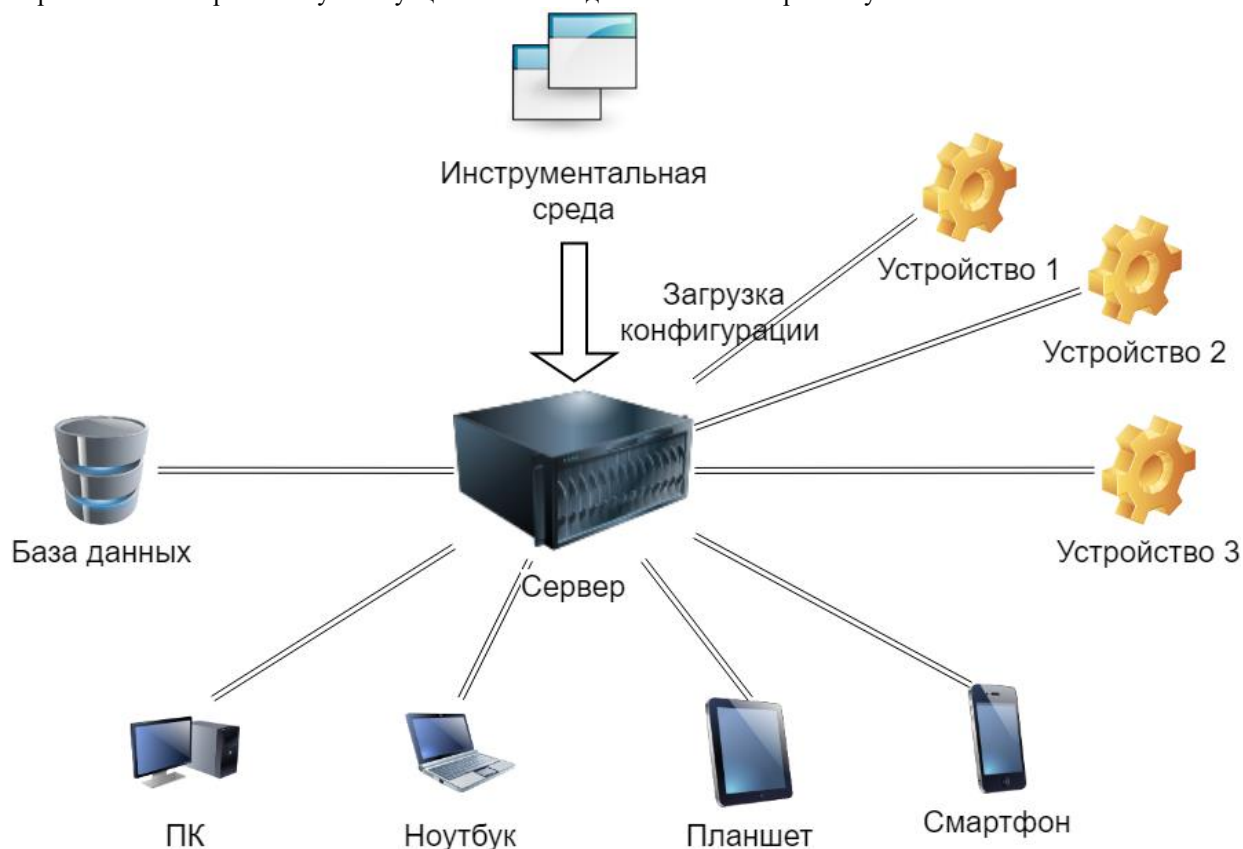


Рисунок 1 – Схема программно-аппаратного комплекса

Интерфейс для АРМ

Автоматизированное рабочее место [3], оснащённое несколькими мониторами, может показывать проект целиком, со всеми подробностями и даже элементами декора (см. рис. 1). Экран АРМа по возможности должен быть не только максимально практичным, но и эстетически привлекательным для оператора. SVG-элементы схемы, стрелочные приборы, слайдеры, журналы и тренды на весь экран – всё это допустимо при проектировании дизайна интерфейса для большого экрана стационарного рабочего места. Разрешающая способность монитора в совокупности с ресурсами ПК или некоторого другого оборудования позволяют выводить большие и подробные мнемосхемы на одном экране без необходимости навигации между окнами.



Рисунок 2 – Мнемосхема на мониторах

Интерфейс для планшетов

На планшете использование окон с множеством мелких элементов неудобно, т.к. затрудняет навигацию. Чтобы вместить больше полезной информации на экране планшета, мнемосхема для стационарного ПК разбивается на несколько окон меньшего размера и заполненности (см. рис. 2). На рисунке видно, что у окон появилась навигация, которая на больших экранах не была необходима. В журнале было сокращено число столбцов, остались только самые важные. Размер окна можно масштабировать жестами, однако при необходимости можно такое масштабирование запретить.

Ввиду аппаратных ограничений на мобильных устройствах (размер дисплея, вычислительные мощности), окна должны быть не перегружены информацией и иметь достаточно большие относительно размеров экрана кнопки для удобного нажатия. Несмотря на то, что мощности современных телефонов и планшетов сопоставимы с настольным ПК – питание от батареи по-прежнему требует экономного расходования ресурсов. Слишком «тяжелый» интерфейс значительно ограничивает мобильность, т.к. устройство приходится заряжать несколько раз за смену. Также следует учитывать, что события на сенсорных экранах отличаются от событий мыши. Сенсорный экран открывает новые возможности, такие как мультисенсорное масштабирование, навигация с помощью жестов, но исключает использование любых событий, связанных с передвижением мыши, таких как подсвечивание элементов и всплывающие подсказки. Ввод данных с всплывающей виртуальной клавиатуры будет не таким удобным, как с физической.

Восстановление	Событие	Время квитирования
12-11-2017 14:06:09	Ускорение ТНЭП при включении выключателя	
12-11-2017 14:05:54	Ускорение ТО при включении выключателя	
12-11-2017 14:05:46	Неисправность цепей управления	12-11-2017 14:05:50
12-11-2017 14:05:45	Неисправность цепей управления	
12-11-2017 14:05:44	Неисправность цепей управления	
12-11-2017 14:05:34	Пружина не заведена	12-11-2017 14:05:37
12-11-2017 14:05:26	Блокировка включения и отключения	12-11-2017 14:05:30
12-11-2017 14:05:20	Неисправность обогрева	
12-11-2017 14:05:13	Неисправность обогрева	
12-11-2017 14:05:12	Неисправность обогрева	
12-11-2017 14:04:59	Неисправность цепей оптертока	12-11-2017 14:05:05
12-11-2017 14:04:41	Сработала защита от непереключения фаз	12-11-2017 14:04:54
12-11-2017 14:02:49	Срабатывание РТО	
12-11-2017 14:02:41	Фиксация положения выключателя	
12-11-2017 14:02:39	Фиксация положения выключателя	
12-11-2017 14:02:24	Ускорение ТНЭП при включении выключателя	12-11-2017 14:02:27
12-11-2017 14:02:17	Пружина не заведена	
12-11-2017 14:01:41	Пружина не заведена	
12-11-2017 14:01:33	Неисправность обогрева	
12-11-2017 14:01:09	Неисправность обогрева	
12-11-2017 14:00:49	Сработала защита от непереключения фаз	
12-11-2017 14:00:41	Сработала защита от непереключения фаз	
12-11-2017 14:00:24	Неисправность выключателя	
12-11-2017 14:00:22	Неисправность цепей управления	
12-11-2017 13:59:58	Пружина не заведена	12-11-2017 14:00:04
12-11-2017 13:59:45	Блокировка включения и отключения	12-11-2017 13:59:46
12-11-2017 13:59:30	Неисправность обогрева	12-11-2017 13:59:32
12-11-2017 13:59:25	Низкое давление элгзада	
12-11-2017 13:59:16	Низкое давление элгзада	12-11-2017 13:59:24
12-11-2017 13:59:11	Неисправность цепей оптертока	12-11-2017 13:59:12
12-11-2017 13:59:06	Сработала защита от непереключения фаз	12-11-2017 14:00:35
12-11-2017 13:46:12	Сработала защита от непереключения фаз	
12-11-2017 13:45:58	Неисправность цепей управления	
12-11-2017 13:45:43	Неисправность цепей оптертока	
12-11-2017 13:45:24	Неисправность цепей оптертока	
12-11-2017 13:44:45	Срабатывание 4 ступени МТЗ	
12-11-2017 13:44:38	Ускорение ТНЭП при включении выключателя	
12-11-2017 13:44:23	Неисправность выключателя	
12-11-2017 13:44:18	Низкое давление элгзада	
12-11-2017 13:43:25	Низкое давление элгзада	12-11-2017 13:43:27
12-11-2017 13:43:15	Неисправность цепей оптертока	
12-11-2017 13:43:05	Сработала защита от непереключения фаз	
12-11-2017 13:38:39	Сработала защита от непереключения фаз	
12-11-2017 13:38:34	Сработала защита от непереключения фаз	

Температура наружного воздуха (северная сторона)	6.42 °C
ОПУ. Температура в помещении РЦ	25.68 °C
ОПУ. Температура в серверной	20.47 °C
ОПУ. Температура в помещении оперативного персонала	24.24 °C
ОПУ. Температура в помещении связи	24.03 °C
ОПУ. Температура в помещении 1сш РУСН	19.57 °C
ОПУ. Температура в помещении 2сш РУСН	17.05 °C
ЗРУ 10кВ. Температура в помещении 3сек. ЗРУ	11.47 °C
ЗРУ 10кВ. Температура в помещении 4сек. ЗРУ	14.46 °C
Температура в помещении насосной	9.35 °C
Температура в помещении очистных	14.47 °C

Температура	Журнал	График	Схема	Состояния
Температура	Журнал	График	Схема	Состояния

Рисунок 3 – Мнемосхема на планшете

Интерфейс для смартфонов

Экран смартфона по размеру ещё меньше планшета, поэтому возникла необходимость оставить только основную информацию (см. рис. 3). Обычно мобильными устройствами пользуются операторы, которые не привязаны к одному месту [4], а постоянно передвигаются между объектами. В таком случае их не интересуют информационные сообщения и красота дизайна, им важны только тревоги с высоким приоритетом, при которых нужно срочно принимать меры. Также на смартфонах нет разъема для проводного интернета, поэтому для соединения с сетью будет использоваться Wi-Fi или мобильный интернет. Трафик в таком случае будет



иметь значение, как и объем загружаемых данных, необходимый для работы.

Рисунок 4 – Мнемосхема на смартфоне

Решение проектировать систему с разделением представления на полную и мобильную версию принимает инженер. Но от программиста зависит на сколько хорошо проект будет работать на разных устройствах и в разных условиях.

Современная реализация ЧМИ на основе веб-технологий должна быть гибкой и универсальной, на основании следующих критериев:

- быстрой в загрузке и обновлении конфигурации / развёртывании, чтобы в случае непредвиденных ситуаций / замены оборудования оперативно восстановить рабочий режим (время простоя в исключительных ситуациях особенно критично для бизнеса);
- быстрой в рабочем режиме для своевременного отклика оператора на тревогу;
- надёжной, отображать только актуальную и правдивую информацию (при искажении данных производство может остановиться);
- потреблять стабильный объём оперативной памяти, так как в случае утечек и превышения лимита браузер прекращает выполнение скриптов;
- занимать разумный объём места на диске, экономно потреблять трафик (на счету каждый мегабайт и память у микроконтроллера не бесконечна).

Заключение

Учитывая богатое разнообразие условий эксплуатации этим критериям может быть выдан различный приоритет, но проще всего рассмотреть наихудший вариант — оператора со старым почти разряженным

телефоном на объекте в глубинке с плохой сотовой связью. Ему будет необходим максимально простой и надёжный вариант с минимальным достаточным интерфейсом.

При адаптации мнемосхем под различные экраны увеличивается срок разработки, но управление на каждом отдельном устройстве становится более удобным.

W3C делает все шаги в этом направлении, постоянно дополняя стандарт с учётом особенностей новых устройств. Для использования возможностей различных мобильных устройств уже не нужно писать для них отдельные приложения, это работает и из браузера, причём с небольшими потерями в производительности, т. к. JavaScript движки постоянно обновляются, да и сам стандарт ECMAScript не стоит на месте. Однако, это не отменяет необходимость учитывать особенности устройств при разработке конкретных интерфейсов.

Разработка SCADA систем не стоит на месте, стараясь не отставать от современного Интернета с его множеством сервисов, карт, распознавания и генерации речи, красочной анимированной графикой. Но всё это нужно реализовывать очень гибко, чтобы основополагающие функции мониторинга и управления, ради которых предприятия автоматизируются, не пострадали.

Литература

1. MDNWebdocs. HTML5 [Электронный ресурс] : сайт. - Электрон. дан.. 2018. – Режим доступа: <https://developer.mozilla.org/ru/docs/HTML/HTML5>. - Загл. с экрана.
2. Национальная библиотека им. Баумана. IEC 61131-3 [Электронный ресурс] : сайт. - Электрон. дан. - Москва. – 2016. - Режим доступа: https://ru.bmstu.wiki/IEC_61131-3. - Загл. с экрана.
3. АСУ ТП. Scada [Электронный ресурс] : сайт. - Электрон. дан.. – 2018. - Режим доступа: <https://automation-system.ru/main/62-about-scada/scada.html>. - Загл. с экрана.
4. Master дистанционного управления [Электронный ресурс] / В. В. Хлюдинский, Г. Л. Веселуха // Информатизация и системы управления в промышленности. - 2018. - № 3(75) . – Режим доступа: <http://isup.ru/articles/2/13043/>. - Загл. с экрана.
5. AboutW3C [Электронный ресурс] : сайт. - Электрон. дан.. – 2018. - Режим доступа: <https://www.w3.org/Consortium/>. - Загл. с экрана.

Савкова Д.Г. Обзор проблем реализации человеко-машинного интерфейса в системах SCADA.
В данной статье описана базовая схема работы автоматизированного объекта, рассмотрены особенности человеко-машинного интерфейса на разных устройствах, перечислены проблемы, с которыми может столкнуться инженер при автоматизации объекта и оператор при использовании уже налаженной системы

Ключевые слова: *человеко-машинный интерфейс, SCADA, автоматизация, графический интерфейс, браузер, АРМ, мобильное устройство.*

Savkova Daria. Survey of HMI implementation problems in SCADA systems.
This article describes the basic operation of an automated object, discusses the features of the human-machine interface on different devices, enumerates automation and supervising problems

Key words: *human-machine interface, SCADA, automation, GUI, browser, automated workplace, mobile device.*

Применение ERP-технологий в стратегическом управлении предприятием

Гвоздюкова С.Н., Суворова Е.Ю.

Луганский национальный университет имени Тараса Шевченко
hvozdiukova@mail.ru, suvorova.itstep@gmail.com

Гвоздюкова С.Н., Суворова Е.Ю. Применение ERP-технологий в стратегическом управлении предприятием. В данной статье ставится задача рассмотреть основные возможности новых информационных технологий в управлении организацией. В частности, рассматривается классификация ERP-систем и особенности их внедрения. На основе проведенного анализа выделены позитивные и негативные стороны внедрения информационных систем, а также представлена концепция уровней ответственности системы управления при внедрении ERP-технологии.

Ключевые слова: информационная система, управление, SEM-система, ERP-система, технология, программное обеспечение.

Введение

В современных экономических условиях, когда многие предприятия вынуждены балансировать между рентабельностью и увеличивающейся конкуренцией, убытки предприятия от не эффективного управления могут стать особенно опасными, и, как правило, приводят к банкротству. Особенно это касается крупных предприятий, в частности холдингов, где управление рассредоточено между многими отделами и дочерними предприятиями.

За прошедшее десятилетие многие руководители холдингов осознали необходимость системного подхода к управлению, заменяя командные методы более современными технологиями, основанными на внедрении информационных систем. Системный подход, основанный на иерархии информационных технологий, позволит при наличии большого объема взаимосвязанной информации, получить прочную основу для принятия управленческих решений, особенно это касается рискованных операций.

В начале эпохи промышленного использования информационных технологий процесс автоматизации предприятий сводился к тому, что различные его службы покупали программы, решающие отдельные задачи, у разных компаний. Использование отдельных модулей автоматизации разных процессов не могло составить целостной картины происходящего на предприятии. В связи с этим в последнее десятилетие наметился высокий спрос на информационные системы, претендующие на роль управленческих, то есть системы, которые интегрируют всю информацию, способные проанализировать ее и помочь в принятии решений.

Проблемы автоматизации управления предприятием рассматривались в научных публикациях отечественных и зарубежных ученых и специалистов по информационным технологиям: Ф.Ф. Бутынец [1], А.В. Клименко [2], В.Д. Шквир [3] и др. Но, несмотря на такое разнообразие изучения и разработки подходов к автоматизации управления предприятием, недостаточное количество теоретиков и практиков рассматривают вопрос о сущности управленческих информационных систем и унификации их классификационных признаков.

Если сравнивать значение канала стратегической информации и информационных технологий, то последние открывают дополнительные преимущества, а именно: позволяют разрабатывать новые стратегии, создавая совершенно новые источники конкурентного преимущества, способствуют многочисленным новшествам в сфере товаров и услуг. В тоже время почти все компьютерные системы сегодня можно рассматривать как формы поддержки управления.

Изложение основного материала

Термином «управленческие информационные системы» традиционно называют системы сообщений и составления отчетов, которые отличаются от «интеллектуальных» систем. Они выполняют стандартные бизнес-процессы, бухгалтерский учет, финансовые отчеты, платежные сведения, размещения заказа и т.д. По сравнению с традиционными неавтоматизированными методами эти системы дешевые, эффективные и широко распространены. Сейчас они стали основной частью любого бизнес-плана. Чтобы оставаться конкурентоспособным, бизнес нуждается в эффективной управленческой информационной системе, которая может предоставлять информацию, необходимую определенным людям в определенное время.

Проанализируем определение понятия «управленческая информационная система» в трудах отечественных ученых. В.Д. Шквир водит понятие «экономической информационной системы». Под ним понимают систему, предназначенную для хранения, поиска и выдачи экономической информации по запросу пользователей [3].

Ф.Ф. Бутынец приводит определение компьютерной информационной системы предприятия, под которой понимает совокупность экономико-математических методов и моделей, технических, программных, технологических средств и решений, а также специалистов, предназначенной для обработки информации и принятия управленческих решений [1].

Почти такое же определение приводит О.В. Клименко: «компьютерная информационная система предприятия – совокупность информации, методов, моделей, технических, программных, технологических средств и решений, а также специалистов, которые занимаются обработкой информации и принятием управленческих решений в пределах предприятия» [2].

Подытожив сказанное, можно отметить, что управленческие информационные системы разработаны для обеспечения корпоративных руководителей и старших менеджеров, информация в рамках информационной системы им действительно необходима для мониторинга, контроля, принятия решений и администрирования. В коммерческих организациях эти системы обычно добавляют к существующим организационным системам, они предоставляют руководителям возможность прямого доступа ко всей финансовой и коммерческой информации, а также к стратегическому планированию.

Управленческие информационные системы дают руководителям возможность систематически искать или изучать более детальные элементы внутренней или внешней информации, отслеживать тенденции или просто просматривать файлы для создания общей картины текущей ситуации. Управленческие информационные системы проектируют таким образом, чтобы они были особенно удобными в использовании. Некоторые системы позволяют пользователю диктовать сообщения на экран и посылать голосовую почту. Разработки ведутся в направлении гипертекстовой коммуникации, которая сочетает текст с голосом и видеоизображением.

Сейчас стремительно развиваются системы планирования ресурсов предприятия (ERP), которые представляют собой новое поколение информационных систем, разработанных для решения текущих задач бизнеса с учетом множественных функциональных областей крупных корпоративных предприятий. Системы ERP обеспечивают крупные организации высоко интегрированными решениями, опирающимися на использование общих баз данных. То есть все внутренние и внешние связи независимо от функциональной деятельности предприятия хранятся в единой базе. Таким образом, данные о деловых операциях вводятся в систему один раз, влияние этих операций на бизнес-процессы можно оценить немедленно, а соответствующие отчеты получить сразу.

Большая часть современных ERP-систем основана на архитектуре клиент-сервер, при этом используется общедоступная технология реляционных баз данных как для управления данными, так и для генерации отчетов.

В таблице 1 представлен набор функций и их характеристики, которые можно считать типичными для программных продуктов класса ERP.

Таблица 1 - Функции ERP-систем

Функции	Характеристика
Управление проектами	Планирование заданий проекта и ресурсов, необходимых для его реализации
Ведение конструкторских и технологических спецификаций	Определение состава конечного продукта, а также материальных ресурсов и операций, необходимых для его изготовления
Управление спросом и формирование планов продаж и производства	Прогнозирование спроса и планирования выпуска продукции
Планирование потребностей в материалах	Определение объемов всех видов материальных ресурсов (сырья, материалов, комплектующих), необходимых для использования производственного плана, а также сроков поставок, разметов партий и прочее
Управление запасами и закупочной деятельностью	Организация ведения договоров, реализация схемы закупок, обеспечение учета и оптимизация складских запасов
Планирование производственных мощностей	Контроль наличия их загруженности, включая укрупнение планирования и детальное планирование, вплоть до рабочих мест
Финансовые функции	Финансовый учет, управленческий учет, оперативное управление финансами

Пересылка данных при этом осуществляется с помощью стандартных коммуникационных протоколов. Открытая архитектура систем ERP означает, что ERP может работать на различных платформах, что позволяет устанавливать систему где угодно, не заботясь о структурах и географическом положении отдельных участников бизнеса.

Как правило, ERP-системы ориентированы на одну или несколько отраслей и типов производства: серийное (машиностроение), мелкосерийное (авиация, тяжелое машиностроение), дискретное (металлургия, химия), непрерывное (нефте- и газодобычи). По масштабам ERP-системы принято относить к большим интегрированным информационным системам управления, поскольку они обеспечивают полную функциональность для большого количества пользователей.

мм таблице 2 представлены основные преимущества и недостатки использования ERP-систем на предприятии.

Таблица 2 –Преимущества и недостатки использования ERP-систем в управлении предприятием

Преимущества	Недостатки
<ul style="list-style-type: none"> - помогают снизить расходы на внутренние операции; - заставляют компанию оптимизировать бизнес-процессы; - хранения данных в одной, общей для предприятия, базе позволяет проводить целый ряд аналитических исследований; - намного упрощается задача построения хранилищ данных, когда данные могут быть получены от одной системы; - значительная экономия при управлении операционными системами и центрами данных, достигается путем введение единых стандартов во всех отделах организации. 	адаптация систем ERP требует переоценки, а возможно, и полной перестройки бизнес-процессов, часто требует другой квалификации кадров и преобразований в организационной структуре.

Анализируя таблицу не возможно не подчеркнуть, что преимуществ от внедрения ERP-системы, несомненно, больше.

Информационные системы можно классифицировать по различным признакам: по сфере деятельности, по уровню автоматизации процессов управления, по степени централизации обработки информации, по степени интеграции функций [3]. До недавнего времени в литературе можно было встретить и классификацию управленческих информационных систем по типу используемых вычислительных средств.

А.В. Клименко автоматизированные системы управления предприятием распределяет на три большие группы. К системам первой группы относят простые, так называемые «коробочные», продукты, реализующие небольшое число бизнес-процессов организации. Ко второй группе относят системы среднего класса, которые отличаются большей глубиной и широтой охвата функций. К высшему классу относят системы, которые отличаются высоким уровнем детализации хозяйственной деятельности предприятия [2].

Для целей выбора можно обобщить основные характеристики современных управленческих информационных систем, которые применяют для управления предприятиями в странах СНГ (табл.3).

Таблица 3 – Основные характеристики управленческих информационных систем

Характеристика	Локальные системы	Интегрированные системы		
		Малые	Средние	Крупные
Название системы	1С:Бухгалтерия, Бэст, Акцент, Парус-Бухучет	1С: Предприятие, Галактика, Инфософт,Scala, Concorde	Platinum	SAP R/3
Сложность внедрения	1-2 месяца	По этапам, 4 месяца и больше	По этапам, 6-9 месяцев	Сложное по этапам, 9-12 месяцев
Соотношение расходов (лицензия на управленческую информационную систему / внедрение / оборудование и системное программное обеспечение)	1 / 0,5 / 2	1 / 1 / 1	1 / 2 / 1	1 / 1-5 / 1

Впервые понятие организации основанной на информации было предложено П. Друкером [4]. Данный тип предприятия устанавливает интегрированные в масштабе всего предприятия операционные среды, которые ориентированы как на индивидуальных пользователей, так и на рабочие группы. В подобном типе предприятий ERP-системы используют как каркас для реорганизации работы, учетно-аналитических подразделений (рис.1).

	1980 гг. Вертикальные (функциональные) структуры	1990 гг. Глобальные матричные структуры	2000 гг. Глобальные сетевые структуры
Обработка транзакций	Традиционные системы	Интегрированные ERP-системы	Гибридные системы
Направление управления	Управление доходами и	Управление, ориентированное	Управление сетью

Рисунок 1 – Процесс перехода от вертикально-интегрированных организаций к сетевым и функционально интегрированным бизнес-структурам

Рассмотрим на конкретном примере механизм перечисления, реализованный в SAP (рис. 2).

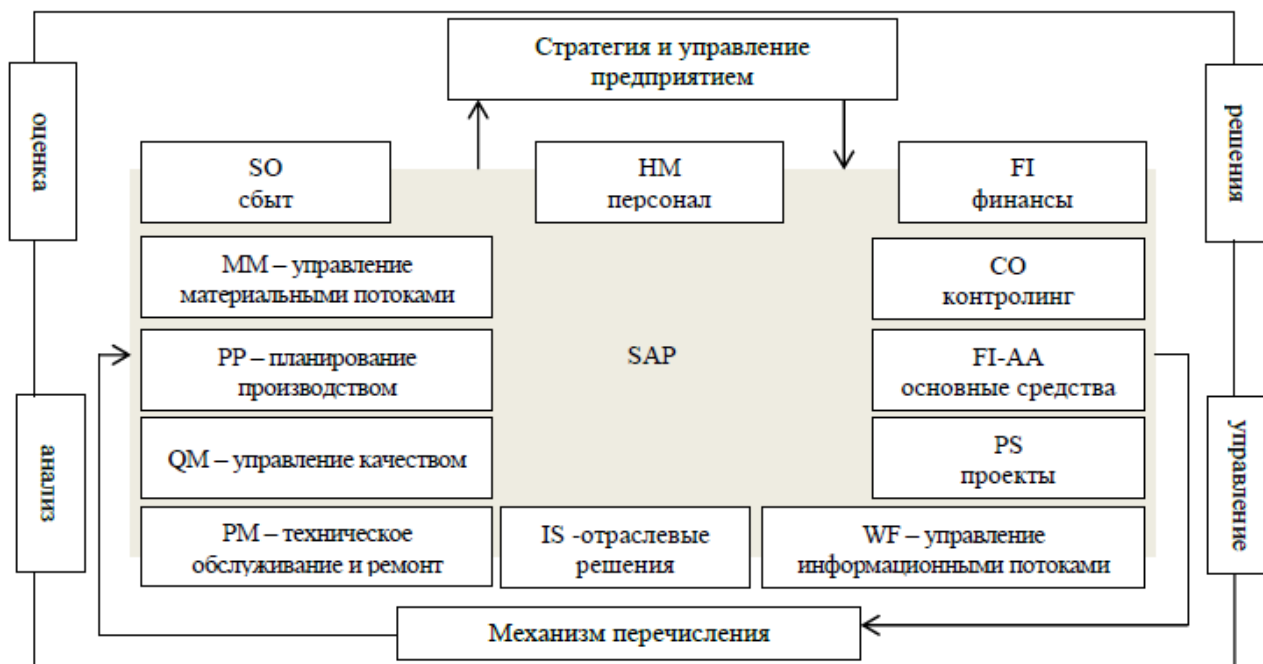


Рисунок 2 – Концепция «перечислительный механизм»

Система настроена таким образом, что начальная информация трансформируется в соответствии с целями. Разместив начальную информацию между процессами и данными в промежуточной зоне «Перечислимый механизм», получают различные оценки (например, чтобы обеспечить соответствие национальным и международным стандартам бухгалтерского учета) для самых разных целей (Финансовой отчетности и определения рентабельности продуктов).

В такой системе данные промежуточной зоны дают информацию для аналитических программ, а это уже логические, а не физические ERP-системы, которые лишены недостатков, связанных с существованием традиционных межсистемных границ [5].

Основа концепции ERP – автоматизация процессно-ориентированного предприятия, поэтому отбор процессов для внедрение на промышленных предприятиях имеет большое значение. Команда, ответственная за

отбор, принимает решение в зависимости от того, насколько легко ERP-система может автоматизировать наиболее важные для предприятия процессы, и насколько гибкой будет система при внедрении других, необходимых предприятию, процессов. Выбор процессов предусматривает систематическую компиляцию всех процессов и вариантов процессов, превалирующих на предприятии. Эти процессы могут принадлежать любому организационному уровню, независимо от того, компьютеризированные они или нет.

Среди зарубежных систем автоматизированного управления система SAP, состоящая из прикладных модулей, которые поддерживают различные бизнес-процессы, вызывает наибольший интерес у руководителей и владельцев предприятий.

Задачи предприятия, которые относятся к архитектуре информационных потоков, могут быть присвоены различным уровням. Каждому уровню присваиваются различные задачи и области ответственности (рис.3).

Системы для управления технологическим оборудованием и для сбора данных с производственных площадок и цехов относятся к первому, самому низкому уровню. На нем устанавливается прямое соответствие между микропроцессорными системами или программируемыми логическими контроллерами и производственными станками и оборудованием. Область ответственности таких систем ограничена одной производственной единицей, поэтому на этом уровне может работать большое количество систем.



Рисунок 3 – Концепция уровней ответственности системы управления промышленным предприятием

Второй уровень включает автоматизированные системы управления технологическими процессами. Эти системы управляют производственной областью. Системы и производственные единицы на этом уровне также управляют процессами на основе оптимизированных математических моделей и отслеживают материалы, проходящих через область оценки на уровне подразделения.

Третий уровень систем отвечает за управление производством на большой площади, например, на всем предприятии. Особенности ERP заставляют уделять повышенное внимание планированию последовательности обработки и отслеживанию запасов на всех переделах. Отдельные материалы могут учитываться и контролироваться на всем протяжении производственного процесса. На этом уровне ИТ-системы носят название системы управления производством или MES-системы.

На четвертом уровне ведется планирование производства, управление качеством, отгрузки и транспортировки готовой продукции, техническое обслуживание и ремонт оборудования. На четвертом уровне традиционно размещаются ERP-системы, такие как программные решения SAP «Управление ресурсами предприятия» (mySAP ERP) или «Управление логистической сетью» (mySAP SCM).

Процесс создания единого информационного пространства предполагает реорганизацию информационных потоков и ориентацию их на:

- удовлетворение потребности предприятия в информации для поддержки внутрифирменного процесса принятия решений и их реализации (учет, анализ, планирование и контроль);
- удовлетворение потребности внешних структур в информации согласно законодательных предписаний или для самопрезентации предприятия.

Важность этих задач непосредственно связана с ключевыми целями промышленных холдингов, такими как:

- повышение инвестиционной привлекательности бизнеса за счет экономической прозрачности предприятия;
- повышение конкурентоспособности на внешнем и внутреннем рынках за счет производственной прозрачности предприятия и улучшения качества продукции;
- повышение эффективности производства за счет применения современных технологий управления;
- повышение уровня и устойчивости управления предприятием и его производствами.

Для достижения этих целей предприятиями выделяются значительные инвестиции. Большая их часть должна быть направлена на проекты, направленные на решение вышеуказанных задач средствами информационных технологий, в частности через проекты по внедрению информационных систем класса ERP, систем управления производством (MES - Manufacturing Execution System), объединение локальных информационных систем, и автоматизированных систем управления технологическим процессом в единый интегрированный комплекс управления (рис. 4). Для промышленных предприятий при внедрении систем класса ERP и консолидации на их основе финансовых потоков все более актуальными становятся вопросы оптимизации производства и повышения качества управления производственными процессами в рамках функционирования учетно-аналитической системы управления (ОАСУ).



Рисунок 4 – Место MES-системы в структуре промышленных холдингов

Создание комплексных интегрированных многопользовательских систем управления с высокими требованиями к достоверности и защищенности информации - очень сложная задача. Для ее решения необходимы значительные профессиональные ресурсы и тщательное соблюдение технологий проектирования и управления сложными проектами [6].

По мнению академика М. Чумаченко «... информационная система трансформируется в информационно-управляющую систему, которая определена как совокупность методов, математического обеспечения и автоматизированных технических средств по сбору, переработке и выдаче управленческой информации и стандартных управленческих решений разным взаимосвязанным системам» [6].

Процесс внедрения MES-системы связан с реализацией на предприятиях следующих функций: оперативное управление производством; контроль производственных графиков; выполнение производственных графиков; мониторинг производственного процесса; управления историческими данными; управление качеством; мониторинг материалов; управления складом.

Интеграция производства с ОАСУ достигается, с одной стороны, за счет предоставления пользователям ERP-систем возможности диспетчеризации производства - оперативного визуального контроля за ходом производственного процесса в системе ключевых показателей эффективности, а с другой – путем передачи сводных производственных отчетов в режиме реального времени.

Процесс исследования бизнес-процессов предполагает возможное наличие локальных проблемных областей, негативно влияющих на процедуры выполнения процессов и выполнения целей внедрения. Обнаруженные проблемные области позволяют установить основные тенденции и, опираясь на них, определить пути ликвидации проблемных областей. На первом этапе исследования на основе общих признаков следует объединить проблемы в однородные группы.

Вывод

Итак, можно констатировать наличие проблем, связанных с дублированием и избыточностью функций. Необходимо отметить, что в ходе внедрения ERP-систем практически всегда возникает необходимость оптимизации бизнес-стратегии промышленных предприятий по реорганизации процессов, методов учета, контроля и анализа.

При внедрении информационных систем на промышленных холдингах нуждается в реорганизации работа учетно-аналитических подразделений и логики работы систем, передающих информацию в систему, и как следствие большинство из информационных систем, которые уже используются, могут быть замещены.

При этом применение современных информационных технологий в управлении остается одним из главных инструментов эффективности бизнеса. Цель применения управленческих информационных систем заключается в том, чтобы упорядочить информационные потоки, предоставить эффективный доступ к данным менеджеров всех уровней для принятия мотивированных управленческих решений.

Литература

Бутынец, Ф.Ф. Информационные системы бухгалтерского учета: учебник / Ф.Ф. Бутынец, С.В. Ивахненко, Т.В. Давидюк, Т.В. Шахрайчук; под ред. проф. Ф.Ф. Бутынец. – 2-е изд., – Житомир : ПП Рута, 2002. – 544 с.

Клименко, О.В. Информационные системы и технологии в учете: учебное пособие / О.В. Клименко. – Киев : Центр учебной литературы, 2008. – 320 с.

Шквир, В.Д. Информационные системы и технологии в учете: Учебное пособие / В.Д. Шквир, А.Г. Завгородний, О.С. Высочан. – Киев : Знание, 2006. – 439с.

Классики менеджмента / Под ред. М. Уорнера; [Пер. с англ. под ред. Ю. Н. Каптуревского] – СПб. : Питер, 2001. – 1168 с.

Рид С. Финансовый директор как интегратор бизнеса / Седрик Рид, Ханс-Дитер Шойерман и группа SAP ERP Financials. [Пер. с англ.] – М. : Альпина Бизнес Букс, 2007. – 397 с.

Чумаченко Н. Г. Проблемы учета и экономического анализа в управлении промышленным производством: автореф. дис. на соискание науч. степени док. экон. наук: спец. 08.601 «Бухгалтерский учет и анализ хозяйственной деятельности» / Н. Г. Чумаченко. – Москва, 1970. – 38 с.

Hvozdiukova S.N., Suvorova E.Yu. The use of ERP-technology in the strategic management of the organization. This article sets the task to consider the main possibilities of new information technologies in the management of an organization. In particular, the classification of ERP-systems and features of their implementation are considered. Based on the analysis, positive and negative aspects of the implementation of information systems were identified, and the concept of levels of responsibility of the management system during the implementation of ERP-technology was presented.

Key words: information system, management, SEM-system, ERP-system, technology, software.

УДК 658.2

Многоагентные системы полуавтоматического 3D проектирования изделий машиностроения

Евгеньев Г.Б.
МГТУ им. Н.Э. Баумана, Москва
g.evgenyev@mail.ru

Евгеньев Г.Б. Многоагентные системы полуавтоматического 3D проектирования изделий машиностроения. В настоящее время в мире происходит четвертая промышленная революция, получившая наименование «Индустрия 4.0». Целью этой революции является создание цифровых производств, которые обеспечивали бы кардинальное повышение производительности и качества производства. Работа посвящена изложению теории и практики создания головного компонента таких систем – многоагентной методологии генерации систем полуавтоматического типового вариантного проектирования машиностроительных изделий.

Ключевые слова: Индустрия 4.0, цифровые производства, многоагентные системы, автоматизированные системы проектирования.

Введение

По всеобщему признанию сейчас в мире происходит четвертая промышленная революция (4ПР) для обозначения которой широко используется термин Индустрия 4.0. Суть этой революции заключается в том, что материальный мир соединяется с виртуальным, в результате чего рождаются новые киберфизические комплексы, объединенные в одну цифровую экосистему. Революция 4ПР означает все большую автоматизацию всех процессов и этапов производства, включая цифровое проектирование изделия.

Настоящая статья посвящена изложению методов искусственного интеллекта в этой области.

1. Пирамида знаний

Традиционное проектирование изделий включает стадии, определенные стандартом: эскизное, техническое и рабочее проектирование. Эскизный проект должен содержать принципиальные решения, дающие общее представление об устройстве и принципы работы изделия. Технический проект включает окончательные технические решения, дающие полное представление об устройстве разрабатываемого изделия и исходные данные для разработки рабочей документации. В результате рабочего проектирования формируется документация, обеспечивающая возможность изготовления изделия в соответствии с техническими требованиями.

Фундаментом для построения интеллектуальных систем являются базы знаний. При построении их структуры целесообразно воспользоваться многовековым опытом создания материальных изделий. На рис.1 представлена пирамида знаний, соответствующая описанным выше принципам конструирования изделий.

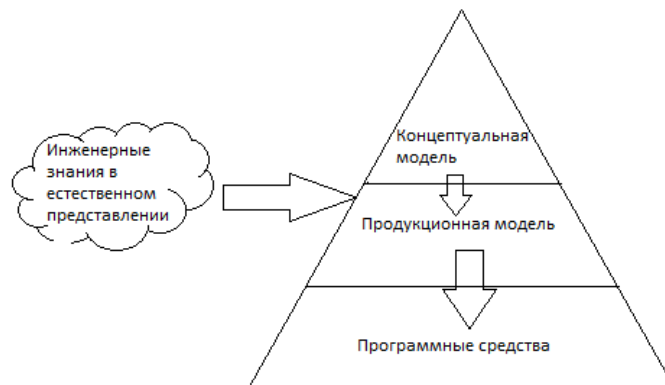


Рисунок 1. Пирамида знаний

2. Концептуальные и производственные модели

На рис.2 приведен пример статической концептуальной модели многоагентной системы типового вариантного проектирования нерегулируемых редукторных приводов.

Модель, представленная на рис.2 имеет непосредственное отношение к морфологическим методам проектирования, при которых сначала определяется пространство поиска, которое должно включать искомое решение, а затем это пространство сужается для поиска этого решения. Пространство поиска называется морфологическим множеством, а процесс определения этого пространства – морфологическим анализом. Поиск решения представляет собой морфологический синтез.

На рис.2 представлено морфологическое И/ИЛИ дерево в нотации UML, необходимое для идентификации. Для построения таких моделей необходимо выделить классификационные признаки, определить их значения и составить их систему, определив сочетаемость их значений.

В информационных технологиях объект представляет собой некоторую сущность в компьютерном пространстве, обладающую определенным состоянием и поведением, имеющую заданные типы значений свойств (атрибутов) и операций над ними (методов).

Важнейшей методической основой для реализации современных интеллектуальных систем является теория многоагентных систем (МАС). Любой агент представляет открытую систему, помещенную в некоторую среду, которой является проект, формируемый в базах данных. Агент является разновидностью объектов, отличающаяся номенклатурой свойств. Если в UML свойства объектов делятся на два класса – видимые, помечаемые знаком +, и невидимые со знаком -, то у агента видимые свойства делятся на входные для метода со знаком (+<) и выходные со знаком (->) (рис.2). В этой связи агент представляет собой объект-функцию.

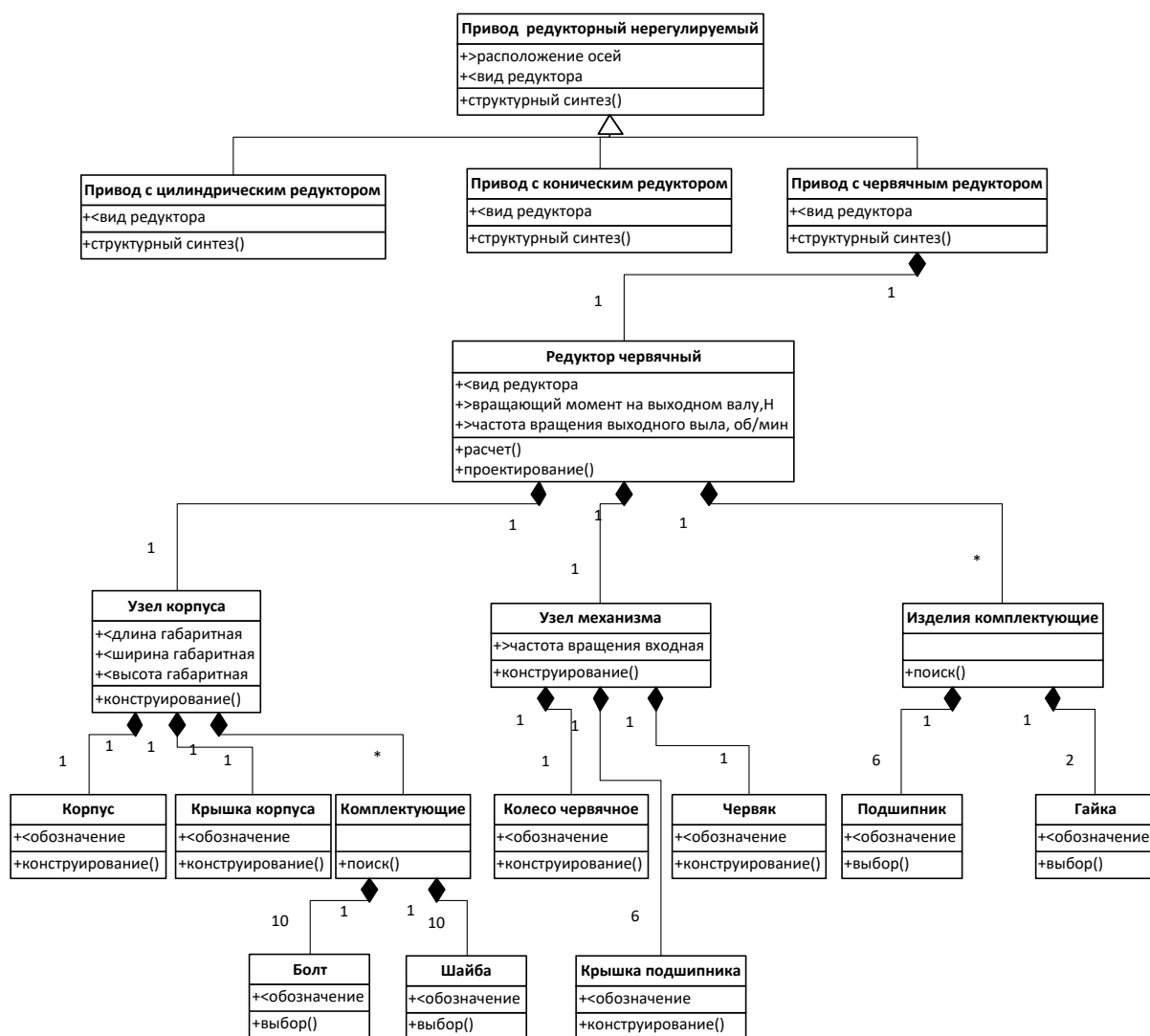


Рисунок 2. Концептуальная модель

При проектировании изделий используются две процедуры: процедура структурного синтеза и процедура параметрического синтеза. Верхние два уровня концептуальной модели на рис.2. связаны с процедурой структурного синтеза, осуществляемого морфологическими методами. Для этого в связках типа ИЛИ,

соединяющих родительский объект (Привод редукторный нерегулируемый) с потомками, представляющими разновидности родительского объекта, нужно выбрать одну разновидность. Это осуществляется методом структурного синтеза родительского объекта. В данном случае он представляет собой таблицу, приведенную ниже.

Расположение осей входного и выходного валов	Тип редуктора
параллельное	Редуктор цилиндрический
пересекающееся	Редуктор конический
скрещивающееся	Редуктор червячный

Для процедур параметрического синтеза целесообразно использовать модель агента в виде функционального блока IDEF0 (рис.3).

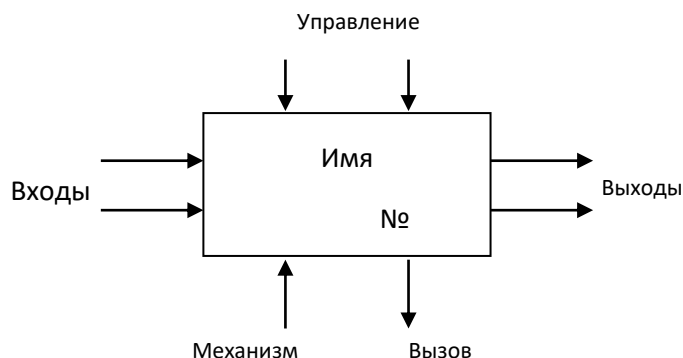


Рисунок 3. Функциональный блок IDEF0

Функциональный блок IDEF0 имеет все необходимые компоненты информационного объекта: состояние его определяется текущими значениями входных, выходных и управляющих атрибутов, а поведение его обеспечивается работой механизма и вызываемых процедур.

Для выполнения операций параметрического синтеза разработана методология экспертного программирования [Евгеньев и др., 2015]. В ней функциональные блоки рассматриваются как продукционные правила и носят название модулей знаний (МЗ).

Ниже приведено внешнее представление МЗ на языке деловой прозы. Этот модуль производит расчет передаточного числа.

МЗ: "PR3" - Расчет числа передаточного

Предусловия запуска

имя	наименование	тип	условие
VidSE	Вид CE	STRING	Редуктор
nt	Частота вращения на выходе, об/мин	REAL	(0,.)

Входные свойства

имя	наименование	тип	значение
nh	Частота вращения на входе, об/мин	REAL	
nt	Частота вращения на выходе, об/мин	REAL	

Механизм - Формула

$$Ur_z = nh/nt$$

Выходные свойства

имя	наименование	тип	значение
Ur_z	Передаточное число	REAL	

Механизмы модулей знаний обеспечивают реализацию всех функций, которые могут потребоваться при формировании баз знаний. В частности для генерации 3D моделей разработан соответствующий механизм [3]. Модули знаний, представляющие собой элементарные порождающие системы, автоматически объединяются в структурированные порождающие системы. Моделью структурированных порождающих систем являются семантические сети МЗ представляющие собой ациклический ориентированный граф (рис.4) [1].

Если при структурном синтезе используются статические диаграммы агентов, соответствующие диаграммам классов UML, то для параметрического синтеза используются динамические диаграммы агентов, соответствующие диаграммам состояний UML.

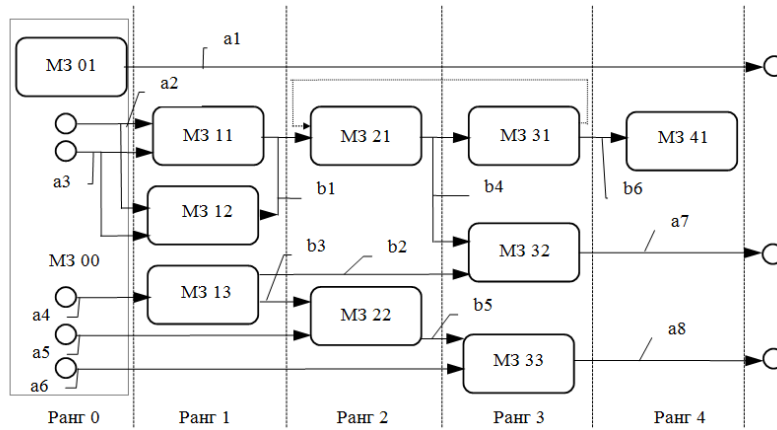


Рисунок 4. Комплексный динамический агент

На рис.4 приведена диаграмма комплексного динамического агента, сгенерированного из множества модулей знаний. Такой агент представляет собой ранжированную семантическую сеть МЗ, являющуюся, по сути, алгоритмом вычисления выходных переменных на основе входных переменных ранга 0.

При экспертном программировании с помощью системы СПРУТ-ЭКСПРО формирование МЗ, их трансляция с получением на одном из традиционных языков объектных или исполняемых модулей и тестирование проводятся как одна операция. После получения необходимого набора МЗ осуществляется генерация метода, использующего подмножество созданных модулей. Метод генерируется с помощью того же языка, что и МЗ, и представляет собой скомпилированную реализацию решателя для заданного набора МЗ.

В процессе работы метода формируется протокол, реализующий объяснительную функцию базы знаний.

3. Программные средства реализации технологии

На рис.5 приведен интерфейс инструментальной системы Sprut-X, с помощью которой диаграммы концептуальных моделей, типа представленной на рис.2, преобразуются в программные средства многоагентной системы проектирования [2]. Работа ведется в интерактивном режиме, в результате которого в окне инспектора формируется графический эквивалент диаграммы, содержащей упорядоченное множество взаимосвязанных агентов, их свойств и методов.

На рис.5 показано начало формирования агента «Редуктор цилиндрический двухступенчатый», который представляет собой наряду с агентом «Редуктор цилиндрический одноступенчатый» разновидность родительского агента «Редуктор цилиндрический». Агенты имеют свойство «Mt» – «Момент крутящий на выходном валу, Нм» и метод «RsCilRed» – «Проектирование цилиндрических редукторов»

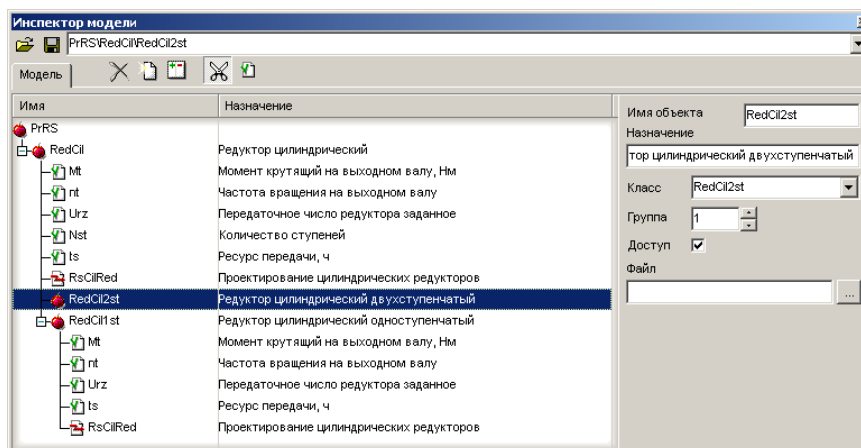


Рисунок 5. Инструментальное средство генерации концептуальных моделей

На рис.6 приведена общая модель работы многоагентной системы полуавтоматического проектирования и конструирования изделий машиностроения на примере проектирования двухступенчатого редуктора с генерацией 3D модели изделия.

Слева внизу приведено отображение экрана ввода технических требований к проектируемому редуктору. Здесь задаются величины момента вращающего на выходном валу, частота вращения на выходе, ресурс передачи и количество изделий для проектирования технологических процессов.

В верхнем ряду отображены экраны запросов, задаваемых пользователю в процессе проектирования. Они включают выбор электродвигателя из наличия двигателей в базе данных, выбор вида зуба колес, выбор варианта термообработки колеса и шестерни по ступеням, выбор материала колеса и шестерни по ступеням и назначение свойств материала шестерен и колес. С помощью информации, полученной с помощью этого диалога, производится управление процессом проектирования.

Справа внизу отображена 3D модель спроектированного изделия.

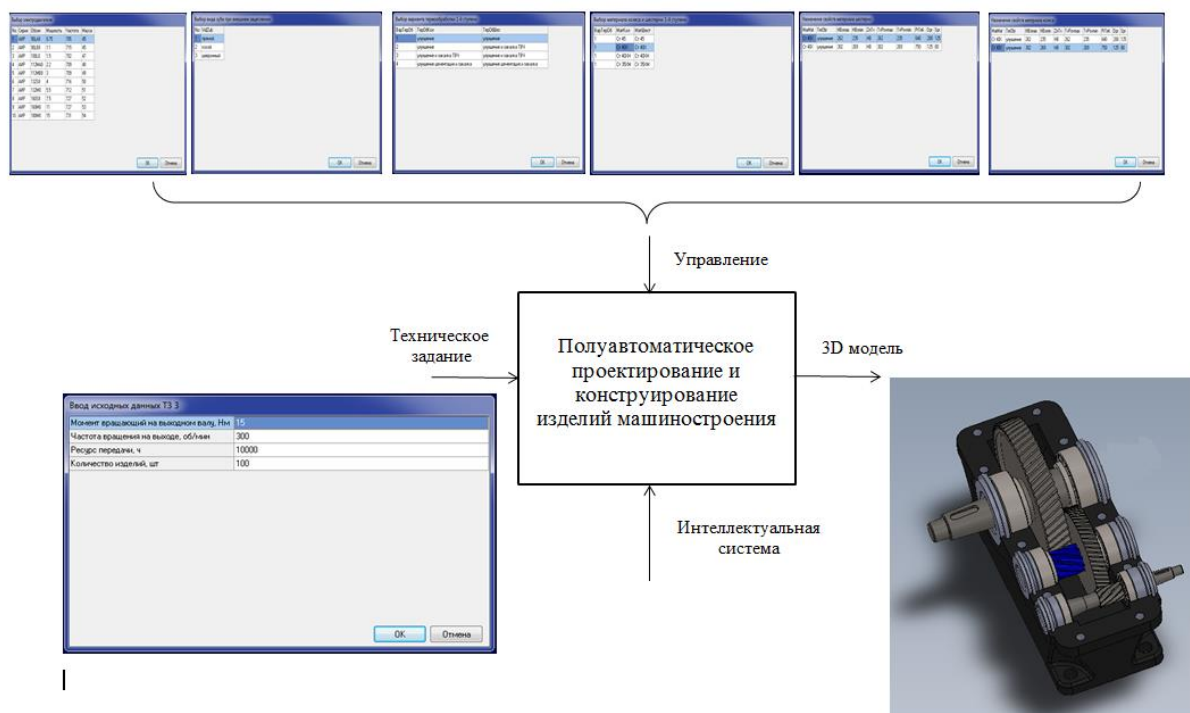


Рисунок 6. Общая модель работы системы

Заключение

Разработана методология полуавтоматического типового вариантного 3D проектирования изделий машиностроения. На базе этой методологии с использованием отечественных программных средств реализованы системы проектирования редукторных нерегулируемых электромеханических приводов.

Литература

1. Евгенов Г.Б. Интеллектуальные системы проектирования. – М.: Издательство МГТУ им. Н.Э. Баумана, 2012.
2. Основы автоматизации технологических процессов и производств : учебное пособие : в 2 т. / [Г.Б. Евгенов и др.] ; под ред. Г.Б. Евгенова – Москва: Издательство МГТУ им. Н.Э. Баумана, 2015.
3. Интеллектуальные системы полуавтоматического проектирования и быстрого прототипирования изделий машиностроения. // Евразийский Союз Ученых (ЕСУ) # 9 (18), 2015/ ТЕХНИЧЕСКИЕ НАУКИ. С. 19-25.

Евгенов Г.Б. Многоагентные системы полуавтоматического 3D проектирования изделий машиностроения. В настоящее время в мире происходит четвертая промышленная революция, получившая наименование «Индустрия 4.0». Целью этой революции является создание цифровых производств, которые обеспечивали бы кардинальное повышение производительности и качества производства. Работа посвящена изложению теории и практики создания головного компонента таких систем – многоагентной методологии генерации систем полуавтоматического типового вариантного проектирования машиностроительных изделий.

Ключевые слова: Индустрия 4.0, цифровые производства, многоагентные системы, автоматизированные системы проектирования.

G.B. Evgenev. Multy-agent systems semy-automatic 3D designing machine-building products Now in the world there is a fourth industrial revolution, received the name "Industry 4.0". The purpose of this revolution is the creation of digital productions, which would provide a radical increase in productivity and quality of production. The work is devoted to the theory and practice of creating the main component of such systems – the multi-agent methodology for the generation of semi-automatic standard-variant design of engineering products

Keywords: Industry 4.0, digital production, multi-agent systems, automated design systems.

Концептуальные каркасы онтологий самоструктурируемых открытых информационных систем

Жевнерчук Д.В.

Нижегородский государственный технический университет им. Р.Е. Алексеева
zhevnerchuk@yandex.ru

Жевнерчук Д.В. Концептуальные каркасы онтологий самоструктурируемых открытых информационных систем. В статье обсуждается подход к онтологическому моделированию информационных систем, предложенные концептуальные каркасы онтологий самоструктурируемых открытых информационных систем могут быть применены при разработке алгоритмического обеспечения средств интеллектуальной поддержки их структурно-параметрического синтеза.

Ключевые слова: концептуальный каркас, онтология, синтез, самоструктурирование, техническая самоорганизация, открытая информационная система.

Введение

Структура информационных систем может быть описана совокупностью фильтрующих и нефильтрующих компонентов, взаимодействующих через стандартизированные интерфейсы [1,2]. Рассматривая метасистему, включающую информационную систему и источники внешнего информационного воздействия, можно заметить, что информационные процессы, а также требования и ограничения представляются в форме фильтрующих компонентов и многокомпонентных интероперабельных структур (МИС) [2], благодаря чему достигается единообразие формализации воздействий внешней среды и внутренних компонентов открытых информационных систем (ОИС). В свою очередь это позволяет формировать многокомпонентные структуры, элементами которых являются как преобразователи информации, так и информационные процессы, требования, ограничения, выраженные в компонентной форме.

Постановка задачи

Синтез ОИС включает два основных этапа. На первом этапе выполняется сопряжение компонентов в пределах одного уровня (синтез многокомпонентных структур), включающее следующие шаги:

- синтез интерфейсов;
- синтез пар блоков, сопряженных через интерфейсы;
- синтез комплексных блоков (блоки, определяемые последовательностями сопряженных блоков, причем у комплексных блоков входной интерфейс совпадает с входным интерфейсом первого блока последовательности, а выходной интерфейс - с выходным интерфейсом последнего блока последовательности сопряженных блоков).

На втором этапе происходит сопряжение компонентов, принадлежащих различным уровням (синтез многокомпонентных интероперабельных структур). На этом этапе выполняются следующие шаги:

- синтез фильтрующих интерфейсов;
- фильтрация многокомпонентных структур в пределах межуровневых связей;
- формирование многокомпонентных структур на основе оставшихся в результате фильтрации компонентов.

Синтез целевой ОИС является результатом композиции интероперабельных структур. Базовой моделью, отражающей структурные и параметрические аспекты МИС, является граф, и реализация выделенных шагов сводится к известным решениям задач [3,4,5]:

- изменения структуры подграфов (добавление, изменение, удаление ребер и узлов);
- поиск подграфов;
- кластеризация на одном графе;
- кластеризация на множестве графов.

В докладе обсуждаются основополагающие онтологии, описывающие этапы синтеза многокомпонентных интероперабельных структур, которые могут быть применены для построения интеллектуальных алгоритмов структурного синтеза открытых информационных систем. Поскольку переход к

многокомпонентным структурам осуществляется в автоматическом режиме, то можно говорить о самоструктурируемых ОИС.

Принцип самоструктурирования открытых информационных систем

Известно, что открытые информационные системы обладают свойствами расширяемости по многокомпонентным структурам. Дополнительно введем требование, согласно которому ОИС должна обеспечивать сопряжение своих компонентов, являющихся преобразователями информации и в условиях изменения требований и ограничений внешней среды быть способна к реструктуризации.

Принцип самоструктурирования, направленный на обеспечение расширяемости ОИС по многокомпонентным структурам, представлен следующими основными положениями:

1. Целевая открытая информационная система описывается посредством преобразователей информационных потоков, представленных в форме компонентов, обладающих стандартизированными интерфейсами, а также информационными связями между ними, в форме правил сопряжения через их интерфейсы.
2. Отдельные аспекты поведения ОИС на разных уровнях описаны в виде, совокупности параметров, режимов, определяемых комбинациями параметров и диапазонами их значений.
3. Процесс формирования многокомпонентных структур является иерархическим, причем каждый уровень иерархии описывается базовым набором типов объектов и семантическими связями между ними, являющимся каркасом уровня иерархии.
4. Выделены следующие уровни формирования многокомпонентных структур:
 - уровень блочного представления и стандартизации;
 - уровень компонентного представления;
 - уровень представления многокомпонентными структурами.
5. Переход к следующему уровню выполняется на основе системы фактов и аксиом, определяющих локальные связи между структурами текущего уровня:
 - блоки, точки доступа, свойства, стандартизированные свойства;
 - стандартизированные свойства, точки доступа, интерфейсы;
 - интерфейсы, компоненты;
 - компоненты, многокомпонентные структуры.

В результате перехода к следующему уровню формируются новые структуры меньшей размерности, позволяющие вводить новые локальные связи, а также факты и аксиомы их поддержки.

Реализация принципа самоструктурирования ОИС основывается на онтологическом подходе. Для построения целевых онтологий воспользуемся понятием концептуального каркаса онтологии, введенным в работе [6].

Онтологическое моделирование самоструктурируемых открытых информационных систем

Согласно сформулированному принципу самоструктурирования выделим три уровня онтологического представления открытых информационных систем:

- уровень блочного представления и стандартизации (уровень 1);
- уровень интерфейсов и компонентного представления (уровень 2);
- уровень многокомпонентных структур (уровень 3).

На уровне блочного представления и стандартизации осуществляется ввод первичной информации о блоках информационных систем, их точках доступа и связанных с ними свойствах. Кроме того, на этом уровне осуществляется формализованное представление спецификаций, их параметров и областей ограничений. На рис. 1 представлен каркас онтологии уровня 1. В табл. 1, 2 представлено описание основных концептов, ролей и аксиом.

Отметим, что на рисунке связи а) между свойствами блоков и параметрами спецификации; б) между доменными ограничениями свойств блоков и доменными ограничениями параметров спецификации определяют роли *specified*.

В результате выполнения аксиомы на первом уровне автоматически формируются роли, отражающие отношения стандартизации между зарегистрированными блоками спецификациями, индивиды доменных ограничений, которые связываются со свойствами блоков, а также роли, выделяющие стандартизированные свойства.

Стандартизированные блоки и свойства, далее, становятся доступны на следующем уровне, где происходит формирование слоя компонентов, обладающих интерфейсами, определяющими каналы обмена информацией между компонентами

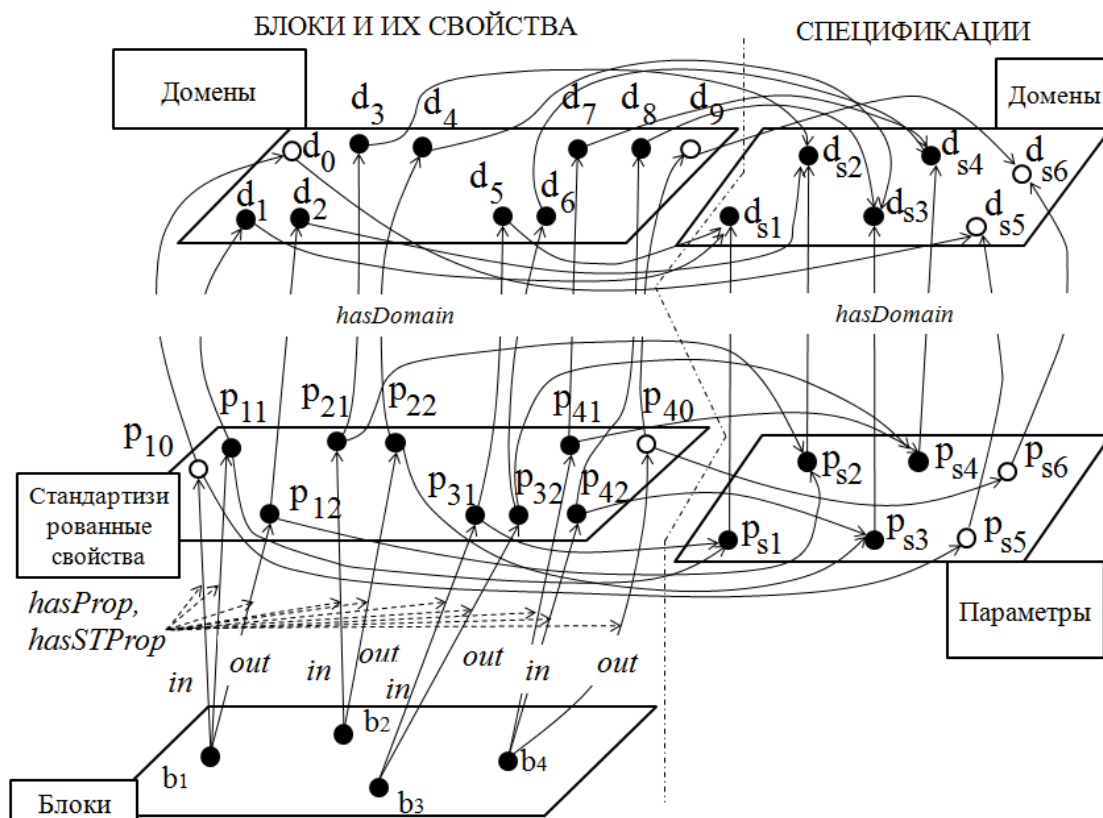


Рисунок 1 – Концептуальный каркас онтологии самоструктурируемых ОИС (Уровень 1)

Таблица 1 – Концепты и роли онтологии самоструктурируемых ОИС (Уровень 1)

№ пп	Формула	Тип	Комментарий
1	b^{new}	входные данные	обозначение нового блока
2	$p_1^{b^{new}}, \dots, p_n^{b^{new}}$	входные данные	Обозначение новых свойств нового блока
3	$in(b^{new}, p_i^{b^{new}})$	входные данные	обозначение связи i -го свойства блока с точкой доступа блока типа вход
4	$out(b^{new}, p_i^{b^{new}})$	входные данные	обозначение связи i -го свойства блока с точкой доступа блока типа выход
5	$specified(p_i^{b^{new}}, p_{qs})$	входные данные	i -ое свойство нового блока определяется q -ым параметром, определенным в s -ой спецификации
6	d^{new}	входные данные	новая область определения свойства блока
7	$hasDomain(p_{qs}, d_{ks}^{standard})$	входные данные	q -ый параметр, определенный в s -ой спецификации определен на k -ом домене спецификации s
8	$hasSTProp(b^{new}, p_i^{b^{new}})$	результат обработки аксиомы	обозначение факта, что свойство $p_i^{b^{new}}$ является стандартизированным
9	$hasDomain(p_i^{b^{new}}, d^{new})$	результат обработки аксиомы	i -ое свойство нового блока определено на новом домене
10	$specified(d^{new}, d_k^{standard})$	результат обработки аксиомы	новый домен определяется k -ым доменом спецификации s

Таблица 2 – Аксиома онтологии самоструктурируемых ОИС (Уровень 1)

№ пп	Формула	Псевдокод
1	$specificated(p_i^{b^{new}}, p_{qs}) \cap$ $hasDomain(p_{qs}, d_k^{standard}) \rightarrow$ $hasDomain(p_i^{b^{new}}, d^{new}) \cap$ $specificated(d^{new}, d_k^{standard}) \cap$ $hasSTProp(b^{new}, p_i^{b^{new}})$	<p>Если</p> <p>[</p> <p>(i-ое свойство нового блока определяется q-ым параметром, определенным в s-ой спецификации)</p> <p>И</p> <p>(q-ый параметр, определенный в s-ой спецификации определен на k-ом домене спецификации s)</p> <p>],</p> <p>то</p> <p>[</p> <p>(i-ое свойство нового блока определено на новом домене)</p> <p>И</p> <p>(новый домен определяется k-ым доменом спецификации s)</p> <p>И</p> <p>(i-ое свойство нового блока является стандартизированным)</p> <p>]</p>

На уровне 2 происходит формирование связи точек доступа блоков с доменами напрямую, минуя слой свойств. Отображение пары доменов на один элемент интерфейса происходит в случае, если они определяются общим доменом спецификации и через связанную с ним точку доступа существует канал обмена данными между двумя и более блоками. Для блоков, обладающих интерфейсами, формируется индивид компонент, кроме того, компоненты, способные обмениваться информацией через интерфейсы, дополнительно связываются посредством роли, которая определяет направление передаваемой информации. На рис. 2 представлен каркас онтологии уровня 2. В табл. 3, 4 представлено описание основных концептов, ролей и аксиом.

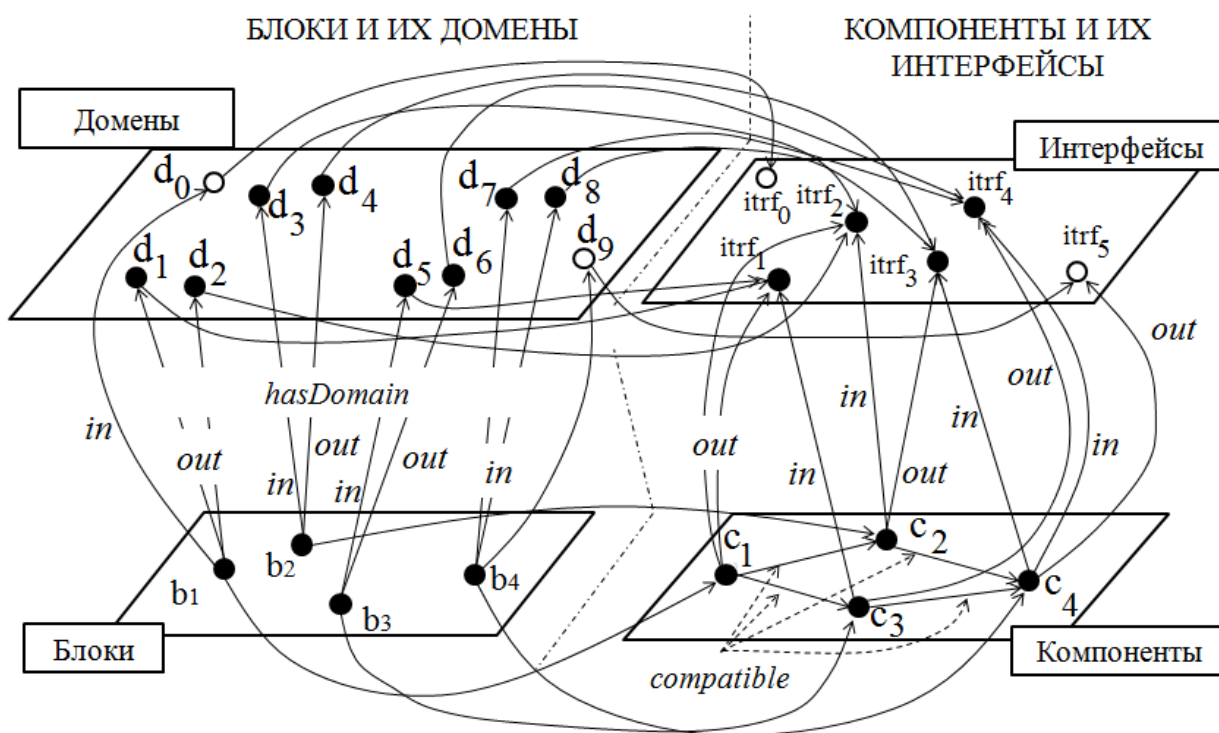


Рисунок 2 – Концептуальный каркас онтологии самоструктурируемых ОИС (Уровень 2)

Таблица 3 – Концепты и роли онтологии самоструктурируемых ОИС (Уровень 2)

№ ПП	Формула	Тип	Комментарий
1	$in(b^{new}, d^{P_i})$ $out(b^{new}, d^{P_i})$	результат обработки аксиомы	домен d^{P_i} связан с точкой доступа типа вход/выход нового блока b^{new}
2	$hasDomain(b^{new}, d^{P_i})$	результат обработки аксиомы	блок b^{new} поддерживает домен d^{P_i}
3	$in(c_i^{new}, itrif^{new})$ $out(c_j^{new}, itrif^{new})$	результат обработки аксиомы	новый интерфейс $itrif^{new}$ связан с точкой доступа типа вход/выход нового компонента c_i^{new}
4	$based(c_i^{new}, b_i)$	результат обработки аксиомы	новый компонент c_i^{new} основывается на блоке b_i
5	$compatible(c_i, c_j)$	результат обработки аксиомы	компонент C_i сопряжен с компонентом C_j в направлении от C_i к C_j

Таблица 4 – Аксиомы онтологии самоструктурируемых ОИС (Уровень 2)

№ ПП	Формула	Псевдокод
1	а) Добавление концептов и ролей, связанных ассоциацией доменов с блоками (исключение слоя параметров, привязанных к точкам доступа типа «выход» блоков) $out(b^{new}, p_i^{b^{new}}) \cap$ $hasDomain(p_i^{b^{new}}, d^{P_i}) \rightarrow$ $out(b^{new}, d^{P_i}) \cap$ $hasDomain(b^{new}, d^{P_i})$	Если [(<i>i</i> -е свойства блока связано с точкой доступа типа выход нового блока b^{new}) И (<i>i</i> -ое свойство нового блока определено на домене d^{P_i})], то [(домен d^{P_i} связан с точкой доступа типа выход нового блока b^{new}) И (блок b^{new} поддерживает домен d^{P_i})]
	б) Добавление концептов и ролей, связанных ассоциацией доменов с блоками (исключение слоя параметров, привязанных к точкам доступа типа «вход» блоков) <i>строится аналогично</i>	
2	<i>Добавление концептов и ролей, связанных с формированием интерфейсов и компонентов</i>	
	а) домены, определяемые общим доменом спецификации, привязаны к точкам доступа типа вход и выход разных блоков, причем для них еще не созданы интерфейсы и компоненты	

	$in(b_i, d_k^{b_i}) \cap out(b_j, d_q^{b_j}) \cap$ $specified(d_k^{b_i}, d^s) \cap$ $specified(d_q^{b_j}, d^s) \rightarrow$ $specified(itrf^{new}, d^s) \cap$ $in(c_i^{new}, itrf^{new}) \cap$ $out(c_j^{new}, itrf^{new}) \cap$ $based(c_i^{new}, b_i) \cap$ $based(c_j^{new}, b_j)$	<p>Если</p> <p>[(домен $d_k^{b_i}$ связан с точкой доступа типа вход нового блока b_i)</p> <p>И (домен $d_q^{b_j}$ связан с точкой доступа типа вход нового блока b_j)</p> <p>И (домен $d_k^{b_i}$ определяется доменом спецификации s)</p> <p>И (домен $d_q^{b_j}$ определяется доменом спецификации s)],</p> <p>то</p> <p>[(новый интерфейс $itrf^{new}$ определяется доменом спецификации s)</p> <p>И (интерфейс $itrf^{new}$ связан с точкой типа вход компонента c_i^{new})</p> <p>И (интерфейс $itrf^{new}$ связан с точкой типа выход компонента c_j^{new})</p> <p>И (новый компонент c_i^{new} основывается на блоке b_i)</p> <p>И (новый компонент c_j^{new} основывается на блоке b_j)]</p>
<p>б) домены, определяемые общим доменом спецификации, привязаны к точкам доступа типа вход и выход одного и того же блока, причем для него не определены интерфейсы и компоненты (строятся аналогично)</p>		
3	<p>Добавление концептов и ролей, связанных с попарным сопряжением компонентов</p>	
<p>а) Сопряжение двух разных компонентов</p>		
	$in(c_i, itrf) \cap out(c_j, itrf) \rightarrow$ $compatible(c_i, c_j)$	<p>Если</p> <p>[(интерфейс $itrf$ связан с точкой типа вход компонента C_i)</p> <p>И (интерфейс $itrf$ связан с точкой типа выход компонента C_j)],</p> <p>то</p> <p>[(компонент C_i может быть сопряжен с компонентом C_j)]</p>
<p>б) Сопряжение одного компонента с самим собой</p>		
	$in(c_i, itrf) \cap out(c_i, itrf) \rightarrow$ $compatible(c_i, c_i)$	<p>Если</p> <p>[(интерфейс $itrf$ связан с точкой типа вход компонента C_i)</p> <p>И (интерфейс $itrf$ связан с точкой типа выход компонента C_i)],</p> <p>то</p> <p>[(компонент C_i может быть сопряжен с компонентом C_i)]</p>

Синтез многокомпонентных структур осуществляется посредством поиска подструктур построенной структуры compatible-связности компонентов. На рис. 3 представлен каркас онтологии уровня 3. В табл. 5, 6 представлено описание основных концептов, ролей и аксиом.

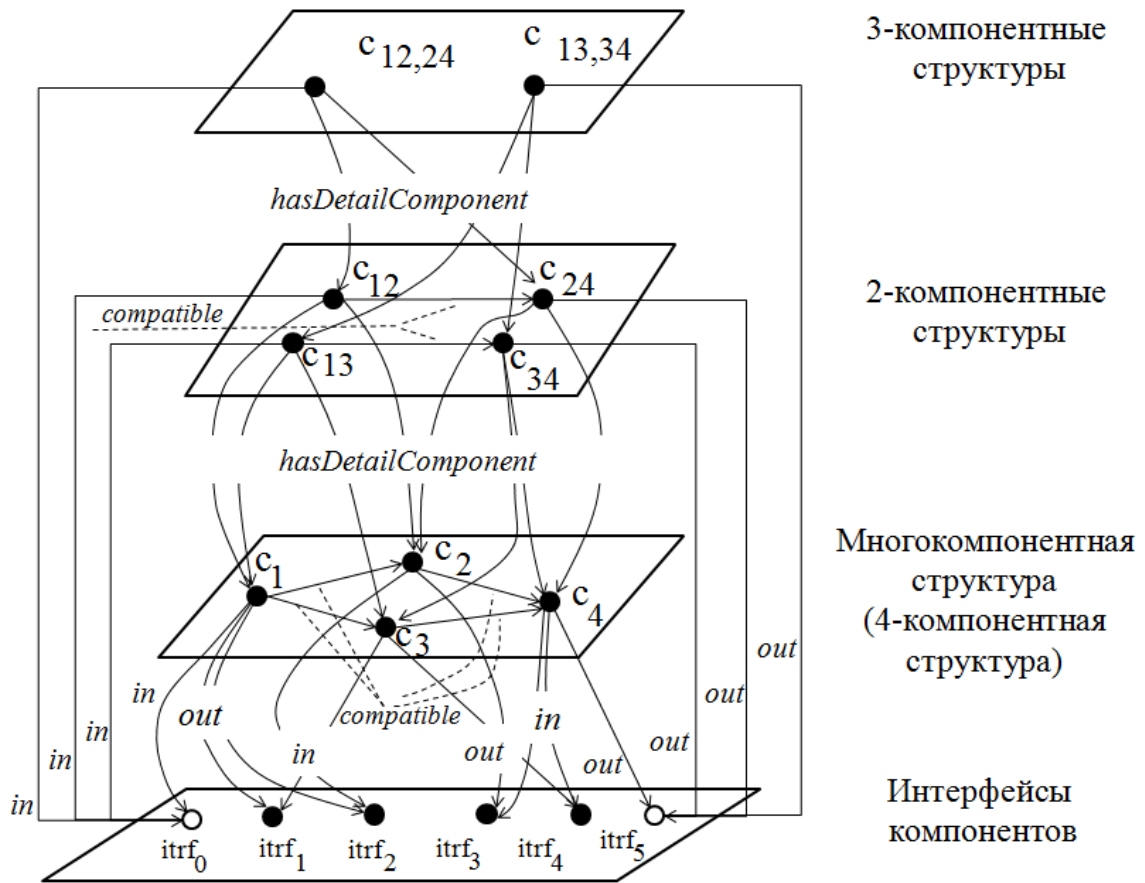


Рисунок 3 – Концептуальный каркас онтологии самоструктурируемых ОИС (Уровень 3)

Таблица 5 – Роли онтологии самоструктурируемых ОИС (Уровень 3)

№ пп	Формула	Тип	Комментарий
1	$hasDetailComponent(c_i^r, c_j^{r-1})$	результат обработки аксиомы	i -ый компонент на r -ом уровне иерархии связан с j -ым компонентом на $r-1$ уровне иерархии.

Таблица 6 – Аксиома онтологии самоструктурируемых ОИС (Уровень 3)

№ пп	Формула	Псевдокод
1	$compatible(c_i^{r-1}, c_j^{r-1}) \cap in(c_i^{r-1}, itrf_i) \cap out(c_j^{r-1}, itrf_j) \rightarrow hasDetailComponent(c_k^r, c_i^{r-1}) \cap hasDetailComponent(c_k^r, c_j^{r-1}) \cap in(c_k^r, itrf_i) \cap out(c_k^r, itrf_j)$	<p>Если</p> <p>[</p> <p>(компонент c_i^{r-1} сопряжен с компонентом c_j^{r-1} в направлении от c_i^{r-1} к c_j^{r-1})</p> <p style="text-align: center;">И</p> <p>(интерфейс $itrif_i$ связан с точкой доступа типа вход компонента c_i^{r-1})</p> <p style="text-align: center;">И</p> <p>(интерфейс $itrif_j$ связан с точкой доступа типа вход компонента</p>

	c_j^{r-1}], то [(k-ый компонент на r-ом уровне иерархии связан с i-ым компонентом на r-1 уровне иерархии) И (k-ый компонент на r-ом уровне иерархии связан с j-ым компонентом на r-1 уровне иерархии) И (интерфейс itr_f_i связан с точкой доступа типа вход компонента c_k^r) И (интерфейс itr_f_j связан с точкой доступа типа выход компонента c_k^r)]
--	---

Заключение

Предложена иерархия онтологических каркасов, содержащих базовые семантические конструкции, аксиомы локальных межструктурных связей каждого уровня иерархии. В их число входят каркасы, описывающие уровни представления: исходных и стандартизированных блоков и их свойств, интерфейсов и компонентов, многокомпонентных структур.

Такая организация онтологических аспектов самоструктурируемых открытых информационных систем позволяет организовать иерархический вычислительный процесс, на каждом уровне которого формируются новые конструкции, обладающие более высоким уровнем абстракций. При переходе к новому уровню сокращается число параметров, требующих обработки, например, на уровне формирования многокомпонентных структур нет необходимости оперировать такими концептами как блоки, точки доступа, свойства.

Литература

1. Жевнерчук Д.В. Синтез открытых информационных систем с использованием алгебраических структур как моделей / Д.В. Жевнерчук, Л.С. Ломакина // *Фундаментальные исследования*. – 2017. – № 10. – С. 29-33
2. Жевнерчук Д.В., Ломакина Л.С. Алгебраические аспекты и структурно-параметрический синтез открытых информационных систем // *Труды Конгресса по интеллектуальным системам и информационным технологиям «IS&IT'18»*. – 2018. – Т.2. – с. 133-141
3. Старостин, Н.В., Панкратова М.А. Генетические алгоритмы решения задачи отображения графа [Текст] Н.В. Старостин, М.А. Панкратова // *Вестник Нижегородского университета им. Н.И. Лобачевского*: № 5-1.– Н.Новгород: Изд-во ННГУ им. Н.И. Лобачевского, 2013. – С. 204-209.
4. Батищев, Д.И. Гибридный подход к решению экстремальных задач на графовых структурах [Текст] / Д.И. Батищев, Н.В. Старостин // *Известия СПбГЭТУ «ЛЭТИ»*. Серия «Информатика, управление и компьютерные технологии», № 3. – С.Петербург, 2002. –С.10-17
5. Старостин, Н.В. Аспекты программной реализации гиперграфов [Текст] / Н.В. Старостин, А.В. Филимонов // *Труды Нижегородского государственного технического университета им. Р.Е. Алексева*. Серия «Информационные технологии», т. 56. – Н.Новгород, 2005. – С. 80-89.
6. Кулинич, А.А. Концептуальные «Каркасы» онтологий в поддержке принятия решений в условиях неопределенности [Электронный ресурс] /А.А. Кулинич, – Режим доступа <http://www.masters.donntu.edu.ua/2013/fknt/bilykn/library/4.htm>. - Загл. с экрана.

Жевнерчук Д.В. Концептуальные каркасы онтологий самоструктурируемых открытых информационных систем. В статье обсуждаются подход к онтологическому моделированию информационных систем, предложенные концептуальные каркасы онтологий самоструктурируемых открытых информационных систем могут быть применены при разработке алгоритмического обеспечения средств интеллектуальной поддержки их структурно-параметрического синтеза.

Ключевые слова: концептуальный каркас, онтология, синтез, самоструктурирование, техническая самоорганизация, открытая информационная система.

Zhevnerchuk D.V. Conceptual frameworks of ontologies of self-structured open information systems. The article discusses approach to the ontological modeling of information systems, the proposed conceptual frameworks of ontologies of self-structured open information systems can be applied in the development of algorithmic support tools for intellectual support of their structural-parametric synthesis.

Key words: conceptual framework, ontology, synthesis, self-structuring, technical self-organization, open information system.

Анализ существующих подходов и перспектив развития интеллектуальных генераторов программного кода

Захаров А.С., Жевнерчук Д.В.

Нижегородский Государственный Технический Университет им. Р.Е. Алексева

4 Lukalex.nnov@gmail.com

Захаров А.С., Жевнерчук Д.В. Анализ существующих подходов и перспектив развития интеллектуальных генераторов программного кода. В работе представлен анализ технологий синтеза программного кода, а также осмысление и понимание процесса эволюции генераторов программного кода (ГПК). Показаны возможные перспективы развития средств интеллектуальной поддержки генераторов кода.

Ключевые слова: компиляция, кодогенерация, программный код, шаблоны, скаффолдинг

Введение

В настоящее время ведутся активные исследования в области синтеза программного кода. Данное направление развивается в соответствии с одним из свойств теории вычислимости - полноты по Тьюрингу, на основании которого делается вывод о том, что программа может написать другую программу [1].

Это мощная идея, которая не оценена по заслугам. Программы, пишущие программы являются важным элементом при автоматизации разработки программного обеспечения. К сожалению, слишком часто для многократного использования кода применяются операции копирования и вставки, как в больших, так и в малых масштабах. В тех случаях, когда повторяется один и тот же базовый код, но с мелкими изменениями в реализации, помещение общего кода в шаблон с дальнейшей генерацией служит прекрасным способом сберечь время и усилия, потраченные на механическую работу.

Используя общий шаблон для написания похожих по своему содержанию частей исходного кода, разработчик может уберечься от ошибок, связанных с опечатками, глаз разработчика "замыливается" и многие, даже самые явные ошибки в тексте могут быть пропущены. С ростом сложности и гетерогенности программных систем количество такого шаблонного кода постоянно увеличивается, причем это касается как алгоритмической части, так и частей, связанных с сопряжением различных программных модулей, встраиваемых компонентов и приложений.

Цель работы: Целью данной работы является анализ существующих технологий синтеза программного кода, их всестороннее осмысление и понимание процесса эволюции генераторов программного кода. А также выявление связи алгоритмов генерации программного кода с различными моделями описания предметных областей. Согласно сформулированной цели был проведен анализ развития генераторов программного кода, в том числе их алгоритмов и архитектур, обеспечивающих единое информационное пространство между машинным и человеко-ориентированным пониманием программного кода. Решение задач позволило оценить перспективы дальнейшего развития выбранного направления исследований.

Виды кодогенераторов

В современном мире различают два принципиально различных вида кодогенерации:

- генератор является физически отдельной программой, которая преобразует код на читаемом для человека языке (язык программирования высокого уровня) в язык понятный для компьютера (машиноориентированный язык или объектный код).
- код генерируется на языке понятном для человека из заранее имеющихся моделей

Если попытаться охарактеризовать современный уровень развития компьютерных и информационных технологий, то первое, на что следует обратить внимание - это возрастающая сложность не только отдельных физических и программных компонентов, но и лежащих в основе этих технологий концепций и идей. Языки программирования представляют собой средство описания вычислений для людей и машин. Современный мир зависит от языков программирования, так как все программное обеспечение написано на том или ином языке программирования. Однако для того чтобы запустить программу на компьютере или любом другом вычислительном устройстве необходимо преобразовать ее в форму, которая выполняется на компьютере. Программные комплексы, которые выполняют такое преобразование, называются компиляторы [6].

Переход от первых компиляторов до кодогенерации

Исторический аспект проблемы генерации кода в основном изложен в виде ключевых дат создания тех или иных программных средств. Большая часть этих данных относится к компиляции кода и компиляторам.

С разных позиций проблему синтеза программного кода рассматривали отечественные и зарубежные ученые, среди которых А.В. Леоненков, А.Ю. Молчанов, А.В. Бойченко, Martin Fowler, David Schreiber, Thomas Allen Akin, Perry L. Miller и др.

История компиляторов как преобразователей (и в том числе синтезаторов) кода, берет начало с группы разработчиков первого компьютера под руководством Джона Моучли (Мокли). Идеи, представленные Моучли, сильно повлияли на одну из программисток его группы. Грейс Мюррей Хоппер вместе уже со своей группой ученых занялась разработкой системы, которая могла бы транслировать программы, написанные на "человеческом" языке в "машинные" программы. Хоппер назвала эту программу компилятором. Компилятор Хоппер получил широкую известность благодаря инновационным на тот момент методам. В ее компиляторе преобразовывался весь код сразу, в отличие от аналогов (таких как например "Краткий код") где происходила построчная компиляция.

Кроме того, немало важным является тот факт, что у Хоппера получилась именно компилирующая, а не интерпретирующая программа. Последняя должна находиться в памяти постоянно во время выполнения основной программы. Т.е. программы должны работать параллельно, находясь в одной памяти (и естественно занимая ее). Для компилирующей программы нахождение в памяти не нужно. Один раз откомпилированная программа может запускаться сама без присутствия сторонних программ.

Основоположником информатики в СССР, в частности раздела автоматизации программирования, является Алексей Андреевич Ляпунов, первым предложивший рассматривать программу как последовательность чередующихся этапов, на которых выполняется некая обработка данных. Этап, Ляпунов, предложил назвать оператором, а схемой счета — совокупность операторов и логических условий. Схема и совокупность спецификаций каждого оператора — это программа [6]. Такой взгляд на программу в 50-х был революционным и сразу стал основой первых трансляторов (компиляторов) или программирующих программ, как их тогда называли.

Методология процедурно-ориентированного программирования

Появление первых электронных вычислительных машин или компьютеров ознаменовало новый этап в развитии техники вычислений [4]. Казалось, достаточно разработать последовательность элементарных действий, каждое из которых преобразовать в понятные компьютеру инструкции, и любая вычислительная задача может быть решена. Эта идея оказалась настолько жизнеспособной, что долгое время доминировала над всем процессом разработки программ.

Основой данной методологии разработки программ являлась процедурная или алгоритмическая организация структуры программного кода. Это было настолько естественно для решения вычислительных задач, что ни у кого не вызывала сомнений целесообразность такого подхода. Исходным понятием этой методологии являлось понятие алгоритма, под которым, в общем случае, понимается некоторое предписание выполнить точно определенную последовательность действий, направленных на достижение заданной цели или решение поставленной задачи.

Разработка больших систем потребовала декомпозиции процедур на более мелкие фрагменты. Отдельные части крупных процедур, автономно оформленные как единицы компиляции, были названы модулями. Модули, в свою очередь, содержали более мелкие процедуры и функции – особого вида небольшие процедуры со своим способом вызова и обязательным возвратом значения (результата выполнения процедуры). Были разработаны методы структурного программирования, которые помогали в процедурно-ориентированной методологии осуществлять разбиение всей программы на модули и более мелкие фрагменты – процедуры и функции (70-80е годы).

Методология объектно-ориентированного программирования

Со временем ситуация существенно изменилась. Реальная трудоемкость начальных этапов проектирования новой программы и программирования стала все больше превосходить плановую. Сроки разработки всё более сложного программного обеспечения стали затягиваться. Коллективные разработки стали сложным процессом, возникли серьезные трудности управления проектами в условиях динамически меняющихся требований заказчиков. Языки спецификаций программ были сложны и мало понятны обычному заказчику непрограммисту. Возникла настоятельная необходимость в изменении самой методологии программирования и проектирования. Основные принципы нового подхода к разработке программного обеспечения были сформулированы в середине 80-х в виде объектно-ориентированного программирования (ООП) [6].

Важной особенностью классов является возможность их организации в виде некоторой иерархической структуры, которая по внешнему виду напоминает схему классификации понятий формальной логики. В этой

связи следует заметить, что каждое понятие в логике имеет некоторый объем и содержание. При этом под объемом понятия понимают все другие мыслимые понятия, для которых исходное понятие может служить определяющей категорией или главной частью. Содержание понятия составляет совокупность всех его признаков или атрибутов, отличающих данное понятие от всех других [3].

Принцип разделения модуля на интерфейс и реализацию отражает суть наших представлений об окружающем мире. В интерфейсной части указывается вся информация, необходимая для взаимодействия с любыми другими объектами. Реализация скрывает или маскирует от других объектов все детали, не имеющие отношения к процессу взаимодействия объектов

В соответствии с разделением кода на логические блоки и совокупности объектов, объединенные в классы, на процесс генерации кода можно взглянуть под другим углом. Можно генерировать не только код, который будет выполняться на машине, но и код, который может быть вставлен в “человеческий код”, т.е. код на высокоуровневом языке программирования. В дальнейшем этот код может быть добавлен к написанному вручную коду, модифицирован и скомпилирован, так же, как и любой другой код.

Благодаря тому, что данный подход отражает суть наших представлений о мире, а код пишется на понятном для человека языке, разработчик (да и не только он) может легко проверить данный код на наличие логических ошибок. Таким образом мы можем смоделировать поведение программы еще до момента ее компиляции и даже выполнения. Этот факт ускоряет поиск ошибок, а также позволяет устранять их в разы быстрее.

Работа с объектами и изменение их свойств гораздо проще протестировать. При тестировании отдельных модулей можно сосредоточиться на одной функциональности и провести тестирование гораздо тщательней, нежели чем тестировать большие блоки, плохо связанного кода.

Современная языки программирования с точки зрения генерации кода

23 мая 1995 года Джеймсом Гослингом был разработан язык для программирования бытовых электронных устройств который был назван Oak (“Дуб”). В дальнейшем язык был переименован в Java и стал использоваться для написания как клиентских приложений, так и серверного программного обеспечения. Такое название было дано в честь марки кофе, именно поэтому на логотипе изображена чашка с кофе [10].

Основной особенностью языка стал тот факт, что исходный код транслируется в байт-код выполняемый виртуальной машиной Java (JVM). Согласно спецификации JVM - это программа, которая обрабатывает байтовый код и передает инструкции оборудованию как интерпретатор. В рамках интерпретатора код анализируется построчно в отличии от компиляции, при которой, как уже было сказано, весь код транслируется в машинный.

Как уже было показано выше, такой подход, казалось бы, не слишком удачен. необходимо держать программу интерпретатор в памяти машины. Однако с тех пор прошло достаточно много времени и электронно- вычислительные машины стали гораздо более производительными и имеют достаточно много памяти. Кроме того, процесс написания таких компиляторов так же претерпел существенные изменения, был накоплен большой опыт по их созданию. Усовершенствовались и средства написания таких интерпретаторов.

До недавнего времени исполняемый на JVM код должен был быть написан на высокоуровневом языке программирования Java. Но в последнее время возникло много переосмыслений этого факта. Разработчикам необходима была гибкость в процессе разработки программного обеспечения. Компания JetBrains после пятилетней работы выпустила первый релиз языка Kotlin (Котлин — это небольшой остров в Финском заливе недалеко от Санкт-Петербурга) [11]. Особое влияние на новый язык и принцип его работы испытал от спецификации CLI.

Переосмысление процесса генерации кода (Метапрограммирование)

Разработчики переосмыслили процесс создания программ и начали рассматривать его в совокупности с уже известным и зарекомендовавшим себя процессом компиляции. Таким образом можно генерировать программный код на понятном для человека языке и потом итоговый код компилировать в исполняемый машинный код.

В данный момент именно данная тема является актуальной. В настоящее время, все более и более актуальной задачей становится создание программного продукта, позволяющего автоматизировать процесс написания прикладного программного обеспечения.

Метапрограммирование — это парадигма построения кода информационной системы с динамическим изменением поведения или структуры в зависимости от данных, действий пользователя или взаимодействия с другими системами [7]. Иными словами, метапрограммирование — вид программирования, связанный с созданием программ, которые порождают другие программы как результат своей работы (в частности, на стадии компиляции их исходного кода), либо программ, которые меняют себя во время выполнения (само модифицирующийся код).

Генерация кода на основе шаблонов

Для того чтобы сократить время для разработки стартовой (начальной) части проекта используются автоматизированные системы (АС) генерации программного кода. При этом подходе код программы не пишется вручную, а создается автоматически программой-генератором на основе другой программы.

Такой подход приобретает смысл, если при программировании вырабатываются различные дополнительные правила (более высокоуровневые парадигмы, выполнение требований внешних библиотек, стереотипные методы реализации определенных функций и пр.). При этом часть кода (или данных) теряет содержательный смысл и становится лишь механическим выполнением правил [10]. Когда эта часть становится значительной, возникает мысль задавать вручную лишь содержательную часть, а остальное добавлять автоматически. Это и прodelывает генератор.

Наиболее известным инструментом для быстрого создания стартового проекта является скаффолдер Yeoman. Скаффолдинг – это метод метапрограммирования, при котором проект генерируется на базе анализа требований разработчика [5]. Основная функция мета-программирования – это использование программ для создания других программ. Yeoman базируется на трех основных компонентах: менеджер зависимостей пакетов, необходимый для загрузки, обновления и удаления дополнительных пакетов (bower); инструмент для сборки java-script проектов с использованием заранее написанных задач (grunt); базовое приложение, отвечающее за генерацию базы для нового приложения (yo). Настройка всего проекта выполняется одной командой, которая запускает один из более 5500 генераторов кода, описанных на сайте [5].

Для разработчиков на платформе dotNet компания Microsoft разработала генератор кода Text Template Transformation Toolkit или T4 [2] на базе шаблонов. Для использования данного генератора необходимо написать файл-шаблон в формате *.tt. Существует большое количество готовых шаблонов для генерации стартового проекта под необходимые требования, однако из-за отсутствия единого репозитория поиск этих шаблонов крайне затруднен. Кроме того, T4 используется исключительно для генерации только C# кода. [1].

В стеке Java-технологий для генерации рутинного повторяющегося программного кода для решения таких типовых задач, как создание кода для конструкторов классов с большим количеством полей, создание и редактирование JavaDocs, геттеров/сеттеров, управление кодом коллекций и сборщиков композиционных объектов, управление кодом для тестирования также применяются многочисленные генераторы. Одни встраиваются в интегрированные среды разработки, другие являются самостоятельными системами, поставляемыми в виде, например, jar файлов: FreeBuilder, Proto. Интеграция внешних генераторов зачастую требует написания своего специфического рутинного кода, исчисляемого десятками строк.

В мире Web-разработки известны надстройки, называемые «синтаксический сахар», которые способствуют сокращению количества строк исходного кода за счет его генерации. К ним можно отнести SugarJS, CoffeeScript [7]. Подобные надстройки транслируют код, написанный на новых языках программирования, в код, написанный на языке Java-script. Самостоятельные шаблонизаторы, такие как Underscore templates, HandleBars, позволяют хранить параметрически настраиваемые шаблоны html страниц и их фрагментов на сервере, и по мере необходимости генерировать html код и пересылать его клиенту.

Итак, можно сделать выводы о том, что современные подходы и средства генерации кода не эффективны при разработке современных распределенных автоматизированных систем (АС), решающих задачи корпоративного, стратегического уровня, поскольку:

- способны генерировать код только для одного файла (класса);
- в большинстве случаев ориентированы на выполнение генерации по известным шаблонам;
- код для шаблона пишется самим разработчиком;
- в общем случае внесение изменений в системы генерации кода невозможна без перекомпиляции системы;
- Требуют установки стороннего программного обеспечения для генерации программного кода.

Генераторы кода лучше всего применимы в тех проблемных областях, в которых выполняются небольшие по объему и одинаковые по сути операции. Потенциально генераторы могут применяться в любых областях (задачах) где можно автоматизировать процесс работы с данными. Наиболее близкими к этому описанию являются задачи генерации кода проекта, взаимодействующего с базой данных, имеющего сходные классы и структуры для объектов, приложения для тестирования программного кода или написания документации по существующей программе.

Генерация кода на основе графического пользовательского интерфейса (GUI)(UML диаграммы)

Одним из самых распространенных методов написания кода на базе GUI (графического интерфейса) является метод построения кода на базе UML диаграмм.

Сейчас UML является языком широкого профиля, это — открытый стандарт, использующий графические обозначения для создания абстрактной модели системы, называемой UML-моделью. Он используется для визуализации проектирования и документирования. Для генерации исходного кода

программы используется диаграмма классов. Диаграмма классов (Class diagram) — статическая структурная диаграмма, описывающая структуру системы, демонстрирующая классы системы, их атрибуты, методы и зависимости между классами.

Так как язык строго типизирован, то на основе диаграмм можно сгенерировать каркас для проекта. Сейчас разработано много инструментальных систем, в том числе Rational Rose, осуществляющих кодогенерацию из UML в MS Visual C++, Java, Delphi, Power Builder, MS Visual Basic и др.

Генераторы программного кода на основе семантических сетей

В работах [1,4,6] представлен другой вариант генерации программного кода. Описанный в статьях метод основан на применении аппарата семантических сетей.

Общая схема генерации программного кода включает в себя два этапа. На первом этапе происходит шаблонизация и конфигурирование. На вход шаблонизатора передаются конструкции языка и изменяемые поля, на выходе же получается шаблон. Для процесса конфигурирования входными параметрами являются модели предметной области, на выходе получается итоговая конфигурация. Модели предметной области формализуют структурные и функциональные аспекты автоматизируемой предметной области и являются источником используемых в автоматизируемой системе имен, типов, ограничений. На втором этапе из полученных шаблона и конфигурации происходит рендеринг итогового программного кода.

Данная схема представлена на Рис.1. В данной системе шаблоны программного кода хранятся в особом формате внутри семантической сети. Шаблоны могут быть добавлены любым разработчиком программного обеспечения по особым правилам и в особом формате. В дальнейшем система сама будет выбирать, какой из шаблонов использовать, а разработчику не придется писать самостоятельно одни и те же шаблоны, которые постоянно используются в их работе. Все шаблоны будут находиться в единой сети и постоянно пополняться в случае изменения стандартов языка.

В общем случае данную методику можно использовать не только для генерации программного кода, будущих приложений, но также и для генерации тестов или каркаса документации, не только для кода, но и для других предметных областей.

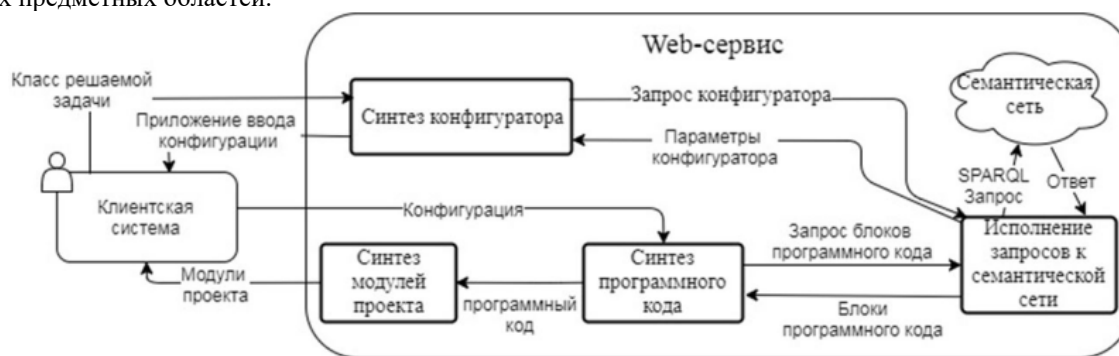


Рис. 1 Схема описания методики синтеза исходного кода проекта

В работе [1] представлена обобщенная модель генератора программного кода (ГПК) РОЮ. Данная модель содержит базовые концепты, формализующие рендер и конфигуратор программного кода, как подклассы компонента ГПК, с которым связаны точки входа/выхода потоков данных.

Кроме того, рендер и конфигуратор содержат шаблоны, которые реализуются с помощью простых атрибутов. Данные атрибуты представляют собой строки, которые содержат код программных конструкций (шаблоны текста) и параметры-подстановки. Данные параметры имеют однозначное соответствие со значениями дескрипторов соответствующих изменяемых полей. Изменяемые поля могут быть сгруппированы с помощью концепта, задающего интерфейс.

Ожидается, что развитие подхода к построению генераторов программного кода на основе семантических сетей позволит обеспечить ГПК свойствами интероперабельности и адаптивности и решить большинство проблем имеющихся в реализациях существующих ГПК.

Заключение

В настоящее время ведутся активные исследования в области синтеза программного кода. Данная работа направлена на методологический анализ процесса синтеза программного кода на основе исторических фактов.

Генераторы кода лучше всего применимы в тех проблемных областях, в которых выполняются одинаковые по сути операции и потенциально генераторы могут применяться в тех задачах, в которых можно автоматизировать процесс работы с данными. Примерами таких задач являются задачи генерации кода проекта и каркаса для документации. Так, например, процесс генерации программного кода проекта, который взаимодействует с базой данных и имеет сходные классы и структуры для объектов, может быть

автоматизирован, так как создается по одинаковому шаблону для каждого из объектов. При этом, количество возможных ошибок увеличивается в разы: неправильные формулы, остаточные данные и так далее. Использование генераторов в этих случаях заметно сокращает количество ошибок.

В ходе выполнения работы выяснилось, что придуманные еще в 80-х годах прошлого века алгоритмы работают до сих пор. Безусловно, они претерпевают изменения, связанные с адаптацией к текущим реалиям, но их основа остается неизменной. Алгоритмы, которые были придуманы в самом начале пути развития процесса кодогенерации и оказались неиспользуемыми по каким-либо причинам, сейчас показывают свою значимость, например, транслятор для Java кода.

Сегодня генераторы программного кода являются комплексными сложными системами, многоцелевыми инструментами, способными на основании моделей предметных областей строить программный код, понятный человеку, а это, в свою очередь, требует привлечения методов и средств поддержки принятия решения и искусственного интеллекта.

Литература

1. Жевнерчук Д.В. Семантическое моделирование генераторов программного кода распределенных автоматизированных систем / Жевнерчук Д.В. Захаров А.С // Журнал "Труды НГТУ им. Р.Е. Алексеева" Нижний Новгород: НГТУ: 2018 №1 (120) - С. 23-30
2. А.Ю. Молчанов. Системное программное обеспечение: Учебник для вузов. 3-е изд./ А.Ю. Молчанов // Учебник для вузов – с.130
3. Питер, 2009. Захаров В.Н. История информатики в России: Информатика-неограниченные возможности и возможные ограничения. / Захаров В.Н., Подловченко Р.И. // М:Наука, 2003.
4. Жевнерчук Д.В., Исследование интероперабельности систем мониторинга вычислительных процессов и ресурсов / Жевнерчук Д.В., Захаров А.С., Родионов П.А. // Журнал "Труды НГТУ им. Р.Е. Алексеева" Нижний Новгород: НГТУ: 2014 №1 (103) - С. 66-72
5. Yeoman: The web's scaffolding tool for modern webapps. 2017. [Электронный ресурс] // URL: <http://yeoman.io/>.
6. Zhevnerchuk D.V. Semantic Modeling Of The Program Code Generators For Distributed Automated Systems / Zhevnerchuk D.V., Lomakina L.S., Surkova A.S., Zakharov A.S // IV International Research Conference "Information Technologies in Science, Management, Social Sphere and Medicine". – (ITSMSSM 2017) . P 266-270
7. Википедия - свободная энциклопедия. 2018. [Электронный ресурс] // URL: <https://ru.wikipedia.org/wiki/>.
8. Н. А. Криницкий Г. А. Миронов Г. Д. Фролов. Программирование / под ред. М. Р. Шура-Бура. М. : Государственное издательство физико-математической литературы, 1963.
9. Евгеньев, Г. Б. Интеллектуальные системы проектирования: учеб. пособие / Г. Б. Евгеньев. – М.: Изд-во МГТУ им. Н.Э. Баумана, 2009. – 334 с.
10. Клыкков, Ю. И. Ситуационное управление большими системами / Ю. И. Клыкков. – М.: Энергия, 1974. – 134 с.
11. Kelly, S. Domain-Specific Modeling: enabling full code generation / S. Kelly, J. P. Tolvanen // Hoboken. – New Jersey, USA: Wiley-IEEE Computer Society Press. – 2008. – 444 p.

Захаров А.С., Жевнерчук Д.В. Анализ существующих подходов и перспектив развития интеллектуальных генераторов программного кода. В работе представлен анализ технологий синтеза программного кода, а также осмысление и понимание процесса эволюции генераторов программного кода. Показаны возможные перспективы развития средств интеллектуальной поддержки генераторов кода.

Ключевые слова: компиляция, кодогенерация, программный код, шаблоны, скаффолдинг

Zakharov A.S., Zhevnerchuk D.V. Analysis of existing approaches and prospects for the development of intelligent program code generators. The article provides an analysis technologies for the synthesis of software code, as well as understanding the process of evolution of program code generators. The possible prospects for the development of intellectual support for code generators are shown.

Key words: compilation, code generation, program code, patterns, scaffolding

Стандарты этики в искусственном интеллекте*

Ройзензон Г.В.

Институт системного анализа ФИЦ ИУ РАН, Москва

МФТИ, Долгопрудный, Московская область

МЭИ, Москва

5 rgv@isa.ru

Ройзензон Г.В. Стандарты этики в искусственном интеллекте. В 2016 IEEE выступила с глобальной инициативой в области этики искусственного интеллекта (этически обусловленное проектирование). В работе рассмотрены проекты стандартов IEEE этики ИИ. Проанализированы основные проблемы формализации понятия этики в ИИ. Рассмотрены основные типы рисков, связанные с внедрением в повседневную жизнь современных технологий, использующих ИИ. Проведен критический анализ различного математического инструментария, позволяющего формализовать понятие этики в ИИ. Предложены механизмы для успешного решения поставленной задачи. Приведены примеры решения практических задач.

Ключевые слова: искусственный интеллект, этика, стандарты IEEE, вербальный анализ решений

Введение

В 2016 году структура IEEE (Institute of Electrical and Electronics Engineers – Институт инженеров электротехники и электроники), параллельно с еще несколькими организациями (например, ЮНЕСКО), выступила с глобальной инициативой в области этики искусственного интеллекта (далее, ИИ) [10, 21, 33]. Важность предпринятых IEEE усилий определяется ее фактически ведущей ролью в сообществе ученых и инженеров в электротехнике, электронике, информационных технологиях, телекоммуникации и т.д., что обязательно окажет самое непосредственное влияние на разработчиков технологий ИИ. В результате усилий IEEE был разработан документ «Ethically Aligned Design» («Этически обусловленное проектирование») [33]. В документе IEEE отражены основные актуальные угрозы и риски [3], связанные с внедрением автономных систем на базе ИИ.

В соответствии с работами [3, 13] под измерением риска понимают определение опасности от того или иного источника (вида деятельности) для индивидуума или группы. Отметим основные четыре подхода, которые используются для измерения риска. Первый подход хорошо известен как инженерный. В рамках данного подхода основные усилия направлены на сбор статистических данных о поломках, авариях, связанных с утечкой вредных веществ в окружающую среду [8]. Инженерный подход ориентирован на количественный расчет вероятности поломок, отказов и других нежелательных событий. Второй подход принято называть модельным. Данный подход предполагает моделирование процессов, которые могут спровоцировать различные нежелательные последствия (аварии и т.п.) [8]. Третий подход к измерению риска известен как экспертный [3, 13, 17]. Такой подход к оценке риска применяют, когда возникают определенные сложности при использовании инженерного и модельного подходов. Четвертый подход измерения риска, известен как социологический [2, 13, 26]. В рамках данного подхода предполагается измерить восприятие населением и его отдельными группами того или иного риска.

Рассмотрим несколько небольших примеров, связанных с оценкой риска при внедрении в повседневную жизнь различных технологий ИИ. По причине перераспределения рабочей силы на постсоветском пространстве после распада СССР в Москве сосредоточено несколько сотен тысяч водителей автотранспорта (таксисты, водители маршруток, троллейбусов, автобусов и т.п.). Соответственно, если использование беспилотных автомобилей будет набирать современные темпы, то буквально через несколько лет значительная часть водителей – выходцев из стран ближнего зарубежья, останется без работы. Большинство этих людей другими профессиями быстро овладеть не смогут. Можно предположить, что для такого мегаполиса как Москва (как, впрочем, и для любого другого) такое развитие ситуации может легко привести к социальному взрыву, если предварительно не проделать глубокий анализ всех возможных последствий от внедрения тех или иных технологий ИИ (фактически нужно оценить социально-экономические риски).

Рассмотрим еще пример с моногородом (а это город, как хорошо известно, который сформировался и развился вокруг какого-то крупного промышленного предприятия). Предположим, на этом предприятии работает 20 тыс. человек. При этом используется технология, которая связана с достаточно низкой (по

*Работа выполнена при финансовой поддержке РФФИ (проекты № 16-07-00865, 16-29-12901, 16-29-12878).

современным меркам) производительностью труда, относительно высоким процентом брака готовой продукции и наносит достаточно существенный экологический ущерб. Есть новая технология, использующая ИИ, которая сразу решает большинство перечисленных проблем (рост производительности труда, уменьшение экологического ущерба и т.п.) и внедрение которой может снова сделать этот промышленный гигант конкурентным на мировом рынке. Но внедрение этой новой технологии предполагает увольнение с предприятия 18 тыс. человек, что также грозит моногороду социальными потрясениями, т.к. других предприятий, в которые можно было бы устроить работников по специальности, в городе попросту нет. Что делать в этой ситуации? Если, вообще ничего не менять, то предприятие разорится, т.к. станет совершенно не конкурентоспособным и вообще все 20 тыс. людей останутся без работы. Если внедрить самую передовую технологию, использующую ИИ, то тоже значительную часть сотрудников нужно будет уволить, но предприятие сохранится и в городе можно будет, например, улучшить экологию, а на вырученные от поступления в бюджет региона налоговые средства организовать новую программу по переподготовке кадров и т.п.

Поэтому, важность внедрения стандартов этики ИИ IEEE состоит в том, что очевидно эта проблема является междисциплинарной. При ее решении напрямую затрагиваются не только различные чисто этические проблемы, но и стоит задуматься о различных социально – экономических последствиях (рисках) от внедрения в нашу повседневную жизнь различных технологий, использующих ИИ.

В июле 2017 года была образована Российская рабочая группа IEEE по вопросам этики ИИ в составе: Готовцев П. М. (руководитель), Карпов В. Э., Овсянникова Е. Е. (секретарь) и Ройзензон Г. В. [22]. Основные цели группы:

представлять мнения и предложения российских ученых, участвующих в работах над документами IEEE: Ethically Aligned Design;

информировать российских ученых о результатах деятельности рабочей группы IEEE по созданию документа IEEE: Ethically Aligned Design;

привлекать российских ученых к исследованиям по тематике «Этика систем искусственного интеллекта».

В рамках инициативы IEEE предполагается разработка серии стандартов (см. табл. 1), применение которых, скорее всего, станет обязательным для всех специалистов и организаций, занятых в создании различных продуктов в той или иной степени использующих технологии ИИ. Кроме того, со своей стороны, организации и специалисты, использующие технологии ИИ, должны в самое ближайшее время сформулировать и внести свои предложения по выработке дополнительных условий соответствия упомянутым стандартам. Таким образом, в процедуре выработки стандартов должны принять участие несколько сторон (например, IEEE, разработчики и научное сообщество).

Основная цель данной работы — обзор существующего математического инструментария, который может быть использован для формализации понятия этики в ИИ. В частности, особый интерес представляют подходы, позволяющие оценивать на соответствие определенным требованиям (этическим нормам, критериям, стандартам и т. п.) те или иные технологии, использующие ИИ.

Прежде чем рассматривать конкретные математические подходы для решения поставленной задачи важно проанализировать сформулированные к настоящему моменту основные определения для используемых терминов.

Таблица 1. Проекты стандартов IEEE

Код проекта стандарта IEEE	Оригинальное название	Перевод
P7000	Model Process for Addressing Ethical Concerns During System Design	Проект стандарта модельного процесса для решения этических проблем при проектировании систем
P7001	Transparency of Autonomous Systems	Проект стандарта прозрачности автономных систем (АС)
P7002	Data Privacy Process	Проект стандарта обеспечения конфиденциальности данных
P7003	Algorithmic Bias Considerations	Проект стандарта учета необъективности алгоритма
P7004	Standard for Child and Student Data Governance	Проект стандарта управления данными детей и студентов
P7005	Standard for Transparent Employer Data Governance	Проект стандарта для прозрачного управления данными работодателя
P7006	Standard for Personal Data Artificial Intelligence (AI) Agent Standard for	Проект стандарта для интеллектуального агента управления персональными данными

	Personal Data Artificial Intelligence (AI) Agent	
P7007	Ontological Standard for Ethically Driven Robotics and Automation Systems	Проект онтологического стандарта робототехники и систем автоматизации, управляемых на основе этики
P7008	Standard for Ethically Driven Nudging for Robotic, Intelligent and Autonomous Systems	Проект стандарта для учета этических принципов в робототехнике, интеллектуальных и автоматизированных системах
P7009	Standard for Fail-Safe Design of Autonomous and Semi-Autonomous Systems	Проект стандарта проектирования безотказных автономных и полуавтономных систем
P7010	Wellbeing Metrics Standard for Ethical Artificial Intelligence and Autonomous Systems	Проект стандарта метрик благосостояния для систем искусственного интеллекта и автономных систем, действующих на основе этических принципов
P7011	Standard for the Process of Identifying and Rating the Trustworthiness of News Sources	Проект стандарта определения и оценки надежности новостных источников
P7012	Standard for Machine Readable Personal Privacy Terms	Проект стандарта для машиночитаемых правил конфиденциальности личных данных
P7013	Inclusion and Application Standards for Automated Facial Analysis Technology	Проект стандарта включения и применения технологии автоматизированного анализа состояния лица

Этика – философская дисциплина, исследующая вопросы морали и нравственности.

Относительно определения научного направления ИИ дело обстоит несколько сложнее. Нужно признать, что какого-то одного устоявшегося и единодушно принятого научным сообществом определения к настоящему моменту не выработано. Разработано огромное количество различных определений ИИ, сравнение и анализ которых явно выходит за рамки представленной работы. По мнению автора настоящей статьи под ИИ понимается группа методов и подходов, которые ориентированы на решение слабоструктурированных задач. Разумеется, представленное определение не претендует на истину в последней инстанции.

Прежде всего, нужно упомянуть классические фундаментальные работы по этике как отечественных, так и зарубежных специалистов, которые оказали существенное влияние на современные представления. В частности, можно отметить работы: Апресяна Р. Г. [1]; Гусейнова А. А. [4]; Дробницкого О. Г. [6]; Кропоткина П. А. [11]; Фролова И. Т. и Юдина Б. Г. [28]; Швейцера А. [30]; Шпемана Р. [31] и ряда других.

В истории науки можно привести несколько ярких примеров того как развитие новых технологий и вопросы этики приводили к весьма существенным противоречиям и столкновениям мнений политиков, общественных организаций, ведущих мировых ученых и т.п. При этом, важно обратить внимание на то, что в вопросах этики в рамках научно-технического прогресса инициатива исходила зачастую именно от самих ученых. В 1955 году Б. Рассел озвучил одну из первых таких инициатив (знаменитый манифест Рассела-Эйнштейна). Инициатива Б. Рассела положила начало широко известному теперь Пагуошскому движению за мир и разоружение [28], поскольку к середине 1950-х годов угроза всеобщей ядерной катастрофы стала очевидной и появилась необходимость мобилизовать авторитетнейших ученых (Ф. Жолио-Кюри, Л. Полинг, А. Эйнштейн и др.) для преодоления сложившейся критической ситуации. Еще одним таким ярким примером может служить инициатива американского генетика П. Берга, который в 1974 году предложил наложить мораторий на эксперименты с рекомбинантной ДНК для того чтобы иметь возможность оценить все риски и последствия использования этой новой технологии [28].

Основные выводы, которые можно сделать из подобных инициатив ведущих мировых ученых, заключаются в том, что развитие и неразумное использование тех или иных технологий может угрожать существованию человечества (например, ядерное оружие). Безусловно, повсеместное внедрение различных технологий ИИ также сопряжено с определенными опасностями и рисками [3], что во многом и стало причиной инициативы по разработке проектов специальных стандартов в области этики ИИ со стороны IEEE (см. Табл. 1).

Еще один важный вывод, который можно здесь сделать, заключается в том, что, с одной стороны, развитие тех или иных опасных технологий без учета вопросов этики, может привести человечество к совершенно катастрофическому результату. С другой стороны, если «загнать» развитие современных технологий в слишком жесткие рамки это станет причиной замедления темпов научно-технического прогресса (очень характерный пример с уже упомянутыми технологиями геной инженерии [28]).

Кроме того, есть еще один важный аспект, связанный с инициативой IEEE по разработке стандартов. К настоящему моменту на рынке появилось огромное количество продуктов, которые претендуют на то, что в них в той или иной степени используются технологии ИИ. В действительности при тщательном изучении оказывается, что представленные продукты вообще не имеют никакого отношения к рассматриваемой предметной области. Таким образом, предстоящая возможная достаточно жесткая сертификация продуктов,

использующих технологии ИИ, заставит многих недобросовестных производителей задуматься, прежде чем осуществлять различные маркетинговые действия (иными словами голословно заявлять, что в их продукции используются технологии ИИ). Соответственно разработка рассматриваемых стандартов может послужить определенным фильтром, т.к. многие производители продукции много раз подумают, стоит ли им заявлять, что их продукция использует технологии ИИ, зная, что им предстоит обязательная непростая процедура сертификации.

Анализ представленных классических работ по этике, а также некоторые выводы авторов работы [10] позволяют констатировать, что вопросы соотношения этики и ИИ коренным образом отличаются от того, что понимается, например, под этическими проблемами генных технологий, информатики, естествознания и т.п. Это отличие определяется тем, что в ИИ этические вопросы ближе к пониманию этики в философском или социогуманитарном смысле, и связаны эти этические аспекты прежде всего с тем, что они касаются вопросов поведения и принятия решений. Соответственно данный аспект существенно влияет на выбор математического инструментария для формализации понятия этики в ИИ.

1. Краткое описание проектов стандартов IEEE этики ИИ

Приведем более подробное описание целей проектов стандартов этики ИИ, представленных в табл. 1.

P7000. Инженеры, технологи и другие заинтересованные стороны проекта нуждаются в методологии для выявления, анализа и согласования этических проблем конечных пользователей с самого начала жизненного цикла систем и программного обеспечения. Цель этого стандарта - обеспечить практическое применение такого типа методологии проектирования на основе ценностей, которая демонстрирует, что концептуальный анализ ценностей и обширный анализ возможности их соблюдения могут помочь в улучшении требований этической составляющей в системах и жизненных циклах программного обеспечения. Этот стандарт предоставит инженерам и технологам полезный инструмент, выстраивающий процессы управления инновациями, подходы к разработке ИС и методы разработки программного обеспечения с целью минимизации этического риска для организаций, заинтересованных сторон и конечных пользователей.

P7001. Ключевой проблемой автономных систем (АС) является то, что их работа должна быть прозрачной для широкого круга заинтересованных сторон по ряду причин. (i) Прозрачность важна для пользователей, поскольку она формирует доверие к системе, предоставляя простой способ понять, что и почему делает система. Если взять в качестве примера робота, ухаживающего за пожилыми и больными людьми, то прозрачность означает, что пользователь может быстро понять, что может делать робот в разных обстоятельствах или, если робот должен сделать что-то неожиданное, пользователь должен иметь возможность спросить робота «почему это только что было сделано»? (ii) Валидация и сертификация прозрачности АС важны, поскольку они раскрывают процессы, происходящие в системе, для проведения проверки. (iii) Если происходят несчастные случаи, АС должна быть прозрачной при расследовании аварии; внутренний процесс, который привел к аварии, должен легко прослеживаться. После несчастного случая (iv) адвокатам или другим свидетелям-экспертам, которым может потребоваться предоставление доказательств, прозрачность необходима для демонстрации своих доказательств. Наконец, (v) для революционных технологий, таких как беспилотные автомобили, необходим определенный уровень прозрачности для более широких кругов общества, чтобы повысить доверие общественности к технологии. Для дизайнеров стандарт предоставит руководство для собственной оценки прозрачности в процессе разработки и даст механизмы для повышения прозрачности (например, необходимость защищенного хранения данных датчиков и данных о внутреннем состоянии аналогично регистратору данных полета или «черному ящику»).

P7002. Цель этого стандарта сформировать единый общий методологический подход, который определяет методы управления вопросами конфиденциальности в процессах жизненного цикла систем / программ.

P7003. Этот стандарт предназначен для того, чтобы обеспечить отдельным лицам или организациям, создающим алгоритмы, в основном предназначенные для автономных или интеллектуальных систем, ориентированных на сертификацию методологий, четко сформулированную отчетность и ясность в отношении того, как и с какой целью работают алгоритмы, как происходит оценка и как это влияет на пользователей и заинтересованные стороны, использующие указанный алгоритм. Сертификация в соответствии с этим стандартом позволит разработчикам алгоритмов сообщать пользователям и регулирующим органам, что в разработке, тестировании и оценке алгоритма использовались современные передовые методы, что, в свою очередь, позволит избежать необоснованных дифференцированных последствий для пользователей.

P7004. Этот стандарт предназначен для того, чтобы предоставить организациям, работающим с данными о детях и учащихся, процессы, ориентированные на управление, и сертификацию, гарантирующие прозрачность и подотчетность их действий в том, что касается безопасности и благополучия детей, их родителей, учебных заведений, в которых дети зарегистрированы, и сообщества и общества, где они проводят время, как в автономном режиме, так и в режиме онлайн. Стандарт также разработан для помощи родителям и преподавателям в понимании того факта, что большинство людей могут быть недостаточно технически грамотны, чтобы знать основные проблемы использования данных, но при этом должны быть

соответствующим образом проинформированы о защищенности данных о своих детях, и иметь в своем распоряжении инструменты и услуги, которые обеспечивают надлежащие возможности для выбора данных о семье на основе контента и полученной ранее информации.

P7005. Этот стандарт предназначен для предоставления организациям набора четких рекомендаций и сертификатов, гарантирующих, что они хранят, защищают и используют данные сотрудников этически и прозрачным способом. Он также предназначен для того, чтобы помочь работодателям понять, что большинство людей могут не обладать достаточными техническими навыками, чтобы понимать основные проблемы использования данных, но, тем не менее, должны быть соответствующим образом проинформированы о безопасности данных своих сотрудников, и иметь в своем распоряжении инструменты и услуги, которые обеспечивают надлежащие возможности для выбора информации, которой они обмениваются на рабочем месте, на основе контента и ранее полученной информации. Этот стандарт, разработанный по модели законодательства ЕС относительно правил защиты персональных данных GDPR (General Data Protection Regulation), будет сформулирован как разновидность «GDPR для работников». Он будет гарантировать, что работники, сталкивающиеся с широко распространенными проблемами автоматизации, связанными с потенциальной потерей рабочих мест, сохраняют контроль и управление своей персональной информацией, которая напрямую является базовым активом их идентичности и жизни, вне зависимости от того, получена ли она в результате мониторинга производственного процесса или из хранилища персональных данных.

P7006. С появлением ИИ и его развитием существует риск того, что решения, принимаемые в результате общения машин друг с другом, будут делаться на основе входных данных, имеющих характер «черного ящика» без прозрачности этих входных данных для людей. Для создания ИИ на основе этических принципов индивидуумам потребуются средства для управления и определения ценностей, правил и входной информации, которые направляли бы развитие персонализированных алгоритмов и ИИ. Им понадобится агент, который сможет представлять их индивидуальные права и возможности в системе общих социальных норм, этики и прав человека и который к тому же будет предвидеть этические последствия обработки данных и помогать их смягчить. Такой подход позволит людям безопасно организовывать и распространять свою личную информацию на машиночитаемом уровне и даст возможность персонализированному ИИ действовать в качестве их представителя при принятии решений машинами. Важнейшей целью создания этого стандарта является информирование государственных и коммерческих субъектов о том, что в их интересах предпочтительнее создание механизмов, предназначенных для индивидуальных пользователей, для обучения персональных агентов ИИ для преодоления этой асимметрии и гармонизации использования персональных данных в будущем.

P7007. Стандарт устанавливает набор определений и их взаимосвязей, которые обеспечат развитие робототехники и систем автоматизации в соответствии с мировыми теориями в области этики и морали с упором на приведение инженерных сообществ к пониманию того, как рационально разрабатывать такие системы и гармонично их внедрять. Эти определения обеспечат точное взаимопонимание экспертов в различных областях, включая робототехнику, автоматизацию и этику. Использование онтологий для представления знаний в любой области имеет несколько преимуществ, которые включают: а) формальное определение понятий конкретного домена в представлении, независимом от языка, то есть вне зависимости от конкретного языка программирования, однако с формальным описанием для реализации на целевом языке; б) инструменты для анализа понятий и их взаимоотношений в поисках несогласованности, неполноты и избыточности; в) язык для использования в процессе коммуникации между роботами от разных производителей; и т.п.

P7008. Стандарт для этически направленных роботизированных, интеллектуальных и автономных систем устанавливает набор определений и функций и взаимосвязей между ними с преимуществами, зависящими от культурных аспектов пользователей (благополучие, здоровье и т.п.), что позволит развить робототехнику, интеллектуальные и автономные системы в соответствии с мировыми теориями этики и морали с упором на приведение инженерных сообществ к пониманию того, как рационально разрабатывать такие системы и гармонично их внедрять. Этот стандарт вместе с определениями обеспечит точное взаимопонимание мировых экспертов в различных областях, включая робототехнику, автоматизацию и этику.

P7009. Этот стандарт устанавливает практическую техническую основу конкретных методологий и инструментов для разработки, внедрения и использования эффективных надежных механизмов в автономных и полуавтономных системах. Стандарт включает (но не ограничивается этим): четкие процедуры измерения, тестирования и сертификации способности системы надежно работать в интервале от слабых до сильных нагрузок, а также инструкции по ее усовершенствованию в случае неудовлетворительной работы. Стандарт служит основой для разработчиков, а также для пользователей и регуляторных органов в части создания устойчивых к отказам механизмов надежным, прозрачным и ответственным образом.

P7010. Стандарт метрики благополучия для этического искусственного интеллекта и автономных систем позволяет программистам, инженерам и технологам эффективнее анализировать, как именно продукты и услуги, которые они создают, могут повысить благополучие людей, на основе более широкого набора показателей, нежели только рост и производительность. Сегодня системы с датчиками распознавания эмоций

количественно оцениваются в первую очередь по своей экономической ценности на рынке за пределами сферы их применимости в определенных областях (психология и т. д.). Хотя часто понимается, что этические соображения для интеллектуальных и автономных систем могут препятствовать инновациям из-за введения нежелательного регулирования, тем не менее, без метрик, оценивающих психическое и эмоциональное здоровье, как на уровне индивидуума, так и общества, инновации невозможно количественно оценить. Внедрение и использование этих показателей для программистов и технологов означает, что благосостояние человека - помимо экономического роста - может быть измерено и улучшено.

P7011. Целью стандарта является устранение негативных последствий неконтролируемого распространения ложных новостей путем предоставления открытой системы легких для понимания рейтингов. При этом стандарт должен помогать в восстановлении доверия к определенным поставщикам новостей, и надлежащим образом дискредитировать других, подавлять стремление к публикации ложных новостей и пропагандировать пути улучшения для провайдеров, стремящихся к этому. Целью стандарта является репрезентативный образец набора новостных сообщений для формирования осмысленной и точной шкалы рейтингов.

P7012. Цель стандарта - предоставить людям средства для выставления своих собственных условий, касающихся личной конфиденциальности, способами, которые могут быть прочитаны, признаны и подтверждены машинами, которые управляются из сетевого мира. В более формальном смысле цель стандарта заключается в том, чтобы дать индивидуумам возможность быть главными в соглашениях с другими сторонами - в основном компаниями - выступающими в качестве вторых сторон. Следует обратить внимание, что целью настоящего стандарта является не рассмотрение политик конфиденциальности, поскольку они предполагают наличие только одной стороны и не требуют наличия соглашения. Для соблюдения условий необходимо наличие соглашения, а для соблюдения политики конфиденциальности оно не требуется.

P7013. Исследования продолжают демонстрировать, что ИИ, который используется для автоматизированного анализа лиц, подвержен тенденциозности, которая может усугублять человеческие предубеждения и систематически дискриминировать людей по признаку пола, этнической принадлежности, возраста и других факторов. Цель стандарта - предоставить руководящие принципы для разработки и сравнительной оценки автоматизированной технологии анализа лица для смягчения демографической и фенотипической предвзятости и дискриминации. Разделы и протоколы отчета, прописанные в настоящем стандарте, служат для повышения прозрачности этой автоматизированной технологии с тем, чтобы разработчики и лица, принимающие решения, могли сравнивать доступные варианты для выбора наиболее подходящей технологии на основе целевых групп населения и предполагаемых вариантов использования. Учитывая значимость биометрических данных, получаемых из анализа человеческого лица, стандарт также определяет в общих чертах надлежащее и ненадлежащее использование автоматизированного лицевого анализа на основе точности ценностей, установленных мировым сообществом.

2. Формальные методы

Проблема формализации этических норм включает в себя две основные задачи. Первая – это создание форм представлений норм, вторая – выбор соответствующего математического аппарата для работы с этими формами: сопоставления, измерения, анализа и т. д. Нечеткая, многозначная или вероятностная логика – это достаточно глубоко проработанные области, доведенные, вообще говоря, до уровня практически применимых технологий. Здесь гораздо важнее определиться с качественным уровнем представления параметров систем ИИ и этических норм. Важно отметить, что проблема формализации этических норм тесно связана с более общей задачей, а именно: с формализацией гуманитарного знания [16, 24, 26, 27].

Кроме того, в рамках этики ИИ требуется разработка новых норм, таких, например, как гуманность (как машины влияют на наше поведение и взаимодействие), сингулярность (как мы сможем контролировать сложную «умную» систему), безопасность [29] и т. п. Таким образом, не всегда соответствие тем или иным нормам можно свести к классическим «да» и «нет». Поэтому здесь актуально рассмотрение и использование различных неклассических логик (например, многозначных, темпоральных [7]), механизма многокритериальной классификации, вероятностных подходов и нечеткой логики, теории решеток [24] и т.п. Рассмотрим некоторые перечисленные подходы подробнее.

2.1. Булева алгебра

Концепция формализации различных этических понятий активно развивается на протяжении последних десятилетий. В качестве пионерской работы по исследуемому вопросу важно упомянуть книгу В. А. Лефевра «Алгебра совести» [16]. В этой книге есть целая глава, которая посвящена вопросам этики и возможным аспектам, связанным с формализацией этого понятия. Для решения рассматриваемой задачи в основном используется математический аппарат булевой алгебры. Это имеет как положительные, так и определенные отрицательные стороны. К положительным можно отнести то, что булева алгебра к настоящему моменту очень хорошо развита, есть множество приложений, программных библиотек для самых разных инструментальных

средств и т. п. К отрицательным можно отнести то, что не всегда различные этические проблемы (в том числе и относящиеся к ИИ) можно строго разделить на «белые» и «черные» [18], а механизм булевой алгебры зачастую предполагает именно такой подход. В работе Д. А. Поспелова [18] для преодоления данной проблемы вводится понятие «кольцевых» шкал, что для решения задачи формализации этики ИИ является весьма оригинальным и перспективным подходом.

2.2. Многозначные логики

В рамках развития различных неклассических парадигм в ИИ, в частности, подхода многозначных логик важно упомянуть работы отечественных специалистов, в частности А. С. Карпенко [9], В. К. Финна [26], О. П. Кузнецова [12], В. Б. Тарасова [25], В. Н. Вагина [5] и др. Использование многозначных логик для формализации понятия этики ИИ также сопряжено с определенными сложностями. В частности, переход от трехзначной логики к четырехзначной может потребовать кардинальной «переделки» соответствующих математических конструкций, что фактически означает необходимость решения указанной задачи заново.

2.3. Теория вероятностей и нечеткая логика

В этом смысле использование вероятностного аппарата и нечеткой логики [34] для формализации понятий этики ИИ [32] является весьма интересным подходом, так как нечеткую логику можно считать неким обобщением многозначной логики [25]. К известным особенностям применения нечеткой логики можно отнести определенные проблемы при построении функций принадлежности: разные способы построения таких функций приводят к разным результатам (неустойчивость методов нечеткой логики относительно исходных данных).

2.4. Вербальный анализ решений

Еще одним возможным подходом для формализации понятия этики ИИ является использование методов вербального анализа решений (ВАР) [14].

Группа методов ВАР опирается на использование достижений различных научных дисциплин: когнитивной психологии (измерения, операции получения информации при построении решающего правила, поэтапное построение решающего правила); прикладной математики (обоснование выдвинутого правила, методов получения и проверки информации на непротиворечивость); теории организаций (получение объяснений); компьютерных наук (диалог человек-компьютер). Разработанные в рамках этого подхода методы принятия решений позволяют при анализе вариантов сложных решений органично сочетать качественную и количественную информацию об альтернативах, суждения экспертов и предпочтения лиц, принимающих решения, объективные и субъективные факторы, характерные для проблемной ситуации.

Например, применительно к обозначенной проблеме формализации этики в ИИ, с использованием методов ВАР, возможна следующая постановка задачи. Если в рамках этики ИИ разработать некоторый перечень норм, то степень соответствия той или иной норме можно рассматривать как задачу многокритериальной порядковой классификации [14, 17]. Соответственно, на основе анализа таких норм этики ИИ мы должны будем принять решение о том, что либо нормы соблюдены, либо есть некоторое несущественное их нарушение, либо наблюдается какой-то заметный отход от принятых норм и т. п. То есть нам будет нужно отнести определенную совокупность оценок по каждой из норм к некоторому классу решений (категории).

К положительным сторонам использования методов ВАР, прежде всего, можно отнести то, что к исходным данным не применяются никакие операции по их переводу в количественную форму. Известно, что перевод вербальных измерений в «цифру» зачастую весьма субъективен и не имеет строгого математического обоснования. Кроме того, методы ВАР позволяют получить объяснения принятых решений (интерпретация результата) в терминах предметной области, здесь – в терминах описания норм этики ИИ. В качестве недостатков методов ВАР можно отметить большие трудозатраты эксперта или лица, принимающего решения, при работе в признаковом пространстве большой размерности. В этом случае необходимо применять различные методы снижения его размерности [19, 20].

Таким образом, можно констатировать, что к настоящему моменту разработан широкий спектр инструментальных средств, опирающихся на самые различные математические конструкции, позволяющие успешно решить задачу формализации этических норм в ИИ.

3. Пример: оценка надежности новостных источников и проект стандарта P7011

Последнее десятилетие характеризуется беспрецедентными масштабами ведения различных информационных войн. Проект стандарта IEEE этики ИИ P7011 (оценка надежности новостных источников) ориентирован на устранение негативных последствий неконтролируемого распространения поддельных («фейковых») новостей путем предоставления открытой системы оценок. Очевидно, что «фейковые» новости

являются одним из важнейших инструментов информационных войн, поэтому любые инициативы, направленные на «оздоровление» информационного пространства представляют определенный интерес. Если в самое ближайшее время не предпринять соответствующих усилий по предотвращению распространения «фейковых» новостей, это может привести к практически полной потере доверия конечных потребителей новостного контента к масс-медиа (телевидение, радио, интернет-ресурсы и т.п.). Потеря аудитории, в свою очередь, приведет к оттоку рекламодателей от производителей медиаконтента. А уже результатом этого будет фактическое разорение и банкротство большинства медиа-холдингов. Поэтому во внедрении стандарта этики ИИ P7011 в первую очередь заинтересованы сами производители медиаконтента. Для решения указанной задачи, очевидно, могут быть использованы методы лингвистической семантики и семантического анализа текстов (семантические технологии Web). Тем не менее, в современных условиях при распространении «фейковых» новостей необходимо проводить анализ и других компонентов (например, видеоряд, фотографии, звук и т.п.), т.к. они также могут быть фальсифицированы. Поэтому разработка специальной системы критериев для оценки надежности новостных источников может, в определенных условиях, также рассматриваться и как система этических норм в рамках стандарта этики ИИ P7011. Можно предложить следующую модельную систему критериев: 0. «Кричащий» заголовок (оценки: 0. Заголовок новости является «Кричащим»; 1. Нормальный заголовок); 1. Характеристика источника (оценки: 0. Подозрительный источник (много рекламы, странный дизайн, URL и т.п.); 1. Нормальный источник); 2. Неверная дата публикации (оценки: 0. Дата публикации очевидно не соответствует появлению новостного сообщения; 1. Правильная дата); 3. Подложные фотографии (оценки: 0. Очевидно поддельные фотографии; 1. Достоверность фотографий вызывает определенные сомнения; 2. Нормальные фотографии); 4. Обилие ошибок (оценки: 0. Много орфографических и синтаксических ошибок; 1. Есть некоторое количество орфографических и синтаксических ошибок; 2. Ошибок нет); 5. Давление на жалость (оценки: 0. Присутствует давление на жалость; 1. Давления на жалость нет). Классы решений: А. Определенно фейковая новость. В. Есть подозрение, что новость является фейковой. С. Новость не является фейковой.

Рассмотрим более подробно применение метода ВАР ЦИКЛ [14, 17], предназначенного для построения многокритериальной порядковой классификации с привлечением ЛПР или эксперта, применительно к проблеме формализации понятия этики ИИ. В результате работы метода формируется верхние и нижние границы классов решений (фактически это наборы решающих правил), которые позволяют классифицировать любые наборы оценок по критериям (в нашем случае, этических норм).

Построенная классификация выглядит следующим образом: Класс А (верхняя граница): (000000); Класс А (нижняя граница): (100010) (001120) (000220) (001201) (001111) (000211) (001021) (000121); Класс В (верхняя граница): (010000) (101000) (100100) (001210) (100020) (100001) (001121) (000221); Класс В (нижняя граница): (111200) (101210) (111120) (100220) (111101) (111021) (110121) (101121) (011221); Класс С (верхняя граница): (110210) (101220) (100201) (111111); Класс С (нижняя граница): (111221).

Нужно отметить, что если бы в исходной системе критериев использовались только бинарные оценки на шкалах критериев, и новости нужно было разделить только на два класса (категории), то можно использовать, например, математический аппарат булевой алгебры, т.е. применить подход, предложенный В. А. Лефевром [16] или методы расшифровки булевых функций [23], а также более эффективный с вычислительной точки зрения, по сравнению с методом ВАР ЦИКЛ, метод ВАР ДИФКЛАСС [15] для решения задачи формализации понятия этики ИИ.

Заключение

В представленной статье рассмотрена инициатива IEEE по разработке проектов стандартов этики ИИ. Предложены возможные определения этики и ИИ. Проанализирована особенность применения некоторых норм этики применительно к разработке и использованию современных технологий ИИ. Представлен перечень проектов стандартов этики ИИ IEEE (P7000 – P7013), что говорит о достаточно широком спектре проблем, с которыми в самое ближайшее время столкнутся разработчики различных систем ИИ. Фактически это означает, что разработчики систем ИИ уже сейчас должны начать подготовку к прохождению различных процедур сертификации (т.е. соответствия их продукции, использующей технологии ИИ, указанным стандартам ИИ). В этой связи критический анализ различного математического инструментария, позволяющего формализовать понятия этики ИИ, будет способствовать разработке понятных и прозрачных «правил игры». К формальным методам, использование которых весьма востребовано для поставленной в статье задачи, можно отнести следующие: булева алгебра, многозначные логики, нечетная логика и теория вероятностей, а также методы ВАР. Благодаря тому, что к настоящему моменту все перечисленные концепции весьма развиты, можно достаточно оптимистично смотреть на решение задачи формализации этики ИИ. В качестве примера рассмотрена формализация понятия этики для проекта стандарта P7011 (оценка надежности новостных источников) с использованием метода ВАР ЦИКЛ. Из нерешенных задач необходимо отметить, что для каждого из проектов стандартов IEEE необходима разработка индивидуальных систем критериев (фактически это некий перечень норм) и соответствующих шкал оценок, по которым можно будет принять решение о степени соответствия той или иной технологии ИИ предложенным этическим стандартам.

Литература

1. Апресян, Р. Г. Этика: учебник / Р. Г. Апресян. — М.: КноРус, 2017.
2. Бек, У. Общество риска. На пути к другому модерну / У. Бек. — М.: Прогресс-Традиция, 2000.
3. Бритков, В. Б. Междисциплинарный подход к анализу риска / В. Б. Бритков, Г. В. Ройзензон // IX Московская международная конференция по исследованию операций (ORM2018). Труды. В двух томах / Под ред. Ф. И. Ерешко. — Т. 2. — М.: МАКС Пресс, 2018. — С. 340–345.
4. Гусейнов, А. А. Античная этика / А. А. Гусейнов. — М.: URSS, 2017.
5. Достоверный и правдоподобный вывод в интеллектуальных системах / В. Н. Вагин, Е. Ю. Головина, А. А. Загорянская, М. В. Фомина; Под ред. В. Н. Вагина, Д. А. Поспелова. — М.: Физматлит, 2008.
6. Дробницкий, О. Г. Моральная философия: Избранные труды / О. Г. Дробницкий. — М.: Гардарики, 2002.
7. Еремеев, А. П. Темпоральные модели на основе логики ветвящегося времени в интеллектуальных системах / А. П. Еремеев, И. Е. Куриленко // Искус. интеллект и принятие решений. — 2011. — № 1. — С. 14–26.
8. Интеллектуальные системы поддержки принятия решений в нештатных ситуациях с использованием информации о состоянии природной среды / В. А. Геловани, А. А. Башлыков, В. Б. Бритков, Е. Д. Вязилов. — М.: Эдиториал УРСС, 2001.
9. Карпенко, А. С. Развитие многозначной логики / А. С. Карпенко. — 3-е изд. — М.: URSS, 2016.
10. Карпов, В. Э. К вопросу об этике и системах искусственного интеллекта / В. Э. Карпов, П. М. Готовцев, Г. В. Ройзензон // Философия и общество. — 2018. — № 2(87). — С. 84–105. — DOI: 10.30884/jfio/2018.02.07.
11. Кропоткин, П. А. Этика: Избранные труды / П. А. Кропоткин. — М.: Политиздат, 1991.
12. Кузнецов, О. П. Неклассические парадигмы в искусственном интеллекте / О. П. Кузнецов // Известия РАН. Теория и системы управления. — 1995. — № 5. — С. 3–23.
13. Ларичев, О. И. Проблемы принятия решений с учетом факторов риска и безопасности / О. И. Ларичев // Вестник АН СССР. — 1987. — Т. 57, № 11. — С. 38–45.
14. Ларичев, О. И. Вербальный анализ решений / О. И. Ларичев. — М.: Наука, 2006.
15. Ларичев, О. И. Система ДИФКЛАСС: построение полных и непротиворечивых баз экспертных знаний в задачах дифференциальной классификации / О. И. Ларичев, А. А. Болотов // Научно-техническая информация. Серия 2. Информационные процессы и системы. — 1996. — № 9. — С. 9–15.
16. Лефевр, В. А. Алгебра совести / В. А. Лефевр. — М.: «Когито-Центр», 2003.
17. Метод многокритериальной классификации ЦИКЛ и его применение для анализа кредитного риска / А. А. Асанов, О. И. Ларичев, Г. В. Ройзензон и др. // Экон. и мат. методы. — 2001. — Т. 37, № 2. — С. 14–21.
18. Поспелов, Д. А. «Серые» и/или «черно-белые» / Д. А. Поспелов // Прикладная эргономика. Специальный выпуск «Рефлексивные процессы». — 1994. — № 1. — С. 29–33.
19. Ройзензон, Г. В. Способы снижения размерности признакового пространства для описания сложных систем в задачах принятия решений / Г. В. Ройзензон // Новости искус. интеллекта. — 2005. — № 1. — С. 18–28.
20. Ройзензон, Г. В. Синергетический эффект в принятии решений / Г. В. Ройзензон // Системные исследования. Методологические проблемы. Ежегодник / Под ред. Ю. С. Попкова, В. Н. Садовского, В. И. Тищенко. — №36. 2011-2012. М.: УРСС, 2012. — С. 248–272.
21. Ройзензон, Г. В. Проблемы формализации понятия этики в искусственном интеллекте / Г. В. Ройзензон // Шестнадцатая национальная конференция по искусственному интеллекту с международным участием (КИИ-2018). Труды конференции. В 2-х томах. — Т. 2. — М.: РКП, 2018. — С. 245–252.
22. Российская рабочая группа IEEE по вопросам этики ИИ. — 2017. — Режим доступа: <http://ecai.raai.org/>.
23. Соколов, Н. А. Оптимальная расшифровка монотонных булевых функций / Н. А. Соколов // Журнал вычислительной математики и математической физики. — 1987. — Т. 27, № 12. — С. 1878–1887.
24. Таран, Т. А. Булевы модели рефлексивного управления в ситуации выбора / Т. А. Таран // Автоматика и телемеханика. — 2001. — № 10. — С. 103–117.
25. Тарасов, В. Б. От многоагентных систем к интеллектуальным организациям: философия, психология, информатика / В. Б. Тарасов. — М.: Эдиториал УРСС, 2002.
26. Финн, В. К. Интеллектуальные системы и общество: Сборник статей / В. К. Финн. — М.: КомКнига, 2006.
27. Фоминых, И. Б. О формализации гуманитарного знания / И. Б. Фоминых // Одиннадцатая национальная конференция по искусственному интеллекту с международным участием (КИИ-2008). Труды конференции. — Т. 1. — М.: Ленанд, 2008. — С. 133–141.
28. Фролов, И. Т. Этика науки: Проблемы и дискуссии / И. Т. Фролов, Б. Г. Юдин. — М.: URSS, 2016.
29. Цыгичко, В. Н. Безопасность критических инфраструктур / В. Н. Цыгичко, Д. С. Черешкин, Г. Л. Смолян. — М.: URSS, 2019.
30. Швейцер, А. Культура и этика / А. Швейцер. — М.: Прогресс, 1973.
31. Шпеман, Р. Основные понятия морали / Р. Шпеман. — М.: Московский философский фонд, 1993.
32. Шрейдер, Ю. А. Проблема неполного добра в модели ценностной рефлексии по В. А. Лефевру / Ю. А. Шрейдер, Н. Л. Мухелишвили // Системные исследования. Методологические проблемы. Ежегодник / Под ред. Д. М. Гвишиани, В. Н. Садовского. — № 25. 1997. М.: УРСС, 1997. — С. 213–224.
33. IEEE. Ethically Aligned Design. — 2016. — Режим доступа: <https://ethicsinaction.ieee.org>.

34. Zadeh, L. A. Fuzzy sets / L. A. Zadeh // Information and Control. — 1965. — Vol. 8, no. 3. — P. 338–353.

Ройзензон Г.В. Стандарты этики в искусственном интеллекте. В 2016 IEEE выступила с глобальной инициативой в области этики искусственного интеллекта (этически обусловленное проектирование). В работе рассмотрены проекты стандартов IEEE этики ИИ. Проанализированы основные проблемы формализации понятия этики в ИИ. Рассмотрены основные типы рисков, связанные с внедрением в повседневную жизнь современных технологий, использующих ИИ. Проведен критический анализ различного математического инструментария, позволяющего формализовать понятие этики в ИИ. Предложены механизмы для успешного решения поставленной задачи. Приведены примеры решения практических задач.

Ключевые слова: искусственный интеллект, этика, стандарты IEEE, вербальный анализ решений

Royzenzon G.V. Ethical standards in artificial intelligence. In 2016, IEEE launched a global initiative in the field of ethics of artificial intelligence (ethically aligned design). The paper reviewed the draft standards of IEEE ethics AI. The main problems of the formalization of the concept of ethics in AI are analyzed. The main types of risks associated with the introduction of modern technologies using AI into the daily life are considered. A critical analysis of various mathematical tools to formalize the concept of ethics in AI has been carried out. The proposed mechanisms for the successful solution of the problem. Examples of solving practical problems are given.

Keywords: artificial intelligence, ethics, IEEE standards, verbal decision analysis

Экспертно - аналитический метод оценки качества конкурирующей продукции

Скрыпникова М.К.

Московский университет связи и информатики.

E-mail: skrmarina@yandex.ru

Скрыпникова М.К. Экспертно - аналитический метод оценки качества конкурирующей продукции. Поставлена задача разработки экспертно-аналитической системы оценки качества конкурирующей продукции. Рассмотрены методические вопросы такой оценки. Предложены оценочные показатели для экспертной оценки. Намечены основные этапы работы, которые в основном связаны с разработкой очередной ветки соответствующего программного продукта.

Ключевые слова: статья, программная инженерия, экспертный метод, оценка качества, теория вероятностей, математическая статистика.

Введение

Проблема выбора конкурирующей продукции является очень актуальной. Приобретаемое оборудование обладает достаточно разнообразными возможностями выбора, предлагаемых как традиционными организациями, так и поставщиками из других областей техники, в том числе конверсионными предприятиями, совместными и зарубежными фирмами. При этом предложение подчас превышает спрос.

Такое положение осложняет задачу выбора оптимального (рационального) варианта решения из нескольких возможных. В связи с этим проводятся конкурсы, тендеры и другие подобные мероприятия, помогающие решению возникшей проблемы с помощью отработанных процедур. Однако при этом оставалась, по существу, нерешенной задача обоснованного сравнения технического уровня конкурирующего оборудования.

Наличие сотен и тысяч различных параметров и показателей у отдельных устройств и систем в целом делает практически невозможным их качественное сравнение по принципу "подобный с подобным, лучше - хуже" [1,2]. Специалисты, которым поручено провести сравнение, в этих условиях на основе личных пристрастий и суждений обычно неявно выделяют несколько, по их мнению, важнейших показателей и по ним проводят сравнение. У разных специалистов суждения, естественно, разные, поэтому результаты сравнения получаются несовпадающими. В действие вступают иные, "не технические" соображения, и предпочтение, в конце концов, может быть отдано именно им. Таким образом, обоснованность результатов сложных и дорогостоящих конкурсных мероприятий в значительной мере снижается.

"Не техническими" соображениями, конечно, пренебрегать нельзя, но эффективность выбора можно повысить, если дополнить существующие методы качественного анализа и сопоставления конкурирующих разработок количественными способами оценки их технического уровня.

В связи с вышеизложенным предлагается разработать экспертно - аналитический метод и соответствующее программное обеспечение оценки качества и технического уровня конкурирующей продукции, совместное использование которых позволила бы повысить адекватность принимаемых решений при выборе той или иной продукции. Кроме того, решение данной задачи повышает оперативность, достоверность и объективность выбора [6].

Основным преимуществом количественной оценки является возможность обоснованного предпочтения одного изделия другому в том случае, когда (как это чаще всего и бывает) некоторые показатели лучше у первого изделия, а некоторые - у второго. Применяемые обычно подходы не позволяют в таких ситуациях выносить объективное решение. Кроме того, использованный метод дает возможность сопоставлять изделия и по показателям, не имеющим количественной меры (например, уровень апробации изделия на объектах или степень его проработанности).

Предлагаемый метод основан на том, что комплексный показатель технического уровня [4,5], по которому сравнивались конкурирующие изделия, определялся для каждого i -го изделия как взвешенное среднее арифметическое оценок q_{ij} единичных показателей качества и технического уровня

$$T_i = \sum_{j=1}^n (s_j q_{ij})$$

где s_j - коэффициент весомости j -того показателя качества; n -число показателей.

Основными этапами работы при этом являются:

1. Сбор технических показателей и показателей качества конкурирующей продукции и обобщение их для анализа компетентными организациями (например, показателей надежности и т.д.).
2. Создание экспертной группы из разработчиков, изготовителей и потребителей рассматриваемого оборудования.
3. Сбор экспертных данных и расчет весовых коэффициентов выбранных показателей изделия с помощью программного обеспечения, которое будет разработано в рамках данной работы.
4. Сбор экспертных данных по параметрам оборудования и обработка с помощью соответствующего программного обеспечения.
5. Результатами работы программного обеспечения по пунктам 3 и 4 являются соответствующие протоколы, в последнем в конечном итоге представлены ранжированные по степени предпочтительного использования варианты конкурирующей продукции.
6. Правильность работы программного обеспечения контролируется пропуском тестовой программы в ручном или автоматическом режиме.

Сбор технических показателей конкурирующей продукции и обобщение их для анализа компетентными организациями.

Первый и один из наиболее ответственных этапов - формирование номенклатуры показателей качества и технического уровня оборудования. Главная проблема состоит в противоречии между стремлением охватить все основные функциональные, конструктивные, производственные и эксплуатационные характеристики оборудования и необходимостью так ограничить число показателей, чтобы номенклатура была обозримой и удобной для экспертных оценок. Поэтому ряд характеристик предполагается объединить в укрупненные показатели и одновременно по каждому такому показателю должна быть подготовлена подробная расшифровка, чтобы иметь единую для всех экспертов трактовку содержания позиций номенклатуры.

Если в конкурирующих изделиях имеется свое оригинальное программное обеспечение-ПО, то в этом случае особое значение должно придаваться:

- концепции организации ПО, способам разработки прикладного ПО, организации баз данных, хранению программ, загрузки кода, интерпретации входного языка, генерации видеопрограмм, архивов, ведомостей, решению вычислительных задач и т.д.;
- наличию оболочек баз знаний, средств прогнозирования поведения конкурирующих изделий при наличии внештатных ситуаций (отказов, сбоев) и формирования советов оператору;
- сведениям об используемых технологических языках программирования;
- возможностям корректировки ПО пользователем (степени открытости ПО), возможности расширения состава задач;
- степени защиты от несанкционированного доступа и сведениям о тестировании ПО сторонними организациями, компетентными в рассматриваемой области конкурирующих изделий.

В качестве примера базового варианта предлагается следующая структура и содержание номенклатуры показателей для конкурирующих изделий .

Показатели целесообразно разбить на группы, чтобы облегчить экспертизу и последующую обработку результатов. Прежде всего, выделяются классификационные и оценочные показатели. Классификационные показатели предполагается должны включать в себя следующие позиции:

- предназначение данного оборудования;
- режим функционирования оборудования: информационный, локально-автоматический, советующий вместе с локально-автоматическим, полностью автоматический;
- уровень интегрированности оборудования.

Сами по себе классификационные показатели не участвуют в оценке технического уровня, но они необходимы для обеспечения корректности сопоставления различного оборудования, указывая на область применения.

Оценочные показатели предлагается разделить на две основные группы: характеристики оборудования и рейтинг поставщика. В свою очередь в характеристиках оборудования целесообразно выделить подгруппы: функциональных показателей (четыре единичных показателя), технической эффективности (два единичных показателя) и надежности (два единичных показателя), а в рейтинг поставщика включить шесть единичных показателей. Полная номенклатура оценочных показателей приведена ниже.

Оценочные показатели.

1. Характеристики изделий

1.1. Функциональные:

1.1.1. Уровень информационной мощности изделия;

- 1.1.2. Полнота выполнения информационно-вычислительных функций;
- 1.1.3. Полнота выполнения сервисных функций.
- 1.2. Технической эффективности:
 - 1.2.1. Уровень основных характеристик комплекса технических средств конкурирующих изделий;
 - 1.2.2. Уровень программного обеспечения .
- 1.3. Надежности:
 - 1.3.1. Уровень аппаратной надежности ;
 - 1.3.2. Уровень самодиагностики ;
- 2. Рейтинг поставщика
 - 2.1. Степень проработанности изделия;
 - 2.2. Степень компетентности фирмы-разработчика в области использования изделия;
 - 2.3. Уровень апробации конкурирующего изделия на объектах;
 - 2.4. Уровень производства у изготовителя и поставщика;
 - 2.5. Степень отработанности сервисного обслуживания;
 - 2.6. Степень комплектности поставок технических средств изделия, ПО и ЗИП для объектов;

Создание экспертной группы из разработчиков, изготовителей и потребителей рассматриваемого оборудования.

Пользователь на основании нормативного документа, регламентирующего порядок работы при решении данной задачи, выполняет следующие работы:

- формирует экспертную группу;
- последовательно (разница во времени примерно от 10 дней до месяца в зависимости от степени быстроты организации выполнения экспертных процедур) составляет первую (для расчета весовых коэффициентов) анкету и вторую (для расчета показателей качества и технического уровня) и рассылает их членам экспертной группы;
- собирает, систематизирует и анализирует информацию, содержащуюся в возвращенных анкетах, заполняет входные формы.

Сбор экспертных данных и расчет весовых коэффициентов выбранных параметров с помощью соответствующего программного обеспечения.

Следующий необходимый этап - экспертное определение весомости каждого из оценочных показателей в общем техническом уровне изделия. При этом предполагается решить несколько задач:

- создание организационного обеспечения экспертиз данного класса;
- разработка и апробация алгоритмического, математического и программного обеспечения статистической обработки материалов экспертизы;
- апробация и коррекция разработанной номенклатуры показателей;
- определение коэффициентов весомости оценочных показателей.

Оценку весомости эксперты должны выполнить последовательно в несколько этапов: сначала сопоставить вклады в общий уровень укрупненных групп, т.е. назначить оценки s_g весомости характеристик изделий и рейтинга поставщика [4,5]. Затем оценить коэффициенты весомости s_p трех подгрупп 1.1, 1.2 и 1.3 внутри группы 1 характеристик изделия и далее - коэффициенты весомости s_{jp} единичных оценочных показателей внутри каждой подгруппы и внутри группы рейтинга поставщика. Это позволит при каждой оценке предъявлять эксперту лишь от двух до шести показателей, что облегчит задачу экспертизы.

Итоговые оценки весомости единичных показателей могут быть определены по формулам

$$s = s_g s_p s_{jp} \text{ для первой группы}$$

$$\text{и } s_j = s_g s_{jp} \text{ для второй.}$$

Коэффициенты весомости s_g групп, s_p подгрупп в группе, s_{jp} единичных показателей внутри подгруппы и итоговых показателей s_j удовлетворяют условиям нормирования, т.е. сумма их значений внутри каждого массива равна единице.

Обработка результатов анкетирования экспертов будет включать, прежде всего, статистический анализ достоверности и непротиворечивости оценок каждого коэффициента весомости.

Сбор экспертных данных по параметрам оборудования и обработка с помощью программного обеспечения.

Для определения уровня единичных оценочных показателей каждого изделия экспертам может быть предложена следующая пятибалльная шкала [3]:

- 5 - исключительно высокий уровень;
- 4 - очень хороший уровень;
- 3 - достаточно хороший уровень;
- 2 - удовлетворительный уровень;
- 1 - неудовлетворительный уровень.

Экспертам должны быть даны возможности указывать и любые промежуточные оценки.

Практикуемые обычно методы оценки той или иной разработки фактически тоже являются балльными. Но они используют лишь двухбалльную шкалу ("годен" или "не годен") и дают сразу интегральную оценку всей разработке, а не последовательно ее отдельным показателям, что снижает обоснованность принятого решения.

Обработка результатов экспертизы должна заключаться в определении средних по всем экспертизам значений любого из оценочных показателей каждого изделия и в проверке их статистической достоверности и согласованности.

Результатами работы программного обеспечения по пунктам 3 и 4 являются соответствующие протоколы.

Разрабатываемая в рамках настоящей работы диалоговая программная система (ДПС) это организационно-техническая система, элементами которой являются компоненты математического, лингвистического, программно-технического, информационного и организационного обеспечений. В ее состав входят четыре подсистемы: обучения, расчета весовых коэффициентов, расчета показателей качества, а также тестирования системы. Каждая подсистема ДПС состоит из компонентов, объединенных общей целевой функцией.

Основными признаками, по которым классифицируются компоненты, является их принадлежность к определенному виду обеспечений (математическому, программному и т.д.).

Компоненты математического обеспечения представляют собой документы, содержащие описания методов и алгоритмов решения отдельных задач.

Каждый компонент программного обеспечения включает в себя документ с текстами программ, программы на магнитном носителе, эксплуатационные документы [7,8].

Компонент технического обеспечения представляет собой комплекс технических средств, необходимых для обеспечения функционирования системы.

Компоненты информационного обеспечения - это информационные массивы и описания массивов, применяемых при решении задач.

Компонент методического обеспечения представляют собой инструкцию по эксплуатации системы, регламентирующую порядок проведения работ с ее использованием.

Все этапы количественной оценки технического уровня обеспечены диалоговой программной системой, которая:

- готовит комплекты документации для сбора данных о характеристиках конкурирующих изделий и для анкетирования экспертов;
- оформляет номенклатуру показателей качества и их расшифровку;
- выполняет статистическую обработку экспертных оценок и определяет их достоверность и согласованность;
- выдает результаты экспертизы в виде протоколов и таблиц;
- формирует и хранит базы данных по техническим характеристикам различного оборудования, номенклатуре показателей и коэффициентам их весомости, результатам проведенных экспертиз технического уровня.

Правильность работы программного обеспечения контролируется пропуском тестовой программы в ручном или автоматическом режиме.

В режиме работы разрабатываемой в рамках настоящей работы диалоговой программной системы (ДПС) при расчете весовых коэффициентов или оценочных показателей каждого изделия пользователь будет иметь возможность выполнить настройку, просмотр и редактирование списка экспертов, просмотр и редактирование таблиц с исходными данными, просмотр и корректировку протокола с результатами расчетов (только в части подлежащего оформлению).

В режиме НАСТРОЙКА пользователь, работая с каталогами системы (список экспертов, таблицы исходных данных), настраивает ее на выполнение расчетов. При подготовке списков экспертов и таблиц с экспертными оценками у пользователя имеется возможность ввода нового массива или использования в качестве основы массива из каталога. Для просмотра и/или распечатки протокола работы системы, полученного ранее, пользователю высвечивается каталог протоколов работ.

После выполнения расчетов пользователь будет иметь возможность просмотреть протокол работы системы, включающий исходные данные и результаты расчетов. Допускается возможность корректировки протокола с целью надлежащего оформления записей (за исключением цифровых данных).

Выводы

Предложенный дифференцированный подход к сопоставлению конкурирующих изделий позволяет не только обоснованно выявить предпочтительные разработки, но и указать те показатели, по которым данное изделие имеет преимущество перед конкурентами или уступает им и в какой мере.

Область применения данной методики и диалоговой программной системы может быть распространена на любые аппаратно-программные изделия или аналогичные объекты.

Литература

1. Азгальдов Г.Г. Количественная оценка качества (Квалиметрия). Библиография М.: Изд-во стандартов, 1971. — 176 с.
2. Бестужев-Лада И.В. - Рабочая книга по прогнозированию.: М.: Мысль, 1982. – 430 с.
3. Калашников А.А., Ринкус Э.К., Скрыпников С.Н., Скрыпникова М.К., Новодережкин В.П. Оценка технического уровня программно- технических комплексов АСУ ТП энергоблоков ТЭС.: М.: Журнал «Теплоэнергетика», 10.1996. - 12 с.
4. Вентцель Е.С. Исследование операций. Задачи, принципы, методология: М.: Юстиция, 2018. – 192 с.
5. Вентцель Е.С. Теория вероятностей. Учебник. М.: Юстиция, 2018.- 650 с.
6. <http://mylektsii.ru/7-114412.html>Режим доступа: mylektsii.ru - Лекции - 2015-2018 год.
7. Щупак Ю.А. Win32 API. Эффективная разработка приложений: СПб.: Питер, 2007. - 572 с.
8. Бьерн Страуструп. Язык программирования C++: М.: Бином.2015.- 1136 с.

Скрыпникова М.К. Экспертно - аналитический метод оценки качества конкурирующей продукции. Поставлена задача разработки экспертно- аналитической системы оценки качества конкурирующей продукции. Рассмотрены методические вопросы такой оценки. Предложены оценочные показатели для экспертной оценки. Намечены основные этапы работы, которые в основном связаны с разработкой очередной ветки соответствующего программного продукта.

Ключевые слова: статья, программная инженерия, экспертный метод, оценка качества, теория вероятностей, математическая статистика.

Создание баз знаний для поддержки принятия решения в задачах экспертизы промышленной безопасности

Юрин А.Ю., Берман А.Ф., Николайчук О.А., Дородных Н.О.

Институт динамики и теории управления им. В.М. Матросова СО РАН (ИДСТУ СО РАН)

Иркутский национальный исследовательский технический университет (ИрННТУ)

iskander@icc.ru

Юрин А.Ю., Берман А.Ф., Николайчук О.А., Дородных Н.О. Создание баз знаний для поддержки принятия решений в задачах экспертизы промышленной безопасности. Предлагается подход к разработке баз знаний для интеллектуальных систем поддержки принятия решений в задачах экспертизы промышленной безопасности. Предлагаемый подход основан на специализации модельно-управляемого подхода (ModelDrivenDevelopment, MDD) и включает использование онтологий и концептуальных моделей для представления вычислительно-независимых моделей, предметно-ориентированной нотации для создания логических правил, CLIPS (C Language Integration Production System) в качестве языка программирования для баз знаний. Представленные модели и алгоритмы реализуются в исследовательской программной системе: Personal Knowledge Base Designer. Результаты работы в виде баз знаний используются в Иркутском научно-исследовательском и проектно-институте химического и нефтяного машиностроения (ИркутскНИИХимМаш) для прототипирования баз знаний в задачах экспертизы промышленной безопасности.

Ключевые слова: база знаний, модельно-управляемый подход, специализация, логические правила, экспертиза промышленной безопасности.

Введение

Деградация оборудования во многих отраслях опережает темпы его модернизации и замены, что обуславливает актуальность повышения его безопасности. Особенно это касается нефтеперерабатывающего, нефтехимического и химического оборудования. При этом для определения возможных опасностей и соответствующих мер во избежание катастрофических отказов необходимо его полное обследование. В связи с этим важно минимизировать затраты на техническое обслуживание и ремонт, в том числе на процессы мониторинга, диагностирования и прогнозирования технического состояния, составляющих процесс проведения экспертизы промышленной безопасности (ЭПБ).

Разработка и использование методов и средств искусственного интеллекта, в частности экспертных систем (ЭС), позволяет значительно повысить эффективность ЭПБ. В частности, процедура ЭПБ состоит из следующих основных этапов:

1. Планирование работ для ЭПБ.
2. Анализ технической документации.
3. Формирование карты исходных данных.
4. Разработка программы ЭПБ.
5. Техническое диагностирование.
6. Анализ (в том числе интерпретация) результатов диагностирования.
7. Расчет прочности и остаточного ресурса.
8. Принятие решений по ремонту.

Анализ рассмотренных этапов ЭПБ показал, что реализация этапов 4, 6 и 8 требует обработки большого объема слабо формализованной информации. При этом эффективность обработки можно повысить с помощью ЭС, которые позволяют:

- интерпретировать эксплуатационные условия и параметры,
- обосновать программу технической диагностики,
- интерпретировать диагностические параметры.

В этом случае предлагается автоматизировать эти шаги с помощью продукционных ЭС.

Обзор программного обеспечения и подходов к обслуживанию (включая мониторинг, диагностику и обнаружение отказов) разнородного технического оборудования [1-4] показал слабое использование ЭС для ЭПБ и существование «частных» решений для отдельных типов оборудования.

Основным элементом ЭС является база знаний (БЗ), включающая в себя набор систематизированных

знаний, описывающих закономерности предметной области. Проблема совершенствования процесса создания БЗ и ЭС является актуальной и может быть решена различными способами: совершенствованием подходов или созданием специализированного программного обеспечения [5, 6]. Одним из направлений в этой области является использование принципов когнитивной (визуального) моделирования и проектирования, а также подходов на основе порождающего программирования [7], в частности, модельно-управляемого подхода (Model-DrivenDevelopment, MDD) и его модификации, например, архитектуры, управляемой моделью (Model-DrivenArchitecture, MDA) [8-10].

MDA/MDD - это подход в программной инженерии, который рассматривает создание программного обеспечения на основе преобразований и интерпретаций информационных моделей. Существуют примеры применения MDA/MDD для разработки приложений баз данных (например, Bold for Delphi [10]), агентно-ориентированных приложений мониторинга [11], экспертных систем и др. [12-14].

Цель данной работы - описать специализацию и применение MDA/MDD-подхода для прототипирования продукционных БЗ в задачах, связанных с выделенными выше этапами ЭПБ. Предлагаемая специализация основана на принципах MDA/MDD (типах моделей и этапах создания приложений), но эти принципы реализованы в контексте проектирования интеллектуальных систем. В частности, предлагается:

- использование онтологий и концептуальных моделей (в виде диаграмм классов UML или концепт-карт) для представления вычислительно-независимой модели (как правило, эта модель не используется при разработке приложений на основе MDA/MDD);

- использование предметно-ориентированной нотации, а именно RVML (Rule Visual Modeling Language) [15] для повышения наглядности создания платформо-независимых моделей (RVML предназначен для моделирования логических правил);

- использование CLIPS (C Language Integration Production System) в качестве целевой платформы для реализации БЗ.

Предложенный подход реализован в виде исследовательского программного обеспечения, которое используется в Иркутском научно-исследовательском и проектном институте химического и нефтяного машиностроения (Иркутский химмаш) для прототипирования БЗ в задачах ЭПБ [16].

Предлагаемая специализация MDA/MDD

Основные типы моделей и этапы создания приложений согласно MDA/MDD. В соответствии с принципами MDA/MDD [8-10], разработанное программное обеспечение представлено в виде следующих информационных моделей, определяющих его архитектуру, функции и особенности реализации:

- вычислительно-независимой (Computation-IndependentModel, CIM) - эта модель скрывает любые детали реализации, в большинстве случаев содержит только описания требований;

- платформо-независимая (Platform-IndependentModel, PIM) - эта модель скрывает некоторые особенности реализации и содержит платформо-независимые элементы;

- платформо-зависимая (Platform-SpecificModel, PSM) - эта модель описывает особенности реализации приложения, которые зависят от конкретной программной платформы.

В связи с этим процесс разработки программного обеспечения представляет собой последовательный переход от абстрактных моделей к конкретным с их последовательным преобразованием и генерацией программных кодов и спецификаций: CIM → PIM → PSM → CODE.

Преобразование моделей является одним из основных принципов MDA/MDD подхода и может рассматриваться с разных точек зрения. В частности, существует два типа преобразования: модель-модель (M2M); модель-текст (M2T) и текст-модель (T2M), также преобразования классифицируются на эндогенные и экзогенные, вертикальные и горизонтальные [17].

Таким образом, при разработке программного обеспечения необходимо реализовать следующую последовательность экзогенных горизонтальных преобразований: M2M-преобразование для CIM → PIM; M2M-преобразование для PIM → PSM; M2C-преобразование для PSM → CODE.

В настоящее время существуют некоторые тенденции для реализации модельных преобразований: с помощью графовых грамматик (graphrewriting); с использованием языков трансформации моделей (например, QVT, ATL, TMRL [18] и т. д.); с использованием декларативных и процедурных языков программирования [19], а также языков для преобразования XML-документов (например, XSLT).

В рамках данной работы было решено использовать подход прямого манипулирования [20] для описания трансформаций и разработать «частное» («ad-hoc») решение.

Предлагаемый подход. Специализация процесса разработки MDA/MDD для БЗ представлена следующей последовательностью шагов, основанных на [21]:

Шаг 1: построение модели предметной области, содержащей основные понятия и отношения. На этом шаге конечный пользователь создает CIM. Эта модель может быть реализована в виде онтологии или диаграммы классов UML [22-23].

Эффективность этого шага можно повысить путем повторного использования существующих концептуальных моделей, созданных в CASE-средствах (например, Protégé, IHMCCmapTools и

IBMRationalRoseEnterprise). Большинство программ, поддерживающих MDA/MDD подход (например, BoldforDelphi) не поддерживают этого шага. В случае проектирования интеллектуальных систем этот шаг соответствует этапу концептуализации знаний.

Шаг 2: построение PIM, которая описывает логические правила, сформированные в результате преобразования CIM (CIM → PIM). В процессе преобразования CIM понятия преобразуются в шаблоны фактов и элементы правил (например, условия и действия), а причинно-следственные связи преобразуются в шаблоны правил.

Визуальное моделирование является одним из основных аспектов MDA/MDD подхода, который традиционно использует UML. Следует отметить, что в случае создания специализированного программного обеспечения разработчики используют различные расширения UML [24], что, в свою очередь, позволяет учесть некоторые особенности конкретной предметной или проблемной области, архитектуры, языков программирования или формализмов. Для представления логических правил в данной работе использовано проблемно-ориентированное расширение UML, а именно RuleVisualModelingLanguage (RVML) [15].

Шаг 3: построение PSM, которая позволяет учитывать особенности определенного языка представления знаний (например, CLIPS), такие как приоритеты правил и значения слотов «по умолчанию». PSM является результатом преобразования PIM (PIM → PSM).

Шаг 4: генерация кода БЗ или спецификации. На этом шаге выполняется интерпретация RVML диаграмм (PSM → CODE). Основными результатами интерпретации являются программные коды и спецификации для интерпретатора.

Шаг 5: тестирование полученных программных кодов в специальном программном обеспечении (в интерпретаторе).

Следует отметить, что конечный пользователь (эксперт или системный аналитик) разрабатывает только CIM, PIM и часть PSM.

Описанная последовательность шагов почти совпадает со «стандартной» методологией MDA/MDD, но содержание шагов переопределено в контексте создания продукционных БЗ.

Модели. Опишем модели, используемые в специализированном MDA/MDD-подходе.

CIM включает описание онтологии предметной области и онтологии экспертных систем, основанных на правилах, и может быть представлен в виде метамодели MOF (MetaObjectFacility) см. рис. 1).

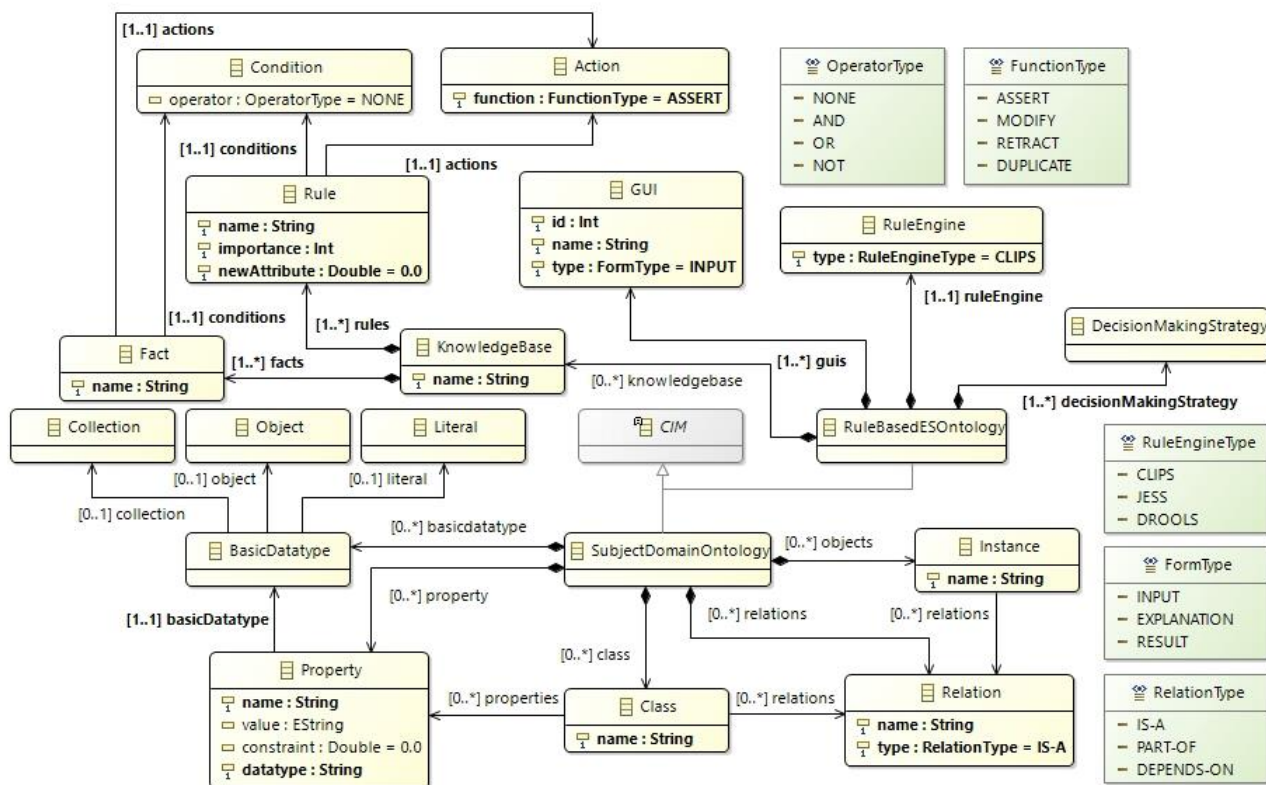


Рисунок 1 – Метамодель CIM (фрагмент).

PIM включает два элемента, предназначенных для представления и моделирования архитектуры ЭС и логических правил БЗ (см. рис 2).

RVML используется для создания PIM и PSM для БЗ. Эта нотация обеспечивает механизм описания

причинно-следственных связей на довольно абстрактном уровне (для PIM). Кроме того, спецификация некоторых элементов нотации (таких как приоритет (важность) и коэффициент уверенности) предоставляет возможность создания PSM, в частности, для CLIPS.

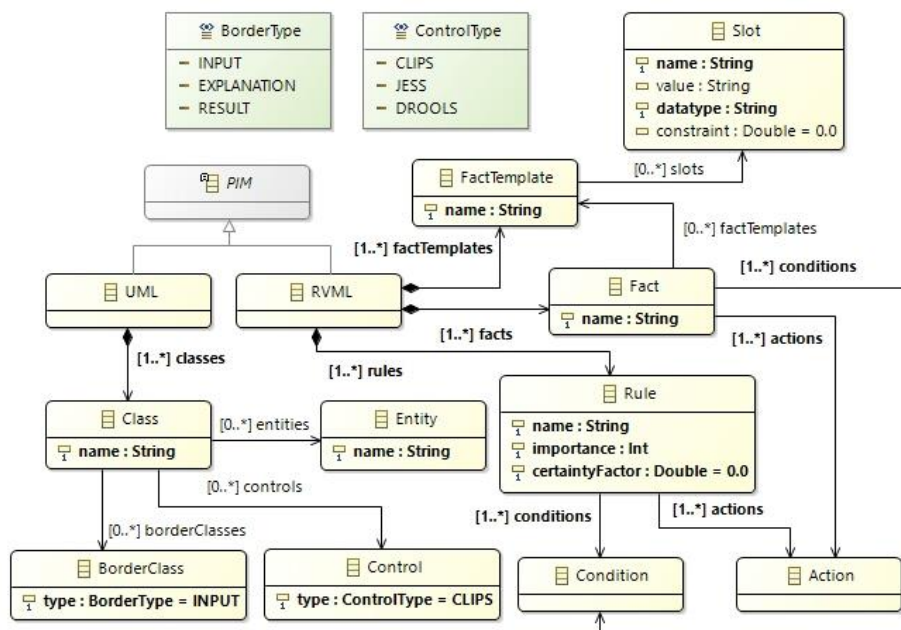


Рисунок 2 – Мета модель PIM (фрагмент).

Преобразования моделей основаны на сравнении элементов метамодели CIM, PIM, PSM и программной платформы. В нашем случае соответствие элементов могут быть представлены в виде таблицы (Таблица 1) [21].

Таблица 1 – Соответствие элементов для CIM, PIM, PSM и CLIPS (фрагмент)

Элементы онтологии (CIM)	Элементы БЗ (PIM, PSM)	CLIPS (платформа)
Project (name, description)	Knowledge base (name, description)	-
Class (name, description)	Template (name, description)	deftemplate
Object(name, description)	Fact (name, description)	defacts
Method	-	-
Property	Slot (description, value)	slot
Property value	Slot value	default
Property type	Slot type	type
Relationship	Rule (nodal element)	defrule

Преобразования моделей реализованы с помощью языка программирования общего назначения (Object Pascal) в исследовательском программном обеспечении - PersonalKnowledgeBaseDesigner (PKBD) [25].

Создание база знаний для задач экспертизы промышленной безопасности

Одной из задач требующей разработки БЗ, является задача анализа результатов диагностирования и прогнозирования возможных деградационных процессов. В качестве теоретической основы для этой задачи предлагается использовать модель динамики технического состояния (МДТС) из [26]. Данная МДТС отражает факторы этапов проектирования, изготовления и эксплуатации технического оборудования, приводящие к его деградации. Значение этих факторов (например, условия эксплуатации, ошибки персонала и производственные дефекты) определяют возможные технические состояния и их структуру (включая переходы между состояниями).

Рассмотрим проектирование базы знаний для задачи прогнозирования деградационных процессов.

Шаг 1: построение CIM с использованием МДТС для нефтехимического оборудования (синтез полиэтилена). Эта модель включает классы состояний (начальная дефектность, повреждение, разрушение, отказ), которые описываются набором параметров со значениями.

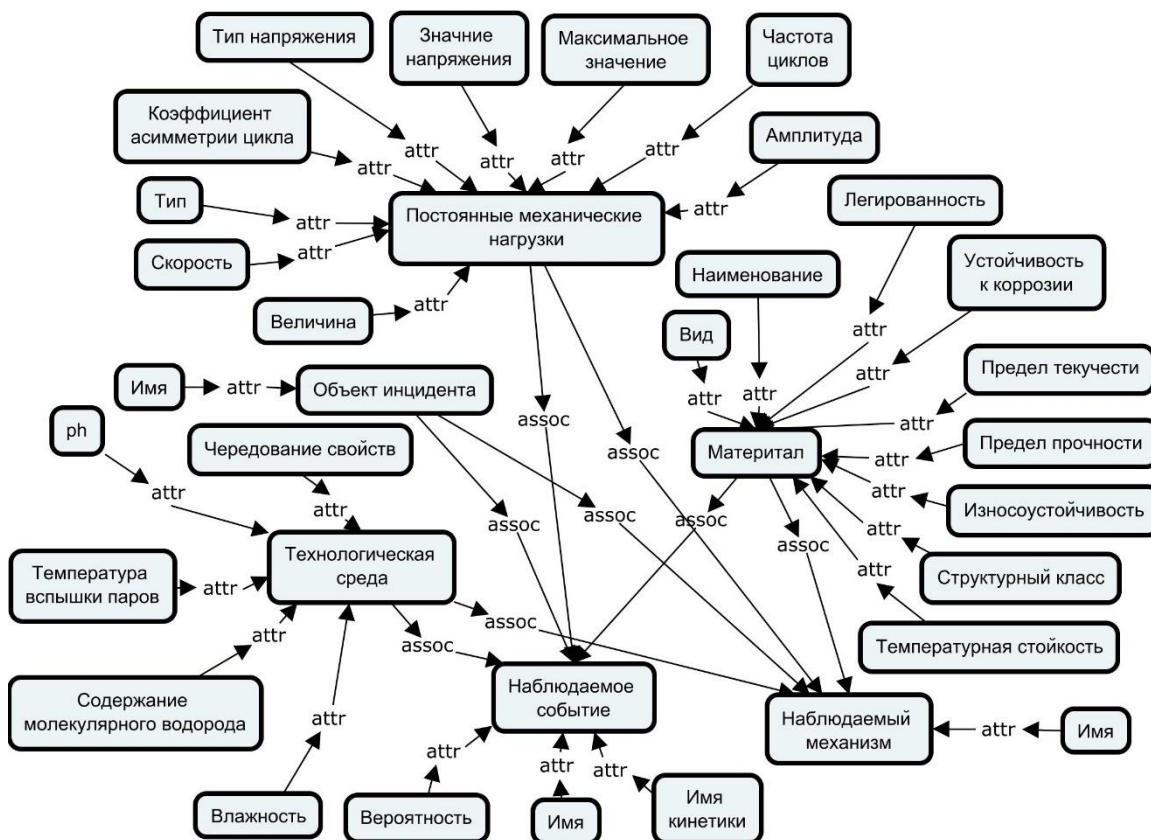


Рисунок 3 – CIM для ЭПБ (фрагмент).

Основным результатом шага является модель предметной области в виде онтологии (или концепт-карты), содержащая понятия и отношения, например (см. рис.3): механизм деградационного процесса, конструкционный материал, технологическая среда и др.

В частности, получены следующие основные шаблоны фактов: объект инцидента, свойства объекта, технологическая наследственность, теплообменная технологическая среда, постоянные механические нагрузки, технологическая среда, материал, дефекты изготовления, наблюдаемое событие, наблюдаемый механизм, наблюдаемая кинетика, наблюдаемое повреждение, наблюдаемое разрушение и т.д.

Получены следующие основные шаблоны правил:

- ЕСЛИ постоянные механические нагрузки И технологическая среда И материал И дефекты изготовления ТО наблюдаемое событие И наблюдаемый механизм
- ЕСЛИ наблюдаемое событие ТО наблюдаемая кинетика
- ЕСЛИ наблюдаемая кинетика И наблюдаемое событие ТО наблюдаемое событие
- ЕСЛИ наблюдаемое повреждение И наблюдаемое событие ТО наблюдаемое разрушение И наблюдаемое событие

...

На основе данных зависимостей пользователь определяет конкретные правила, например:

ЕСЛИ постоянные механические нагрузки (частота цикла = высокая)

И технологическая среда (рН = активная; свойства чередование = да)

И материал (вид = сталь; легированность = низколегированная сталь)

И дефекты изготовления (технологическая наследственность ?id-th-m1)

ТО наблюдаемое событие (имя = механизм коррозионной усталости) и (наблюдаемый механизм (имя = коррозионная усталость)

Примеры моделей RVML, соответствующих разработанным специфическим правилам, представлены в [23].

Были рассмотрены следующие деградационные процессы: коррозионное растрескивание, механическое разрушение, водородное охрупчивание и коррозионное растрескивание. БЗ для коррозионного растрескивания содержит 14 шаблонов фактов, 12 шаблонов правил, 4 начальных (исходных) фактов и 20 конкретных правил.

Шаг 3: построение PSM включает специализацию PIM с учетом особенностей определенного языка программирования, в частности, CLIPS. В данном случае в RVML были определены значения по умолчанию для свойств фактов (слотов) в шаблонах фактов и коэффициенты уверенности для конкретных правил.

Шаг 4: создание программного кода и спецификаций, включая:

- код CLIPS:

- спецификации ЭС для интерпретатора, который обеспечивает генерацию пользовательского интерфейса для создания, чтения, обновления и удаления (CRUD) элементов базы знаний.

Шаг 5: тестирование базы знаний осуществляется экспертом с помощью логических выводов в РКВД.

После успешного тестирования БЗ используется в системе поддержки принятия решений в задачах ЭПБ (см. рис. 4) [16].

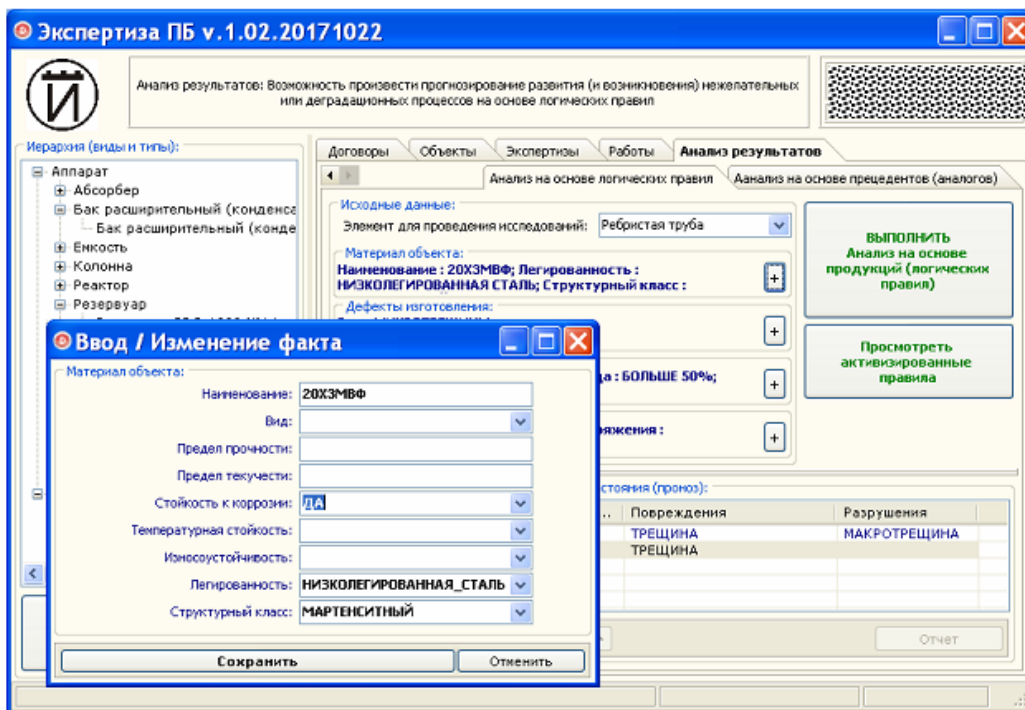


Рисунок 4 – Пример интерфейса программного обеспечения для ЭПБ.

Оценка эффективности предлагаемого подхода

Применимость и эффективность предложенного подхода оценивалась косвенным и прямым методами. Основными критериями оценки являлось использованное время и семантическая адекватность разработанных БЗ.

Семантическая адекватность определялась на основе экспертных оценок специалистов Лаборатории информационно-телекоммуникационных технологий исследования техногенной безопасности ИДСТУ СО РАН и ИркутскНИИХимМаш.

Необходимость использования косвенного метода была обусловлена сложностью и высокой стоимостью модельных экспериментов, оценивающих реальное время разработки БЗ (когда разработка выполняется стандартным методом, а затем предлагаемым). Косвенный метод позволяет получить теоретические оценки времени разработки БЗ в соответствии с [5]. При этом этап концептуализации был разделен на три подэтапа: идентификацию проблемы, извлечение знаний и их структурирование (Табл. 2).

Таблица 2 – Использованное время (косвенный метод оценки)

Этапы создания БЗ	Стандартный метод (нед.)	Специализация MDA/MDD (нед.)
1. Идентификация проблемы	1-2	0.75-1.5
2. Извлечение знаний	4-12	3-8
3. Структуризация	2-4	1.5-3
4. Формализация	4-8	3-6
5. Реализация (кодификация)	0.5	-
6. Тестирование	1-2	1-2
Итого:	12.5-27.5	9.25-20.5

Подход, основанный на специализации MDA/MDD, показал некоторое сокращение времени разработки БЗ за счет исключения этапа реализации (кодификации) (используется автоматическая генерация кода) и исключения (или сокращения времени участия) инженера по знаниям на этапах 1-4.

Прямой метод был основан на оценке времени выполнения учебных заданий студентов факультета Кибернетики (сейчас Института высоких технологий) Иркутского национального исследовательского университета (ИрНТУ) [22] стандартным и предлагаемым методами. По результатам оценки был сделан вывод о сокращении времени разработки при использовании специализации MDA/MDD-подхода, особенно при создании БЗ с простыми правилами и на основе моделей предметной области с небольшим числом (до 8) сущностей.

Заключение

В статье описывается специализация и применение MDA/MDD-подхода для разработки производственных БЗ для задач ЭПБ. Специализация включает: использование онтологии в качестве CIM, использование RVMML для создания PIM и PSM. CLIPS выбран в качестве целевого языка программирования БЗ. Рассмотрены предложенная методология, модели, преобразования и пример применения.

Предлагаемый подход ориентирован на непрограммирующих пользователей: экспертов и системных аналитиков, которые способны разработать только две информационные модели: CIM (онтология) и PIM (модель производственной базы знаний). При этом существует возможность автоматизировать создание PIM на основе анализа концептуальных моделей (UML диаграмм классов, концепт-карт и диаграмм Исикавы) [22,23]. Согласно MDA/MDD-подходу, другие модели и их преобразования либо интегрируются в программное обеспечение, которое реализует подход, либо создаются автоматически вплоть до этапа тестирования.

Преимущества предлагаемого подхода по сравнению со стандартным методом [5,6] заключаются в следующем:

- значительное сокращение времени реализации и минимизация ошибок программирования за счет автоматической генерации кода;
- сокращение времени на этапы идентификации, концептуализации и формализации за счет использования онтологии и когнитивной графики.

Предложенный подход был использован при создании БЗ для оценки причин деградационных процессов нефтехимического оборудования [16].

Работа выполнена при частичной поддержке РФФИ (проекты: 18-07-01164, 18-08-00560, 18-37-00006).

Литература

1. Lee, J. Modern computer-aided maintenance of manufacturing equipment and systems: Review and Perspective / J. Lee // *Computers & Industrial Engineering*. – 1995. – Vol.28, No.4. – P. 793-811.
2. Ruiz, D. Fault diagnosis support system for complex chemical plants / D. Ruiz, J. M. Nougues, L. Puigjaner // *Computers & Chemical Engineering* – 2001. – Vol. 25, Issue 1. – P. 151-160.
3. Venkatasubramanian, V. A review of process fault detection and diagnosis: Part I: Quantitative model-based methods / V. Venkatasubramanian, R. Rengaswamy, K. Yin, S.N. Kavuri // *Computers & Chemical Engineering*. – 2003. – Vol. 27, No. 3. – P. 293-311.
4. Wang, H.C. A hybrid expert system for equipment failure analysis / H.C. Wang, H.S. Wang // *Expert Systems with Applications*. – 2005. – Vol.28. – P.615-622.
5. Джексон, П. Введение в экспертные системы. Пер. с англ. М: Вильямс, 2001. – 624 с.
6. Giarratano, J.C., Riley, G. *Expert Systems: Principles and Programming*, 2005.
7. Чернецки, К., Айзенкер, У. Порождающее программирование: методы, инструменты, применение / Пер. с англ. СПб: Питер, 2005. – 736 с.
8. Frankel, D. *Model Driven Architecture: Applying MDA to Enterprise Computing*. New York: Wiley, 2003.
9. Kleppe, A., Warmer, J., Bast, W. *MDA Explained: The Model Driven Architecture: Practice and Promise*, 1st Edition. Addison-Wesley, 2003.
10. Грибачев, К.Г. *Delphi и Model Driven Architecture. Разработка приложений баз данных*. СПб: Питер, 2004. – 352 с.
11. Gascueña, J.M. Model-to-model and model-to-text: looking for the automation of VigilAgent / J.M. Gascueña, E. Navarro, A. Fernández-Caballero, R. Martínez-Tomás // *Expert Systems*. – 2014. – Vol.31, No. 3. – P.199-212.
12. Canadas, J. InSCo-Gen: A MDD Tool for Web Rule-Based Applications. *Web Engineering* / J. Canadas, J. Palma, S. Tunes // *LNCS*. – 2009. – Vol. 5648. – P. 523-526.
13. Distanto, D. Model-Driven Development of Web Applications with UWA, MVC and JavaServer Faces. *Web Engineering* / D. Distanto, P. Pedone, G. Rossi, G. Canfora // *LNCS*. – 2007. – Vol. 4607. – P. 457-472.
14. Nofal, M. Developing Web-Based Semantic Expert Systems / M. Nofal, K.M. Fouad // *International Journal of Computer Science*. – 2014. – Vol.11, No. 1. – P. 103-110.
15. Юрин, А.Ю. Нотация для проектирования баз знаний производственных экспертных систем / А.Ю. Юрин // *Объектные системы*. – 2016. – №12. – С. 48-54.
16. Берман, А.Ф. Поддержка принятия решений на основе производственного подхода при проведении

- экспертизы промышленной безопасности / А.Ф. Берман, О.А. Николайчук, А.Ю. Юрин, К.А. Кузнецов // Химическое и нефтегазовое машиностроение. – 2014. – №11. – С.28-35.
17. Mens, T. ATaxonomyofModelTransformations / T. Mens, P.V. Gorp // ElectronicNotesinTheoreticalComputer Science. – 2006. – Vol.152. – P. 125-142.
18. Дородных, Н.О. Model Transformations for Intelligent Systems Engineering / Н.О. Дородных, С.А. Коршунов, Н.Ю. Павлов, Д.Ю. Сопп, А.Ю. Юрин // Открытые семантические технологии проектирования интеллектуальных систем. – 2018. – Т.2. – №8. – С.77-81.
19. Берман, А.Ф. Интеллектуальная информационная система анализа отказов / А.Ф. Берман, О.А. Николайчук, А.Ю. Юрин // Проблемы машиностроения и надежности машин. – 2012. – № 4. – С. 88-96.
20. Czarnecki, K. Feature-based survey of model transformation approaches / K. Czarnecki, S. Helsen // IBM Systems Journal. – 2006. – Vol.45. – No.3. – P.621-645.
21. Грищенко, М.А. Применение модельно-управляемого подхода для создания производственных экспертных систем и баз знаний / М.А. Грищенко, Н.О. Дородных, О.А. Николайчук, А.Ю. Юрин // Искусственный интеллект и принятие решений. – 2016. – №2. – С. 16-29.
22. Дородных, Н. О. Формирование баз знаний производственного типа на основе UML-моделей / Н. О. Дородных, А. Ю. Юрин // Информатика и кибернетика. – 2016. – № 3(5). – С.44–50.
23. Дородных, Н.О. Использование концепт-карт для автоматизированного создания производственных баз знаний / Н.О. Дородных, А.Ю. Юрин // Программные продукты и системы. – 2017. – №4. – С. 658-662.
24. Miguel, M. PracticalExperiencesintheApplicationofMDA / M. Miguel, J. Jourdan, S. Salicki // LNCS. – 2002. – Vol.2460. – P.128-139.
25. Дородных, Н.О. Система программирования производственных баз знаний: Personal Knowledge Base Designer / Н.О. Дородных, М.А. Грищенко, А.Ю.Юрин // Открытые семантические технологии проектирования интеллектуальных систем. – 2016. – №6. – С. 209-212.
26. Berman, A.F.. A methodology for the investigation of the reliability and safety of unique technical systems / A.F.Berman, O.A. Nikolaychuk, A.Yu.Yurin, A.I. Pavlov // Part O: Journal of Risk and Reliability. – 2014. – Vol. 228. – no. 1. – P. 29–38.

Юрин А.Ю., Берман А.Ф., Николайчук О.А., Дородных Н.О. Создание баз знаний для поддержки принятия решений в задачах экспертизы промышленной безопасности. Предлагается подход к разработке баз знаний для интеллектуальных систем поддержки принятия решений в задачах экспертизы промышленной безопасности. Предлагаемый подход основан на специализации модельно-управляемого подхода (ModelDrivenDevelopment, MDD) и включает использование онтологий и концептуальных моделей для представления вычислительно-независимых моделей, предметно-ориентированной нотации для создания логических правил, CLIPS (C Language Integration Production System) в качестве языка программирования для баз знаний. Представленные модели и алгоритмы реализуются в исследовательской программной системе: Personal Knowledge Base Designer. Результаты работы в виде баз знаний используются в Иркутском научно-исследовательском и проектно-институте химического и нефтяного машиностроения (ИркутскНИИХимМаш) для прототипирования баз знаний в задачах экспертизы промышленной безопасности.

Ключевые слова: база знаний, модельно-управляемый подход, специализация, логические правила, экспертиза промышленной безопасности.

Yurin Aleksandr, Berman Aleksandr, Nikolaychuk Olga, Dorodnykh Nikita About. Developing knowledge bases for decision support in tasks of industrial safety expertise. An approach for the development of knowledge bases for intelligent decision support systems in the tasks of industrial safety expertise is proposed. The proposed approach is based on the specialization of model-driven approach (model Driven Development, MDD) and includes the use of ontologies and conceptual models to represent computational-independent models, domain-specific notation to create logical rules, CLIPS (C Language Integration Production System) as a programming language for knowledge bases. The presented models and algorithms are implemented in the research software: Personal Knowledge Base Designer. The results of the work in the form of knowledge bases are used in the Irkutsk Research and Design Institute of Chemical and Petroleum Engineering (IrkutskNIHimMash) for prototyping knowledge bases in the tasks of industrial safety expertise.

Key words: knowledge base, model-driven approach, specialization, logical rules, industrial safety expertise.

Нейросетевой подход к моделированию корпоративных сетей

Юсков И.О.

Московский технический университет связи и информатики

Yus_92@mail.ru

Юсков И.О. Нейросетевой подход к моделированию корпоративных сетей В данной работе рассмотрен вопрос повышения эффективности построения и автоматизации процессов в создании и улучшении корпоративных сетей. Был проведен обзор литературы по данным вопросам, предложены варианты модели построенной на искусственной нейронной сети и учитывающая различные факторы. Была описана методика моделирования с применением искусственной нейронной сети, обучающаяся на основе проведенных экспериментальных данных исследований. В результате построенных моделей была изучена перспективность использования нейросетевого подхода к моделированию процесса.

Ключевые слова: Нейросетевой подход, корпоративная сеть, моделирование

В современных условиях перед любой развивающейся компанией, является она производственным предприятием или коммерческой организацией, всегда стоит задача дальнейшего развития. При этом, как правило, затрагиваются основные составляющие компании, и в первую очередь связь – как связующий элемент всех сторон бизнеса.

При широкой географии распределения компании актуальной становится проблема быстрого, надежного и экономичного обмена информацией между филиалами. Для решения этого вопроса операторы связи предлагают объединять удаленные друг от друга офисы в единое защищенное информационное пространство – корпоративную сеть с полным спектром новых телекоммуникационных услуг.

Проблема повышения эффективности создания и автоматизации процессов корпоративных сетей стоит остро в потребностях современного мира. Сложность ее решения связана с тем, что процесс различных взаимосвязанных факторов, влияющих как на ход процесса, так и на конечный результат. Поэтому мониторинг и определение оптимальных режимов является актуальной технической и экономической задачей в корпоративных сетях. Значение данной задачи возрастает в связи с внедрением новых технологий и их взаимодействием с уже существующими системами. Эффективность оптимизации как процедуры выбора оптимальных технологических условий корпоративной сети требует высокого качества всех составляющих ее частей. В связи с этим повышение качества используемых моделей является одной из актуальных задач повышения эффективности корпоративной сети. Требования, предъявляемые современным бизнесом к телекоммуникационным услугам, непрерывно растут. Компании операторов связи, которые раньше других смогут дать заказчикам самые оптимальные решения, удовлетворить все их требования и предложить новые дифференцированные услуги, такие как пакетная телефония, электронная коммерция, хостинг приложений и поддержка мультимедиа, добьются финансового успеха, получив новые источники доходов, и повысят прибыльность своего бизнеса

Корпоративная сеть - это сложная система, включающая тысячи самых разнообразных компонентов: компьютеры разных типов, начиная с настольных и кончая мейнфреймами, системное и прикладное программное обеспечение, сетевые адаптеры, концентраторы, коммутаторы и маршрутизаторы, кабельную систему. Основная задача системных интеграторов и администраторов состоит в том, чтобы эта громоздкая и весьма дорогостоящая система как можно лучше справлялась с обработкой потоков информации, циркулирующих между сотрудниками предприятия и позволяла принимать им своевременные и рациональные решения, обеспечивающие выживание предприятия в жесткой конкурентной борьбе. А так как жизнь не стоит на месте, то и содержание корпоративной информации, интенсивность ее потоков и способы ее обработки постоянно меняются. Последний пример резкого изменения технологии автоматизированной обработки корпоративной информации у всех на виду - он связан с беспрецедентным ростом популярности Internet в последние 2 - 3 года

Описанные выше системы дают возможность управлять паролями для одного или двух-трех приложений через встраивание к соответствующим клиентам. Именно в этом подходе и кроется существенный недостаток таких систем — отсутствие интеграции со всем спектром используемых решений и их клиентов. Как правило, эти системы не интегрируются в информационную систему всего предприятия и не могут быть выбраны средней или крупной компанией как единственное решение. А стоимость внедрения, поддержки и сопровождения нескольких решений одновременно ставит под сомнение перспективность такого подхода в целом.

Решения, реализующие единый доступ (single sign-on, SSO), лишены описанных выше недостатков. В таких решениях пользователь, прошедший процедуру аутентификации, получает доступ ко всем возможным для него приложениям и сетевым ресурсам. Пока пользователь не покинет систему, ему нет необходимости повторять процедуру аутентификации в каком-либо приложении. Решения SSO повышают удобство работы пользователя в корпоративной гетерогенной среде. Пользователь просто не замечает прохождения процедур аутентификации, и вся среда становится более дружелюбной. Главным недостатком SSO-решений является то, что все системы и решения должны доверять одному сервису аутентификации.

Решения, реализующие технологию SSO, известны уже много лет. Как пример, можно рассмотреть протокол Kerberos, использующийся в Unix-системах и поддерживаемый Microsoft, начиная с платформы Windows 2000. Для успешной аутентификации пользователь должен ввести свое имя и пароль. Приложение, использующее Kerberos, взаимодействует через интерфейс с репозиторием, где производится аутентификация пользователя, и при успешном прохождении пользователю выдается «билет». Последующим приложениям достаточно обратиться к выданному «билету», чтобы разрешить или отклонить доступ. Такая модель достаточно сильна, если все приложения в информационной системе доверяют службе, выдающей «билеты».

Современные сети в большинстве случаев гетерогенны. Данная модель может быть также применена в гетерогенной сети. При первом обращении пользователь аутентифицируется в SSO-системе. В дальнейшем при обращении к какому-либо приложению или ресурсу система SSO отыскивает имя и пароль пользователя для требуемой системы и выполняет процедуру аутентификации.

Системы SSO дают большой ощутимый экономический выигрыш для компании благодаря сокращению простоя пользователя при невозможности входа в систему. Системы SSO сокращают и время на саму аутентификацию в системе. Хотя этого времени и недостаточно при успешном входе в систему, учитывая многократность выполнения таких процедур, общее выигранное время может быть существенно. Наиболее существенно экономится время администраторов, поскольку запросы с просьбой «сбросить пароль» существенно сократятся. Учитывая, что именно такие запросы обычно составляют до 25%, общий выигрыш может быть существенен.

Системы SSO могут значительно повышать безопасность информационной системы в целом. Если ранее политики, накладываемые на применяемый пароль, все же зависели от человеческого фактора — пользователю приходилось помнить довольно сложный пароль, то теперь системы SSO могут сами генерировать, менять пароль, причем пароли, как и парольные политики для разных систем, могут быть разные

В этой статье представлены примеры расчетов на графических средствах пакета Scilab и примеры их для визуализации.

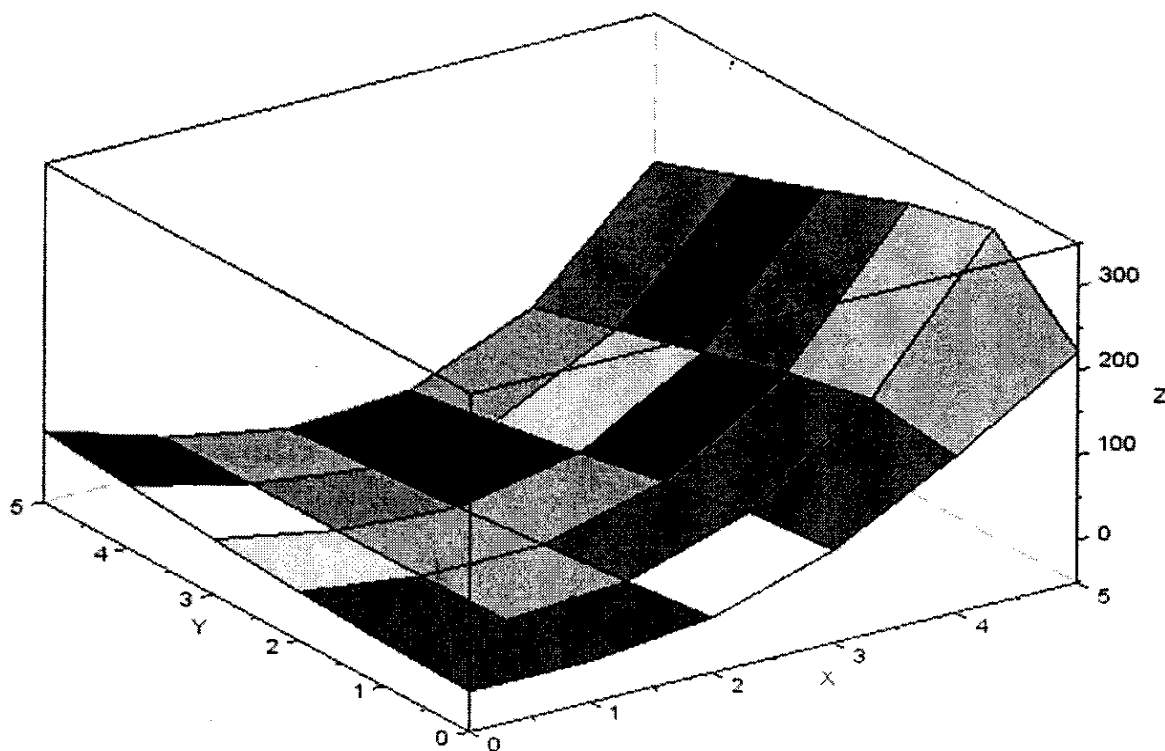


Рисунок 1— График, построения первой модели

Однако следует учесть, что система SSO, использующая для разных приложений и задач один пароль, может ощутимо снизить безопасность всей информационной системы в целом, так как в данном случае общая стойкость системы будет определяться «самым слабым звеном». Причем в этом случае политики, накладываемые на подобный пароль, не позволяют решить проблему в целом. Более сильная политика уменьшает уязвимость системы, но не устраняет ее. Учитывая, что атаки на такой «золотой пароль» могут вестись, вовсе не ломая систему, а косвенными путями, например, анализируя сетевой трафик, применение единого пароля в серьезных системах SSO неприемлемо. Поэтому, администрируя систему SSO, следует опасаться задавать единый пароль для всех систем.

Другой проблемой, с которой может столкнуться специалист, внедряющий SSO-систему, может оказаться физическая доступность для злоумышленника брошенного или случайно оставленного рабочего места сотрудника, уже успешно прошедшего аутентификацию. Поскольку по ходу работы не потребуется вводить дополнительные пароли или подтверждать свою подлинность, такая брошенная рабочая консоль может стать серьезной дырой в системе. Существует несколько решений этой проблемы. Большинство систем SSO предлагают блокировать консоль рабочего места по истечении определенного времени неактивности пользователя. Другие системы предлагают блокировку консоли при отключении электронного ключа или смарт-карты. Администратору системы не стоит пренебрегать подобными настройками.

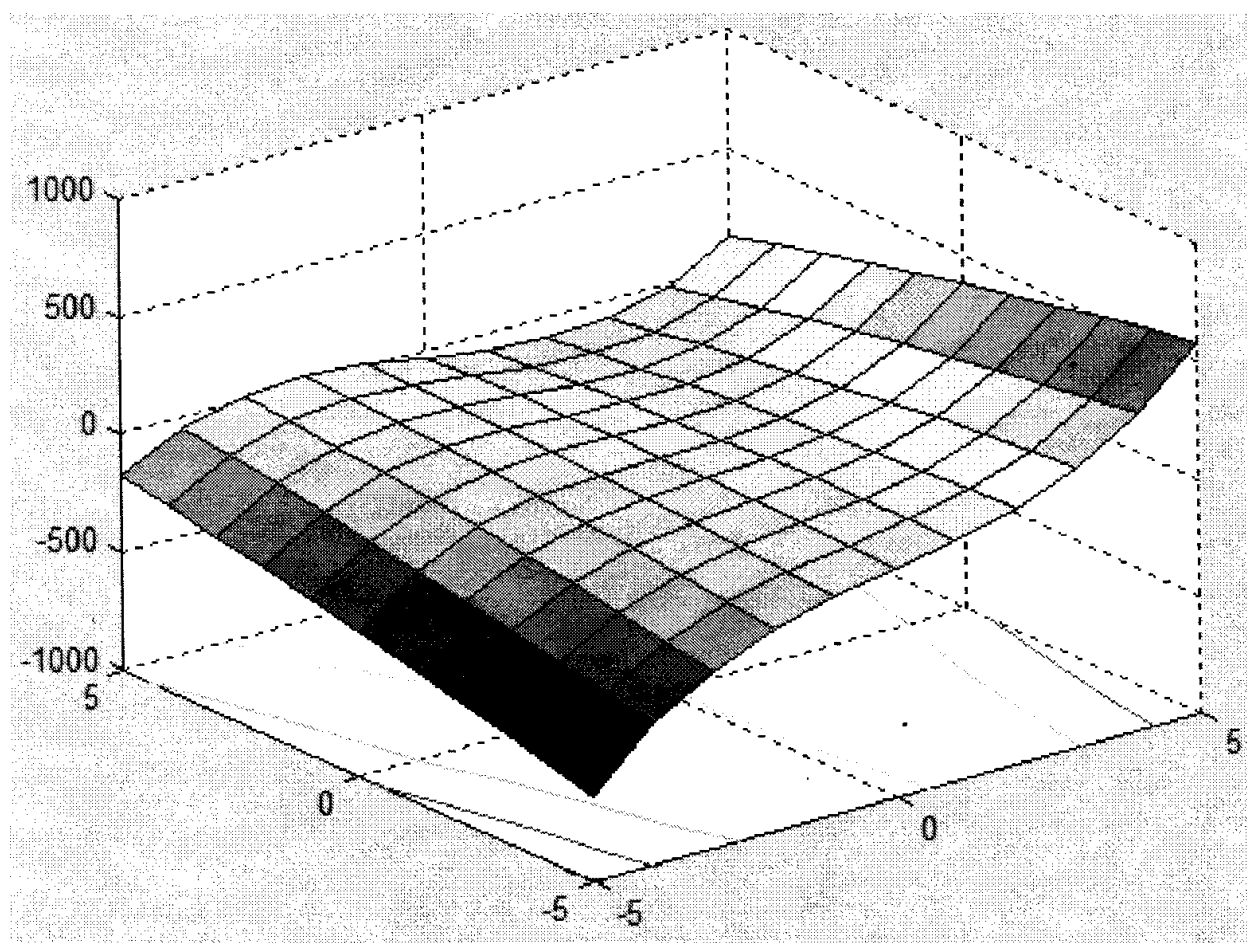


Рисунок 2— График, построения второй модели

Описанные выше решения так или иначе связаны с паролем. Эти системы заменяют пароль, записывают в защищенное хранилище, накладывают политики, синхронизируют и сбрасывают. Однако все они так или иначе привязаны к одному и тому же фактору — паролю. Поэтому хотя пользователь будет иметь двухфакторную аутентификацию и реальный пароль будет сохранен в надежном хранилище, все же часть уязвимостей системы и атак, например, построенных на ловле пароля по сети, потенциально остается, и злоумышленники могут воспользоваться такими лазейками. Ни для кого не секрет, что уже сейчас существует класс систем, реализующих аутентификацию пользователя при помощи цифрового сертификата и личного ключа, базирующихся на технологии PKI. Такие системы реализуют новый подход к построению аутентификации в целом и делают принципиально недоступными старые технологии атаки на систему.

Одними из важных элементов технологии РКІ являются центр сертификации (ЦС), реализующий выпуск и отзыв сертификатов, и хранилище сертификатов, решающие доступность открытого ключа и списка отозванных сертификатов. При этом предполагается, что пользователь хранит свой личный ключ в защищенном хранилище — смарт-карте или электронном ключе, доступ к которому невозможен без знания PIN-кода. Большинство таких систем предлагает собственную реализацию всех компонентов системы, за исключением защищенного хранилища. Безусловно, это может быть удобно, если вы пользуетесь только одной такой системой.

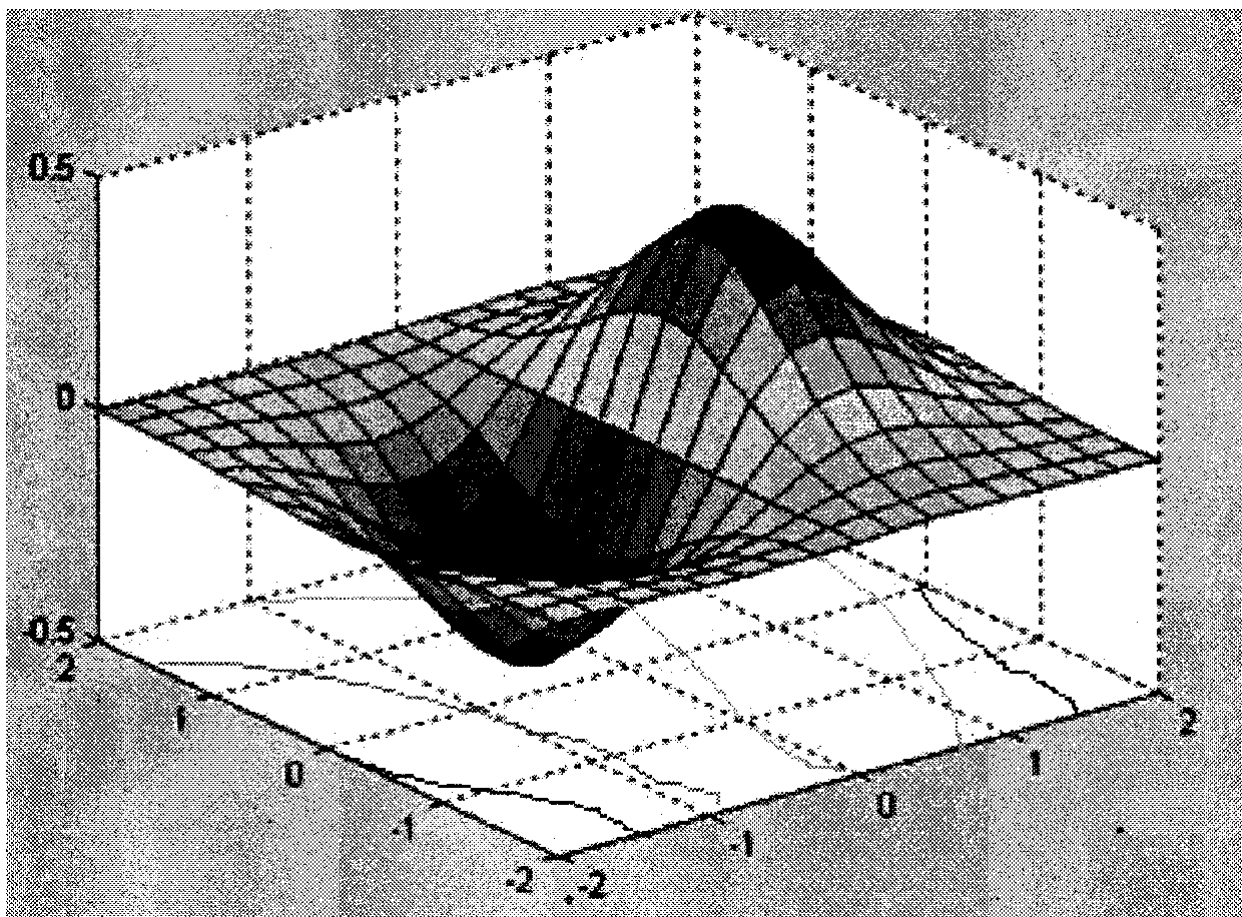


Рисунок 3— График, построения финальной модели

Описанные выше решения так или иначе связаны с паролем. Эти системы заменяют пароль, записывают в защищенное хранилище, накладывают политики, синхронизируют и сбрасывают. Однако все они так или иначе привязаны к одному и тому же фактору — паролю. Поэтому хотя пользователь будет иметь двухфакторную аутентификацию и реальный пароль будет сохранен в надежном хранилище, все же часть уязвимостей системы и атак, например, построенных на ловле пароля по сети, потенциально остается, и злоумышленники могут воспользоваться такими лазейками. Ни для кого не секрет, что уже сейчас существует класс систем, реализующих аутентификацию пользователя при помощи цифрового сертификата и личного ключа, базирующихся на технологии РКІ. Такие системы реализуют новый подход к построению аутентификации в целом и делают принципиально недоступными старые технологии атаки на систему. Одними из важных элементов технологии РКІ являются центр сертификации (ЦС), реализующий выпуск и отзыв сертификатов, и хранилище сертификатов, решающие доступность открытого ключа и списка отозванных сертификатов. При этом предполагается, что пользователь хранит свой личный ключ в защищенном хранилище — смарт-карте или электронном ключе, доступ к которому невозможен без знания PIN-кода. Большинство таких систем предлагает собственную реализацию всех компонентов системы, за исключением защищенного хранилища. Безусловно, это может быть удобно, если вы пользуетесь только одной такой системой.

Хорошим примером, иллюстрирующим преимущества и недостатки такого подхода, может быть система аутентификации по смарт-карте в сети Microsoft Windows 2000. Данная технология носит название SmartCard-Logon. Для реализации решения необходимо установить и настроить сервер сертификатов — он будет осуществлять выпуск и отзыв сертификатов. При этом открытые части всех выписанных сертификатов будут публиковаться в службе каталога — Active Directory, а личные ключи — в смарт-картах. На первый взгляд решение просто и относительно легко для внедрения в корпоративной сети. Но это решение применимо, если вся сеть построена на серверах Windows 2000 и в качестве рабочих мест используются станции только Windows 2000 или Windows XP. Однако в большинстве корпоративных информационных систем это не так — сети гетерогенны, что не позволяет реализовать данное решение. Другая важная проблема для реализации подобного решения — отсутствие целесообразности иметь на всех рабочих местах операционные системы Windows 2000 или XP и защищенный вход по смарт-карте, так как не все пользователи работают с критически важными данными. Также данное решение не позволит пройти аутентификацию пользователю, чье рабочее место не подключено к сети. Это создает невозможность применения данного решения для сотрудников компании, работающих вне офиса.

При необходимости добавить в такую систему аутентификацию на основе биометрии, одноразовых паролей или любого другого устройства, отличного от поддерживаемого стандартно Microsoft, специалист по безопасности столкнется с проблемами совместимости стандартного клиента Microsoft с решениями третьих производителей. Как правило, расширение подобного клиента строится на недокументируемых возможностях и скорее наносит ущерб безопасности. Еще одной серьезной проблемой будет являться отсутствие единой точки администрирования. Часть пользователей, использующих пароли и смарт-карты для входа в сеть, придется администрировать при помощи стандартных утилит Microsoft, а другая часть пользователей будет недоступна для администрирования с помощью этих утилит. Отсутствие единой точки администрирования пользователей сильно усложнит администрирование системы и может привести к серьезным ошибкам. Поэтому при выборе системы крайне важно иметь единую точку администрирования, особенно систем с большим количеством пользователей. Учитывая проблемы, описанные выше, данное решение, базирующееся на стандартных возможностях Windows 2000 и условно-бесплатное для владельцев этих систем, не получило широкого распространения.

Исходя из описанного выше примера, можно представить, каким требованиям должна соответствовать система, предоставляющая инфраструктуру управления аутентификацией (AMI), для удовлетворения большинства современных потребностей среднего и крупного предприятия. Данная система безусловно должна базироваться на технологии PKI, при этом крайне желательно, чтобы такая система могла не только иметь свой центр сертификатов, но и интегрироваться с внешним центром сертификатов. В первом случае это важно, если в корпоративной сети в текущий момент отсутствует центр сертификатов или компания не предполагает задействовать внешние центры сертификатов, например, по причине дороговизны. Во втором случае, если уже в организации существуют какие-либо решения на PKI-технологии, то внедряемая система не должна ломать устоявшиеся решения. Внедряемая система инфраструктуры управления аутентификацией должна плотно интегрироваться с какой-либо службой каталога, где и будут размещаться публичные части системы. Учитывая, что большинство корпоративных систем гетерогенно, такой каталог не должен быть привязан к одной из платформ, а представлять кроссплатформенное решение.

Одним из обязательных требований к каталогу, на котором будет базироваться AMI, должна быть возможность работы по LDAP, протоколом, по сути ставшим стандартом для взаимодействия многих приложений со службой каталога. Желательна полная поддержка спецификации LDAP v3, что сделает возможным SASL (Simple Authentication and Security Layer) аутентификацию в службе каталога. Еще одним из желательных требований также является поддержка градуированной (ступенчатой) аутентификации на уровне службы каталога. Указанным выше требованиям полностью соответствуют такие каталоги, как Sun ONE и eDirac-toy, — эти каталоги полностью поддерживают спецификацию LDAP v3 и градуированную (ступенчатую) аутентификацию. Однако eDirectory имеет преимущества в силу своей кроссплатформенности и низкой стоимости.

Система инфраструктуры управления аутентификацией должна сама являться кроссплатформенным продуктом, что позволит легко внедрять такую систему в гетерогенной информационной системе предприятия. Поскольку данная система представляет «ворота аутентификации», то она должна поддерживать многофакторную аутентификацию, построенную на различных методах — от простого пароля до смарт-карт и биометрии. Безусловно, в такой системе важную роль будет играть поддержка разных вендоров производителей оборудования и решений для аутентификации. Такая поддержка должна быть построена на принципах модульности и открытых интерфейсов взаимодействия между системой и аппаратными решениями. Например, взаимодействие с широким классом смарт-карт и электронных ключей может быть построено на основе стандарта PKCS#11. Такой подход позволит с легкостью переходить на современные устройства и при этом не потребует переписывания программных модулей.

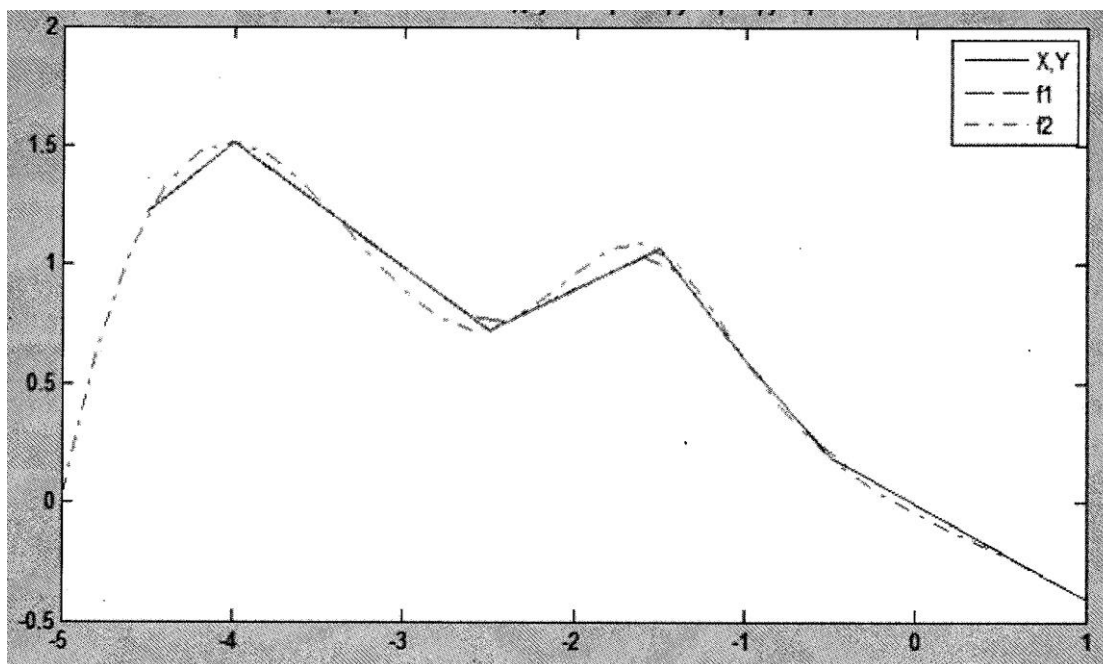


Рисунок 4— График, показатель нагрузки на оборудование

Заключение

Выпускная квалификационная работа магистра посвящена созданию методов стандартизации измерений в корпоративных сетях. В работе проведен анализ уже существующих предложений для различных видов устройств, созданных для различных задач и используемых в различных корпоративных сетях, их систематизации и приведению к единому стандарту. Определены тенденции развития электросвязи в целом. Классифицированы современные интегральные процессы. Один из разделов посвящен уже существующим системам систематизации отдельных аспектов или некоторых типов корпоративной сети. Были рассмотрены их недостатки.

Литература.

1. Богомолова Н.Е., Маликова Е.Е., Сети подвижной сотовой связи. Методические указания и контрольные задания 2011 г.
2. Гольдштейн Б.С., Соколов Н.А., Яновский Г.Г. Сети связи.
3. Данилов А.Н., Цифровые сети интегрального обслуживания. Методические указания и контрольное задание 2010 г.
4. Степанов И.В., Системы беспроводного широкополосного абонентского доступа., учебное пособие 2011 г.
5. Ю. А. Зингеренко Оптические цифровые телекоммуникационные системы и сети синхронной цифровой иерархии
6. Семенова Т.И., Шакин В.Н. Основы работы с математическим пакетом Scilab. Учебное пособие и практикум: МТУСИ. – М., 2017. -129 с.[Электронный ресурс] – Режим доступа: http://www.mtuci.ru/structure/library/catalogue/download.php?book_id=1834

Юсков И.О. Нейросетевой подход к моделированию корпоративных сетей В данной работе рассмотрен вопрос повышения эффективности построения и автоматизации процессов в создании и улучшении корпоративных сетей. Был проведен обзор литературы по данным вопросам, предложены варианты модели построенной на искусственной нейронной сети и учитывающая различные факторы. Была описана методика моделирования с применением искусственной нейронной сети, обучающаяся на основе проведенных экспериментальных данных исследований. В результате построенных моделей была изучена перспективность использования нейросетевого подхода к моделированию процесса.

Ключевые слова: Нейросетевой подход, корпоративная сеть, моделирование

Внедрение программных продуктов группы компаний CSOFT в учебный процесс ДонНТУ и промышленность ДНР

Джура С.Г., Чурсинов В.И., Якимишина В.В.

Донецкий национальный технический университет
ovs@donntu.org, a.chursinova@gmail.com, yvsm@list.ru

Джура С.Г., Чурсинов В.И., Якимишина В.В. Внедрение программных продуктов группы компаний CSOFT в учебный процесс и промышленность ДНР. В данной работе проанализирован опыт внедрения в учебный процесс и промышленность ДНР программных продуктов группой компаний CSOFT (РФ), сотрудничество с которой успешно развивается в новейшее время. Одновременно проанализированы существующие методы расчета систем электроснабжения, определены их достоинства и недостатки, сделаны предложения по их совершенствованию на базе разработок ученых ДонНТУ, а также сделан вывод о целесообразности предложения о внедрении их в программный продукт указанной группы компаний CSOFT (РФ).

Ключевые слова: инновационные методы, система электроснабжения, CSOFT

Введение

Донецкий национальный технический университет (ДонНТУ) имеет большой опыт в реализации международных проектов, в том числе и по созданию авторизованных центров ведущих фирм мира, в частности в области электроэнергетики [1]. Так в ДонНТУ успешно функционируют центры таких мировых промышленных лидеров электроэнергетики как Siemens, Schneider-Electric, Muller-Electric [2]. В связи с начавшейся гражданской войной и историческим выбором Донбасса на воссоединение с Большим Русским Миром, а также ориентацией Донецкой Народной Республики на стандарты Российской Федерации, ДонНТУ подписал договор о сотрудничестве с ведущей российской группой компаний CSOFT, которая осуществляет консалтинг и внедрение комплексных решений в области систем автоматизированного проектирования (САПР), технологической подготовки производства (ТПП), документооборота и геоинформационных систем (ГИС) [3]. Большая часть ее решений базируется на уникальном сочетании мировых и отечественных разработок от CSOFT Development, «Нанософт», Bentley, Oracle, SEA Technology, Siemens, Canon, Contex, Ose и других ведущих компаний [3]. По договору о сотрудничестве группа компаний CSOFT в 2016 году предоставила ДонНТУ учебные лицензии на 9 программных продуктов по проектированию электроснабжения промышленных предприятий (каждая по 30 рабочих мест). Это большая помощь во время гражданской войны. Студенты и преподаватели ДонНТУ сердечно благодарят компанию CSOFT и активно внедряют ее продукты в учебный процесс [4-6]. Очень важно, что продукты компании CSOFT основаны и реализованы, в том числе и на ГОСТ-ах РФ и методиках расчетов ведущих вузов России. Здесь есть большие перспективы для сотрудничества, так как на электротехническом факультете ДонНТУ разработаны свои уникальные методики электротехнических расчетов, уровень которых подтвержден защищенными кандидатскими и докторскими диссертациями. В 2017 году компания CSOFT нашла возможность расширить сотрудничество с ДонНТУ и выделила еще 15 видов программного продукта для реализации на других факультетах (это такие факультеты как горный, горно-геологический, физико-металлургический и факультет экологии и химической технологии). Уже проведены семинары с представителями промышленных предприятий, которые дали самые лучшие отзывы о программном продукте компании CSOFT и его важности и пользы для внедрения в ДНР [7].

Постановка задачи

Целью написания данной статьи является аргументация позиции ДонНТУ по вопросу сотрудничества с компанией CSOFT. Мировой опыт проектирования говорит о том, что известна аксиома программного обеспечения о том, что нельзя создать сложный продукт абсолютно свободный от недостатков.

И дело тут даже не в самих ошибках (известные разработчикам устраняются в новых версиях), а в стремлении в идеалу – совершенствованию самих методов расчета, а также о том, что ДонНТУ есть свои разработки, которые целесообразно предложить нашим российским коллегам, чтобы быть не только потребителем в данном случае программного продукта, а и тем, кто может внести посильный вклад в его

совершенствование. Таким образом, проанализируем современные методы расчета систем электроснабжения и сделаем вывод о целесообразности применения того или иного метода, а также место ученых ДонНТУ в этом процессе. В документации самой компании CSOFTунаходим такие строки, с которыми полностью согласны, что в перспективе нужно (они сами ставят себе такие задачи):

- совершенствование алгоритмов автовыбора оборудования;
- создание более дружелюбного интерфейса и мастеров работы по каждому продукту;
- учет конструктивного аспекта при создании модели;
- автоматическая сборка схемы из сохраненных фрагментов;
- прямое документирование результатов с использованием AutoCAD.

Авторы полагают, что к этим направлениям нужно добавить еще и расширение методов расчета. Мы полагаем и будем аргументировать это ниже, что в дополнение к имеющимся (каждый из которых имеет свои достоинства и недостатки) нужно добавит и методы, разработанные в ДНР. Предлагаем принять это как гипотезу статьи.

Анализ существующих методов расчета

Из известных публикаций с нашей точки зрения наиболее интересным сравнительным анализ проблемы дан в работе [8].

Таблица 1. Сравнительные методы расчета систем электроснабжения.

Метод расчета	Формула	Пояснения
По электроемкости продукции	$P_p = \sum \varepsilon_i M_i / T_m$	M_i, ε_i - объем и электроемкость продукции i -го вида; T_m - годовое число часов использования
По общегодовому электропотреблению	$P_p = K_m A / T_g$	K_m - среднегодовой коэффициент максимума; A - общегодовое электропотребление; $T_g = 8760$ - число часов в году
По удельным мощностям нагрузок	$P_p = \gamma F$	γ - удельная плотность нагрузки; F - площадь предприятия, района, цеха
По среднегодовому коэффициенту спроса K_c	$P_p = K_c P_{уст}$	$P_{уст}$ - сумма установленных мощностей; K_c – коэффициент спроса
Метод упорядоченных диаграмм	$P_p = K_m K_i P_{уст}$	$P_{уст}$ - сумма установленных мощностей; K_m – коэффициент максимума (табл. 1.8); K_i – коэффициент использования

Каждый из указанных методов имеет свои достоинства и недостатки. Методы приведены в порядке их исторического принятия. Каждая конкретная задача определяет и выбор метода ее решения. Разработчик должен владеть разными методами расчета и знать когда их применить. Расчет нагрузок на разных уровнях электроснабжения проводится различными методами в зависимости от исходных данных и требований точности. Обычно расчет ведут от низших уровней к высшим (по напряжению). Однако при проектировании крупных предприятий иногда целесообразно вести расчеты от верхних уровней к нижним. В этом случае рекомендуют пользоваться комплексным методом расчета. За основу обычно берут информационную базу аналогичного предприятия (по профилю). При этом сначала решают вопросы электроснабжения предприятия в

целом, заметем его комплекса цехов, а также всех потребителей, питающихся от одной распределительной подстанции [8].

Группа компаний CSOFT пользуется методиками расчета ОАО «Институт Теплоэлектропроект» [9]. Видно, что указанный институт уделяет большое внимание развитию методов проектирования и обмену опытом и методиками расчетов. Пресс-служба ОАО «Институт Теплоэлектропроект» сообщает, что 3-6 июня 2013 года специалисты ОАО «Институт Теплоэлектропроект» И.И. Шабанов и С.Н. Абрамов приняли участие в обучающем семинаре «Сравнительная оценка подходов к энергосбережению в России и Австрии: опыт применения механизмов энергосервисных контрактов». По результатам семинара институтом получен сертификат Российского Энергетического Агентства. С российской стороны организаторами выступали: Российское энергетическое агентство Министерства энергетики Российской Федерации при поддержке Федеральной энергосервисной компанией (ФЭСКО), с австрийской: Австрийское энергетическое агентство (АЕА) и Ассоциация энергосервисных компаний Австрии (ДЕСА) [9].

Этот семинар позволил изучить австрийский опыт развития рынка энергосервисных услуг и стал платформой для взаимодействия и обмена информацией между ведущими энергосервисными компаниями России и Австрии. В рамках семинаров и круглых столов были затронуты темы законодательного обеспечения энергосервисной деятельности в России, нормативного и методического обеспечения энергосервиса (России), методология оценки эффекта от реализации энергосберегающих мероприятий в соответствии с международным протоколом измерений и верификации эффекта энергосбережения IPMVP (Австрия), международный проект EINSTEIN, примеры передового опыта энергетических перформанс-контрактов и много другое [9].

Открытое обсуждение методик расчетов на форуме проектировщиков электрических и слаботоковых сетей [10] подтверждает, что программный комплекс CSOFT позволяет решать такие задачи как определение расчетных токовых нагрузок для всех элементов распределительной сети различными методами (по коэффициентам загрузки электродвигателей и коэффициентам одновременности максимумов в узлах сети, по методу института «Теплоэлектропроект», по коэффициентам расчетной мощности в соответствии с «Указаниями по расчету электрических нагрузок РТМ 36.18.32.4-92»);

В программе наряду с расчетом мгновенного режима сети осуществляется и определение расчетных токов в элементах сети.

Для определения расчетных нагрузок в программе на выбор предоставляется 3 метода, которые можно поменять в команде "Общие свойства модели":

- 1) По умолчанию принимается метод учитывающий коэффициенты загрузки оборудования и коэффициенты одновременности в узлах (K_z и K_o), этот метод может быть эквивалентен методу расчета по коэффициенту спроса;
- 2) Метод института «Теплоэлектропроект» (ТЭП), рекомендуется для собственных нужд ТЭС;
- 3) Метод института «Тяжпромэлектропроект» (ТПЭП), рекомендуется для машиностроительных предприятий, ремонтных мастерских и т.п.

Результаты расчета по каждому методу так же можно документировать с применением шаблона в MS Word [10].

Мы полагаем, что метод кафедры ЭПГ ДонНТУ, который мы рассмотрим ниже можно внести в меню как один из вариантов и сравнить его применение с указанными. Мы уже это сделали в учебных проектах и получили, что результаты расчетом по программе CSOFT несколько завышены, что может быть полезно для развития производства, если оно выполнено по таким результатам, но может быть и не экономно, если такового не планируется.

Если сравнить программный продукт компании CSOFT с конкурентами, то Конкурентов у EnergyCS Электрика не так уж много, но есть [10]:

Сравнение EnergyCS Электрика (CSOFT) с конкурентными программами [10]:

1. Продукт Winelso (Русская промышленная компания [11]).

Недостатком можно считать тот факт, что оформление документа совмещено с созданием модели, а также повышенная трудоемкость во время разработки схемы. Предлагаются упрощенные расчеты. Не применимо для энергообъектов и сложных случаев производств. Не все расчеты раскрыты.

Достоинством является простота оформления выходной документации, так как она создается на этапе ввода модели.

2. Программа Guexpert (Московский энергетический институт [12])

Недостатком является тот факт, что программа написана для MS DOS. Выполняет только расчет токов короткого замыкания для одного заданного режима. Расчет хороший с учетом всех требований, но только расчет токов короткого замыкания. Есть оценка селективности в отдельном модуле. Выбора оборудования нет. Расчет постоянного тока в отдельном модуле.

К достоинствам можно отнести более широкую подготовку данных, так как задача существенно более узкая.

3. Программа Pixta от компании Shneider Electric [13].

Недостатком является ориентация на одного производителя оборудования. Нет соответствующих интерфейсов.

Достоинством является получение прямое решение на множестве оборудования одной фирмы.

Промежуточные выводы

Таким образом на сегодня можем констатировать, что программный продукт компании CSOFT весьма конкурентоспособен и с нашей точки зрения является наилучшим. Это мнение подтверждает вышеприведенный анализ. Однако, всегда есть пути совершенствования, об этом написано и в документации самой компании CSOFT.

На кафедре ЭПГ ДонНТУ выпущено учебное пособие по электротехническим САПР CSOFT [14]. На рис. 1. Представлены логотипы протестированных продуктов, результаты расчетов по которым методами кафедры ЭПГ.



Рисунок 1. Логотипы протестированных программ компании CSOFT.

Решение задачи и предложения от ДонНТУ.

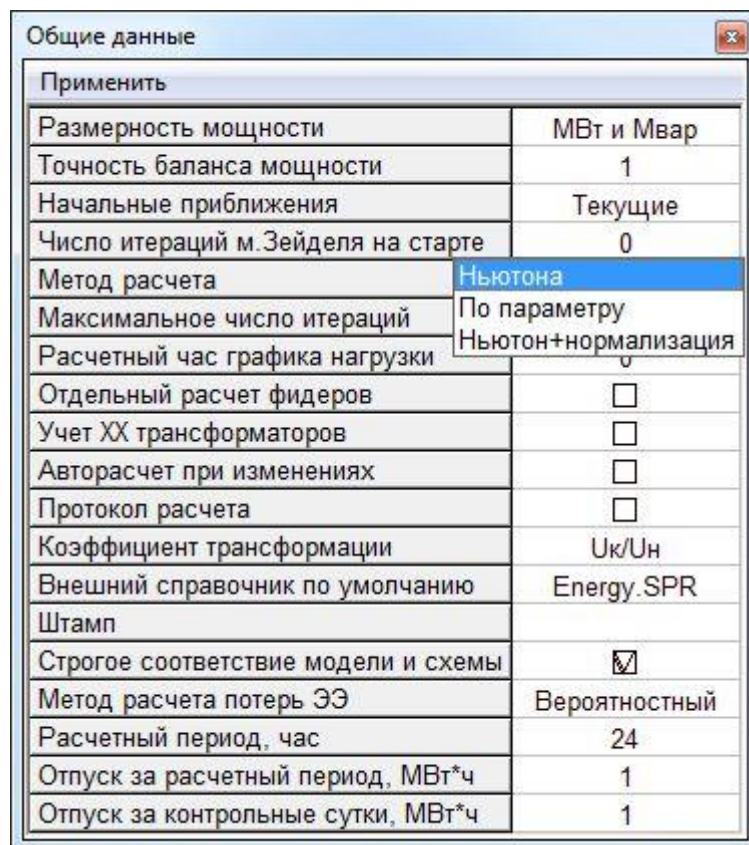
На кафедре электроснабжения промышленных предприятий (ЭПП) ДонНТУ системно ведется совершенствование методов расчета систем электроснабжения [15-17].

Так, согласно [17] в основу действующих Указаний [18] положен метод упорядоченных диаграмм (УД), предложенный профессором Г.М. Каяловым [19,20]. Метод сыграл значительную роль в становлении теории нагрузок и практики расчетов. Однако стремление к уменьшению объема исходной информации с целью упрощения расчетов обусловило принципиальную неточность метода упорядоченных диаграмм. В статье рассматриваются методы расчета, которые могут быть использованы для разработки новых Указаний по определению электрических нагрузок. Аргументируются недостатки действующих указаний. При этом раскрывается физический смысл предлагаемых методов, но не сама методика практических расчетов (исходные справочные данные, расчетные кривые, примеры расчетов, программа вычислений на ЭВМ). Приводится инерционный и квадратичный метод и метод имитации, а также метод эквивалентного параметра корреляционной функции [17].

Авторы приходят к выводу о том, что принятое в действующих Указаниях понятие расчетной нагрузки по нагреву в виде «инерционного» максимума десятиминутной тепловой нагрузки отражает физику задачи. Понятие же «кумулятивного» максимума получасовой средней нагрузки из ПУЭ [18] приводит к противоречащим смыслу результатам.

Метод УД сыграл положительную роль в развитии теории электрических нагрузок, но в настоящее время требует замены из-за существенного завышения расчетных нагрузок. Определение расчетных электрических нагрузок по нагреву целесообразно выполнять точным методом имитации, допуская применение инженерных методов: инерционного и квадратичного.

Таким образом предлагаем пилотный проект: ввести в следующую версию программ CSOFT указанные методики в дополнение к уже действующим. Пример такого выбора может быть в программе, как приведено на рис. 2.



Общие данные	
Применить	
Размерность мощности	МВт и Мвар
Точность баланса мощности	1
Начальные приближения	Текущие
Число итераций м.Зейделя на старте	0
Метод расчета	Ньютона
Максимальное число итераций	По параметру
Расчетный час графика нагрузки	Ньютон+нормализация
Отдельный расчет фидеров	<input type="checkbox"/>
Учет XX трансформаторов	<input type="checkbox"/>
Авторасчет при изменениях	<input type="checkbox"/>
Протокол расчета	<input type="checkbox"/>
Коэффициент трансформации	Uк/Un
Внешний справочник по умолчанию	Energy.SPR
Штамп	
Строгое соответствие модели и схемы	<input checked="" type="checkbox"/>
Метод расчета потерь ЭЭ	Вероятностный
Расчетный период, час	24
Отпуск за расчетный период, МВт*ч	1
Отпуск за контрольные сутки, МВт*ч	1

Рисунок 2. Выбор метода расчета нагрузок.

Выводы

Согласно приведенному анализу, логично сделать вывод о целесообразности внесения в новые версии программного комплекса CSOFT новых методик, предложенных кафедрой ЭПП ДонНТУ и провести их

тестирование (как это уже проведено в ДонНТ и отражено в методических указаниях [19,20]) в других вузах и проектных организациях. чтобы сделать объективный вывод о целесообразности внесения изменений в действующие нормативные указания [21], а пока использовать как пилотный проект и аргументированную альтернативу для проектирования систем электроснабжения разной сложности. Так же рады сообщить, что книга [14], изданная в ДонНТУ по внедрению программного продукта группы компаний CSOFT стала Лауреатом Всероссийского конкурса (номинация «лучшая научные издания» [22]) и Евразийского конкурса (номинация «лучшая учебная книга» [23]).

Литература

1. Донецкий национальный технический университет (ДонНТУ). Официальный сайт. - Режим доступа: <http://donntu.org/>
2. Авторизированные центры обучения ДонНТУ. - Режим доступа: <http://donntu.org/centr-obucheniya-schneider-electric>
3. Группа компаний CSOFT. Официальный сайт. - Режим доступа: <http://www.csoft.ru/>
4. Кафедра ЭПГ получила уникальный дар // Официальный сайт ДонНТУ. - Режим доступа: <http://donntu.org/news/id201610131436>
5. Студенты ДонНТУ получили новое программное обеспечение для использования в учебном процессе от российской группы компаний CSOFT // Официальный сайт Министерства образования Донецкой Народной Республики. – Режим доступа: <http://mondnr.ru/vysshee-professionalnoe-obraz/item/889-studenty-donntu-poluchili-novoe-programmnoe-obespechenie-dlya-ispolzovaniya-v-uchebnom-protssesse-ot-rossijskoj-gruppy-kompanij-csoft>
6. Донецкий политехник. Видео-журнал. Выпуск 12 от 15.10.2016. - Режим доступа: <https://youtu.be/hlBCHwgPdYY?t=165>
7. Семинар по презентации программ CSOFT на электротехническом факультете очень впечатлил специалистов ДПИ НИИ «Теплоэлектропроект». // Донецкий национальный технический университет (ДонНТУ). Официальный сайт. - Режим доступа: <http://donntu.org/news/id201711091127>
8. Кабышев А.В., Обухов С.Г. Расчет и проектирование систем электроснабжения: Справочные материалы по электрооборудованию: Учеб. пособие / Том. политехн. ун-т. – Томск, 2005. – 168 с. – Режим доступа: http://portal.tpu.ru:7777/departments/kafedra/espp/literatura/Tab/M_Kabishev_Obuhov_Raschet.pdf
9. Официальный сайт акционерного общества «Институт Теплоэлектропроект». – Режим доступа: <http://www.tep-m.ru/>
10. Форум проектировщиков электрических и слаботочных сетей. Режим доступа: <http://eom.com.ua/index.php?topic=6532.0>
11. Русская промышленная компания. Официальный сайт. Режим доступа: <https://winelso.ru/>
12. Московский энергетический институт. Программа GUEXPERT. Официальный сайт. - Режим доступа: <http://es.mpei.ac.ru/guexpert.htm>
13. Компания «SchneiderElectric». Официальный сайт. – Режим доступа: <https://www.schneider-electric.com/en/product-category/5100-software/>
14. Левшов А.В., Джура С.Г., Бершадский И.А. Введение в электротехнические САПР группы компаний CSOFT. – Донецк : ДОННТУ, 2017. – Режим доступа: http://iic.roerich.com/russian/ovs/_intro_CAD_CSOFT_2017_.pdf
15. Шидловский А.К., Вагин Г.Я., Куренный Э.Г. Расчеты электрических нагрузок систем электроснабжения промышленных предприятий. - М.: Энергоатомиздат, 1992. - 224 с.
16. Куренный Э.Г., Дмитриева Е.Н., Погребняк Н.Н. Совершенствование методов расчета электрических нагрузок // Промышленная электроэнергетика и электротехника. Выпуск 4. Серия КВ, №2079. – 1997. – с. 14-28. Режим доступа: http://masters.donntu.org/2006/eltf/belyaev/library/ist_2.htm
17. Куренный Э.Г., Дмитриева Е.Н., Погребняк Н.Н. Развитие методов расчета электрических нагрузок // «Известия вузов. Электромеханика». 2004. - с. 43-45.
18. Правила устройства электроустановок. – М.: Кнорус, 2015. – 488 с.
19. Методическое пособие и справочные материалы к выполнению курсового проекта по курсу «Электроснабжение промышленных предприятий» (проектирование цеховой электрической сети)/И.А. Бершадский и др. – Донецк: ДонНТУ, 2008. – 99 с.
20. Методическое пособие и справочные материалы к выполнению дипломной работы квалификационного уровня бакалавр по направлению «Электроэнергетика и электротехника» / Сост.: Э.Г. Куренный, А.Д. Коломытцев В.В. Якимишина. – Донецк: ДонНТУ, 2016. – 20 с.
21. Указания по расчету электрических нагрузок. Руководящий технический материал. Проектирование электроустановок. О расчете электрических нагрузок. (технический циркуляр ВНИГМ Тяжпромэлектропроект. № 359-92 от 30 июля 1992 г. – Режим доступа: https://znaytovar.ru/gost/2/RTM_361832492_Ukazaniya_po_ras.html

22. Диплом Всероссийского конкурса на лучшую научную книгу. – Режим доступа: http://www.roerich.com/iic/russian/ovs/_dipl_kniga_2017_.jpg

23. Диплом открытого межрегионального конкурса «Книга года: Сибирь – Евразия – 2018». – Режим доступа: http://iic.roerich.com/russian/ovs/_siberia_.jpg

Джура С.Г., Чурсинов В.И., Якимишина В.В. Внедрение программных продуктов группы компаний CSOFT в учебный процесс и промышленность ДНР. В данной работе проанализирован опыт внедрения в учебный процесс и промышленность ДНР программных продуктов группой компаний CSOFT (РФ), сотрудничество с которой успешно развивается в новейшее время. Одновременно проанализированы существующие методы расчета систем электроснабжения, определены их достоинства и недостатки, сделаны предложения по их совершенствованию на базе разработок ученых ДонНТУ, а также сделан вывод о целесообразности предложения о внедрении их в программный продукт указанной группы компаний CSOFT (РФ).

Ключевые слова: инновационные методы, система электроснабжения, CSOFT

Dzhura S. G., Chursinov V.I., Yakimishina V.V. Introduction of software products of CSOFT group of companies in educational process and the industry of the DPR. In this work experience of introduction in educational process and the industry of the DPR of software products by group of companies the CSOFT companies (Russian Federation), cooperation with which successfully develops in the latest time, is analysed. The existing methods of calculation of systems of power supply are at the same time analysed, their merits and demerits are defined, offers on their improvement on the basis of developments of scientists of DONNTU are made and also the conclusion is drawn on expediency of the offer on their introduction in the software product of the specified CSOFT group of companies (Russian Federation).

Keywords: innovative methods, system of power supply, CSOFT

Проблематика организации проектирования систем при реализации методов программно-информационного обеспечения управления проектами, оценка рисков и способы их устранения

Олейник А.Н.

Министерство образования и науки Донецкой Народной Республики

kitaboy@gmail.com

Горун В.В., Гуридов А.Г.

«Республиканская служба по контролю и надзору в сфере образования и науки»

inform.obrnadzor@mail.ru

Олейник А.Н., Горун В.В., Гуридов А.Г. Проблематика организации проектирования систем при реализации методов программно-информационного обеспечения управления проектами, оценка рисков и способы их устранения. В работе рассмотрены основные проблемы при проведении программно-информационного проектирования в организациях, раскрыты методики и возможные методы решения возможных рисков.

Ключевые слова: проект, система, риски, программное обеспечение, информационные технологии.

Введение

По прогнозам корпорации IDC (International Data Corporation), к 2020 году объемы информации в современном мире достигнет величину в 40 зеттабайт. По сравнению с 2010 годом объем передаваемых и хранимых данных в мире вырос в более чем 50 раз. По данным исследования корпорации процент информации, нуждающейся в защите, постоянно растет, а уровень защиты данных остается по-прежнему на недостаточно высоком уровне.

Рост общего объема информации происходит в основном за счет автоматически генерируемых данных, к 2020 году объем которых должен увеличиться более чем в 15 раз. К 2020 году развивающиеся рынки станут основным поставщиком информации. Именно поэтому организации активно используют программно-информационное обеспечение, позволяющее упростить и ускорить процесс получения, передачи, хранения и обработки данных.

В последующие года радикально изменятся не только места хранения информации, но и типы хранящихся данных в облаках. Согласно прогнозам 46,7% данных в облаках будут относиться не к корпоративной информации, а к сфере развлечений. Остальная часть - данные систем видеонаблюдения, встроенных и медицинских систем, синхронизируемая информация с персональных компьютеров, телефонов и устройств IoT (InternetofThing – интернет вещей).

Численность ИТ-персонала не успевает за быстрым ростом количества сервисов, потребность в управлении которыми неуклонно растет. Корпорация IDC в своих аналитических материалах отмечает явное несоответствие квалификации, отсутствие опыта и нехватку ресурсов ИТ-персонала стремительному росту объемов данных. В следующие годы ИТ-инфраструктура всего мира столкнется со рядом проблем, таких как:

- количество серверов (как физических, так и виртуальных) увеличится минимум в 10 раз;
- объем управляемой ИТ-инфраструктурой информацией вырастет в более чем 50 раз;
- количество единиц информации будет расти на порядки быстрее, чем объем хранимой в них информации, т.к. огромное распространение получают встроенные информационные системы (интеллектуальные датчики в одежде, домах, сооружениях, медицинских устройствах и других сферах жизни человека);
- численность ИТ-персонала, который должен будет контролировать и управлять всей информационной инфраструктурой вырастет всего лишь в 1,5 раза.

Поэтому внедрение облачных вычислений и соответственно интеллектуальных информационных систем позволит в будущем снизить расходы и улучшить эффективность управленческих операций. IDC прогнозирует, что в ближайшие годы около 20% информации будет проходить через провайдеров услуг облачных вычислений, т.е. в течение своего жизненного цикла она какое-то время будет храниться в облаке или обрабатываться в нем.

Актуальность проблематики

Чтобы успешно реализовывать политику управления информационными проектами, для начала нужно проанализировать работу организации и распланировать весь процесс создания проекта. Правильная постановка целей и задач помогает правильно реализовать проект. Если в организации уже функционирует некая программно-информационная система, то внедряемое решение с использованием новых методов и технологий выполняющее те же функции, что и ранее используемое, может значительно повысить эффективность и производительность системы.

Применение проектного управления имеет особое значение и обуславливается такими основными факторами:

- возрастающая сложность проектов и их организация. Современный мир управления проектами намного требовательней используемых старых методов управления. Его сущность заключается в управлении проектами, как объектами, экономическими единицами в частности или бизнесом и целыми сферами в общем.

- значительный рост конкуренции и ускорение научно-технического прогресса (НТП). Непрерывные усложнения и рост потребностей общества приводит к усилению интеграции науки и производства, что в свою очередь требует разработки более эффективных форм и методов управления ими. В условиях рыночной экономики существуют специфические проблемы осуществления проектов: сокращение сроков производства продукции и услуг, увеличение количества уникальных продуктов и услуг, ужесточение временных и стоимостных требований.

Совокупность этих факторов обусловила резкое увеличение интереса к управлению проектами. Это позволяет экономить ресурсы, добиваться высоких темпов роста, снижать издержки производства и делать товары и услуги более конкурентоспособными. Таким образом, все большее число руководителей приходят к пониманию необходимости проектного управления и осознают положительные моменты применения данной технологии. Однако в силу новизны, управление проектами часто интерпретируют в понятиях и связях системного анализа, управления системами, информационных технологий либо путают с такими видами управленческой деятельности, как администрирование и руководство.

К основным факторам, оказывающим влияние на управление проектами, можно отнести время, ресурсы и качество. Так, увеличив затрачиваемые на проект время и ресурсы, можно значительно повысить качество работы.

Для успешного управления проектом необходимо провести ряд мероприятий. Наиболее важными из них являются разделение проекта на отдельные задачи, расстановка приоритетных задач, а также детализированное описание каждой задачи с момента её начала до окончания реализации. В дальнейшем для каждой подзадачи необходимо сформировать и организовать профессиональную команду, определить возможные риски при выполнении задач и разработать планы реагирования и устранения рисков. На завершающей стадии реализации проекта – необходимо провести всестороннее тестирование и проанализировать результат работы, а при необходимости внести соответствующие изменения. На начальных этапах рекомендуется привлекать к обсуждению разработчиков, это поможет сократить время создания проекта.

Управление персоналом

В управлении проектами весомую роль играет управление персоналом, поэтому большая доля успеха проекта зависит от правильного его подбора. При этом стоит учитывать, что часть сотрудников будут предпочитать работать на дому, другим необходим постоянный контроль со стороны руководства проекта. При подборе персонала для работы над проектом стоит внимательно отнестись к информации, размещаемой будущим сотрудником в социальных сетях. Обычно люди размещают о себе всестороннюю информацию, и при детальном ее изучении можно сделать различные выводы об уровне коммуникабельности, умении работать в команде, когда для реализации одной задачи требуется привлечь сразу несколько человек. Для создания сплочённого коллектива рекомендуется проводить мероприятия, направленные на её укрепление, например, корпоративы, совместные работы, командные тренинги. Так же не стоит пренебрегать таким элементом развития команды как различные методы поощрения.

Любая команда, работающая над проектом, состоит из отдельных личностей и в ней нередко случаются конфликтные ситуации между сотрудниками. Для их решения, руководитель проекта может использовать несколько методов:

- 1) принуждение – склонить сторону (стороны) конфликта к своей точке зрения, не интересуясь их мнением. Преимущество такого метода состоит в отсутствии больших временных затрат на решение проблемы;

- 2) сглаживание – такой метод призывает стороны (или одну сторону) конфликта к солидарности и необходимости забыть о проблеме. Недостатки его заключаются в том, что гнев имеет свойство накапливаться и в будущем может вызвать ещё большее количество проблем;

- 3) компромисс – в таком случае обе стороны конфликта принимают точку зрения друг друга и находят решение, устраивающее их всех. Данный метод хорош, однако требует больше времени на решение проблемы.

Руководитель проекта должен сочетать в себе различные способности: от профессиональных знаний до наличия лидерских и организаторских качеств. Он должен уметь:

- 1) создавать условия для работы команды проекта (т.е. решать текущие проблемы и обеспечивать команду всеми необходимыми ресурсами для работы);
- 2) грамотно использовать материальные и людские ресурсы при планировании работ проекта;
- 3) внимательно следить за ходом выполнения работ проекта и своевременно искать новые методы решения поставленных задач;
- 4) регулярно проводить анализ работы коллектива и результатов реализации проекта;
- 5) не допускать в коллективе появления конфликтных ситуаций, а в случае их возникновения оперативно принимать решения по их ликвидации.

Следование выше озвученным методам управления персоналом проекта позволит руководителю может грамотно распоряжаться человеческим ресурсом, управлять настроением внутри коллектива и не только сократит число возможных рисков между членами команды, но и создать благоприятные условия для осуществления эффективной работы над проектом.

Управление проектами

Анализ научного задела в области проектирования социально-экономических процессов, их моделирования и управления позволяет рассматривать проектное управление как ведущую инновационную сферу.

Во-первых, потому, будущее уже становится полноценным предметом построения и проектирования, и будет выполнять функции организации научного знания и связи с практикой.

Во-вторых, проектирование создает связь между теорией и практикой, осуществляет процесс применения теоретических разработок, в котором проект приобретает статус средства преобразовательной деятельности.

В-третьих, управление проектами реализует воплощения потенциальных тенденций в актуальные процессы:

- принятие во внимание перспективности потребностей;
- практические возможности по реализации поставленных целей;
- общественная целесообразность и управляемость создаваемых проектов;
- надежность, эффективность и общая результативность проектов.

Управление проектом – это системное применение методов и инструментов менеджмента в целях достижения намеченных целей согласно с жизненным циклом проекта.

Деятельность по управлению проектом развития региона должна быть [1,2,3,4]:

- ориентирована на повышение благосостояния общества в целом и удовлетворенность заинтересованных лиц в частности.

- направлена на обеспечение эффективности не только в экономической, но и морально-этической части процесса, что в целом требует оценки уровня развития объекта управления и возможности мониторинга его состояния;

- направлена на улучшение процесса коммуникации пространства, взаимодействия между субъектами проекта, что оказывает большое влияние на специфику информационной и социальной среды, в котором реализуется проект.

Каждый проект разрабатывается при условии и с учетом того, что для его реализации создается соответствующая структура по организации и специфическая система управления проектом.

Риски при управлении проектом

При осуществлении управления проектами нельзя обойтись без рисков. Они всегда связаны с неопределенностью. В связи с этим необходимо обращать внимание на два момента: степень неопределенности и причины ее возникновения. Под степенью неопределенности стоит понимать такое состояние объективных условий, при которых проект может быть принят к исполнению, но которое в свою очередь не позволяет предвидеть последствия решений по причине неточности или неполноты имеющейся информации. Степень неопределенности очень значима, в силу того что можно управлять только такими рисками, по которым имеется значимая информация.

Под риском в проектной деятельности стоит понимать то возможное событие, при котором принявший решение субъект, не сможет получить запланированных результатов проекта или его отдельных параметров, имеющих различную оценку. При управлении проектом характеристику риска можно связать с определенными источниками или причинами. В результате их воздействия у него появляются последствия, таким образом риск оказывает влияние на результативность проекта. Ключевыми факторами в определении являются:

- вероятность;
- событие;

- субъект;
- решение;
- потери.

Величина рисков напрямую зависит от величины проекта. Однако, минимизация рисков в начале разработки проектной документации не гарантирует их полное отсутствие. Необходимо уметь не только предугадать все возможные риски, которые могут возникнуть, но и иметь план мероприятий по устранению возникших проблем в кратчайшие сроки. Управление рисками всегда должно быть частью проекта. При проектировании принято выделять следующие основные шаги по управлению рисками:

- 1) идентификация рисков – определение факторов, в связи с которыми он может проявиться;
- 2) анализ рисков – определение последствий риска для проекта;
- 3) составление плана по реагированию на возникшие риски – составление мер для снижения влияния последствий от возникших рисков.

При ведении проекта рекомендуется регистрировать риски (т.е. вести их учёт), таким образом можно упростить процесс их анализа.

Риски принято разделять на контролируемые и неконтролируемые. Контролируемым принято считать риск, на который можно создать план реагирования, то есть продолжить проект. В случае, если план реагирования составить невозможно, к тому же последствия от риска слишком велики, то ресурсов на реализацию проекта может не хватить и лучше остановить проект. Так же очень важным фактором при управлении проектом является наличие умения правильно расставить приоритеты, исходя из возможных последствий.

Грамотное управление проектными рисками часто приносит пользу организации, т.к. помогают сэкономить ресурсы и время на реализацию проекта, а также повысить его качество.

Системы управления проектами

Наиболее важным элементом управления системой является то, что эффективная система управления проектом предполагает наличие двух неотъемлемых составляющих: индивидуальной компетенции менеджеров и корпоративной зрелости организации в области проектного менеджмента [5].

При реализации крупных проектов рекомендуется использовать системы управления проектами (т.е. программное обеспечение, автоматизирующее процесс управления проектами). На данный момент существует множество систем для управления проектами, но вся суть их работы сводится к следующему:

- 1) контролировать ход проекта;
- 2) оптимально распределить ресурсы;
- 3) контролировать нагрузку сотрудников;
- 4) контролировать финансовые расходы;
- 5) ускорить обмен информацией между работниками.

Однако наряду с большим функционалом использование таких систем имеет и свои недостатки:

- 1) высокая стоимость покупки программного продукта;
- 2) необходимость регулярно обновлять (покупать) новую лицензию для программного продукта;
- 3) ошибки в программном продукте (использование устаревшей версии);
- 4) короткий срок поддержки программного продукта разработчиками;
- 5) дорогостоящее внедрение программного продукта из-за высоких системных требований к комплектующим персонального компьютера;
- 6) неудобный или непонятный интерфейс программного продукта;
- 7) защита данных от несанкционированного доступа;
- 8) содержание подпрограмм, не имеющих отношение к выполнению поставленной задачи.

В современном мире большинство организаций решают проблемы программно-информационного обеспечения нанимая в штат ИТ-персонал, который занимается созданием программных продуктов непосредственно для решения задач организации, в которую был принят.

При наличии квалифицированного ИТ-персонала в организации возможно решить ряд проблем, таких как:

- высокая стоимость покупки программного продукта;
- необходимость регулярно обновлять (покупать) новую лицензию для программного продукта;
- ошибки в программном продукте (использование устаревшей версии);
- короткий срок поддержки программного продукта разработчиками;
- неудобный или непонятный интерфейс программного продукта;
- содержание подпрограмм, не имеющих отношение к выполнению поставленной задачи.

Использование Web-технологий позволяет защитить данные от несанкционированного доступа, а сама технология не требует высоких системных требований к комплектующим персонального компьютера. Таким образом возможно решить следующие проблемы: дорогостоящее внедрение программного продукта из-за

высоких системных требований к комплектующим компьютерного оборудования; защита данных от несанкционированного доступа. При этом доступ к таким системам осуществляется с любого компьютера, для этого не требуется установка дополнительных подпрограмм для их работы.

При создании проекта рекомендуется сразу определиться с применяемым интерфейсом и описать логику работы программы. Можно заранее, на этапе проектирования, провести анкетирование пользователей по выявлению пожеланий к организации интерфейса системы и ее функциональным возможностям.

В период тестирования проекта не рекомендуется сразу исправлять все ошибки и недостатки. Исправлению подлежат лишь критические ошибки, которые либо оказывают влияние на работу системы, либо блокируют работу пользователя. Широкое применение при автоматизации учёта и сбора ошибок, получили методы сбора, передачи и анализа специализированных отчетов - логов системы. При активном использовании системы наиболее подробное логирование действий пользователя даст возможность для скорейшего устранения возникших проблем при функционировании системы. Исправление текущих и мало критичных по уровню функционала ошибок рекомендуется исправлять за определенный период (день, неделя, месяц), собрав информацию об их возникновении, проанализировав, расставив приоритеты, а лишь затем ведя поэтапное исправление согласно принятым приоритетам. Такие исправления следует проводить так называемыми текущими обновлениями или кумулятивными (накопительными). Текущие обновления принято использовать для оперативного исправления критических ошибок, влияющих на безопасность системы, ее корректную работу или исправляющие функциональные ошибки системы, выпускать их по мере устранения таких ошибок. Кумулятивные обновления напротив реализуют большое число незначительных изменений в системе или же изменения, коренным образом затрагивающие ее функционал. Такие обновления принято выпускать привязываясь к определенным временным периодам (месяц, квартал, полугодие, год).

Заключение

Современные тенденции развития проектного управления постоянно повышают требования, предъявляемые при проектировании системы управления. Задачи, реализуемые в этих проектах, становятся с каждым новым витком все сложнее и требовательнее к организации процесса реализации и функционирования. Данное направление, в совокупности с принимающим увеличенные темпы роста web-ориентированных большим количеством сервисов, делает рассмотренные аспекты создания и управления такими системами очень важными и актуальными в современной системе управления проектами. В данной статье были рассмотрены основные задачи, с которыми сталкиваются специалисты информационных технологий, занятые в проекте. Рассмотрены специалисты ИТ технологий, как одни из ключевых фигур командной работы по достижению результата. Также рассмотрены основные необходимые знания, без которых работа над такими проектами с благоприятным результатом просто невозможна. Выделены наиболее важные теоретические и практические навыки идентификации, анализа вероятных угроз применяемые специалистами при проектировании в современных реалиях и рассмотрены основные методики реагирования на возникающие неблагоприятные события. Потому данная тема не только на сегодняшний день, но и в обозримом будущем является и будет являться актуальной и острой. А работа над способами и методами реализации таких проектов даст толчок к развитию не только информационных технологий, но и изменению мирового социума в целом.

Использование системы управления проектами позволяет организации контролировать процесс управления проектом, что в свою очередь помогает снизить ресурсы на его реализацию и делать операции более прозрачными.

Следовательно, система управления проектом - это единый неразрывный комплекс организационных, методических, технических, программных и информационных мер и средств, которые направлены на создание, реализацию и повышение эффективности процессов и управления проектами.

В целях обеспечения эффективной реализации таких процессов, необходимо проводить соответствующие действия, направленные на оптимизацию и ускорение реализации проекта:

- 1) следует определить количество, последовательность и характер операций, составляющих этот процесс;
- 2) провести адаптацию всех операций, применяя современные методики и доступные технические средства;
- 3) выделить оптимальные условия для создания и всего времени жизненного цикла процесса реализации задач проекта.

Для повышения эффективности управления важную роль играет уровень техногенности процесса. Именно современные технологии в совокупности с требованиями времени позволяют и даже вынуждают вести постоянное развитие методологии управления процессами, как неотъемлемую часть развития самих систем управления проектами.

Литература

1. Азаров Н.Я., Ярошенко Ф.А., Бушуев С.Д. Инновационные механизмы управления программами развития. – К: Саммит-Книга», 2011
2. Инновационные механизмы управления программами развития / Азаров Н.Я., Ярошенко Ф.А., Бушуев С.Д. – К: «Саммит-Книга», 2011
3. Леньков Р.В. Социальное прогнозирование и проектирование : [учеб.пособие] / Р.В. Леньков. – М.: ЦСП и М, 2013
4. Науковий супровід, моніторинг та оцінка ефективності соціальних проектів / О.О.Яременко, О.Р.Артюх, О.М.Балакірева та ін. – К.:ДЦССМ, 2002. – С.5-102
5. Бушуев С.Д., Ярошенко Р.Ф. Управління інноваційними програмами на основі застосування системи знань P2M та технології Кайзен // Науковий вісник Міжнародного гуманітарного університету. – 2010.- № 2. – С.9-12
6. Руководство к Своду знаний по управлению проектами PMBoKGuide. - PMIStandardsCommittee. - Москва, 2004
7. Управление социальными проектами: терминологический словарь / сост. Ю. Э. Комлев. – Оренбург: ОГИМ, 2009. – 22 с
8. КССП-1.1 Евразийский стандарт управления проектами (корпоративная версия). Расширение для социальных проектов. Версия 1.1/080908 / Евразийский Центр Управления Проектами. Коллегия Аналитиков. – Москва, 2008. – 30 с. - [Электронный ресурс] // Режим доступа: URL: <http://www.collegian.ru/files/ESUP.pdf>
9. Чурилов А. А. Понятие, роль и актуальность проектного управления в России // Молодой ученый. — 2013. — №3. — С. 299-301. — URL <https://moluch.ru/archive/50/6423/> (дата обращения: 13.10.2018)

Олейник А.Н., Горун В.В., Гуридов А.Г. Проблематика организации проектирования систем при реализации методов программно-информационного обеспечения управления проектами, оценка рисков и способы их устранения. В настоящей статье осуществлен краткий обзор базовых принципов и методов, применяемых для управления современными проектами. Тенденции развития проектного управления постоянно повышают значение данного компонента системы. Рассмотрены задачи специалистов информационных технологий, занятых в проекте, как ключевых фигур командной работы по достижению результата проектной задачи, а также рассмотрены основные знания, в которых они нуждается при работе с такими проектами. Наиболее же важными являются практические навыки идентификации, анализа вероятных угроз и методика реагирования на возможные вызовы неблагоприятных событий. Потому данная тема на сегодняшний день является актуальной и острой.

Ключевые слова: проект, система, риски, программное обеспечение, информационные технологии.

Oleynik Alexey, Gorun Valentina, Guridov Alexander. Problems of organization of system design in the implementation of methods of software and information support for project management, risk assessment and ways to eliminate them. This article provides a brief overview of the basic principles and methods used to manage modern projects. Development trends of project management are constantly increasing the value of this component of the system. The tasks of information technology specialists employed in the project, as key figures of teamwork to achieve the result of the project task, are considered, and the basic knowledge they need to work with such projects are considered. The most important are practical skills of identification, analysis of probable threats and methods of responding to possible challenges of adverse events. Therefore, this topic is relevant and acute today.

Keywords: project, system, risks, software, information technology.

Модульный подход к представлению моделей методик проектирования в интеллектуальных САПР

Григорьев А.В.

Донецкий национальный технический университет
grigorievalvl@gmail.com

Григорьев А.В. Модульный подход к представлению методик проектирования в интеллектуальных САПР. В работе определены общие принципы модульного построения базы знаний для методик проектирования в интеллектуальных САПР, а так же ряд конкретных методов представления знаний и методов организации логического вывода в рамках предлагаемого общего подхода.

Ключевые слова: методика проектирования, модель, интеллектуальные САПР, модуль, представление знаний, логический вывод.

Введение

Эффективность процессов проектирования в любых предметных областях в настоящее время во многом определяется наличием и степенью развития средств, позволяющих проектировщику формализовать и повторно использовать экспертные методики для проектирования тех или иных изделий [1-6].

Методики проектирования призваны существенно упростить проектировщику решение задачи проектирования, исходя из требований технического задания на изделие (ТЗ). Такие средства позволяют создавать методики проектирования, сохранять их, редактировать и - применять по мере надобности. Эффект в этом случае состоит в сокращении времени и трудоемкости процесса принятия проектных решений. Учитывая актуальность задачи построения инструментальных средств создания моделей методик проектирования, важной представляется задача анализа существующих подходов в этой области, а так же определение перспективных направлений их дальнейшего развития. В настоящее время подходы, применяемые для представления методик проектирования в современных интеллектуальных САПР, можно условно подразделить на три группы: онтологии, алгоритмы проектирования, программный синтез.

Методы, предложенные ранее автором, относятся к программному синтезу [7-22]. Однако, данные работы не содержат систематического изложения понятия методики проектирования.

Цель предлагаемой работы:

- дать более полное, систематическое определение понятия «модель методики проектирования» в технологии мета-эвристических оболочек;
- дать более систематическое описание методов представления знаний и организации процесса вывода для модели методики проектирования» в технологии мета-эвристических оболочек.

1. Определение модели методики проектирования в технологии мета-эвристических оболочек

Дадим систематическое определение понятия «методика проектирования в технологии «мета-эвристические оболочки».

Модель методики проектирования – это:

- 1) Состав базы знаний, как набор модулей знаний для структурных блоков различных типов;
- 2) Алгоритм вывода, построенный как вывод на множестве вложенных модулей знаний, включая:
 - 2.1) Вывод, определенный в рамках единичного модуля знаний для блока некоторого типа, целью которого есть построение:
 - структуры блока, как состава нижележащих блоков, рассматриваемых как нижележащие модули знаний с конкретными постановками задачи дальнейшего синтеза;
 - набора связей между модулями знаний, как средства доопределения внешних границ внутренних блоков, т.е. – модулей знаний;
 - набора внешних границ внутренних подблоков, как постановки задачи для синтеза нижележащего модуля знаний, относящегося к некоторому типу структурных блоков;

В работе [13] дана общая модель методики проектирования и - единичного модуля знаний.

Дадим краткую характеристику этого подхода.

Данный подход:

- работает только с семантикой модели;
- не затрагивает грамматику представления модели блока;
- задает общую структуру модели блока;
- задает структуру множеств, составляющих структуру модуля знаний;
- задает сборник всех аксиом, определяющих требования к составу ИЛИ-узлов и И-узлов И-ИЛИ-дерева при декомпозиции блоков;
- задает сборник возможных типов правил вывода, задающий рамки, общий состав правил вывода, но не сами правила вывода;
- задает конкретный состав правил вывода в заданных рамках (т.е. – наличие правил вывода каждого типа) должен определяться режимом (умный, средний, глупый);
- форма задания правил вывода напрямую определяется конкретным методом представления знаний, который можно определить в рамках данного подхода;
- определяет состав продукций, составляющих единичный модуль знаний,
- отдельный тип продукций рассматривается как отношение между множествами структурных компонент различного типа;

Вывод: подход задает общий подход к заданию методики проектирования, т.е.:

- 1) общую структуру модуля знаний и набор типов продукций;
- 2) общий порядок вывода на наборе модулей знаний;
- 3) общий порядок вывода в рамках единичного модуля знаний;

Различные подходы к представлению знаний позволят доопределить общий подход до уровня модели методики проектирования.

2. Общие методы построения методики проектирования

Определим общую модель модуля знаний и – общий порядок синтеза желаемой структуры блока в рамках единичного модуля знаний.

2.1. Модель структуры блока как механизм деления единичного модуля знаний на отдельные подмодули.

Исходя из предлагаемой структуры модели структурного блока, рассматривается и соответствующий состав подмодулей модуля знаний. А именно:

Блок, как компонент внешней среды:

- внешней границы;

Внутренняя среда:

- внутренней границы (состав подблоков различных типов, обратная граница);
- набор структурных связей между подблоками.
- внешние границы подблоков;

2.2. Общий порядок синтеза желаемой структуры блока в рамках единичного модуля знаний.

Общий порядок синтеза на совокупности подмодулей знаний, составляющих модуль знаний о структуре блока, таков:

- определение состава внешней границы блока на основе совокупности связей во внешней среде;
- автоматический вывод на множестве подмодулей знаний, исходя из совокупности данных, составляющих внешнюю границу блока;
- «ручное» доопределение части недоопределенных ИЛИ-узлов, выполняемое путем ведения диалога с пользователем, осуществляющим прямой выбор требуемой ему альтернативы, как части ввода технического задания на изделие путем определения необходимых структурных отличий блока.

Существуют различные стратегии вывода на данном множестве подмодулей. Они разнятся в способе представления знаний о наборе структурных компонент различных типов, отличающегося :

- количеством подмодулей;
- наличием и формой задания семантических связей между отдельными подмодулями.

Например:

- наличием или – отсутствием подмодуля определенного типа;
- наличием и количеством отношений совместности-несовместности между отдельными подмодулями знаний;
- наличием и количеством отношений зависимости между подмодулями знаний в форме общей части описания структурной модели блока.

3. Отношения зависимости между отдельными подмодулями знаний как И-ИЛИ-деревьями.

Отдельный подмодуль знаний представляет собой онтологию в форме И-ИЛИ-дерева, над ИЛИ-узлами которого определен ряд отношений совместности-несовместности.

Над совокупностью И-ИЛИ-деревьев также может быть определен набор зависимостей.

Зависимости могут задаваться двумя способами:

- 1) в форме отношений совместности-несовместности между ИЛИ-узлами двух И-ИЛИ-деревьев;
- 2) в форме общей части описания структурной модели блока.

Рассмотрим эти формы зависимости детальнее.

3.1. Отношения зависимости между И-ИЛИ-деревьев в форме отношений совместности-несовместности

Над совокупностью И-ИЛИ-деревьев также может быть определен набор отношений совместности-несовместности.

Отношения могут выстаиваться:

- между всеми возможными парами И-ИЛИ-деревьев;
- между подмножеством возможных пар И-ИЛИ-деревьев;

Если имеются все возможные пары отношений, то порядок вывода может быть любой, с любым критерием оптимизации вывода. Например:

- по критерию наибольшего сужения неопределенности, предложенного ранее автором.

Если имеется только определенный набор пар отношений совместности-несовместности между парами И-ИЛИ-деревьев, то тогда порядок вывода определяется как «от уже определенного дерева к еще не определенному».

3.2. Отношения зависимости между И-ИЛИ-деревьев в форме общей части описания структурной модели блока

Отношения между парами И-ИЛИ-деревьев может выстаиваться не только посредством отношений совместности-несовместности, но и - посредством общей части множества используемых элементов структуры одного типа.

Например:

- 1) Внешняя граница однозначно определяет часть внутренней обратной границы блока;
- 2) Внутренняя обратная граница детализируемого блока определяет:

- часть структурных связей, существование которых определяется наличием того или иного компонента внутренней обратной границы;

- 3) Состав структурных связей однозначно определяет:

- состав всех внутренних подблоков;
- все внешние границы всех внутренних подблоков;

4) Состав внутренних подблоков однозначно определяет часть структурных связей, входящих во внутреннюю среду блока.

Специфика вывода связанных подмодулей на основе структурной связи:

- в связи есть идентификация блока (внешней границы), к ней добавляется единичное данное, расширяя тем самым текущее состояние множества внешней границы;

- в этом самом множестве внешней границы уже есть системообразующая часть, не зависящая от условий применения;

- факультативная часть представляет собой И-ИЛИ-дерево; И-ИЛИ-дерево сужается после каждого шага вывода;

- т.о., структурная связь есть правило вывода, трактуемое так «наличие связи порождает необходимые данные у внешних границ связанных блоков»;

- порождаемое данное изменяет (сужает) набор вариантов представления факультативной части; при этом порождаются новые постоянные части внешней границы;

3.3. Некоторые специфические формы представления знаний и организации вывода

Рассмотрим различные способы представления знаний и организации логического вывода в случае отсутствия отношений совместности-несовместности между отдельными подмодулями знаний.

В этом случае вывод определяют только отношений зависимости между подмодулями знаний в форме общей части описания структурной модели блока

Например:

- 1) «Сокращенный» способ синтеза.

Хранится:

- подмодуль внешней границы;
- подмодуль обратной границы;
- подмодуль состава внутренних связей.

Можно хранить среды как наборы связей, создавая И-ИЛИ-дерево, объединяющее воедино различный состав связей для тех или иных прототипов.

В дальнейшем выбранный состав связей однозначно определяет всю внутреннюю структуру блока.

Т.е., исходя из выбранного состава связей, порождается:

- состав внутренних подблоков;
- внешние границы внутренних подблоков.

2) «Обычный» способ синтеза.

Хранятся все необходимые подмодули:

- внутренней границы (состав подблоков);
- набор структурных связей между подблоками;
- внешние границы всех типов подблоков;

Исходя из состава данных, составляющих внешнюю границу блока, последовательно выполняется синтез необходимого состава:

- соответствующий фрагмент внутренней обратной границы детализируемого блока;
- внутренних связей;
- внутренних подблоков, как недостающую часть «внутренней» обратной границы блока;
- внешних границ внутренних подблоков.

Выводы

В работе достигнуты следующие результаты:

- дано более полное, систематическое определения понятия «модель методики проектирования» в технологии мета-эвристических оболочек;
- дано более систематическое описание методов представления знаний и организации процесса вывода для модели методики проектирования» в технологии мета-эвристических оболочек.

По вышеизложенному материалу можно сделать вывод, что задача, поставленная в цели работы, выполнена.

В качестве перспектив следует выделить следующие работы:

- уточнить специфику модельного представления методик проектирования для различных режимов построения базы знаний, связанных с уровнем квалификации проектировщика в инженерии знаний, а именно – эксперта низкого уровня квалификации, не способного вообще строить продукционные системы, эксперта среднего уровня квалификации, способного явно строить И-ИЛИ-деревья, и – эксперта высоко уровня квалификации, способного самостоятельно построить простейшую продукционную базу знаний;
- соотнести специфику данного подхода с альтернативными способами модульного представления знаний.

Литература

1. Gruber T. Towards principles for the design of Ontologies used for knowledge sharing // International Journal of Human-Computer Studies. – 1995. – № 43(5/6). – С. 907-928.
2. Гаврилова Т.А. Базы знаний интеллектуальных систем / Т.А. Гаврилова, В.Ф. Хорошевский. – С-Пб.: Питер, 2000. – 384 с.
3. Нариньяни А.С. Кентавр по имени теон: тезаурус+онтология [Электронный ресурс] / А.С. Нариньяни. - Режим доступа: <http://www.dialog-21.ru/materials/archive.asp?id=6771&y=2001&vol=6077>.
4. Норенков И.П. Разработка систем автоматизации проектирования. М.: МГТУ им. Э.Н.Баумана, 1994. – 207 с.
5. Поспелов Д.А. Ситуационное управление: теория и практика. М.: Наука, 1986. - 288 с.
6. Нариньяни А.С. Недоопределенность в системах представления и обработки знаний. // Известия АН СРСР. Техническая кибернетика. - 1986.- № 5. – С. 3—28.
7. Григорьев А.В. Комплекс моделей САПР как система взаимосвязанных уровней о действительности. // Научные труды ДонНТУ. Серия "Информатика, кибернетика и вычислительная техника", (ИКВТ-2000) выпуск 10. - Донецк, ДонГТУ, 2000. - С. 155—167.
8. Григорьев А.В. Семиотическая модель базы знаний САПР. // Научные труды Донецкого государственного университета. Серия "Проблемы моделирования и автоматизации проектирования динамических систем". Выпуск 10: - Донецк: ДонГТУ, 1999. - С. 30—37.
9. Григорьев А.В. Представление недоопределенности знаний в инструментальной оболочке для построения САПР. // Искусственный интеллект. N 6, 1999 , С. 56—66.

10. Григорьев А.В. Состав параметров адаптации на проблемную область в инструментальной оболочке для автоматизации построения интеллектуальных САПР. В кн. Наукові праці ДонНТУ. Серія „Інформатика, кібернетика та обчислювальна техніка” (ІКОТ-2011). Випуск 14(185).- Донецьк: ДВНЗ «ДонНТУ», 2011.С. 252—261.
6. Григорьев А.В. Пути создания интеллектуальных САПР при различных уровнях квалификации экспертов / А.В. Григорьев // Научно-теоретический журнал «Искусственный интеллект». – 2005. – №3. - С. 758–763.
11. Григорьев А.В. Вербальная модель предметной области для интеллектуальных САПР. Наукові праці Донецького Державного технічного університету. Серія "Обчислювальна техніка та автоматизація". Випуск 20. - Донецьк, ДонДТУ, 2000. - С. 171-180.
12. Григорьев А.В. Семантика модели предметной области для интеллектуальных САПР. Научные труды Донецкого государственного университета. Серія "Информатика, кибернетика и вычислительная техника", (ИКВТ-2000) выпуск 10. - Донецк, ДонГТУ, 2000. - С. 148-154.
13. Григорьев А.В. Содержание некоторых категорий абстракций в теории построения интеллектуальных САПР. В кн. Наукові праці Донецького національного технічного університету / Редкол.: Башков Є.О. та інші. Серія: "Обчислювальна техніка та автоматизація": Випуск 64.: Донецьк: Видавництво ДонНТУ, 2003 - С. 166-178.
14. Григорьев А.В., Каспаров А.А. И/ИЛИ-дерево как средство абстрактного представления знаний. Наукові праці національного технічного університету. Серія «Інформатика, кібернетика та обчислювальна техніка». Випуск 39: Донецьк: ДонНТУ, 2002. - С.36-42.
15. Григорьев А.В. Классификация типов продукций в интеллектуальных САПР / Наукові праці національного технічного університету. Серія «Обчислювальна техніка та автоматизація». Випуск 88. — Донецьк: ДонНТУ, 2005. – с. 99-105.
16. Григорьев А.В. Упорядочивание обликков в семиотической модели САПР /Научно-теоретический журнал «Искусственный интеллект», №4, 2005. – Донецьк: ИПИИ МОН и НАН Украины «Наука и образование», 2005. – с. 465–477.
17. Григорьев А.В. Специфика выполнения теоретико-множественных операций над контекстно-свободными грамматиками в условиях различных форм дополнительных семантических правил в семиотической модели интеллектуальных САПР Научные труды ДонНТУ. Серія «Проблемы моделирования и автоматизации проектирования динамических систем» (МАП – 2006). Випуск 5 (116). – Донецьк: ДонНТУ, 2006. – С. 91—104.
18. Григорьев А.В. Теоретико-множественные операции над грамматиками как механизм работы со знаниями в интеллектуальных САПР. Вісник Східноукраїнського національного університету імені Володимира Даля, N 2(48). Луганск, ВУТУ, 2002. С. 186-194.
19. Григорьев А.В. Алгоритм выполнения теоретико-множественных операций над грамматиками в среде специализированной оболочки для создания интеллектуальных САПР. Наукові праці національного технічного університету. Серія «Проблемы моделирования и автоматизации проектирования динамических систем» (МАП -2002). Випуск 52: Донецьк: ДонНТУ, 2002. - С.83-93.
20. Григорьев А.В. Изобретение как метод автоматизации процесса обучения методикам проектирования в семиотической модели САПР // Информатика и кибернетика. – Д.: ДонНТУ, – 2015. – № 1. - С. 51— 66.
21. Григорьев А.В., Морозова О.В. Анализ существующих способов создания интерфейса «языки формальных спецификаций — проблемно-ориентированные языки». В кн. Серія „Інформатика, кібернетика та обчислювальна техніка” (ІКОТ-2011). Випуск 14(185).- Донецьк: ДВНЗ «ДонНТУ», 2011. - С. 270—275.
22. Григорьев А.В. Ограничение когнитивной сложности моделей. В кн. Прогрессивные технологии и системы машиностроения.: Международный сб. научных трудов. - Донецьк: ДонГТУ, 2000. Випуск. 10 - С. 49-58.

Grigoriev A.V. A modular approach to the presentation of design techniques in intelligent CAD systems. The work defines the general principles of the modular construction of a knowledge base for design techniques in intelligent CAD systems, as well as a number of specific methods for representing knowledge and methods for organizing logical inference as part of the proposed general approach.

Keywords: design methodology, model, intelligent CAD, module, knowledge representation, logical inference.

Модель представления знаний в интеллектуальной надстройке над проблемно-ориентированными САПР

Морозова О.В.

Донецкий Национальный Технический Университет

olmalyavka@gmail.com

Морозова О.В. Модель представления знаний в интеллектуальной надстройке над проблемно-ориентированными САПР. Рассмотрена специфика представления знаний в теории искусственного интеллекта, описаны методы представления знаний в современных инструментальных комплексах и методы представления знаний, применяемых в проблемно-ориентированных САПР. Рассмотреть подход к представлению знаний в проблемно-ориентированном САПР, где эксперта в предметной области является инженером по знаниям.

Ключевые слова: CATIA, AutoCADPlant 3D, Clips, Prolog, Protégé, AVEVA Diagrams, И-ИЛИ дерево, продукция, посылка, следствие, модуль знаний.

Введение

Современные задачи конструирования сложных объектов привели к созданию большого количества универсальных программных средств САПР. Для наиболее продуктивной работы данные САПР повышают свой уровень интеллектуальности за счет включения и использования производственных баз знаний, нейронных сетей и т.д. Эти тенденции получили названия гибридных САПР.

Соединения САПР и искусственного интеллекта привела к тенденции использования эксперта в предметной области как инженера по знаниям. На эксперта в предметной области накладываются дополнительные действия заполнения базы знаний. Актуальной проблемой является то, в каком виде будут представлены данные для дальнейшего их использования в интеллектуальном модуле САПР. Современные интеллектуальные САПР пытаются использовать общие подходы из теории искусственного интеллекта, но с учетом специфики САПР - и возможностей разработчиков баз знаний [1].

Представления знаний в САПР должны быть в виде методики проектирования конкретной предметной области и достоверно описывать каждый объект проектирования со всеми взаимосвязанными компонентами готового решения. База знаний методик проектирования формируется из набора входных компонентов на том или ином уровне абстракции модели представления предметной области.

Целью данной статьи является рассмотреть специфику представления знаний в теории искусственного интеллекта, изучить методы представления знаний в современных инструментальных комплексах и методы представления знаний, применяемых в проблемно-ориентированных САПР. Рассмотреть подход к представлению знаний в проблемно-ориентированном САПР, где эксперта в предметной области является инженером по знаниям.

Общие методы представления знаний в теории искусственного интеллекта

Для использования знаний, вносимых в базу знаний, необходимо решать задачи, связанные с составом элементов представления знаний; организации знаний; моделью представления знаний и обоснованием выбранного метода представления знаний.

В общей теории искусственного интеллекта используется две модели представления знаний, такие как логические (формальные) и эвристические (формализованные) [2].

Представление знаний в логических моделях имеют вид отношений, которые образуются между знаниями и могут быть выражены средствами синтаксической записи формальной системы. Основными знаниями логических моделей являются исчисление предикатов первого порядка и системы правил. Эта модель реализует методы вывода, в которых системы посылок регламентируют вывод результата из набора правил. Они относятся к системам дедуктивного типа. Методы вывода в логических системах - это метод резолюции и обратный метод.

Основными эвристическими моделями можно назвать сетевые модели (семантические сети), фреймовые, производственные и объектно-ориентированные модели. Формализованные модели сводятся к более сложному представлению знаний и их организации, содержат информацию о специфике предметной области.

Продукционная модель представляет собой набор правил типа «ЕСЛИ условие, ТО действие», в общем виде «Посылка и следствие». Отличием данной модели от логической: правила имеют более сложную структуру и содержат конкретную специфику предметной области, которые представлены в виде семантических структур.

Семантические сети или сетевые модели представляют собой связный граф, в вершинах которого находятся объекты (понятия) предметной области, а на ребрах отношения между этими объектами.

Фреймовые модели представляют структуру данных концептуального описания объекта. Знания (информация) хранится в слотах фрейма, данные слоты могут являться отдельными фреймами либо соединяться в иерархическую сеть.

Объектно-ориентированные модели, являются продолжением развития фреймовой модели и оперируют такими понятиями как класс и объект. Данная модель является наиболее развитой в искусственном интеллекте. Для представления знаний в объектно-ориентированной модели в качестве объекта можно использовать конкретные предметы. Объект представляет собой класс одинаковых (однотипных) объектов. Класс, в свою очередь, является объединяющим свойством всех объектов.

Методы представления знаний в современных инструментальных комплексах

Большое количество инструментальных комплексов используют разнообразные методы представления и обработки знаний, основанные на стандартных подходах, такие как Clips, Prolog, Protégé, NexpertObject, ExysCorvid, Jess и т.д. Рассмотрим основные, типичные программные продукты и выделим в них модуль, который отвечает за представление знаний [1].

Язык Clips разработан для создания экспертных систем, ориентированных на специалиста предметной области, а не на инженера по знаниям. Clips может работать с такими моделями представления знаний: фреймовая, продукционная и объектно-ориентированная [3].

Рассмотрим, основные конструкторы Clips:

- определение правил; определение фактов;
- определение шаблона факта;
- определение глобальных переменных;
- определение функций;
- определение модулей (совокупности правил);
- определение классов;
- определение объектов по шаблону, заданному defclass;
- определение сообщений для объектов;
- создание заголовка родовой функции;
- определение метода родовой функции.

Конструкторы Clips не возвращают никаких значений. Помимо конструкторов в Clips присутствуют функции разного назначения, такие как функции ввода-вывода; функции работы с составными величинами работы; функции для работы со строками; процедурные функции, реализующие возможности ветвления и организации циклов.

Правила продукционной модели состоят из предпосылок и следствия. Предпосылки в Clips называется LHS-правилами (left-hand-side-of-rule), а следствия называют RHS-правилами (right-hand-side-of-rule). Для создания базы знаний в Clips необходимо внести список вопросов, по которым возможно построить конкретную предметную область. Так же в базе знаний хранится список вопросов, задаваемый пользователю. В Clips база знаний - это массив фактов. Для обработки базы знаний необходимо составить набор правил, которые будут обращаться к базе знаний. Данные правила соответствуют механизму срабатывания и выборки нужного факта из базы знаний, определенный пользователем.

Каждый элемент Clips выполняет следующие функции: в базе фактов хранятся исходные состояния предметной области; в базе правил хранится информация об операциях, которые необходимо выполнить для решения поставленной задачи.

Clips реализует прямой логический вывод, что обеспечивает высокую степень прогнозирования. При этом, если посылки двух правил одинаковы, то необходимо задавать еще один параметр с определенным числовым значением. Данное значение описывает приоритет выполнения правила.

Среда для разработки Prolog является декларативным языком исчисления предикатов [4]. Под предикатом, в системе Prolog понимают логическую запись от различного числа аргументов. В Prolog предикат - это реализация функции, принимающая значения ложь или истина. База знаний Prolog состоит из набора фактов и правил, из которых строится решение, на основе введенных до этого фактов. То есть Prolog получает решение из логического вывода введенных ранее известных тезисов.

Факты в Prolog изображают объекты и отношения, которые за тем описывают правилами, при которых эти отношения являются истинными. Отношение между объектами и есть факт [4]. Если в естественном языке отношения устанавливаются в предложении, то в Prolog (в логике предикатов), отношения соответствуют простой фразе, состоящей из имени отношения и объекта. В Prolog имеются такое понятие как предикат, это

отношение, аргумент – это объект, который связан с этим отношением. Более сложное представление и понятие в Prolog - это правила. Они позволяют вывести один факт из других фактов. Правило является истинным заключением, если было найдено другое истинное заключение или факт является истинным.

Тело правила состоит из одной или более подцелей. При записи нескольких фактов в Prolog задается вопрос, какие отношения между этими фактами. Таким образом, в системе Prolog создаются запросы, то есть вопросы об этих отношениях в фактах. Запрос выполняется перебором всех фактов, начиная с первого, пока не закончатся факты. Обычно программа на Prolog состоит из четырех основных программных разделов, к которым относятся: раздел предложений; раздел предикатов; раздел доменов; раздел целей.

Механизм логического вывода Prolog основан на сравнении фактов и является обратным логическим выводом. С помощью подбора ответов на запросы он извлекает хранящуюся (известную) информацию.

Еще одним рассматриваемым программным продуктом является Protégé. Он является наиболее распространенной программой по созданию онтологий. Онтологии - это объектно-ориентированная модель, в которой определены классы объектов и взаимодействие между данными классами. Объектно-ориентированная модель представляет собой иерархическую структуру данных, включающую классы объектов взаимодействие, правила и ограничения, введенные для данной предметной области. Protégé - это программа для построения баз знаний на основе онтологий. Онтологии сохраняются в различных форматах, таких как RDF, OWL, XML, FunctionalSyntax, ManchesterSyntax, Turtle. В Protégé можно загружать и сохранять существующие онтологии в формате OWL. Так же можно редактировать классы и их свойства [5].

Основная функция работы Protégé - это правила логического вывода SWRL. Как и любое правило, оно состоит из условной части и вывода, по выполнению данного условия. Каждая часть может состоять из нескольких элементарных логических выражений (атомов).

Каждое логическое выражение - это предикат, описывающий объект онтологии. Предикаты могут представлять собой утверждения о связях; утверждения о принадлежности к классу; утверждения о существовании связи; утверждения о значении свойства-литерала; условие раздельности двух индивидуальных объектов; условие совпадения двух объектов; условие о принадлежности значения переменной определенному типу данных; встроены условия.

В целом можно сказать, что перечисленные выше типичные подходы, модифицированные с учетом специфики САПР, являются той базой, на основе которой строятся представления методик проектирования в САПР.

Методы представления знаний, применяемых в проблемно-ориентированных САПР

Современный рынок располагает огромным количеством проблемно-ориентированных САПР для конструирования сложных объектов промышленного назначения, таких как CATIA V7, AutoCADPlant 3D, Грация, SpurtCAM, T-Flex, Pro/ENGINEER, NX, КОМПАС, SolidWorks и т.д. [1].

На примере типичных проблемно-ориентированных САПР рассмотрим основные методы создания интеллектуальных модулей и их возможности организации работы эксперта по знаниям.

Одним из наиболее популярных и, практически, наиболее развитым САПР для проектирования сложных объектов является CATIA V7 (3DEXPERIENCE). В CATIA имеются все возможные стандартные компоненты проектирования и конструирования объектов, так и интеллектуальную часть для внедрения процесса накопления и эффективного использования знаний. CATIA – KnowledgeAdvisor — модуль, который позволяет пользователям внедрять знания в проектирование и использовать их для оказания помощи в инженерных решениях, чтобы уменьшить ошибки или автоматизировать проектирование для максимальной производительности. Пользователи (эксперты по знаниям) могут внедрять знания в проектирование, такие как формулы, правила и проверки, а также использовать их в случае необходимости. Модуль KnowledgeAdvisor содержит средства интерактивного программирования и отладки содержимого базы знаний [6].

В CATIA знания представляются в виде продукционной модели. Эта модель имеет вид набора взаимосвязанных правил, просчета параметров и проверки ограничений.

Важными преимуществами с точки зрения структурного проектирования:

- система отображения структурной сборки;
- быстрая стадия прототипирования;
- возможность анализа детали в пространстве.

Набор инструментариев CATIA включающий встроенный редактор, объектный словарь, список параметров и окно с перечислением параметров модели, также функции проверки синтаксиса и редактирования, позволяет конструктору самостоятельно строить базу знаний. База знаний CATIA представляет собой набор функций, законов и правил. На основе характеристик объекта данная система выполняет вычисления по заданным формулам и логическим операциям. Основным принципом работы базы знаний является объектно-ориентированный подход, в котором реализованы геометрические или негеометрические характеристики объекта. База данных стандартных компонентов сложных объектов с возможными связями и зависимостями представляет собой набор онтологий для дальнейшего использования.

Подобный подход, но более упрощенный, реализован в САПР одежды Грация.

Еще один САПР для цифрового моделирования промышленных объектов AutoCADPlant 3D [1]. Так же является достаточно сложной и мощной системой проектирования и создания технической документации. Ее функциональные возможности позволяют создавать схемы и трехмерные модели объектов. Дополнительными модулями системы AutoCADPlant 3D, является Visual LISP и AutoLISP. Данные модули позволяют расширить технические возможности системы. Условно их можно назвать «базой знаний», так как в них создается некое подобие единичного модуля знаний системы, который расширяет технические спецификации модели. Данные модули можно использовать многократно (использовать как базовый элемент модели), и расширять его технические и функциональные характеристики.

САПР AVEVA Plant — это набор интегрированных между собой приложений для разработки схемной части проекта и создания 3D модели [7, 8]. В интегрированной линейке технологий AVEVA для проектирования есть три основных категории решений:

- схемные решения;
- 3D проектирование;
- контроль и управление данными и изменениями.

Под категорией схемных решений понимают создание схем, диаграмм, опросных листов, экспликаций технологического оборудования и трубопроводов и прочие документы [7, 8]. Схемы содержат в себе основные характеристики систем автоматизации и технологических объектов, на основе которых впоследствии на стадиях проектирования будут приниматься те или иные инженерные решения. С этой задачей справляется приложение AVEVA Diagrams. Решение AVEVA Diagrams полностью совместимо с AVEVA PDMS, AVEVA Outfitting, AVEVA Schematic 3D Integrator и AVEVA NET

AVEVA Diagrams довольно проста в использовании и процесс создания схем построен на основе Microsoft Visio. В арсенале пользователя имеется библиотека символов, разбитая на категории, которые представляют самые распространённые условные обозначения для схем, используемые в проектировании различных отраслей. Также имеется возможность создания собственных символов.

Задействованные в схемах символы могут иметь как базовые параметры, так и заданные пользователем в режиме запроса значений.

При добавлении символов на технологическую схему в иерархии (если возможно) создается соответствующий элемент. Т.е. система определяет символ как элемент схемы. Также существует возможность вручную определить символ как элемент базы данных (Define) или же разорвать связь с элементов иерархии (Undefine) [7, 8].

Приложение AVEVA Diagrams поддерживает концепцию сессий баз данных Dabacon, реализованную через функции SaveWork / GetWork. При этом, помимо записи изменений в базу данных с помощью SaveWork, параллельно создаются файлы соответствующих типов. Благодаря специальным механизмам соблюдается целостность данных проекта [8].

Таким образом, DABACON — специально разработанная для нужд проектирования объектно-ориентированная база данных. Данные можно разделить на несколько типов:

- стандартные компоненты;
- 3D-модель, которую создают проектировщики;
- 2D-данные — чертежи и спецификации, получаемые на основе 3D-модели.

Для хранения каждого такого типа данных в DABACON предусмотрены отдельные базы данных (БД). Информация хранится в БД DABACON в виде иерархической структуры, или «дерева», где каждому элементу базы отведено конкретное место. У каждого типа баз данных есть свое «дерево», и у каждого элемента БД существует определенный набор свойств или атрибутов. Хранение данных в виде дерева элементов, а также наличие у каждого элемента атрибутов существенно упрощает и ускоряет работу с проектной информацией.

Предлагаемый подход к представлению знаний в проблемно-ориентированной САПР

С целью учесть специфику задачи предлагается выполнить совмещение названных подходов (онтологии, алгоритмы проектирования, программный синтез), как среди названных подходов, так и внутренних методов программного синтеза:

- представление алгоритма проектирования как онтологии специфического вида, построенной на базе и-илидерева;
- применение грамматического подхода к «обычным» алгоритмам проектирования: замена программного кода, содержащего операции построения модели объекта проектирования как геометрической модели, на программный код, описывающий модель на внутреннем языке представления объекта проектирования;
- совмещение подходов эксперта высокой и средней квалификации, определённый в теории программного синтеза: использование режима среднего эксперта при создании онтологии объекта проектирования как алгоритма проектирования в форме дерева решений, рассматриваемого как этапы синтеза модели объекта проектирования.

При рассмотрении современных программных продуктов можно выявить специфику адаптации методов представления знаний для проблемно-ориентированных САПР:

- модель представления знаний как модель методики проектирования;
- физическая семантика предметной области;
- специфика данной предметной области;
- тенденция использования эксперта в предметной области как инженера по знаниям;
- необходимость учета достигнутого уровня интеллектуальности инструментов задания методики проектирования в данном проблемно-ориентированном САПР.

Метод самоинтервью представляет собой методику, при которой эксперт предметной области - проектировщик, выписывает все особенности и специфику заданной предметной области, структурные элементы, которые записывает в порядке иерархии. После чего происходит построение семиотической модели [9-12].

Рассмотрим специфику извлечения знаний, при которой эксперт в предметной области выступает инженером по знаниям, что означает:

- эксперт самостоятельно строит продукционную БЗ;
- знания вводятся в БЗ в виде правил, отражающих опыт данного разработчика в области проектирования некоторого типа устройств;
- набор типов правил порядок их следования, т.е. метод представления знаний выбирается самим экспертом в рамках некоторых ограничений.

Моделью представления знаний в данном подходе является продукционная модель. В данном подходе модель представляет собой совокупность продукций, оперирующих с и-или деревьями посылок и выводов.

Как посылки, так и выводы продукций оперируют областью вывода как динамической областью, представленной и-или деревом. По типу и-или деревьев посылок и выводов продукций можно классифицировать типы продукций, которые могут быть использованы в инструментальной оболочке.

Виды продукций по механизму срабатывания [9, 10, 11]:

- проверяющая продукция – это такая продукция, посылка которой, при срабатывании проверяет наличие данного компонента в области вывода базы знаний;
- не проверяющая продукция – это такая продукция, при которой посылка не проверяется на наличие совпадений в области вывода;
- спрашивающая продукция – это продукция, организующая работу с экспертом предметной области, которая в посылке отправляется вопрос;
- не спрашивающая продукция – это такая продукция, при которой посылка напрямую работает с областью вывода, без участия эксперта в предметной области;
- удаляющая продукция – это продукция, при которой наличие посылки означает удаление следствия из области вывода;
- прибавляющая продукция – это продукция, при которой наличие посылки означает добавление следствия в область вывода.

Для организации построения дерева формальных решений используется процедура выбора из ряда продукций, при срабатывании которых наиболее быстро заканчивается вывод. Для оптимизации построения дерева решений и удовлетворения принципа монотонности, будем использовать правильный порядок следования продукций. Из этого следует, что при организации базы знаний, новый факт влияет на имеющийся результат данного этапа и неизменности результатов предыдущих этапов вывода, при том итоговый результат должен удовлетворять всем условиям, заданным в фактах [2]. При этом предполагается учитывать, что факты могут иметь необратимые последствия в базе знаний.

Вывод

В статье рассмотрена специфика и описаны основные модели представления знаний в теории искусственного интеллекта. Рассмотрены модели представления знаний в современных программных продуктах для создания экспертных систем. Так же, описаны и рассмотрены проблемно-ориентированные САПР с их интеллектуальными системами, в которых представлены модели представления знаний. Определены основные аспекты проблемы:

- для представления методик проектирования необходим весь спектр методов;
- наиболее актуальным является продукционный метод представления знаний, позволяющий задавать модель методики проектирования;
- формой представления методики является инструментальный комплекс автоматизации разработки продукционной базы знаний;
- по структуре данный комплекс имеет форму интеллектуальной надстройки, задающей в комплексе с проблемно-ориентированным САПР гибридную систему, где САПР задает «обычную» компоненту представления знаний о методах проектирования, а надстройка – «интеллектуальную» компоненту методики проектирования.

Описан подхода к представлению знаний методик проектирования в интеллектуальной надстройке над проблемно-ориентированными САПР.

Перспективой дальнейшего развития является описание формальной продукционной модели описания методик проектирования в проблемно-ориентированном САПР.

Литература

1. Морозова О.В. Анализ возможностей и недостатков технологии построения баз знаний в современных САПР // Информатика и кибернетика. – Донецк, 2016. - №2(4).
2. Представление знаний в экспертных системах: учебное пособие / сост. В. А. Морозова, В. И. Паутов. — Екатеринбург : Изд-во Урал. ун-та, 2017. — 120 с.
3. Частиков А.П., Гаврилова Т.А., Белов Д.Л. Разработка экспертных систем. Среда CLIPS – СПб.: БХВ-Петербург, 2003. – 608 с.
4. Братко И. Алгоритмы искусственного интеллекта на языке Prolog, 3-е издание. : Пер. с англ. — М. : Издательский дом "Вильямс", 2004. — 640 с.
5. Муромцев Д.И. Онтологический инжиниринг знаний в системе Protégé: Методическое пособие. — СПб: СПбГУ ИТМО, 2007. — 62 с
6. САТІА: орудие производителя XXI века [Электронный ресурс] – Режим доступа: <http://www.sapr.ru/article.aspx?id=7101&iid=290> - Загл. с экрана.
7. AVEVA Diagrams — разрабатываем технологические схемы, схемы ОВиК и однолинейные схемы электрики [Электронный ресурс]/ САПР и графика – 2010 – Режим доступа: <https://sapr.ru/article/21575> - Загл. с экрана.
8. Объектно-ориентированная база данных DABACON как PLM-решение, разработанное специально для проектирования промышленных объектов [Электронный ресурс] – Режим доступа: <https://www.aveva.com/~media/Aveva/Russian-RU/Publications/Technologies/Dabacon%202011.pdf>- Загл. с экрана.
9. Григорьев А.В., Морозова О.В. Создание интеллектуальной оболочки по методикам проектирования локальных сетей при использовании проектировщика как эксперта по знаниям // Наукові праці ДонНТУ, Серія: «Проблемимоделювання та автоматизації проектування». Вип. №1(10)-2(11). – Донецьк, 2012. – С. 144-155.
10. Григорьев А.В. Методы решения задачи структурного синтеза в интеллектуальных САПР, построенных на основе семиотической модели структур // Наукові праці Донецького національного технічного університету. Серія: “Обчислювальна техніка та автоматизація”. Випуск 19 — Донецьк: ДонНТУ, 2010. — С. 128-141.
11. Морозова О.В., Григорьев А.В. Анализ методов построения экспертных систем в продукционных инструментальных оболочках // Наукові праці ДонНТУ. – Донецьк, 2012. – Вип. 16(204). – С. 180-191.
12. Григорьев А.В., Морозова О.В. Анализ существующих способов создания интерфейса «языки формальных спецификаций — проблемно-ориентированные языки» // Сборник научных трудов донецкого национального технического университета. – Донецк, 2011. – № 14. – С. 270 – 275.

Морозова О.В. Модель представления знаний в интеллектуальной надстройке над проблемно-ориентированными САПР. Рассмотрена специфика представления знаний в теории искусственного интеллекта, описаны методы представления знаний в современных инструментальных комплексах и методы представления знаний, применяемых в проблемно-ориентированных САПР. Рассмотреть подход к представлению знаний в проблемно-ориентированном САПР, где эксперта в предметной области является инженером по знаниям.

Ключевые слова: САТІА, AutoCADPlant 3D, Clips, Prolog, Protégé, AVEVADiagrams, продукция, И-ИЛИ дерево, посылка, следствие, модуль знаний.

Morozova OV Model of knowledge representation in the intellectual superstructure over problem-oriented CAD systems. The specificity of knowledge representation in the theory of artificial intelligence is considered, methods of knowledge representation in modern instrumental complexes, and knowledge representation methods used in problem-oriented CAD systems are described. Consider an approach to knowledge representation in a problem-oriented CAD system, where the subject matter expert is a knowledge engineer.

Keywords: САТІА, AutoCAD Plant 3D, Clips, Prolog, Protégé, AVEVA Diagrams, AND-OR tree, product, package, effect, module of knowledge.

СЕКЦИЯ 5. «КОМПЬЮТЕРНОЕ МОДЕЛИРОВАНИЕ, ГРАФИКА И ОБРАБОТКА ИЗОБРАЖЕНИЙ»

УДК 004.042:519.873

Разработка программного комплекса для анализа эффективности алгоритмов оптимизации при решении задачи подбора конструкции нежестких дорожных одежд

Тайлакова А.А.

Кузбасский государственный технический университет имени Т.Ф. Горбачева
taa.vt@kuzstu.ru

Тайлакова А.А. Разработка программного комплекса для анализа эффективности алгоритмов оптимизации при решении задачи подбора конструкции нежестких дорожных одежд. В статье рассмотрена постановка задачи оптимизации конструкции нежестких дорожных одежд по общей стоимости материалов с учетом нормативных требований. Описана архитектура и основные возможности программного комплекса. Рассмотрены возможности модуля оценки эффективности алгоритмов оптимизации и модуля расчета конструкции нежестких дорожных одежд.

Ключевые слова: программный комплекс, MVC-архитектура, оптимизация, нежесткие дорожные одежды

Введение

Согласно ОДН 218.046-01 [1], задача расчета конструкции нежесткой дорожной одежды сводится к подбору толщин слоев одежды в вариантах, намеченных при конструировании, или к выбору материалов с соответствующими деформационными и прочностными характеристиками при заданных толщинах слоев. Отказ конструкции по прочности физически может характеризоваться образованием продольной и поперечной неровностей поверхности дорожной одежды, связанных с прочностью конструкции (поперечные неровности, колея, усталостные трещины), с последующим развитием других видов деформаций и разрушений (частые трещины, сетка трещин, выбоины, просадки, проломы и т. д.) [2]. Поиск проектного решения, в наибольшей степени отвечающего требованиям ОДН 218.046-01, является достаточно сложной задачей и связан с большими затратами труда высококвалифицированных инженеров-проектировщиков. Дорожные одежды являются одним из наиболее дорогих элементов автомобильных дорог, их экономичность в значительной степени определяет строительную стоимость объекта в целом и эффективность капиталовложений. Применение методов математического моделирования при расчете конструкции нежестких дорожных одежд значительно сокращает сроки проектирования и позволяет исключить возможные ошибки, возникающие при традиционном способе [3].

Обзор существующих программных средств

Опыт применения систем автоматизированного проектирования показывает их высокую эффективность не только с точки зрения сокращения сроков проектирования, но и, что самое важное, с точки зрения повышения качества и обоснованности проектных решений. Рынок IT-технологий предлагает множество программных продуктов класса САПР, которые различаются между собой по комплексности, удобству интерфейса, соответствию сложившимся технологиям проектирования.

САПР АД Robur разрабатывается в научно-производственной фирме «Топоматик» (г. Санкт-Петербург) Реализованный на единой методологической основе, обеспечивает решение комплекса дорожных задач от обработки материалов изысканий до выноса проекта в натуру. Верх покрытия и конструкция дорожной одежды представлены в виде шаблонов. Шаблон – это текстовый файл, создаваемый при помощи любого текстового редактора. Шаблон позволяет описать произвольную конфигурацию поперечного профиля (например, бортовые камни, дренаж и т.д.). Более того, замкнутые контуры внутри шаблона могут быть использованы для подсчета объемов работ. Таким образом, шаблон является универсальным способом представления верхней части дороги.

САПР АД GIP является программным продуктом одного из ведущих дорожных проектных институтов – ОАО «Гипродорнии». GIP – это комплекс специализированных программ, при помощи которых можно производить основную часть работы по проектированию автомобильных дорог. Все программы комплекса используют общие типизированные структуры данных и единые алгоритмы.

САПР АД IndorCAD/Road. Система IndorCAD/Road позволяет проектировать автомобильные дороги всех категорий на стадии их строительства, реконструкции, модернизации и ремонта. В основу идеологии системы положены, в первую очередь, расчетные схемы для реконструкции дорог.

Рассмотренные программные средства, содержащие возможность расчета нежестких дорожных одежд, используют следующие подходы:

1. При проектировании автомобильных дорог расчет нежестких дорожных одежд при кратковременном действии нагрузки выполняется по критериям прочности: упругому прогибу, сопротивлению сдвигу в грунте и в слабо связанных слоях одежды, растяжению при изгибе слоев одежды из грунтов и каменных материалов.

2. Расчет дорожных одежд нежесткого типа предусматривает, как правило, расчетную схему, включающую в себя несколько конструктивных слоев с учетом подстилающего грунта.

3. Слои конструкции располагаются по убыванию (сверху вниз) модулей упругости материалов. Толщина отдельных слоев может быть задана по конструктивным соображениям, и эти слои в процессе поиска оптимального решения остаются неизменными.

4. Оптимальное решение ищется в диапазоне между минимальным и максимальным значением толщины каждого конструктивного слоя при определенном шаге перебора толщин.

5. Минимальная толщина задается не менее значений регламентированных ОДН 218.046-01 [1], а шаг перебора задается из условия производства и технологии выполнения работ по устройству каждого конструктивного слоя, точности устройства каждого слоя, имеющегося в наличии оборудования, возможности обеспечения требуемого уплотнения каждого слоя и требуемой точности при приемке работ [4].

Постановка задачи

Проектирование дорожной одежды представляет собой единый процесс конструирования и расчета дорожной конструкции на прочность, морозоустойчивость с технико-экономическим обоснованием вариантов с целью выбора наиболее экономичного в данных условиях. Из-за большого количества возникающих вариантов возможна оптимизация конструкции нежестких дорожных одежд по общей стоимости материалов с применением методов математического моделирования [4].

Основу математического моделирования составляет триада «модель – алгоритм – программа». На первом этапе вычислительного эксперимента строится модель исследуемого объекта, отражающая в математической форме важнейшие его свойства. Математическая модель исследуется традиционными аналитическими средствами прикладной математики для получения предварительных знаний об объекте. Второй этап связан с выбором (или разработкой) вычислительного алгоритма для компьютерной реализации модели. Изучение математических моделей проводится методами вычислительной математики, основу которых составляют численные методы решения задач. На третьем этапе создается программное обеспечение для компьютерной реализации сформированной модели и разработанного алгоритма [5].

Могут быть предложены следующие оптимизационные модели для решения задачи подбора конструкции нежестких дорожных одежд:

1. Подбор конструкции путем варьирования толщин конструктивных слоев.

2. Подбор конструкции путем варьирования материалов конструктивных слоев.

Наибольший экономический эффект дает проектирование оптимальных дорожных одежд в пространстве (т.е. с оптимизацией транспортной схемы строительства), когда конструкцию оптимизируют по длине трассы автомобильной дороги. Однородная на всей протяженности трассы конструкция нежестких дорожных одежд является наиболее простым с технологической точки зрения и, соответственно, наиболее дешевым вариантом. Из-за различия почвенно-грунтовых и гидрологических условий такой вариант не всегда возможен. В этом случае необходимо разбиение трассы на минимально возможное количество участков с различными конструкциями дорожных одежд.

Целью работы является разработка программного комплекса для анализа эффективности алгоритмов при решении задачи подбора оптимальной конструкции нежестких дорожных одежд с использованием предложенных оптимизационных моделей.

В соответствии с целью поставлены следующие задачи:

1. Разработка модуля расчета конструкции нежестких дорожных одежд в соответствии с требованиями ОДН 218.046-01.

2. Программная реализация предложенных оптимизационных моделей.

3. Реализация алгоритмов оптимизации для предложенных математических моделей.

4. Разработка модуля оценки эффективности алгоритмов подбора оптимальной конструкции нежестких дорожных одежд.

Технологии разработки программного комплекса

В качестве технологии разработки был выбран PHP-фреймворк «Yii». «Yii» – это высокоэффективный, основанный на компонентной структуре MVC-фреймворк для быстрой разработки крупных веб-приложений. Он позволяет максимально применить концепцию повторного использования кода и может существенно ускорить процесс разработки [6].

Шаблон MVC описывает простой способ построения структуры приложения, целью которого является отделение бизнес-логики от пользовательского интерфейса. В архитектуре MVC модель предоставляет данные и правила бизнес-логики, представление отвечает за пользовательский интерфейс, а контроллер обеспечивает взаимодействие между моделью и представлением [4].

В процессе разработки программного комплекса спроектирована и создана база данных [7], в соответствии с которой в программный комплекс вошли следующие модели данных: «Section» – проектируемый участок, «Pokritie», «Osnovanie», «Dopolnit» – конструктивный слой, «Construction» – конструкция нежесткой дорожной одежды.

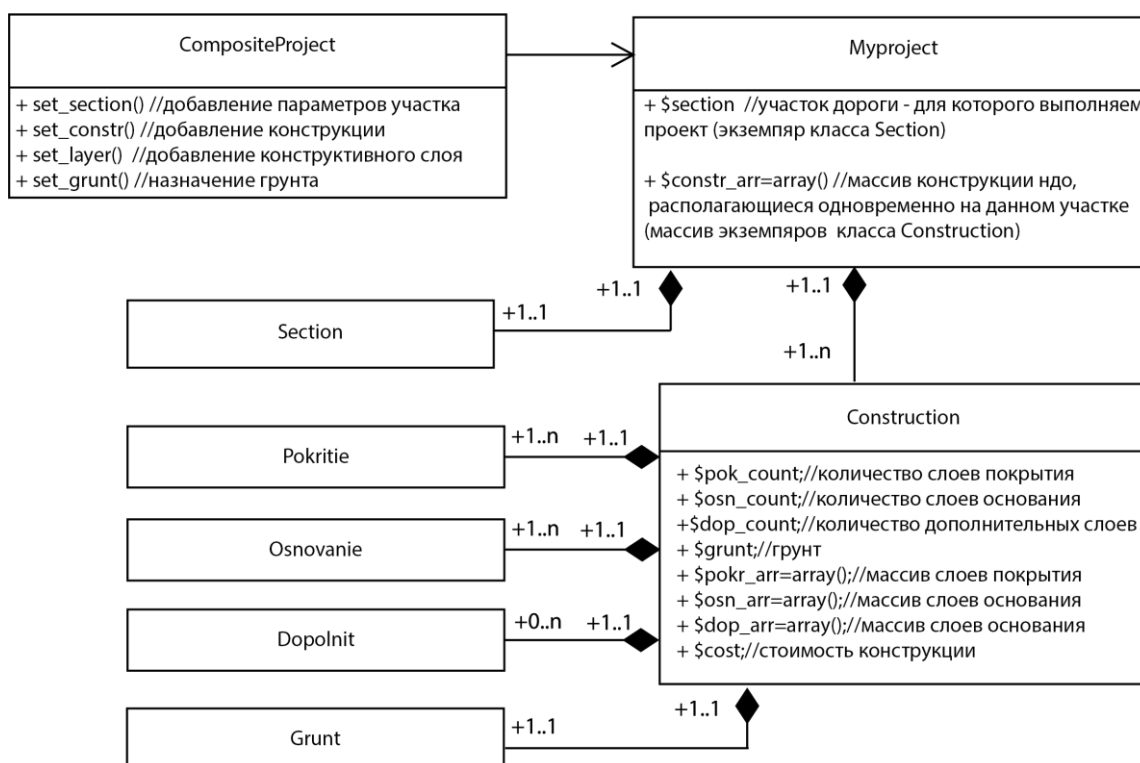


Рисунок 1 – Компоновщик проектируемого участка

Проект участка дороги может быть представлен в виде древовидной структуры, содержащей поле section (экземпляр класса «Section»), не имеющий ответвлений и контейнер const_arr (массив объектов «Construction»), который может содержать некоторое количество конструкций, каждая в свою очередь является контейнером для набора слоев.

Объект «CompositeProject» является объектом-компоновщиком для древовидной структуры проекта – содержит методы для добавления конструкций, конструктивных слоев, назначение грунта и параметров проекта (см. рис.1).

«CompositeProject» – базовый класс для классов компоновщиков проекта. В зависимости от способа создания проекта: на основе данных пользовательских форм; на основе сохраненных ранее данных; путем преобразования данных, полученных в формате JSON – можете быть вызван соответствующий класс-компоновщик, расширяющий класс «CompositeProject» (см. рис.2).

Подбор оптимальной конструкции нежесткой дорожной одежды выполняется на основе различных математических моделей. Соответственно существуют различные вариации каждого из тестируемых алгоритмов. Например, различия в работе генетического алгоритма для различных оптимизационных моделей заключается в способе кодирования хромосом и вычисления целевой функции. «Controller» создает объект для работы с конкретной оптимизационной моделью и передает его в конструктор «Algorithm», «Controller» имеет возможность заменить математическую модель в любой момент, т.к. объект «Algorithm» работает независимо от выбранной математической модели (см. рис.3).

Все тестируемые алгоритмы описаны как методы класса «Algorithm». «Controller» использует объект

класса «Algoritm» как фасад вместо прямой работы с объектами системы.

Программный комплекс позволяет работать с различными оптимизационными моделями расчета конструкции нежестких дорожных одежд:

1. Оптимизационная модель для подбора конструкции путем варьирования толщин конструктивных слоев.
2. Оптимизационная модель для подбора конструкции путем варьирования материалов конструктивных слоев.
3. Оптимизационная модель для подбора конструкции путем варьирования толщин и материалов конструктивных слоев.
4. Оптимизационная модель для подбора развернутой в пространстве конструкции путем варьирования толщин конструктивных слоев.
5. Оптимизационная модель для подбора развернутой в пространстве конструкции путем варьирования материалов конструктивных слоев.
6. Оптимизационная модель для подбора развернутой в пространстве конструкции путем варьирования толщин и материалов конструктивных слоев.

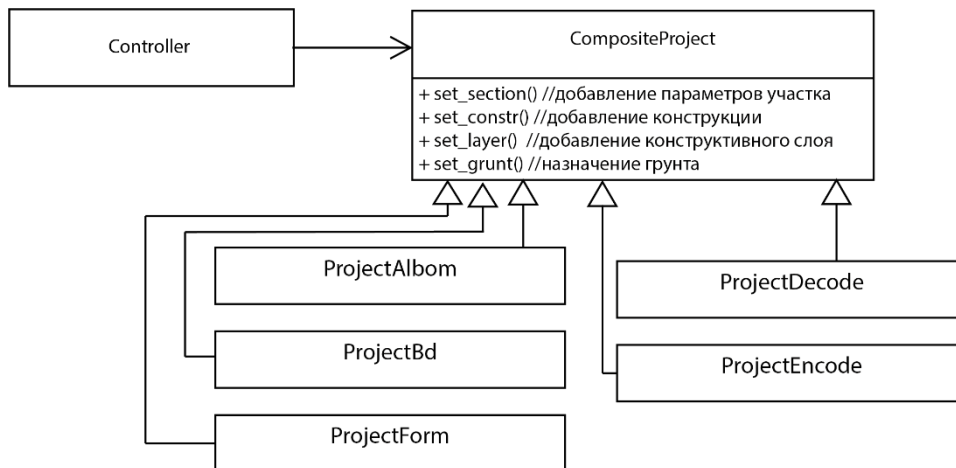


Рисунок 2 – Фабрика компоновщиков проектируемого участка

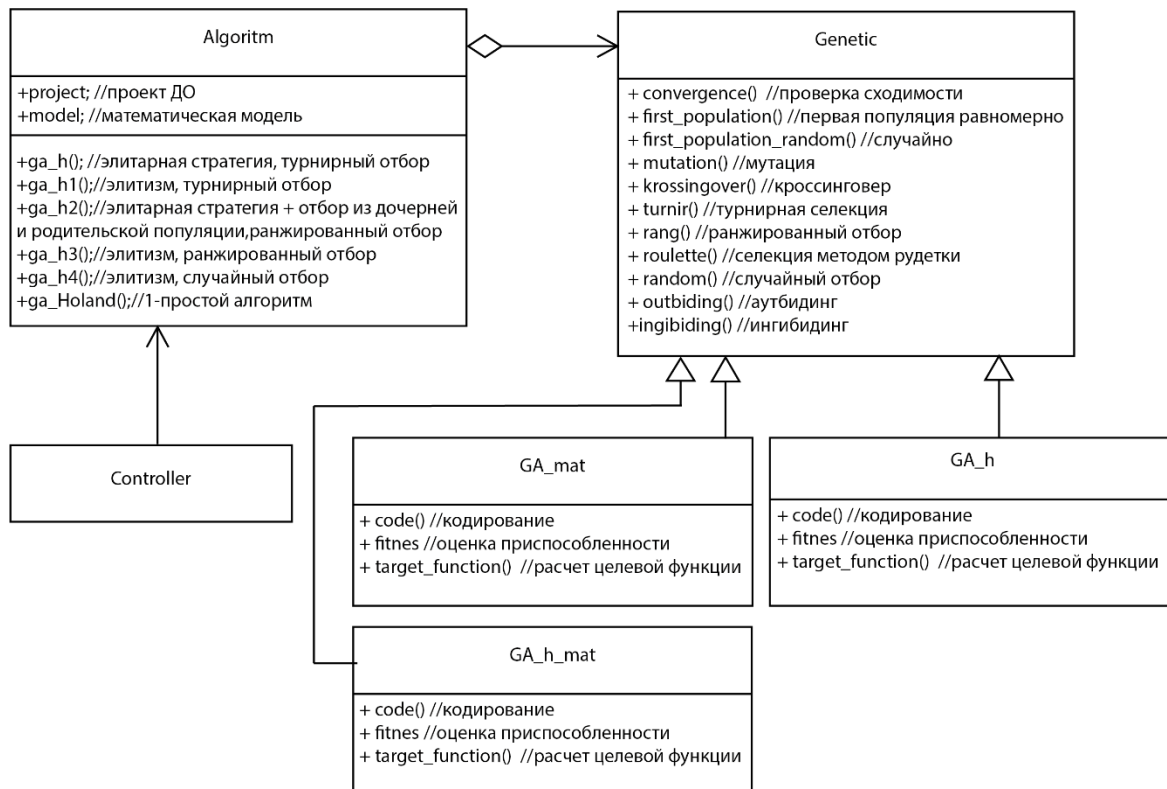


Рисунок 3 – Стратегия работы тестируемого алгоритма

Применение предложенной архитектуры позволяет легко выполнять масштабирование программного комплекса за счет добавления новых алгоритмов, что делает программный комплекс более простым в реализации и сопровождении.

Возможности программного комплекса

Программный комплекс позволяет проводить тестирование алгоритмов подбора оптимальной конструкции нежестких дорожных одежд на основе предложенных оптимизационных моделей в рамках проводимого вычислительного эксперимента.

В связи с нелинейными ограничениями математических моделей использование аналитических методов становится невозможным, таким образом целесообразно применение переборных методов поиска оптимального решения. Любая задача оптимизации может быть сведена к дискретной путем дискретизации интервалов значений переменных и решена с помощью полного перебора, который гарантирует нахождение глобального оптимума, однако возникает проблема неэффективности полного перебора, связанная с его трудоемкостью. Сравнение эффективности переборных методов, различных модификаций генетических алгоритмов, гибридных алгоритмов, параллельных вычислений производится с точки зрения быстродействия и показателя устойчивости алгоритмов. Показателем устойчивости алгоритма является отношение количества удачных запусков к общему числу запуска алгоритма, кроме того, для генетического алгоритма оценивается номер поколения, в котором была остановлена работа алгоритма.

	В	C	D	E	F	G	H	I	J
			Канонический генетический алгоритм	Элитарная стратегия	Равномерный отбор	Турнирный отбор	Отбор с частичной завесой популяции	Гибридный алгоритм	
Надежность			0,015	0,8	0,82	0,4	0,87	0,7	
Среднее время, сек.			8,05	15,37	13,58	2,11	16,96	3,9	
Среднее количество оценки вариантов конструкции			2784	3578,82	4000	691,2	4000	997,6	
Среднее количество обращений к ЦФ			955,89	1221,83	3535,62	86,45	3535,8	47,9	
Среднее поколение получения результата Г.А.			69,6	89,635	100	17,28	100	12,47	

Рисунок 4 – Отчет «Устойчивость алгоритма»

Программный комплекс позволяет выбрать математическую модель, модификацию тестируемого алгоритма и задать число запусков алгоритма. В результате будет сформирован отчет в формате MS Excel (см. рис. 4), а также будет представлен результат работы алгоритма в виде графика.

The screenshot shows a software interface for calculating rigid pavement structure. It includes several sections:

- Parameters of the site:** Includes fields for site name, width, traffic volume, and material properties like modulus of elasticity and Poisson's ratio.
- Construction of rigid pavement structure:** Lists different layers (e.g., surface, base, subgrade) with their respective thicknesses and material types.
- Calculation results:** Displays various calculated values such as fatigue damage, rutting, and stability, along with their respective units and graphical indicators.

Рисунок 5 – Модуль расчета конструкции нежестких дорожных одежд

Программный комплекс включает в себя модуль расчета конструкции нежестких дорожных одежд в соответствии с требованиями ОДН 218.046-01 (см.рис.5). В процессе расчета конструкции на прочность выполняются вычисления по критериям упругого прогиба, сдвигоустойчивости малосвязных конструктивных слоев и подстилающего грунта и расчет на сопротивление монолитных слоев усталостному растяжению при изгибе. Результатом расчета являются значения расчетных коэффициентов по соответствующим критериям. На основании их соотношения с нормативными значениями может быть сделан вывод о соответствии или несоответствии конструкции требованиям прочности. Все рассчитанные варианты могут быть экспортированы на рабочий лист MS Excel.

Выводы

В результате выполнения работы был создан комплекс программ для анализа эффективности алгоритмов при решении задачи подбора оптимальной конструкции нежестких дорожных одежд. Программный комплекс включает в себя модуль для расчета конструкции нежестких дорожных одежд в соответствии с требованиями ОДН 218.046-01, модуль для оценки эффективности алгоритмов подбора оптимальной конструкции нежестких дорожных одежд. Разработанный программный комплекс позволяет проводить вычислительный эксперимент с использованием предложенных оптимизационных моделей. Проведение натуральных экспериментов с целью выявления закономерностей поведения инженерных систем в зависимости от изменения параметров самой системы или внешней среды не всегда представляется возможным. Методы математического моделирования являются мощным инструментом исследования, позволяющим проводить вычислительные эксперименты любой сложности, осуществлять поиск оптимальных параметров системы, выявлять закономерности поведения исследуемого объекта.

Литература

- 1 О отраслевые дорожные нормы. ОДН 218.046-01. Проектирование нежестких дорожных одежд / Минтранс России. – М.: Информавтодор, 2001. – 145 с.
- 2 Тайлакова, А. А. Оптимизационные модели расчета конструкции нежестких дорожных одежд / А. А. Тайлакова, А. Г. Пимонов // Вестник Кузбасского государственного технического университета. – 2015. – № 3. – С. 149 – 155.
- 3 Пуркин, В. И. Основы автоматизированного проектирования автомобильных дорог учебное пособие / МАДИ (ТУ). – М.: 2000. – 141 с.
- 4 Тайлакова, А. А. Web-сервис для поиска оптимальной конструкции нежестких дорожных одежд / А. А. Тайлакова, А. Г. Пимонов // Вестник Кузбасского государственного технического университета. – 2015. – № 6. – С. 160 – 164.
- 5 Пимонов, А. Г. Имитационное моделирование экономических систем / А. Г. Пимонов, С. А. Веревкин. – Кемерово: Кузбас. гос. техн. ун-т им. Т. Ф. Горбачева, 2013. – 138 с.
- 6 Yiiframework. [Электронный ресурс]. – Режим доступа: <http://www.yiiframework.com/doc/guide/1.1/ru/quickstart.what-is-yii> (дата обращения: 29.10.2018).
- 7 Тайлакова, А. А. База данных информационно-вычислительной системы для расчета и оценки стоимости конструкции нежесткой дорожной одежды для автомобильных дорог общего пользования. Свидетельство о государственной регистрации базы данных № 2014621634; заяв. 02.06.2014.

Тайлакова А.А. Разработка программного комплекса для анализа эффективности алгоритмов оптимизации при решении задачи подбора конструкции нежестких дорожных одежд. В статье рассмотрена постановка задачи оптимизации конструкции нежестких дорожных одежд по общей стоимости материалов с учетом нормативных требований. Описана архитектура и основные возможности программного комплекса. Рассмотрены возможности модуля оценки эффективности алгоритмов оптимизации и модуля расчета конструкции нежестких дорожных одежд.

Ключевые слова: программный комплекс, MVC-архитектура, оптимизация, нежесткие дорожные одежды

Taylakova A. A. Development of software for analysis of efficiency of optimization algorithms when solving the problem of selection of design of nonrigid road clothes. The article deals with the formulation of the problem of optimizing the design of non-rigid pavement for the total cost of materials, taking into account regulatory requirements. The architecture and main features of the software complex are described. Possibilities of the module of evaluation of the effectiveness of algorithms of optimization and calculation module design of nonrigid road clothes.

Keywords: software package, MVC architecture, optimization, non-rigid road clothes

Исследование возможности оценки эффективности работы промышленной гидротранспортной системы (ПГТС) системами искусственного интеллекта

Капустин Д. А., Сентяй Р.Н., Швыров В.В.
Луганский национальный университет имени Тараса Шевченко
kap-kapchik@mail.ru

Капустин Д. А., Сентяй Р.Н., Швыров В.В. Исследование возможности оценки эффективности работы промышленной гидротранспортной системы (ПГТС) системами искусственного интеллекта. Рассмотрены вопросы, связанные со снижением энергетических затрат промышленных гидротранспортных систем на примере системы эвакуации золошлаков. Определены основные пути снижения расходов энергии. Представлены расчетные зависимости определения основных параметров гидротранспорта концентрированных гидросмесей мелкодисперсных материалов в условиях образования пространственной структуры. Предложено и обоснован критерий оценки энергоэффективности работы трубопроводных систем, позволяющий выбрать рациональные параметры работы системы при изменении концентрации твердого материала. Определена возможность оценки основных показателей ПГТС системами искусственного интеллекта.

Ключевые слова: гидротранспорт, гидросмесь, энергоэффективность, методика, критерий оценки, искусственный интеллект.

Постановка проблемы

Напорный гидротранспорт промышленных предприятий широко используется для перемещения гидросмеси, разработанных полезных ископаемых, грунтов вскрыши карьеров, отходов обогащения переработки рудных и нерудных материалов; золы и шлака тепловых электростанций. В настоящее время энергоёмкость гидротранспортирования сравнительно высока, поэтому задача снижения энергетических затрат весьма актуальна [2, 5].

Анализ последних исследований и публикаций

Существуют различные критерии оценки работы ГТС, в частности, для определения энергетических затрат используют значения сопротивления движению (потери давления) при перемещении гидросмесей по трубопроводу. Гидравлические сопротивления определяются скоростью и условиями движения гидросмесей (трение частиц между собой и о внутреннюю поверхность труб и т.д.) [5].

Критерием скоростного режима движения для гидросмесей является критическая скорость (скорость начала выпадения частиц на дно трубопровода). Транспортирования ниже этой скорости приводит к образованию ползущего слоя отложений, который вызывает заиливание трубопровода. Такой режим, как правило, не эффективен.

Снижение энергетических затрат возможно достигнуть за счет снижения скорости движения, уменьшением коэффициента трения или применением поверхностно-активных веществ.

При этом необходимо определить критерий оценки эффективности работы гидротранспортной системы при изменении параметров ее работы.

Цель статьи

Определение способов снижения энергетических затрат при транспортировании гидросмесей и критерия оценки эффективности их работы

Изложение основного материала

Одним из способов снижения скорости транспортируемой гидросмеси является уменьшение ее объема, что связано уменьшением доли несущей среды. С увеличением содержания твердого в гидросмеси происходит увеличение ее плотности и изменение реологических характеристик (для мелкодисперсных материалов наблюдается образование пространственных структур [2, 4, 5]). Такие изменения свойств среды приводят к увеличению сопротивления движению и росту затрачиваемой на гидротранспорт мощности, что на первый взгляд нецелесообразно. Однако следует учитывать, что с ростом концентрации транспортируемого

компонента (особенно при малом его первоначальном содержании (до 20%)) наблюдается резкое снижение объемов перемещаемой гидросмеси. В связи, с чем наблюдается снижение затрачиваемой энергии на перемещение единицы твердого компонента и возрастает энергоэффективность гидротранспорта.

В случае перемещения тонкодисперсных сред (золотая гидросмесь и т.д.) или сред с незначительным количеством крупных частиц (золотшлаковая гидросмесь) основные параметры гидротранспорта могут быть рассчитаны по следующим зависимостям:

- диаметр трубопровода

$$D = \sqrt{\frac{4Q_{см}}{\pi u}}, \quad (1)$$

где $Q_{см}$ - расход гидросмеси, м³/с; $u = 1,15u_{кр}$ - скорость транспортирования, м/с.

- критическая скорость [9]

$$u_{кр} = K_u \sqrt{\left(\frac{\left(\frac{\rho_{мс} - 1}{\rho_{см}} \right) \cdot \pi \cdot g}{4} - \frac{2\tau_0}{\rho_{см} \cdot d_u} \right) \left(\frac{D^2}{2D + d_u} \right)}, \quad (2)$$

где K_u - коэффициент определяемый экспериментально в зависимости от массовой концентрации для систем с вертикальными перегибами, для прямолинейного участка равен 1; $\rho_{мс}$ - плотность твердой фазы, кг/м³; $\rho_{см}$ - плотность гидросмеси, кг/м³; τ_0 - начальное напряжение сдвига, Па; D - диаметр трубопровода, м; d_u - размер частицы, м.

После чего значения скоростей $u_{кр}$ и u сравниваются, с учетом коэффициента $1,1 \div 1,15$, если они отличаются более чем на $\varepsilon \leq 0,001$, то скорость транспортирования u принимается равной критической скорости $u_{кр}$ и расчет производится заново (рисунок 1).

- удельные потери напора [8]

$$\Delta i = \lambda_3 \frac{u^2}{2gD}, \quad (3)$$

или по предложенной авторами зависимости

$$\Delta p_k = \frac{D^3}{a_p D^2 + b_p}, \quad (4)$$

где a_p и b_p - коэффициенты, зависящие от скорости транспортирования и концентрации твердой фазы; λ_3 - коэффициент гидравлического сопротивления, определяемый с учетом реологических свойств золотой гидросмеси:

- для ламинарного режима [10]

$$\lambda_3 = (a_\lambda + b_\lambda / \ln Re_3)^2, \quad (5)$$

где a_λ и b_λ - опытные коэффициенты, зависящие от массовой концентрации гидросмеси.

Значения коэффициентов a_λ и b_λ возможно определить по следующим зависимостям [10]

$$a_\lambda = 0,0123 \cdot C - 0,7124, \quad (6)$$

$$b_\lambda = -0,1429 \cdot C + 9,897, \quad (7)$$

где C - массовая концентрация, %.

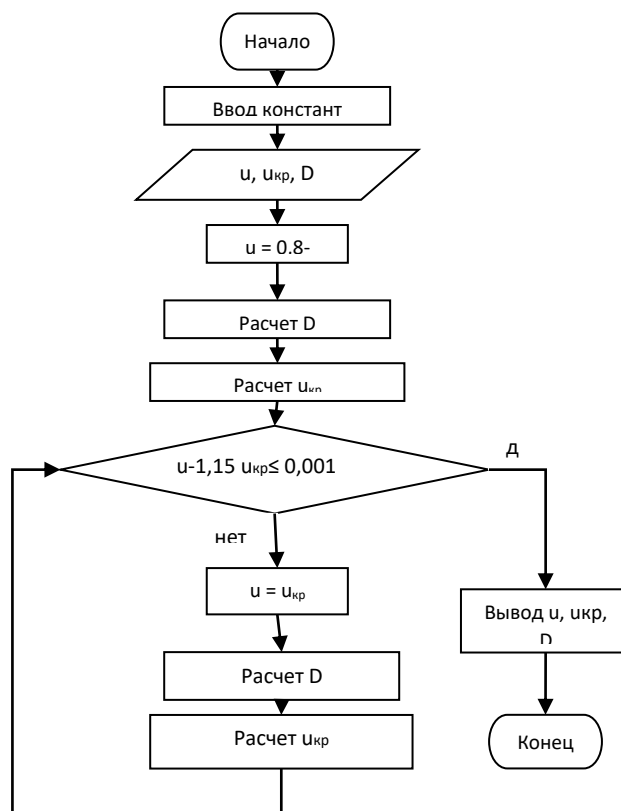


Рисунок 1 - Алгоритм расчета основных параметров гидротранспортной системы

- коэффициент местного сопротивления [10]

$$\zeta_9 = m \cdot \text{Re}_9^n, \quad (8)$$

где $m_{180} = e^{-7,57-4,44 \cdot C^3}$, $n_{180} = \frac{(C-48,5)^2}{320} + 1,89$ – коэффициенты, зависящие от содержания твердой фазы для поворота 180° ;

$m_{90} = 2,99 \cdot 10^{-5} - 4,96 \cdot 10^{-7} C$, $n_{90} = \frac{(C-51)^2}{245} + 1,74$ – для вертикального поворота на 90° .

- значение эффективного числа Рейнольдса [9]

$$\text{Re}_9 = \frac{uD\rho_{см}}{\mu_{см}(C) \left(1 + \frac{\tau_0(C)D}{u\mu_{см}(C)} \right)}. \quad (9)$$

Определяем режим течения гидросмеси.

Для гидросмесейдвигающихся в структурном режиме ламинарный режим сохраняется вплоть до $\text{Re}_9 \leq 1500000 - 1800000$.

- массовая концентрация [7]

$$C = \frac{m_{мс}}{m_{мс} + m_{жс}}, \quad (10)$$

где $m_{мс}$ - масса твердого материала, кг; $m_{жс}$ - масса воды, кг.

- начальное напряжение сдвига [3]

$$\tau_0 = \exp(a + b/C^2), \quad (11)$$

где $a = 5,087$, $b = -14727,383$ - коэффициенты.

- структурная вязкость [3]

$$\mu_{cm} = \exp(c + d / C), \quad (12)$$

где $c = 7,24$, $d = -565,433$ - коэффициенты.

- потребный напор [1, 6, 7]

$$H_{номп} = H_{cm} + iL + \frac{u_{cp}^2}{2g} \sum \xi_{\vartheta}, \quad (13)$$

где $H_{cm} = \Delta z + \frac{p_2}{\rho g}$ - статический напор, м; Δz - полная геометрическая высота подъема гидросмеси, м; $\frac{p_2}{\rho g}$ -

пьезометрическая высота, м; $\frac{u_{cp}^2}{2g} \sum \xi_{\vartheta}$ - потери напора на местных сопротивлениях при течении гидросмеси, м.

Далее строим характеристику гидротранспортной системы и совмещаем ее с характеристикой насоса. Подбор насоса осуществляется таким образом, чтобы рабочая точка находилась в пределах максимального КПД.

Характеристики насосных установок определяются на воде, однако при переходе на гидросмесь они изменяются при этом необходимо пересчитать параметров насосов на гидросмесь [7].

Для перерасчета насосного оборудования для работы на гидросмеси необходимо определить его новую напорно-расходную характеристику, которая для случая транспортирования дисперсных сред определяется зависимостью [9]

$$H_{cm} = b_H - a_H \cdot Q_{жс}, \quad (14)$$

где H_{cm} - напор, развиваемый на смеси соответственно, м. вод. ст; a_H, b_H - коэффициенты, зависящие от плотности (концентрации) транспортируемой среды и определяется экспериментально; $Q_{жс}$ - подача насоса на воде, м³/ч.

Для насосов с ниспадающей характеристикой коэффициенты a_H и b_H могут быть рассчитаны по следующим зависимостям

$$a_H = 0,139 + 3,87 \cdot 10^{-12} \cdot \rho_{cm}^3, \quad (15)$$

$$b_H = 748,5 - \frac{14624,879}{\sqrt{\rho_{cm}}}. \quad (16)$$

- подача насоса на гидросмеси может быть рассчитана [7]

$$Q_{cm} = k_Q \cdot Q_{жс}, \quad (17)$$

где k_Q - экспериментальный коэффициент.

- мощность, развиваемая насосом на гидросмеси

$$N_{cm} = H_{cm} \cdot Q_{cm}, \quad (18)$$

КПД насоса при пересчете на гидросмесь [7]

$$\eta_{cm} = k_{\eta} \cdot \eta_{жс}, \quad (19)$$

где k_{η} - экспериментальный коэффициент, $k_{\eta} = 1 - 0,2S$.

В случае, если не один из серийно выпускаемых насосов или их комбинация не подходит, необходимо регулирование характеристик насосного оборудования.

В качестве критерия оценки энергоэффективности работы гидротранспортной системы следует принимать значение удельной гидравлической мощности насосного оборудования на единицу твердой фазы. Данный критерий выражает долю мощности насосного оборудования, затрачиваемую непосредственно на гидротранспорт, и не зависит от типа насоса. Предложенный подход возможно применить для большинства ПГТС при изменении параметров их работы, связанных с варьированием концентрацией твердого компонента.

Для определения рациональных параметров гидротранспорта, необходимо при постоянном расходе твердого компонента изменить массовую концентрацию и повторить расчет. После чего сравнить величины

затрачиваемых на транспорт удельных гидравлических мощностей насосного оборудования (меньшее значение будет соответствовать более экономичному режиму).

Величину удельной гидравлической мощности на перемещение единицы твердого материала можно вычислить из соотношения (рисунок 2).

$$N_{см.уд} = \frac{N_{см}}{Q_{см} \cdot \rho_{см}} . \quad (20)$$

Анализ зависимости удельной гидравлической мощности на единицу транспортируемого материала позволил установить, что с ростом расхода гидросмеси (скорости транспортирования) наблюдается неравномерное распределение $N_{см.уд}$, резкое ее изменение происходит при малых величинах расхода с дальнейшей стабилизацией темпов роста. С увеличением концентрации твердого компонента наблюдается минимум $N_{см.уд}$ при массовой концентрации твердой фазы 0,55, что особенно проявляется при небольших расходах гидросмеси.

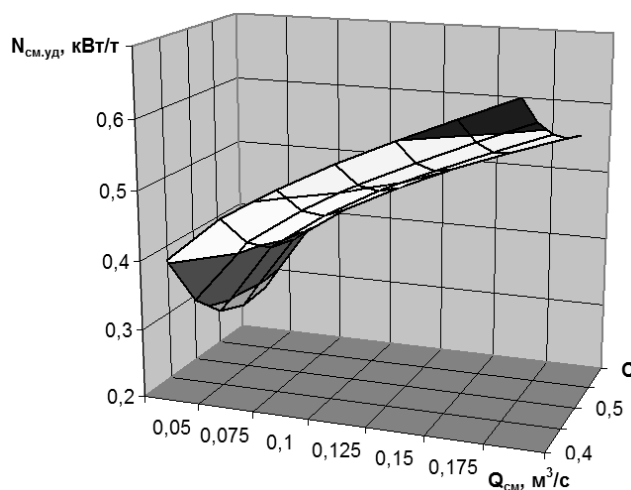


Рисунок 2 - Зависимость удельной гидравлической мощности на единицу твердого материала от расхода гидросмеси $Q_{см}$ (скорости транспортирования) и массового содержания твердой компоненты C

Произведена оценка возможности определения основных параметров работы ПГТС и критерия эффективности работы, а также прогнозирования их изменений средствами нейронных сетей и искусственного интеллекта.

Анализ основных приложений и библиотек для реализации таких систем показал, что рассмотренные пакеты не позволяют численно представить полученные результаты расчета и прогнозирования. А для представления результатов следует применять абстрактное отображение необходимых величин.

Выводы

1. Снижение скорости транспортирования гидросмесей возможно осуществить за счет уменьшения минимального значения критической скорости транспортирования;
2. Резкое снижение критической скорости происходит с ростом содержания транспортируемого материала (особенно мелких фракций) более 50-55%;
3. Критерием оценки энергоэффективности следует считать затраты гидравлической мощности на единицу транспортируемого твердого материала, так как он позволяет однозначно определить наиболее рациональные и целесообразные режимы работы гидротранспортной системы;
4. Выбор самого выгодного в энергетическом плане режима работы ПГТС производится на основании сравнительного расчета значения удельной гидравлической мощности при различных концентрациях твердого компонента.
5. Непосредственное применение систем искусственного интеллекта и нейронных сетей не позволяет получить точных значений искомых параметров, при этом возможно получение результатов при использовании условных символьных величин.

Литература

1. Джваршеишвили А.Г. Системы трубопроводного транспорта горно-обогатительных предприятий/ А.Г. Джваршеишвили – М.: Недра, 1981. – 384 с.
2. Капустін Д.О. Реологічні дослідження концентрованих золевих гідросумішей /Д.О. Капустін //Вісник Східноукраїнського національного університету ім. В.Даля. – Луганськ, 2012. - №3(174). – С. 73-79.
3. Оффенгенден Н. Е. Исследование влияния частиц микронной крупности на характеристики взвесенесущих потоков/ Н.Е. Оффенгенден, Ю. Ф. Власов // Гидравлическая добыча угля. – М.: НИИТЭИуголь, 1969. – № 3. – С. 29-31.
4. Покровская В.Н. Трубопроводный транспорт в горной промышленности/ В.Н. Покровская. - М.: Недра, 1985. – 192 с.
5. Світлий Ю.Г. Гідравлічний транспорт твердих матеріалів/ Ю.Г. Світлий, О.А. Круть. – Донецьк: Східний видавничий дім, 2010. – 268 с.
6. Світлий Ю.Г. Гідравлічний транспорт: монографія/ Ю.Г.Світлий, В.С.Білецький. – Донецьк: Східний видавничий дім, Донецьке відділення НТШ, «Редакція гірничої енциклопедії», 2009. – 436 с.
7. Смолдырев А.Е. Гидро- и пневмотранспорт/ А.Е. Смолдырев. – М.: Металлургия, 1975. – 384 с.
8. Чернецька-Білецька Н.Б. Методика розрахунку основних показників гідротранспорту твердих залишків ТЕС/ Н.Б. Чернецька-Білецька, Є.О. Варакута, Д.О. Капустін// Вісник Східноукраїнського національного університету ім. В.Даля. – Луганськ, 2012. - №5(176). – С. 137-142.
9. Теория и прикладные аспекты гидротранспортирования твердых материалов/ И.А. Асауленко, Ю.К. Витошкин, В.М. Карасик, С.И. Криль, В.Ф. Очеретько. – К.: Наук. Думка, 1981. – 364 с. – Ил. 110. Табл. 16. Список лит.: с. 353 - 362 – (210 назв.).
10. Chernetskaya Natalia. Experimental research of hydrotransporting concentrated residues at solid fuel burning/ Natalia Chernetskaya, Aleksandr Kuschenko, Denis Kapustin // ТЕКА. Commission of motorization and power industry in agriculture. – Lublin-Lugansk, 2012. – Vol. 12, №4. – P. 19-22.

Капустин Д. А., Сентяй Р.Н., Швыров В.В.. Исследование возможности оценки эффективности работы промышленной гидротранспортной системы (ПГТС) системами искусственного интеллекта. Рассмотрены вопросы, связанные со снижением энергетических затрат промышленных гидротранспортных систем на примере системы эвакуации золошлаков. Определены основные пути снижения расходов энергии. Представлены расчетные зависимости определения основных параметров гидротранспорта концентрированных гидросмесей мелкодисперсных материалов в условиях образования пространственной структуры. Предложен и обоснован критерий оценки энергоэффективности работы трубопроводных систем, позволяющий выбрать рациональные параметры работы системы при изменении концентрации твердого материала. Определена возможность оценки основных показателей ПГТС системами искусственного интеллекта.

Ключевые слова: гидротранспорт, гидросмесь, энергоэффективность, методика, критерий оценки, искусственный интеллект.

Kapustin D. A., Sentiay R.N., Shvyrov V.V. Investigation possibility of evaluating effectiveness work of the industrial hydrotransport system (IHTS) by artificial intelligence systems. The issues related to the reduction of energy costs of industrial hydrotransport systems on the example of the evacuation system for ash and slag are discussed. The main ways to reduce energy costs are identified. Calculated dependences of the determination of the main parameters of hydrotransport of concentrated slurries of finely dispersed materials in the conditions of spatial structure formation are presented. A criterion for estimating the energy efficiency of the operation of pipeline systems has been proposed and justified, which makes it possible to select rational parameters for the operation of the system when the concentration of solid material is changed.

Keywords: hydrotransport, slurry, energy efficiency, methodology, evaluation criterion.

Исследование функционирования специализированного вычислительного устройства на основе алгоритмов дельта-преобразований первого порядка с переменным квантом

Пирская Л.В.

Южный федеральный университет
lyubov.pirskaya@gmail.com

Пирская Л.В. Исследование функционирования специализированного вычислительного устройства на основе алгоритмов дельта-преобразований первого порядка с переменным квантом. В работе рассмотрены теоретические положения функционирования специализированного вычислительного устройства на базе итерационного решения СЛАУ с использованием дельта-преобразований первого порядка и переменного кванта. Представлен адаптированный под реализацию в специализированном вычислительном устройстве алгоритм итерационного решения системы линейных алгебраических уравнений на основе дельта-преобразования первого порядка и переменного квантом. Особенностью функционирования специализированного вычислительного устройства на основе данного алгоритма является реализация введения в начале каждого цикла l нового значения переменного кванта, которое отражается на текущем цикле при формировании значений невязки и неизвестной переменной путем их сдвига влево на 1 или 2 разряда. Формирование неизвестных переменных в устройстве осуществляется путем сложения или вычитания знаков квантов первых и вторых разностей переменных, принимающих на каждой итерации значения ± 1 . Указанная особенность нормирования переменных представляет возможности организации вычислительного процесса на основе целочисленного представления данных. На конечном этапе работы алгоритма в устройстве возможно приведение значений переменных к исходному вещественному виду с учетом веса минимального кванта преобразования. С ориентацией на ПЛИС получены сравнительные оценки по аппаратным затратам, быстродействию и комплексная сравнительная оценка эффективности разработанного алгоритма функционирования специализированного вычислительного устройства. В работе для разработанного алгоритма функционирования устройства показана возможность сокращения выполнения одной итерации и итерационного процесса в целом, затрат аппаратных ресурсов и в целом повышения эффективности по сравнению с функционированием специализированного вычислительного устройства на основе метода простой итерации.

Ключевые слова: специализированное вычислительное устройство, дельта-преобразования первого порядка, переменный квант, решение систем линейных алгебраических уравнений

Введение

Исследование вопросов качественного и количественного улучшения показателей вычислительных средств для решения сложных практических задач в режиме реального времени показывает, что при использовании самой современной универсальной вычислительной техники возникают различные трудности: необходимость организации эффективных параллельных вычислительных процессов с решением вопросов сокращения пересылаемых объемов информации, количества одновременно работающих устройств умножения, многоразрядных регистров для хранения информации, упрощения сложных систем обмена информацией, обеспечения высокой производительности вычислителей при обработке информации. Такие проблемы возникают, в частности, при организации параллельного решения в реальном масштабе времени практических задач, сводящихся к задачам вычислительной математики, проектировании бортовых вычислительных систем, специализированных управляющих устройств и т.п. Одним из путей решения данных проблем является создание специализированных вычислительных устройств и систем с учетом проектирования их проблемно-ориентированного назначения и с использованием специальных эффективных методов реализации вычислительного процесса [1].

Необходимость проблемно-ориентированных вычислений возникает, например, при решении в режиме реального времени задачи локальной навигации [2-4], в частности, определения координат летательного аппарата, которая сводится к задаче решения систем линейных алгебраических уравнений. Использование проблемно-ориентированных вычислений в данных условиях позволяют организовать вычислительный

процесс таким образом, чтобы обработка информации на уровне одной итерации выполнялась с достаточной точностью, высоким быстродействием и с предельно большим временным шагом. Тогда представляются возможности предъявлять наиболее низкие требования к производительности вычислительных средств с учетом возможности одновременной реализации других алгоритмов и программ.

Функционирование специализированного вычислительного устройства на основе известных итерационных алгоритмов решения систем линейных алгебраических уравнений (СЛАУ) [5-8] характеризуется значительными затратами аппаратных ресурсов, что связано с необходимостью реализации большого количества устройств многоразрядного умножения коэффициентов и переменных, а также пересылками многоразрядных кодов между уравнениями системы при параллельной реализации. Количество аппаратных и временных ресурсов возрастает, когда необходимо одновременно решать большое количество СЛАУ.

Разработанные алгоритмы для решения систем линейных алгебраических уравнений с постоянными и переменными свободными членами на базе дельта-преобразований первого порядка с постоянным [9-11] и переменным квантом [12-16] при реализации в специализированном вычислительном устройстве позволяют организовывать вычислительный процесс с исключением операций многоразрядного умножения и с получением результата за одну итерацию установившегося процесса. Однако использование постоянного кванта характеризуется значительно большим количеством итераций. Использование переменного кванта позволяет сократить количество итераций.

Сущность данных алгоритмов с переменным квантом заключается в представлении итерационного процесса в виде l итерационных циклов, в каждом из которых осуществляется параллельное для всех уравнений СЛАУ формирование переменных при постоянных по модулю квантах преобразования [14-16].

Проектирование специализированных вычислительных устройств с использованием ПЛИС [17-18] позволяют достигнуть высоких показателей по производительности при минимальных аппаратных затратах в сравнении с известными.

В статье рассматриваются особенности функционирования специализированного вычислительного устройства на основе итерационного решения СЛАУ с использованием дельта-преобразований первого порядка и переменного кванта. Ниже представлены адаптированный под реализацию в специализированном вычислительном устройстве алгоритм, структурная схема и исследование эффективности функционирования данного устройства.

Алгоритм решения СЛАУ, адаптированный под реализацию в специализированном вычислительном устройстве

Предположим, имеется СЛАУ, которая имеет матрицу постоянных коэффициентов и в общем случае переменные свободные члены, удовлетворяющую описанным в работах [5-8] условиям сходимости:

$$BY^*(t) = G(t). \quad (1.1)$$

Преобразуем систему

$$Y^*(t) = AY^*(t) + D(t),$$

и введем невязки $z(t)$ для использования итерационного метода:

$$z(t) = Y(t) - AY(t) - D(t). \quad (1.2)$$

В приведенных системах $B = [b_{rj}]$, $A = [b_{rj} / b_{rr}]$ – матрицы коэффициентов размерности $n \times n$; $G(t)$, $D(t)$ – вектор-столбцы свободных членов системы (в частном случае для системы с постоянными свободными членами $G(t) = G = [g_r]$, $D(t) = D = [g_r / b_{rr}]$); $Y^*(t)$ – вектор-столбец неизвестных системы; $z(t)$, $Y(t)$ – вектор-столбцы невязок и приближенных значений неизвестных; t – независимая переменная; $\det A \neq 0$.

Основываясь на полученных результатах в работах [14-18], запишем алгоритм параллельного решения СЛАУ (1.1) с постоянными свободными членами с использованием дельта-преобразований первого порядка и переменного кванта, ориентированный для построения специализированного вычислителя, представим в следующей разностной форме для i -го шага при начальных условиях $Y_{r01} = 0$, $z_{r01} = -D_r$, $|z_{01}|_{\max}$, $r = \overline{1, n}$, $c_p^* = 2^{-s}$, $s \in N$:

- вычисление значений невязок и неизвестных перед каждым итерационным циклом:

$$z_{r0l} = z_{rR_{\text{int}}^*(1,2)}^{(l-1)} \cdot R_{\text{int}}(1,2); \quad (1.3.1)$$

$$Y_{r0l} = Y_{rR_{\text{int}}^*(1,2)}^{(l-1)} \cdot R_{\text{int}}(1,2); \quad (1.3.2)$$

$$r = \overline{1, n}, l = 1, 2, \dots, P_{\text{int}}(1,2);$$

- определение знаков квантов первых разностей переменных на каждой итерации в циклах:

$$A_{rl} = -\text{sign}(z_{r(i-1)l}); \quad (1.3.3)$$

$$\Delta_{ril} \in \{+1, -1\}; r = \overline{1, n}, i = 1, 2, \dots, R_{\text{int},(1,2)}^*;$$

$$l = 1, 2, \dots, P_{\text{int},(1,2)};$$

- демодуляция:

$$Y_{ril} = Y_{r(i-1)l} + \Delta_{ril}; \quad (1.3.4)$$

- вычисление значений невязок на каждой итерации в циклах:

$$\nabla z_{ril} = \Phi_r(\Delta_{jil}, j = \overline{1, n}); \quad (1.3.5)$$

$$z_{ril} = z_{r(i-1)l} + \nabla z_{ril}; \quad (1.3.6)$$

- условия завершения итерационных процессов в цикле:

$$\left. \begin{array}{l} 1. \text{sign}(z_{rR_{\text{int},(1,2)}^* l}) = -\text{sign}(z_{r(R_{\text{int},(1,2)}^* - 1)l}) \\ \text{или} \\ z_{rR_{\text{int},(1,2)}^* l} = 0; \\ l = 1, 2, \dots, P_{\text{int},(1,2)}, r = \overline{1, n}. \end{array} \right\} \quad (1.3.7)$$

$$\left. \begin{array}{l} 2. \text{sign}(z_{rR_{\text{int},(1,2)}^* l} - \text{sign}(z_{rR_{\text{int},(1,2)}^* l})) = -\text{sign}(z_{rR_{\text{int},(1,2)}^* l}) \\ \text{или} \\ \text{sign}(z_{rR_{\text{int},(1,2)}^* l} - \text{sign}(z_{rR_{\text{int},(1,2)}^* l})) = 0; \\ l = 1, 2, \dots, P_{\text{int},(1,2)}, r = \overline{1, n}. \end{array} \right\} \quad (1.3.8)$$

В алгоритме (1.3) $P_{\text{int},(1,2)}$ – количество итерационных циклов, выполняемых при постоянном по модулю значении кванта, $R_{\text{int},(1,2)}$ – постоянное значение, отражающее изменение кванта преобразования и изменение всех переменных СЛАУ при переходе от цикла к циклу. Значения постоянных величин $P_{\text{int},(1,2)}$ и $R_{\text{int},(1,2)}$ задаются в соответствии с полученными в работах [14-15] соотношениями:

$$P_{\text{int},(1,2)} = \left\lceil \frac{\ln \frac{|z_{01}|_{\max}}{c_p^*}}{\ln R_{\text{int},(1,2)}} \right\rceil \quad (1.4)$$

где c_p^* – вес модуля минимального кванта преобразования на последнем цикле ($c_p^* > 0$), $R_{\text{int},1} = 2$ и $R_{\text{int},2} = 4$.

Количество итераций реального вычислительного процесса в цикле $R_{\text{int},(1,2)}^*$, в соответствии с введенными в работах [14-16] заключениями, может быть большим или меньшим относительно значения $R_{\text{int},(1,2)}$.

В алгоритме (1.3) соотношения (1.3.1) и (1.3.2) отражают введение в начале каждого цикла l перед выполнением соотношений (1.3.3) - (1.3.8) нового значения переменного кванта и иллюстрируют процедуру нормированного на текущем цикле формирования значений z_{r0l} и Y_{r0l} путем их сдвига на 1 разряд при $R_{\text{int},1} = 2$ или на 2 разряда при $R_{\text{int},2} = 4$.

На каждой итерации знаки квантов формируются в виде значений ± 1 и используются в дальнейшем по ходу выполнения алгоритма для вычисления неизвестной переменной Y_{ril} (1.3.4) путем сложения или вычитания единицы от данного текущего значения. Данная особенность нормирования переменных представляет возможности организации вычислительного процесса на основе целочисленного представления данных. По завершении работы алгоритма возможно приведение значений переменных к исходному вещественному виду с использованием веса минимального кванта преобразования c_p^* :

$$Y_{rR_{\text{int},(1,2)}^* P_{\text{int},(1,2)}}^* = Y_{rR_{\text{int},(1,2)} P_{\text{int},(1,2)}} \cdot c_p^*, \quad (1.5)$$

где сочетание $R_{\text{int},(1,2)}^* P_{\text{int},(1,2)}$ в индексе обозначает формирование самого последнего значения неизвестного, т. е. конечного результата.

Соотношения (1.3.7) или (1.3.8) используются в алгоритме (1.3) для формирования момента завершения итерационных процессов в циклах, требующие выполнения итераций в каждом цикле до тех пор, пока по всем уравнениям СЛАУ в цикле одновременно или распределено во времени по итерациям не выполнится хотя бы одно из них.

Архитектура функционирования специализированного вычислительного устройства с использованием дельта-преобразований первого порядка с переменным квантом

На рис. 1 представлена структурная схема функционирования специализированного вычислительного устройства для параллельного итерационного решения СЛАУ на основе алгоритма (1.3) с использованием дельта-преобразований первого порядка и переменного кванта [18].

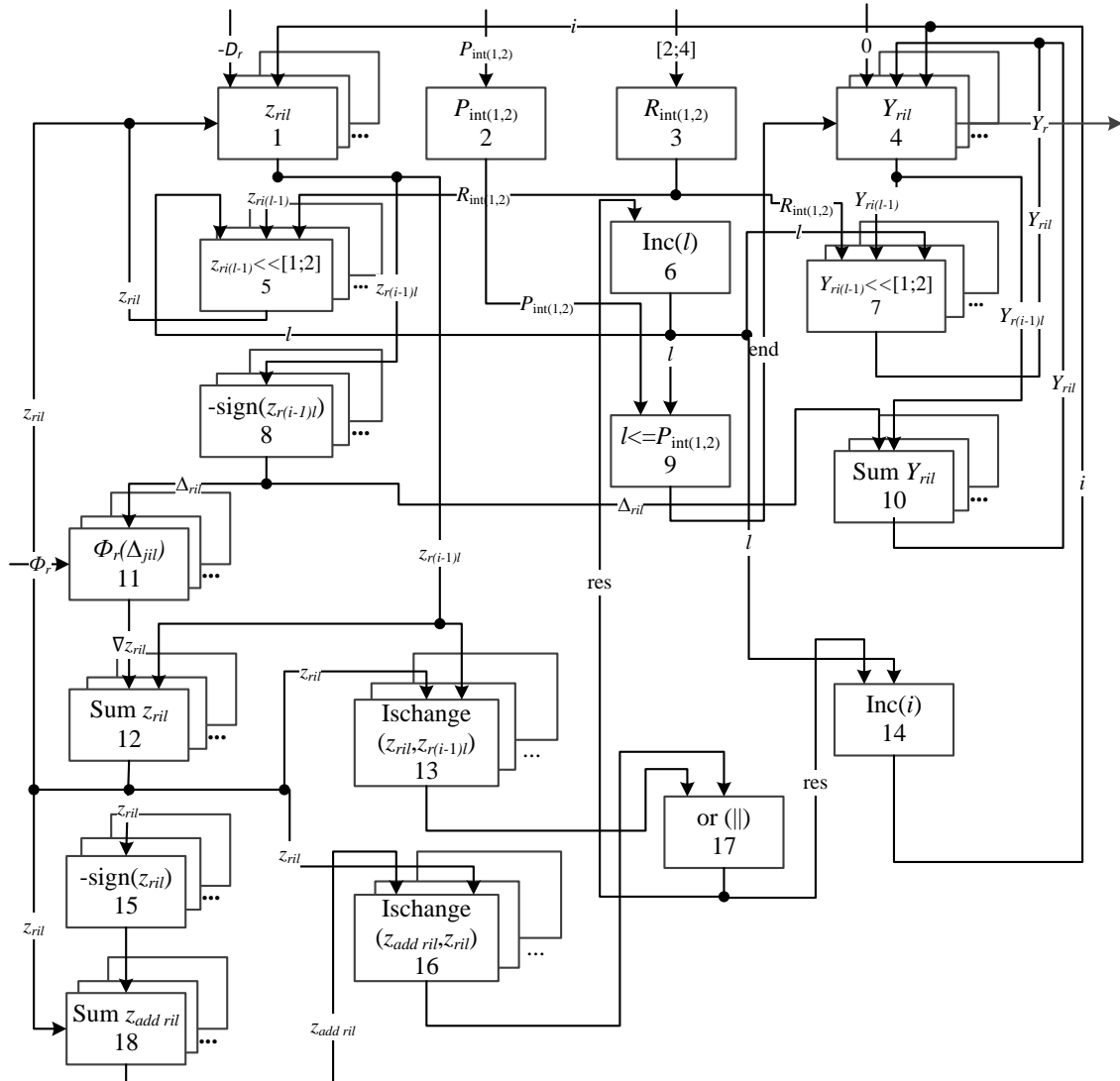


Рисунок 1 – Структурная схема функционирования специализированного вычислительного устройства

Блоки 1, 4 представляют собой r регистров ($r = \overline{1, n}$), в которых содержатся невязки z_{ril} и неизвестные Y_{ril} , $r = \overline{1, n}$ с начальными значениями $z_{r01} = -D_r$ и $Y_{r01} = 0$ соответственно.

Блоки 2, 3 являются регистрами, которые содержат значения количества итерационных циклов, выполняемых при постоянных по модулю значениях квантов, определяющиеся до начала работы алгоритма (1.3) и принимающие одно из двух значений $R_{int(1,2)}$ соответственно.

Блоки 5, 7 представляют собой r регистров сдвига ($r = \overline{1, n}$), реализующие сдвиг значений z_{r0l} (1.3.1) и Y_{r0l} (1.3.2) на 1 разряд при $R_{int,1} = 2$ или на 2 разряда при $R_{int,2} = 4$ в начале каждого цикла l . Полученные значения z_{r0l} и Y_{r0l} после сдвига поступают в регистры блоков 1 и 4, соответственно.

В блоке 6 выполняется работа циклов алгоритма (1.3), условия окончания работы алгоритма (1.3) в целом проверяется в блоке 9 и по выполнению данного условия выводится результат решения СЛАУ из блока 4.

Блоки 8, 15 определяют по соотношениям (1.3.3) знаки квантов первых разностей $z_{r(i-1)l}$ предыдущей итерации и z_{ril} текущей итерации для всех уравнений системы.

В блоке 10 по соотношениям (1.3.4) вычисляются значения неизвестных Y_{ril} в группе сумматоров ($r = \overline{1, n}$), на входы которых поступают знаки квантов первых разностей переменных Δ_{ril} и значения неизвестных Y_{ril} . Далее полученные результаты поступают в блок 4.

На схеме рис. 1 вычисления текущего значения невязки ∇z_{ril} реализуются на основе табличного метода в виде (1.3.5), где

$$\Phi_r(\Delta_{jil}, j = \overline{1, n}) = \Delta_{ril} - \sum_{\substack{j=1 \\ (j \neq r)}}^n a_{rj} \Delta_{ji}$$

и запоминающих устройств ($r = \overline{1, n}$) блока памяти 11 для хранения таблиц. Таблица организована в виде совокупности сумм произведений коэффициентов СЛАУ на кванты преобразования для каждого уравнения. Выборка значений заранее сформированных сумм произведений выполняется на основе совокупности текущих значений $\Delta_{jil}, j = \overline{1, n}$, поступающих из блока 8. Таким образом обеспечивается возможность исключения операции умножения за счет данной организации вычислений и получения результата за один такт.

При большой размерности матрицы А целесообразно разбиение суммы $\Phi_r(\Delta_{jil}, j = \overline{1, n})$ на m блоков:

$$\Phi_r(\Delta_{jil}, j = \overline{1, n}) = \sum_{g=1}^m \Phi_r^*(\Delta_{jil}, j = \overline{1, n})$$

и хранить табличные значения для каждого блока.

Блок 12 является r сумматорами ($r = \overline{1, n}$), где на вход каждого поступают соответствующие уравнениям системы знаки невязки z_{ri} и значения, полученные в блоке 11. Данный блок обеспечивает выполнение соотношения (1.3.6).

В блоках 13 и 16 выполняется проверка условий (1.3.7), (1.3.8) завершения итерационных процессов в l -м итерационном цикле для каждого уравнения СЛАУ. В блоке 17 проверяется сформированность данных моментов для всех уравнений системы либо по (1.3.7), либо по (1.3.8). При успешном выполнении одного из условий, в блоке 6 увеличивается счетчик на единицу, и работа алгоритма организуется на новом итерационном цикле l .

В блоке 14 выполняется подсчет количества итераций в рамках одного итерационного цикла l . Данный счетчик i обнуляется при переходе на каждый следующий цикл l .

Блоки 12, 18 являются группой сумматоров ($r = \overline{1, n}$), в которых вычисляются значения невязок текущих z_{ril} (1.3.6) и дополнительных $z_{add\ ril}$, где

$$z_{add\ ril} = z_{ril} - \text{sign}(z_{ril}).$$

Исследование функционирования специализированного вычислительного устройства на основе разработанного алгоритма

Комплексную эффективность функционирования специализированного вычислительного устройства на основе алгоритма (1.3) с использованием дельта-преобразований первого порядка и переменного кванта по сравнению с использованием метода простой итерации можно рассматривать в виде взаимосвязанной совокупности сравнительных оценок по быстродействию и аппаратным затратам. Учитывая, как показано в работах [14-16], что количество итераций при использовании дельта-преобразований первого и переменного кванта может быть большим или меньшим по сравнению с простой итерацией, при исследовании это количество принято одинаковым для исследуемых методов.

Можно рассматривать реализацию структурной схемы устройств, представленной на рис. 1 с использованием ПЛИС, в частности, типа FPGA. Ресурсные характеристики реализации основных арифметических операций при их аппаратном исполнении существенно неравнозначны. Особенно высокие аппаратные затраты, выражаемые в логических ячейках, требуются для множительных устройств. Операцию умножения будем рассматривать как выполнение умножения с использованием одноктактных аппаратных схем параллельной реализации, а также при расширении разрядной сетки сомножителей – экономичные параллельно-последовательные реализации путем выполнения умножения за несколько тактов по алгоритму «сдвиг с накоплением».

Наиболее просто реализуется структура простого матричного суммирования, которая формирует параллельный умножитель как массив одноразрядных соединенных локальными межсоединениями сумматоров [19]. Данная схема наиболее эффективна на операндах небольшой разрядности (4 и менее), где параллельный

умножитель 4×4 требует для своей реализации 12 сумматоров. При дальнейшем наращивании разрядности матрица одноразрядных сумматоров значительно разрастается, одновременно увеличивается критический путь распространения сигнала переноса, соответственно ограничивается быстродействие, и реализация умножителя становится малорациональной. Таким образом, в данной работе умножители большей разрядности рассматривались как комбинация умножителей 4×4 [19].

Формирование относительных оценок по аппаратным затратам базировалось на учете количества задействованных логических вентилей [20-21]. В соответствии с алгоритмом (1.3) и структурной схемой на рис. 1, решение уравнений выполняется параллельно. Также параллельно реализуются решения уравнений и по методу простой итерации.

Пусть Q – оценка, характеризующая аппаратные затраты алгоритма, $Q_{d.p.1}$ - оценка для алгоритма (1.3), основанного на дельта-преобразованиях первого порядка и переменного кванта, $Q_{p.r.}$ - оценка для метода простой итерации. Для сравнительной аппаратной оценки вводим отношение $\frac{Q_{p.r.}}{Q_{d.p.1}}$.

На основании полученных оценок был построен график зависимости относительных оценок $\frac{Q_{p.r.}}{Q_{d.p.1}}$ от порядка СЛАУ n ($n \geq 3$) при оперировании с 32-разрядными данными и представлен на рис. 2.

Анализ полученных оценок показал, что использование алгоритма, основанного на дельта-преобразованиях первого порядка и переменного кванта, при решении СЛАУ порядка $n=3$ имеет преимущество по аппаратным затратам в $\sim 2,7$ раз при оперировании с 32-разрядными данными по сравнению с методом простой итерации, а с увеличением порядка системы n эффективность увеличивается.

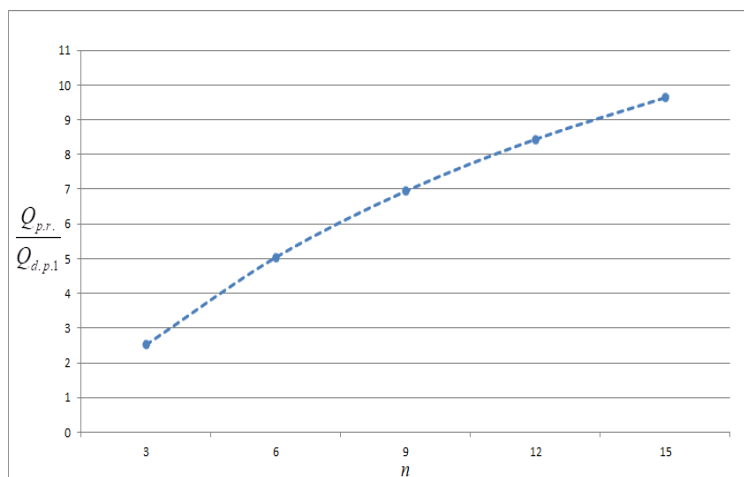


Рисунок 2 - График зависимости относительных оценок $\frac{Q_{p.r.}}{Q_{d.p.1}}$ от порядка СЛАУ n

Оценка по быстродействию производилась в тактах. Временные затраты были рассчитаны в рамках одного прохода по циклу алгоритма, реализующего алгоритмическую последователь действий для всех уравнений системы.

Пусть T – оценка, характеризующая временные затраты алгоритма (единица изменения такт), $T_{d.p.1}$ - оценка для алгоритма (1.3), основанного на дельта-преобразованиях первого порядка и переменного кванта, $T_{p.r.}$ - оценка для метода простой итерации.

На основании полученных оценок был построен график зависимости относительных оценок $\frac{T_{p.r.}}{T_{d.p.1}}$ от порядка СЛАУ n ($n \geq 3$) при оперировании с 32-разрядными данными и представлен на рис. 3.

Анализ полученных оценок показал, что использование алгоритма, основанного на дельта-преобразованиях первого порядка и переменного кванта, при решении СЛАУ порядка $n=3$ имеет преимущество по быстродействию в $\sim 2,5$ раза при оперировании с 32-разрядными данными по сравнению с использованием метода простой итерации, причем с увеличением порядка системы n имеет место увеличение данной эффективности.

Сравнительная комплексная оценка эффективности реализации алгоритмов на основе дельта-преобразований первого и второго порядков и переменного кванта была сформирована как произведение полученных выше относительных оценок по аппаратным затратам и быстродействию системы:

$$E = \frac{Q_{p.r.} \cdot T_{p.r.}}{Q_{d.p.1} \cdot T_{d.p.1}} \quad (1.6)$$

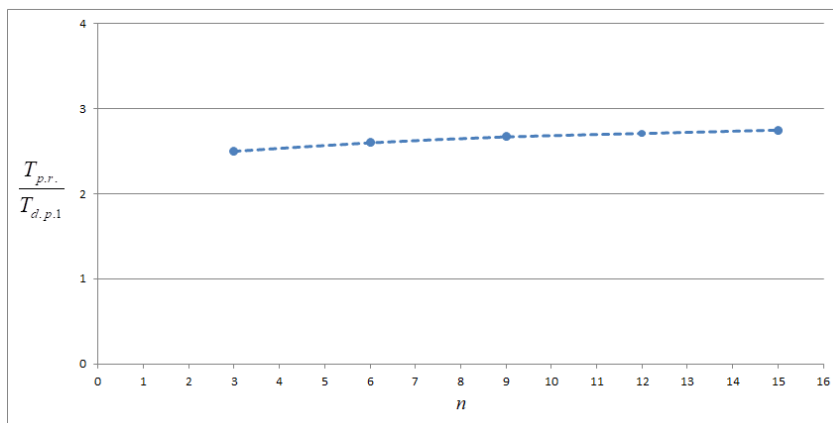


Рисунок 3 - График зависимости относительных оценок $\frac{T_{p.r.}}{T_{d.p.1}}$ от порядка СЛАУ n

В соответствии с (1.6) были получены сравнительные комплексные оценки эффективности. На основании данных результатов был построен график зависимости относительных комплексных оценок эффективности E от порядка СЛАУ n и представлен на рис. 4.

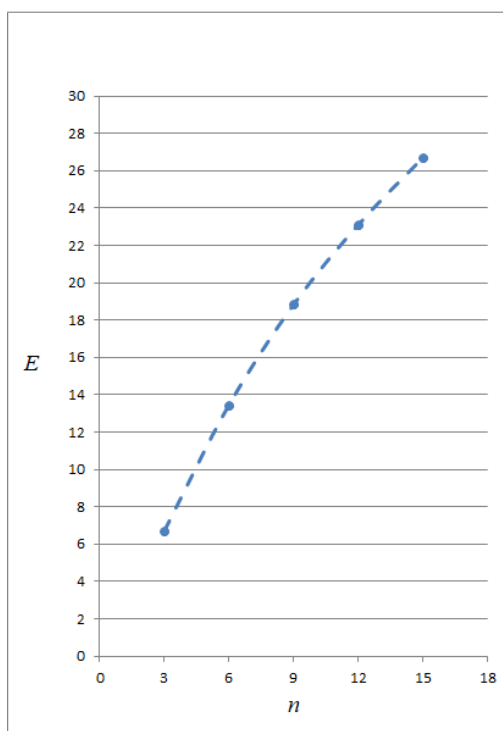


Рисунок 4 - График зависимости комплексных оценок эффективности E функционирования специализированного вычислительного устройства на основе решения СЛАУ от порядка СЛАУ n

В соответствии с данной оценкой E алгоритм, основанный на дельта-преобразованиях первого порядка и переменного кванта, при решении СЛАУ порядка $n=3$ имеет преимущество при оперировании с 32-разрядными данными в $\sim 6,7$ раз по сравнению с использованием метода простой итерации, причем данная оценка с увеличением порядка системы n резко увеличивается.

Заключение

В работе рассмотрены принципы функционирования специализированного вычислительного устройства на базе итерационного решения СЛАУ с использованием дельта-преобразований первого порядка и переменного кванта. Особенностью функционирования специализированного вычислительного устройства на основе данного алгоритма является реализация введения в начале каждого цикла 1 нового значения переменного кванта, которая отражается на текущем цикле при формировании значений невязки и неизвестной переменной путем их сдвига влево на 1 или 2 разряда. Формирование неизвестных переменных в устройстве осуществляется путем сложения или вычитания знаков квантов первых и вторых разностей переменных, принимающих на каждой итерации значения ± 1 . Указанная особенность нормирования переменных представляет возможности организации вычислительного процесса на основе целочисленного представления данных. На конечном этапе работы алгоритма в устройстве возможно приведение значений переменных к исходному вещественному виду с учетом веса минимального кванта преобразования.

С ориентацией на ПЛИС получены сравнительные оценки по аппаратным затратам и быстродействию разработанного алгоритма функционирования специализированного вычислительного устройства. Показана возможность сокращения по сравнению с функционированием специализированного вычислительного устройства на основе метода простой итерации выполнения одной итерации и итерационного процесса в целом в $\sim 2,5$ раз, затрат аппаратных ресурсов в $\sim 2,7$ раз и благодаря этому в целом повышения эффективности функционирования специализированного вычислительного устройства на основе итерационного решения СЛАУ ($n=3$) с использованием дельта-преобразований второго порядка и переменного кванта, причем эти оценки с увеличением порядка системы n увеличиваются.

Данные обстоятельства создают предпосылки, связанные, в частности, с расширением ресурсных возможностей ПЛИС для одновременной реализации в реальном времени сложных задач, в качестве отдельных компонент которых используются СЛАУ.

Результаты, полученные в данной работе, планируется использовать в бортовых системах управления и навигации современными летательными аппаратами. В качестве отдельной компоненты задачи навигации, решаемой в рамках создания бортовых специализированных вычислительных устройств, выступает задача определения координат летательного аппарата, представляемая в виде систем линейных алгебраических уравнений. По результатам предварительных исследований использование представленных в данной работе принципов функционирования специализированного вычислительного устройства в бортовых специализированных устройствах задачи локальной навигации позволит определять координаты летательного аппарата на каждом временном шаге установившегося процесса за одну итерацию и оперировать с практически значимыми временными шагами работы системы при достаточно большой удаленности начала установившегося процесса от начала координат.

Литература

1. Кравченко П.П., Пирская Л.В., Хусаинов Н.Ш. Дельта-преобразования и проблемно-ориентированные вычисления: монография. – Таганрог: Издательство Южного федерального университета, 2016. – 118 с.
2. Н.Ш. Хусаинов, П.П. Кравченко, В.Н. Лугай, С.А. Тарасов, В.В. Щербинин Системы радионавигации современных и перспективных летательных аппаратов. Ч. 1. Методы определения местоположения и автономный контроль целостности: Монография. – Таганрог: Изд-во ЮФУ, 2015. – 118 с.
3. Барабанов О.О., Барабанова Л.П. Математические задачи дальномерной навигации. – М.: ФИЗМАТЛИТ, 2007. – 272 с.
4. Скрыпник, О.Н. Радионавигационные системы воздушных судов: Учебник. – М.: ИНФРА-М, 2014. – 348 с.
5. Березин, И.С. Методы вычислений, том 2. - М.: Наука, 1966. – 632 с.
6. Самарский, А.А. Численные методы. – М.: Наука, 1989. – 432 с.
7. Фаддеев, Д. К. Фаддеева. В.Н. Вычислительные методы линейной алгебры, 4-е изд., стереотип. - СПб.: Лань, 2009. – 734 с.
8. Бахвалов, Н.С. Численные методы [Текст] / Н.С. Бахвалов, Н.П. Жидков, Г.М. Кобельков. - М.: БИНОМ. Лаборатория знаний, 2006. – 632 с.
9. Третьяков, С.И. Алгоритмы работы специализированных процессоров для решения систем уравнений // Кибернетика. – 1978. – №5. – С.34-36.
10. Кравченко, П.П. Оптимизированные дельта-преобразования второго порядка. Теория и применение. Монография. – М: Радиотехника, 2010. – 288 с.
11. Кравченко, П.П. Инкрементные методы решения систем линейных алгебраических уравнений // Многопроцессорные вычислительные структуры. – 1983. – Вып.5(XIV). – С.30-32.

12. Малиновский, Б.Н. Алгоритмы решения систем линейных алгебраических уравнений, ориентированные на структурную реализацию [Текст] / Б.Н. Малиновский, В.П. Боюн, Л.Г. Козлов // Упр. Системы и машины. – 1977. – Вып.5. – С.79-84.
13. Гомозов, О.В. Инкрементные алгоритмы решения систем линейных алгебраических уравнений и архитектура мультипроцессоров на программируемой логике / О.В. Гомозов, Ю.В. Ладыженский // Научные труды ДонНТУ. Серия «Информатика, кибернетика и вычислительная техника». – 2010. – Вып. 12 (165). – С.34-40.
14. Кравченко, П.П. Итерационный метод решения систем линейных алгебраических уравнений, исключая операцию многоразрядного умножения / П.П. Кравченко, Л.В. Пирская // Известия ЮФУ. Технические науки. – 2014. – № 7 (156). – С. 214-224.
15. Pirskaya, L.V. Iterative Algorithm for Solving of Linear Algebraic Equations Systems without Multi-bit Multiplication Operation // Engineering and Telecommunication (EnT), 2014 International Conference on. – IEEE, 2014. – P. 87 – 91.
16. Kravchenko P.P., Pirskaya L.V. The method of organizing the iterative process of the system of the linear algebraic equations solution excluding the multidigit multiplication operation// Biosciences Biotechnology Research Asia December. – 2014. – Vol. 11(3). – P.1831-1839.
17. Пирская Л.В. Об особенностях функционирования специализированного вычислительного устройства для итерационного решения СЛАУ на основе алгоритмов дельта-преобразований // Материалы XXIV научной конференции; Южный федеральный университет «Современные информационные технологии: тенденции и перспективы развития». – Ростов-на-Дону; Таганрог: Издательство Южного федерального университета (25-25 мая 2017). – 2017. – С.159-160.
18. Пирская Л.В. Об оптимизированном по быстродействию функционировании специализированного вычислительного устройства на основе дельта-преобразований первого и второго порядков // Сборник статей II Всероссийской научно-практической конференции молодых ученых, аспирантов, магистрантов и студентов «Информационные системы и технологии: фундаментальные и прикладные исследования». – Таганрог, 2017. – С. 231-232
19. Стешенко, В. Занятие 6. Реализация вычислительных устройств на ПЛИС// Компоненты и технологии. – 2000. – №8. – С. 88-91.
20. Максимов, Н.В. Архитектура ЭВМ и вычислительных систем: учебник. – 3-е изд. / Н.В. Максимов, Т.Л. Партыка, И.И. Попов. – М.: ФОРУМ, 2010. – 512 с.
21. Таненбаум, Э. Архитектура компьютера / Э. Таненбаум, Т. Остин. – 6-е изд. – СПб.: Питер, 2013. – 816 с.

Пирская Л.В. Исследование функционирования специализированного вычислительного устройства на основе алгоритмов дельта-преобразований первого порядка с переменным квантом. В работе рассмотрены теоретические положения функционирования специализированного вычислительного устройства на базе итерационного решения СЛАУ с использованием дельта-преобразований первого порядка и переменного кванта. Представлен адаптированный под реализацию в специализированном вычислительном устройстве алгоритм итерационного решения системы линейных алгебраических уравнений на основе дельта-преобразования первого порядка и переменного квантом. Особенностью функционирования специализированного вычислительного устройства на основе данного алгоритма является реализация введения в начале каждого цикла l нового значения переменного кванта, которое отражается на текущем цикле при формировании значений невязки и неизвестной переменной путем их сдвига влево на 1 или 2 разряда. Формирование неизвестных переменных в устройстве осуществляется путем сложения или вычитания знаков квантов первых и вторых разностей переменных, принимающих на каждой итерации значения ± 1 . Указанная особенность нормирования переменных представляет возможности организации вычислительного процесса на основе целочисленного представления данных. На конечном этапе работы алгоритма в устройстве возможно приведение значений переменных к исходному вещественному виду с учетом веса минимального кванта преобразования. С ориентацией на ПЛИС получены сравнительные оценки по аппаратным затратам, быстродействию и комплексная сравнительная оценка эффективности разработанного алгоритма функционирования специализированного вычислительного устройства. В работе для разработанного алгоритма функционирования устройства показана возможность сокращения выполнения одной итерации и итерационного процесса в целом, затрат аппаратных ресурсов и в целом повышения эффективности по сравнению с функционированием специализированного вычислительного устройства на основе метода простой итерации.

Ключевые слова: специализированное вычислительное устройство, дельта-преобразования первого порядка, переменный квант, решение систем линейных алгебраических уравнений

Pirskaya Liubov Research of the special-purpose calculating unit functioning based on the algorithm of the first order delta-transformations with variable quantum. This paper discusses the theoretical representations of special-purpose calculating unit functioning for the iteration solution of linear systems using the first order delta-transformations and variable quantum. It is considered the algorithm for iterative solution of linear systems based on the first order delta-transformations and variable quantum is adapted for implementation in a special-purpose calculating unit. A special feature of special-purpose calculating unit functioning based on this algorithm is the implementation of the introduction at the beginning of each cycle l of a new variable quantum value that is reflected in the current cycle when the residual values and the unknown variable are formed by shifting them to the left by 1 or 2 bits. Formation of unknown variables in the unit is carried out by adding or subtracting the signs of quantum of the first and second variables differences, taking at each iteration the values ± 1 . This feature of variable normalization represents the possibility of organizing a computational process on the basis of an integer data representation. At the final step of the algorithm operation in the unit, it is possible to bring the values of the variables to the original real form, taking into account the weight of the minimum transformation quantum. With the orientation to FPGA, comparative estimates are obtained for the hardware and time resources of the developed algorithm and comprehensive comparative estimate of the effectiveness for special-purpose calculating unit functioning. In this paper for the developed algorithm of the unit functioning, it is shown that it is possible to reduce the execution of one iteration and the iterative process as a whole, the amount of hardware resources and generally improve the efficiency in comparison with the special-purpose calculating unit functioning based on the simple iteration method..

Keywords: special-purpose calculating unit, first order delta-transformation, variable quantum, linear system solution

Особенности организации объемной визуализации в системах трехмерной компьютерной графики

Зори С.А., Аноприенко А.Я.

Донецкий национальный технический университет
ik.ivt.rec@gmail.com, anoprien@yandex.ru

Зори С.А., Аноприенко А.Я. Особенности организации объемной визуализации в системах трехмерной компьютерной графики. В статье представлен обзор современных принципов организации систем объемного синтеза и 3D- визуализации изображений. Предложены классификации и обобщенные структуры систем 3D- синтеза и визуализации изображений, определены основные особенности выполнения стадий 3D- графического конвейера в них. Рассмотренные системы на сегодняшний день в процессе работы используют специальные процедуры и алгоритмы синтеза и визуализации, не имеющие стандартных средств аппаратной поддержки, что актуализирует дальнейшие исследования и разработку высокопроизводительных вычислительных систем для их реализации.

Ключевые слова: компьютерная графика, объемная 3D- визуализация, 3D- дисплей, графический 3D- конвейер, графическая система 3D- отображения информации, структура.

Введение

Современный этап развития компьютерной 3D - графики и ее самых разнообразных и широко внедряемых повсеместно сфер применения, характеризуется постепенным переходом к, пока еще не ставшей традиционной, пространственной (объемной, 3D -) визуализации [1 - 7].

Актуальными областями практического применения и потенциального использования результатов пространственной визуализации являются в первую очередь интерактивные системы виртуальной реальности, интерактивные графические системы объемно-ориентированного геометрического и кинетического моделирования и визуализации объектов (медицина, архитектура, физика, гео-информация и пр.), 3D-CAD системы, зарождающиеся системы 3D- мультимедиа и 3D- Web визуализации, компьютерные игры и многие другие [1 - 9].

Недостаточный уровень практической реализации таких систем – их сложность, дороговизна, отсутствие стандартизированной аппаратной поддержки, слабая реалистичность создаваемых образов и большое время синтеза (за счет применения в них усложненных трудоемких вычислительных процедур модельной 3D- генерации и 3D- пространственной обработки) - порождает новые направления прикладных исследований по созданию эффективных архитектур вычислительных комплексов для реализации систем пространственной визуализации с использованием технологий объемного отображения [1 - 4].

Задача реалистической пространственной визуализации и объемного отображения

Задача реалистической пространственной визуализации и объемного отображения является общим случаем решения основной задачи трехмерной компьютерной графики, основной задачей которой является синтез изображений (трехмерных, в общем случае) некоторой виртуальной трехмерной сцены (совокупность объектов и некоторой окружающей их обстановки), заданной некоторой моделью, и их визуализация (пространственная, в общем случае) на некотором устройстве отображения (дисплее, мониторе) [10, 1, 3, 4].

Под синтезированным изображением при этом понимается визуальное представление информации, получаемое в результате вычислений в некоторой вычислительной системе (компьютере), для получателя этого изображения (наблюдателя) с учетом специфики устройства отображения и свойств механизмов зрительного восприятия человека.

В настоящее время сложилась следующая укрупненная классификация способов решения основных задач синтеза и визуализации изображений, реализуемых системами компьютерной графики и отображения информации – СКГ/СОИ (рис. 1).

При этом «классическая» компьютерная трёхмерная графика и системы визуализации окружающей обстановки имеют дело с виртуальным трёхмерным пространством отображаемой сцены, изображение которой рассчитывается в зависимости от положения наблюдателя и характеристик освещения сцены и визуализируется

на плоской, двухмерной поверхности экрана устройства отображения. В этом случае эффекты «трехмерности» изображения человек воспринимает за счет сгенерированных расчетами эффектов перспективы, перекрытия объектов, затемнений и затенений.



Рисунок 1 – Укрупненная классификация систем компьютерной графики

Синтез изображения, по сути, состоит в расчете монокулярной проекции сцены на плоскость экрана, а само синтезированное изображение, как и способ его отображения (визуализации), является плоским, двумерным [10, 1, 3, 4, 9].

В отличие от этого, трехмерная пространственная визуализация (рис. 2) предполагает такое компьютерное преобразование модели сцены (пространственный 3D- синтез, рендеринг), которое позволит создать полностью трехмерный визуальный объемный образ (3D- изображение) в специфическом устройстве пространственного отображения (устройстве 3D- визуализации, 3D- мониторе). 3D- изображение (пространственный визуальный образ) затем некоторым способом отображается в экранном пространстве 3D-дисплея, и человек воспринимает его (возможно, с использованием дополнительных средств сопряжения) в объемном виде [1 - 8].

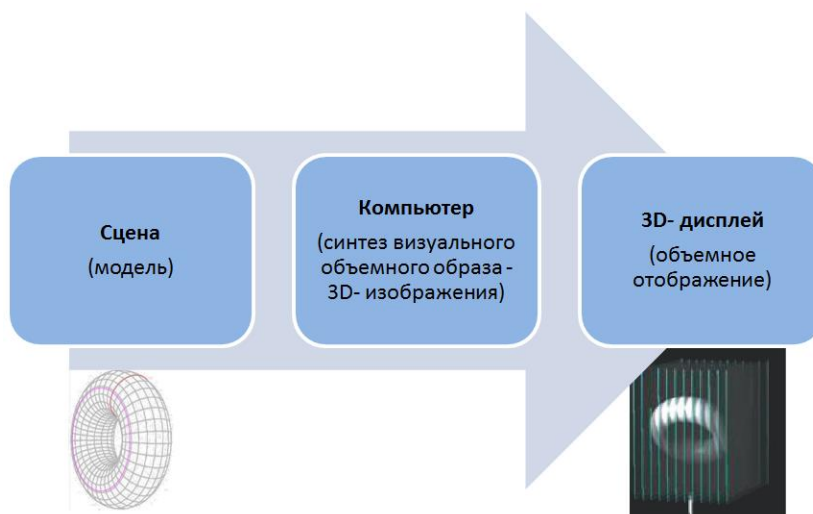


Рисунок 2 – Трехмерная пространственная визуализация

При решении задачи трехмерной пространственной визуализации и организации выполнения процедур пространственного 3D- синтеза принципиальными являются способы задания информации об объектах визуализации (сцене), а также способ пространственной объемной визуализации устройства отображения (3D-дисплея), на котором будет выполняться 3D-визуализация [1 - 4].

Особенности пространственного 3D- синтеза и визуализации изображений

Выполненный анализ литературных источников позволяет произвести классификацию основных способов выполнения пространственного 3D- синтеза и визуализации, приведенный на рис. 3 [1 - 3].

С учетом этого, формальный подход к решению задачи пространственного 3D- синтеза и визуализации, может быть описан следующим образом.

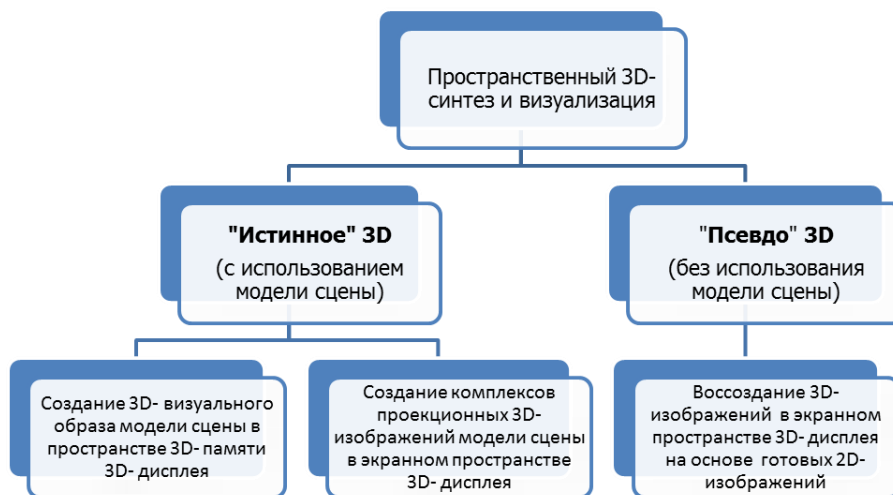


Рисунок 3 – Основные способы выполнения пространственного 3D- синтеза и визуализации

1. «Истинное» 3D. При наличии 3D- модели сцены возможен «полноценный», «истинный» пространственный синтез 3D- изображений и 3D- пространственная визуализация одним из двух следующих способов:

1.1 Синтез (создание) объемного дискретизированного 3D- визуального образа сцены в пространстве 3D-памяти устройства объемного отображения с его последующей 3D- визуализацией на нем (в экранном пространстве 3D- устройства отображения) тем или иным способом (с учетом специфики построения устройства, положения наблюдателя и пр.).

1.2 Синтез 3D- видов (изображений) без создания объемной визуальной модели сцены в памяти устройства пространственного отображения с использованием специфики зрительного аппарата человека – получение комплекса проекционных 2D- изображений сцены в экранном пространстве устройства 3D-отображения и их последующая 3D- визуализация.

2. «Псевдо 3D». 3D- синтез «на основе изображений» (image based 3D- synthesis) - 3D- модель сцены отсутствует, либо не используется для непосредственного получения 3D- изображения. Суть процесса состоит в получении (воссоздании, «реставрации») комплекса проекционных 2D- изображений сцены в экранном пространстве 3D- устройства отображения из полученных заранее (на предыдущем этапе синтеза 2D-изображений по 3D- модели (как в «классической» компьютерной 3D- графике), либо другим способом (фото, видео съемка)) единичных, плоских «видовых» экранных изображений методами их анализа, преобразования и морфинга (view morphing) с дальнейшей 3D- визуализацией на устройстве 3D- отображения.

При формировании трехмерного изображения выделяют этапы описания сцены (сценарная подсистема, сценарный процессор), геометрических преобразований (подсистема геометрии, геометрический процессор), рендеринга (подсистема рендеринга (синтеза), рендер- процессор) и визуализации (подсистема визуализации (отображения), процессор визуализации).

Архитектурная организация систем отображения 3D- объемной информации

Типичная 3D пространственная объемная система отображения информации (3D СОИ) по своей структуре аналогична «классической» графической системе СОИ [1, 3, 7, 9, 11, 12] (рис. 4).

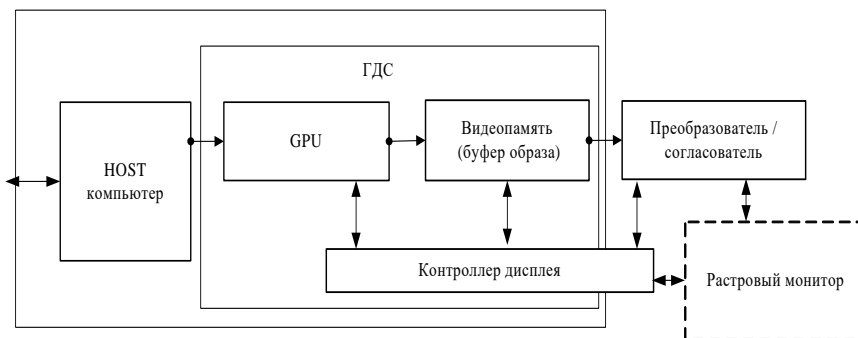


Рисунок 4 – Обобщенная структура «классической» графической СОИ

Традиционная архитектура графической системы включает центральный (главный) компьютер или процессор (HOST Computer) и графическую дисплейную систему (ГДС).

HOST компьютер выполняет функции обмена информацией между ГДС и внешним миром (базовым компьютером или вычислительной сетью), диспетчеризации потоков данных между компонентами рабочей станции и предварительной обработки данных.

Графическая дисплейная система ГДС осуществляет функции формирования (генерации) и модификации наборов данных в памяти изображения (видеопамяти) и управляет режимами вывода графической информации на растровый монитор. Графическая дисплейная система состоит из следующих функциональных компонент:

Видеопамять служит для хранения графических данных в растровой форме.

Графический процессор (Graphics processing unit - GPU) реализует основные функции по синтезу и формированию изображения в видеопамяти. В современных системах графические процессоры, как правило, являются мощными многоядерными параллельными вычислительными системами с аппаратной реализацией рассмотренной выше архитектуры 3D- графического конвейера.

Видеоконтроллер (контроллер дисплея) формирует управляющие сигналы для организации доступа к видеопамяти со стороны графического процессора (возможно, и со стороны HOST процессора), обеспечивает регенерацию буфера образа видеопамяти - части видеопамяти, отображаемой на экран монитора, а также даёт команды на формирование сигналов развёртки для монитора.

В зависимости от способа пространственного 3D- синтеза и визуализации (рис. 1.3) структура 3D СОИ при организации «реального 3D» путем создания 3D- визуального образа модели сцены в пространстве 3D- памяти 3D- дисплея, будет выглядеть, как на рис. 5.

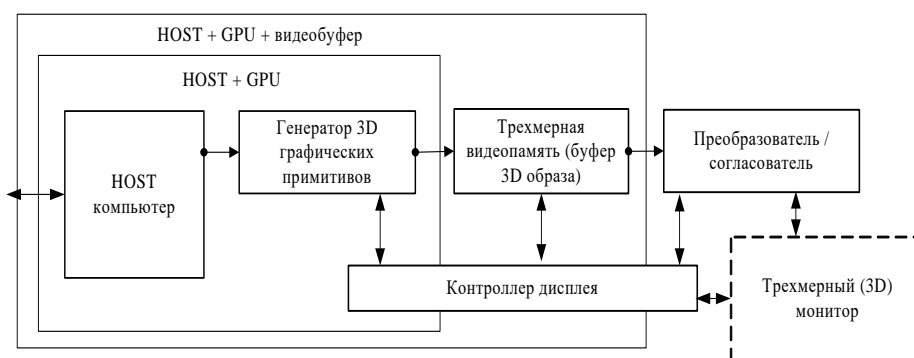


Рисунок 5 – Обобщенная структура объемной 3D СОИ методами создания 3D- визуального образа сцены в трехмерной видеопамяти 3D- дисплея

Особенностью реализации основных этапов синтеза 3D- объемного изображения этим способом по сравнению с классическим графическим 3D- конвейером будет принципиально другая организация вычислительного процесса, связанная с необходимостью выполнения процедур пространственной аппроксимации 3D- моделей объектов сцены 3D- примитивами, созданием и 3D- обработкой дискретизированного пространственного визуального образа модели сцены в пространстве 3D- памяти 3D- дисплея (рис. 6).



Рисунок 6 – Основные этапы синтеза 3D- объемного изображения методами создания 3D- визуального образа сцены в трехмерной видеопамяти 3D- дисплея

Структура 3D СОИ при организации «истинного» 3D способом создания комплекса проекционных 3D- изображений сцены в экранном пространстве 3D- дисплея (фактически комплекса стерео-пар изображений сцены) представлена на рис. 7.

Суть процесса генерации 3D- изображения при этом состоит в использовании особенности организации объемного восприятия зрительным аппаратом человека. Возможна реализация двух вариантов.

1. С использованием особенностей стереоскопического зрения человека, когда каждый глаз наблюдателя видит «свои», различные изображения окружающей обстановки, а мозг «сливает» эти изображения, формируя пространственный объемный образ, воспринимаемый человеком. Таким образом, процесс создания 3D-

объемного изображения представляет собой организацию синтеза комплексов из двух изображений (стереопары) с двукратным применением классических процедур графического 3D- конвейера для генерации двух (или более) независимых проекционных изображений. 3D- визуализация полученных изображений осуществляется путем их преобразования (слияния) специальными устройствами 3D- мониторов и вспомогательного оборудования (шлемов виртуальной реальности, 3D-очков и пр.).

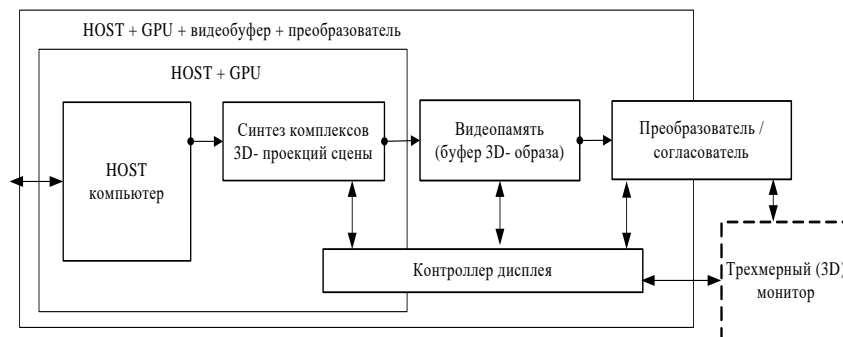


Рисунок 7 - Обобщенная структура 3D СОИ способом создания комплекса проекционных 2D- изображений сцены в экранном пространстве 3D- дисплея

2. С использованием особенностей инерционности зрительного аппарата человека. В этом случае система формирует множество пространственных «срезов» - комплексов проекционных изображений сцены, которые затем последовательно визуализируются с большой скоростью на специальных объемных экранах 3D-монитора, и человек получает ощущение восприятия целостного пространственного объемного образа.

Особенностями реализации основных этапов синтеза 3D- объемного изображения способом создания комплекса проекционных 2D- изображений сцены в экранном пространстве 3D- дисплея по сравнению с рассмотренным выше классическим графическим 3D- конвейером будет определяться выбором метода генерации комплекса проекционных изображений. При использовании традиционных методов растеризации и экранизации процесс реализуется напрямую процедурами графического 3D- конвейера, если используются методы трассировки лучей или другие специализированные методики реалистичного синтеза, то реализация конвейера будет иметь специфические особенности (рис. 8).



Рисунок 8 - Основные этапы синтеза 3D- изображения методом создания комплекса проекционных изображений в экранном пространстве 3D- дисплея

Структура 3D СОИ при организации «псевдо 3D» путем воссоздания комплекса 2D- изображений на экране 3D- дисплея из готовых 2D- изображений фактически не предполагает никаких процедур 3D- расчета сцены, а базируется на принципах синтеза на основе готовых изображений (image based synthesis). При этом структуру 3D СОИ можно представить так, как показано на рис. 9.

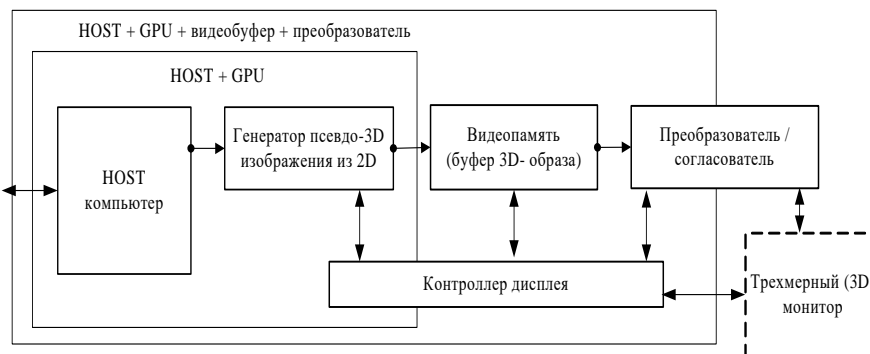


Рисунок 9 - Обобщенная структура 3D СОИ для генерации «псевдо 3D»

Работа системы заключается в организации синтеза (воссоздания, «реставрации») 3D- изображения (обычно 3D-стерео) из имеющегося плоского 2D- изображения.

Методы, применяемые при организации «псевдо 3D»-синтеза не требуют явных геометрических моделей. Термин «невная геометрия» означает, что 3D положения точек не восстанавливаются, а целевое изображение (или некоторый их набор) преобразуется с помощью некоторых манипуляций в 3D- стерео изображение. Для этого используются технологии компьютерного зрения - специфические алгоритмы анализа изображений и создания карт глубины изображения, на основе которых происходит трансформация исходного изображения в стереопару.

Основные этапы синтеза представлены на рис. 10.



Рисунок 10 - Основные этапы синтеза «псевдо 3D-» изображения

Заключение

В заключение необходимо отметить, что существующие сегодня объемные 3D- дисплеи и системы пространственной визуализации на их основе, способные демонстрировать трехмерное изображение, не позволяют пока воспроизводить полноценные и качественные визуально осязаемые копии модели сцены, создаваемой методами трехмерной графики, имеют очень высокую стоимость при фактически единичных произведенных экземплярах для ограниченного круга пользователей.

Кроме того, одной из важных причин малой распространенности устройств на базе объемных технологий является отсутствие стандартизации представления 3D информации для них. Здесь следует отметить, что, в отличие от 2D визуализации, в этой области даже не выработаны определения типовых 3D графических примитивов и алгоритмов их генерации.

Крайне важным, также, является вопрос аппаратной поддержки генерации 3D- изображений для высоко реалистичного синтеза в реальном времени. И если некоторые из рассмотренных способов синтеза 3D- пространственных изображений используют традиционный, ставший стандартом, графический 3D- конвейер, аппаратно реализованный в современных параллельных графических 3D- ускорителях видеокарт, то другие используют специальные процедуры и алгоритмы синтеза, не имеющие стандартных средств аппаратной поддержки, что актуализирует разработку высокопроизводительных вычислительных систем для их реализации.

Литература

1. Bashkov E.A., Zori S.A., Anas-M. Al Oraiqt Spatial Visualization via Real Time 3D Volumetric Display Technologies / Lambert Academic Publishing (LAP). - 2018-01-25 | book. ISBN: 978-620-2-09398-9
2. Zori S.A., Bashkov E.A., Anas-M. Al Oraiqt Specialized Computer Systems for Environment Visualization // Springer 3D Research. – 2018. - Vol. 9, No 2. - 13 P. | DOI: 10.1007/s13319-018-0168-9
3. Зори С.А., Башков Е.А. Реалистичная пространственная визуализация с использованием технологий объемного отображения: Монография / Е.А. Башков, С.А. Зори. – Донецк, ГВУЗ "ДонНТУ", 2014. – 150 с.
4. Е.А. Башков, С.А. Зори Реалистичная визуализация трехмерных объектов и сцен с использованием технологий объемного отображения // Известия ЮФУ. Технические науки. Тематический выпуск "Компьютерные и информационные технологии в науке, инженерии и управлении". – 2012. - №5(130). – С. 133 – 137.
5. Башков Е.А., Зори С.А., Ковальский С.В. Современное алгоритмическое и аппаратное обеспечение виртуальных систем трехмерного моделирования окружающей обстановки // Наукові праці Донецького національного технічного університету. Серія "Інформатика, кібернетика і обчислювальна техніка, ІКВТ – 2009: – Донецьк: ДонНТУ. – 2009. – С. 78 – 89.
6. Blundell B.G., Adam J., Barry G. Blundell The Classification of Volumetric Display Systems // IEEE Transactions on Visualization and Computer Graphics. 2002. – Vol 8, No. 1. – P. 66 – 75.
7. Bahram Javida, Fumio Okano Three – dimensional television, video and display technology / Springer – Verlag Berlin Heidelberg. – 2002. – 505 p.
8. Favalora G.E. 100 million – voxel volumetric display. // Cockpit Displays IX: Displays for Defense Applications. Proc. SPIE—Int'l Soc. for Optical Eng. - 2002. - Vol. 4712. – P. 300 – 312.

9. Display Systems. Design and Applications / Edited Lindsay W. MacDonald, Antony C. Lowe. – Jphn Willey & Sons Ltd., 2009. – 418 p.
10. Foley J.D. Computer graphics principles & practice / Addison-Wesley Longman Publishing Co., Inc. Boston, MA, USA. - 1996. – 1175 p. | book. ISBN: 0-201-12110-7
11. А. Я. Аноприенко, Е. А. Башков, Ю. А. Коба, Р. В. Мальчева Аппаратно – программный комплекс генерации изображений в реальном времени // “Вопросы радиоэлектроники”, серия “ЭВТ 2”. – 1991. – С. 77 – 84.
12. А. Я. Аноприенко, Е. А. Башков, Ю. А. Коба, А. А. Кухтин, Р. В. Мальчева, Т. В. Чухонцева Система синтеза изображений в реальном времени для испытательных стендов // “Гибридные вычислительные машины и комплексы”. –1992. - No 15. – С. 72 – 76.

Зори С.А., Аноприенко А.Я. Особенности организации объемной визуализации в системах трехмерной компьютерной графики. В статье представлен обзор современных принципов организации систем объемного синтеза и 3D- визуализации изображений. Предложены классификации и обобщенные структуры систем 3D- синтеза и визуализации изображений, определены основные особенности выполнения стадий 3D- графического конвейера в них. Рассмотренные системы на сегодняшний день в процессе работы используют специальные процедуры и алгоритмы синтеза и визуализации, не имеющие стандартных средств аппаратной поддержки, что актуализирует дальнейшие исследования и разработку высокопроизводительных вычислительных систем для их реализации.

Ключевые слова: компьютерная графика, объемная 3D- визуализация, 3D- дисплей, графический 3D- конвейер, графическая система 3D- отображения информации, структура.

Zori SA, Anoprienko A.Ya. Features of the volume visualization organization in 3D- computer graphics systems. The article provides an overview of the modern principles organization of 3D volumetric image synthesis and visualization systems. Classifications and generalized structures of 3D- image synthesis and visualization systems are proposed, the main features of the 3D- graphic pipeline stages implementation in them are determined. Now such systems use special procedures and algorithms for 3D- image synthesis and visualization that don't have standard hardware support, which actualizes further researches and development of high-performance computing systems for their implementation.

Key words: computer graphics, 3D volume visualization, 3D display, 3D graphic pipeline, 3D graphics visualization system, structure.

Принципы построения компьютерных моделей многофакторных процессов и явлений методом многомерной интерполяции

Конопацкий Е.В.

Донбасская национальная академия строительства и архитектуры
e.v.konopatskiy@donnasa.ru

Конопацкий Е.В. Принципы построения компьютерных моделей многофакторных процессов и явлений методом многомерной интерполяции. В статье изложены шесть основных принципов составляющих методику компьютерного моделирования многофакторных процессов и явлений методами многомерной интерполяции и аппроксимации, в основу которых заложено конструирование геометрических объектов, инцидентных заданным точкам. Предложенная методика позволяет строить, анализировать и оптимизировать модели многофакторных процессов и явлений с любым, наперёд заданным, количеством факторов, влияющих на функцию отклика.

Ключевые слова: компьютерная модель, многофакторный процесс, многомерная интерполяция, функция отклика, факторы влияния, точечное уравнение, дерево геометрической модели, дуга кривой, отсек поверхности, отсек гиперповерхности.

Введение

В настоящее время существует достаточно большое количество программного обеспечения для компьютерного моделирования. Однако основой любой из программ является аналитическое описание решения поставленной задачи. Например, большинство из существующих компьютерных систем моделирования и расчёта практически в любой отрасли науки и техники основаны на численных методах: метод конечных элементов, метод конечных разностей, метод дискретных элементов и т.п. Также широкое распространение получили системы компьютерного моделирования на основе методов математической статистики (корреляционный анализ, регрессионный анализ и т.п.), которые позволяют выполнить аналитическое описание, анализ, прогнозирование и оптимизацию процессов и явлений, полученных с учётом различных экспериментально-статистических данных. Наряду с выше перечисленным, одним из эффективных инструментов компьютерного моделирования многофакторных процессов и явлений является многомерная интерполяция и аппроксимация. При этом основой компьютерного моделирования может выступать любой математический аппарат, позволяющий решать поставленные задачи. Автор использует в своих исследованиях геометрическое моделирование многофакторных процессов и явлений, реализованное в точечном исчислении (др. название – БН-исчисление [1-3]). Это позволяет найти аналитическую зависимость между любыми, даже совершенно несоизмеримыми и несовместимыми, факторами и функцией отклика и представить её в виде единого параметризованного геометрического объекта. Причём размерность пространства полученного геометрического объекта напрямую зависит от количества факторов, влияющих на функцию отклика. Однако практическое применение полученных точечных уравнений и вычислительных алгоритмов на их основе невозможно без использования вычислительных возможностей современной компьютерной техники. Поэтому геометрическое и компьютерное моделирование в значительной степени дополняют друг друга. Их синтез является залогом успешного решения задач моделирования многофакторных процессов и явлений методами многомерной интерполяции и аппроксимации.

Сама многомерная интерполяция не является чем-то новым. Её важность подчёркивается в работе [4] применительно к использованию для многомерных таблиц в физике и технике. Так примером двухмерной таблицы могут служить таблицы термодинамических функций газов, где независимыми переменными обычно являются температура и плотность. Трёхмерные таблицы составляют и используют значительно реже, но не потому, что таких зависимостей нет, а потому, что таблицы слишком громоздки. Четырёхмерных таблиц практически нет, хотя в физике немало задач с большим числом параметров. Следует отметить, что с появлением современных вычислительных машин, способных оперировать большими объёмами информации, использование таких таблиц, а вместе с ними и многомерной интерполяции, в значительной мере возросла. Кроме того методы многомерной интерполяции и аппроксимации могут быть эффективно использованы для моделирования и анализа экспериментально-статистической информации с возможностью последующей

оптимизации исследуемого процесса или явления.

Научных подходов к решению задач многомерной интерполяции встречается достаточно много. Например, в работе [5] приводится сравнение различных способов многомерной интерполяции, которые включают: интерполяцию многочленом Лагранжа, полиномиальную рекурсивную интерполяцию и рациональную интерполяцию. В работе [6] изучается интерполяция периодической функции многих переменных, заданной в узлах обобщённой параллелепипедной сетки целочисленной решётки. Задача интерполяции функции, заданной на регулярной сетке, для случая большого числа переменных приводится в работе [7]. Решению задачи многомерной сплайн-интерполяции на хаотических сетках с огромным числом интерполяционных точек посвящена работа [8]. Кроме того, задачи многомерной интерполяции и аппроксимации могут быть решены на основе теории случайных функций [9].

Анализируя разнообразие научных подходов к решению задач многомерной интерполяции, можно сделать вывод о наличии как универсальных классических, так и инновационных модифицированных методов решения. Тем не менее, каждый метод в большей степени ориентирован на решение конкретной практической задачи. Применение его к другим задачам может иметь как позитивные, так и негативные последствия. В данной статье излагаются основные принципы компьютерного моделирования и оптимизации многофакторных процессов и явлений, основанных на их геометрическом моделировании, реализованном в БН-исчислении.

Принцип 1. Выбор математического аппарата компьютерного моделирования многофакторных процессов и явлений

Выбор математического аппарата компьютерного моделирования многофакторных процессов и явлений является непростой задачей и в значительной степени зависит как от исходных её условий, так и от ожидаемого результата. При этом необходимо учитывать особенности решения поставленной задачи. В данном случае каждому фактору исследуемого процесса ставится в соответствие параметр его геометрической модели. Соответственно для многофакторных процессов таких параметров будет много (по числу факторов), а многопараметрический геометрический объект должен располагаться в многомерном пространстве. С одной стороны использование многомерного пространства для моделирования многофакторных процессов и явлений является очень перспективным направлением, практически не ограничивающим количество исследуемых факторов, влияющих на функцию отклика, а с другой стороны – возникают сложности восприятия и визуализации, как самого процесса, так и результатов моделирования. Тогда к математическому аппарату предъявляются дополнительные требования, которые заключаются в возможности обобщения всех математических операций и уравнений на многомерное пространство. Причём, не имея возможности использования зрительной наглядности, необходимо прибегать к наглядности логической, основанной на методах обобщения и аналогии, которые и должен обеспечить аппарат моделирования.

Исходя из всего вышеизложенного, для геометрического и компьютерного моделирования многофакторных процессов и явлений был выбран математический аппарат – БН-исчисление [1-3], основанный на инвариантах аффинной геометрии. Все операции с геометрическими объектами, которые находят в БН-исчислении своё аналитическое описание, также основаны на инвариантах аффинной геометрии и включают в себя: простое отношение трёх точек прямой, параллельный перенос и определение точки пересечения двух прямых (с учётом обобщения на многомерное пространство). Все другие операции с геометрическими объектами, какими бы сложными они не были, можно свести к этим трём. Такой подход даёт возможность реализовать очень важное свойство БН-исчисления – покоординатный расчёт, который позволяет представить любой геометрический объект в виде совокупности проекций на оси глобальной системы координат. Это даёт возможность ставить в соответствие каждой оси координат свой фактор влияния и конечно саму функцию отклика. Полученная в такой системе координат геометрическая модель позволяет установить зависимость между функцией отклика и влияющими на неё факторами.

Кроме того, можно выделить и другие преимущества БН-исчисления, как аппарата геометрического моделирования и оптимизации многофакторных процессов и явлений:

1. Точечные уравнения геометрических объектов инвариантны относительно размерности пространства глобальной системы координат. То есть в качестве параметров выбираются такие, которые являются инвариантными относительно параллельного проецирования (например, простое отношение трех точек прямой). Поэтому точечные уравнения справедливы для пространства любой размерности.

2. БН-исчисление позволяет работать в локальном симплексе, а результат получать в глобальном симплексе. Причем, переход от локального симплекса к глобальному осуществляется автоматически.

3. В БН-исчислении каждой геометрической операции соответствует аналитическая операция. Таким образом, БН-исчисление позволяет представить любой геометрический алгоритм построения в аналитическом виде, или в виде точечного уравнения, или в виде вычислительного алгоритма, который, по сути, является упорядоченным множеством точечных уравнений. Поэтому в точечных уравнениях сохраняется наглядный геометрический смысл параметров, который известен из геометрического алгоритма построения. Эта особенность была потеряна для многих параметрических уравнений геометрических объектов, которые были получены аналитически, без использования геометрического алгоритма построения.

4. Точечные уравнения, по своей сути, является символьной записью. Переходя к глобальной декартовой системе координат, точечные уравнения заменяются на систему однотипных параметрических уравнений, количество которых зависит от размерности пространства глобальной системы координат. Эта операция получила название покоординатного расчёта. Точечные уравнения и вычислительные алгоритмы, полученные на их основе, можно считать оптимизированными для распараллеливания потоков и использования многоядерных микропроцессоров, поскольку вычисление каждой отдельной проекции на координатную ось можно выполнять параллельно. При этом вид уравнения остаётся без изменений, меняются только координаты соответствующих точек.

5. В БН-исчислении был разработан специальный метод подвижного симплекса [16], который позволяет конструировать геометрические объекты любой сложности с наперед заданными свойствами. Использование метода подвижного симплекса позволяет установить зависимость между несколькими факторами, что, в свою очередь, позволяет моделировать многофакторные процессы и явления, учитывая не только независимые факторы, но и факторы, которые зависят один от другого.

Принцип 2. Структура и аналитическое описание дерева геометрической модели

В основу геометрического моделирования многофакторных процессов и явлений методом многомерной интерполяции [10] положен простейший принцип принадлежности одного геометрического объекта к другому из начертательной геометрии [11]. Например, прямая линия принадлежит плоскости, если две точки этой прямой принадлежат плоскости. В свою очередь, точка принадлежит плоскости, если она принадлежит прямой, лежащей в этой плоскости. Рассматривая прямую линию, как частный случай кривой, а плоскость – как частный случай поверхности, получим что для того чтобы провести через исходные точки поверхность, нужно организовать их в виде опорных (в начертательной геометрии используется термин направляющих) линий, объединив их впоследствии с помощью образующей. Тогда все точки, принадлежащие опорным линиям, будут принадлежать искомой поверхности, которая, в свою очередь, принадлежит трёхмерному пространству. Причём, образующая также является кривой линией, только проходящей через текущие точки опорных линий искомой поверхности. В результате получим двухпараметрическое множество, один из параметров которого согласовывает движение текущих точек по опорным линиям, а второй – определяет движение текущей точки по образующей линии. В БН-исчислении такой подход получил название – метод подвижного симплекса [12], поскольку образующая кривая поверхности располагается в подсимплексе по отношению к глобальному симплексу, а движение этого подсимплекса по опорным линиям как раз и определяет отсек поверхности. Обобщая этот подход на многомерное пространство, получим следующее утверждение: чтобы гиперповерхность четырёхмерного пространства была носителем наперёд заданных точек, координаты которых соответствуют исходной экспериментально-статистической информации, необходимо чтобы исходные точки принадлежали семейству поверхностей, объединенных в гиперповерхность образующей линией, проходящей через текущие точки семейства поверхностей.

Тогда задача опять сводится к определению кривых линий проходящих через наперёд заданные точки, только таких линий становится больше. Таким образом, формируется дерево модели процесса, которое можно представить в виде некоторой геометрической схемы, по которой впоследствии будет аналитически описан и построен геометрический объект. Аналогичным образом можно построить геометрическую модель процесса, принадлежащую пространству любой размерности. Собственно говоря, размерность пространства напрямую зависит от количества факторов влияющих на процесс и потому получивших название – факторы влияния. Каждому фактору влияния соответствует параметр геометрической модели процесса, а чем больше параметров определяют геометрический объект, тем в более высокоразмерном пространстве он находится.

Так однофакторный процесс или явление геометрически можно представить однопараметрическим множеством точек – линией двумерного пространства, проходящей через наперёд заданные точки (рис. 1), которая определяется следующим точечным уравнением [10]:

$$M = \sum_{i=1}^n M_i p_i(u), \quad (1)$$

где M – текущая точка дуги кривой линии (здесь прямые линии рассматриваются как кривые с нулевой кривизной), проходящей через наперёд заданные точки;

M_i – исходные точки, через которые должна проходить искомая дуга кривой. Координаты этих точек соответствуют исходным экспериментально-статистическим данным;

$p_i(u)$ – функции от параметра u (методика их определения на основе полиномов Бернштейна изложена ниже);

u – текущий параметр, который изменяется от 0 до 1. При этом значение $u=0$ определяет начальную точку дуги кривой M_1 , а $u=1$ – конечную точку дуги кривой M_n .

n – количество исходных точек дуги кривой, соответствующее количеству экспериментов;

i – порядковый номер исходной точки, через которую должна проходить искомая дуга кривой, который изменяется от 1 до n и соответствует порядковому номеру эксперимента.

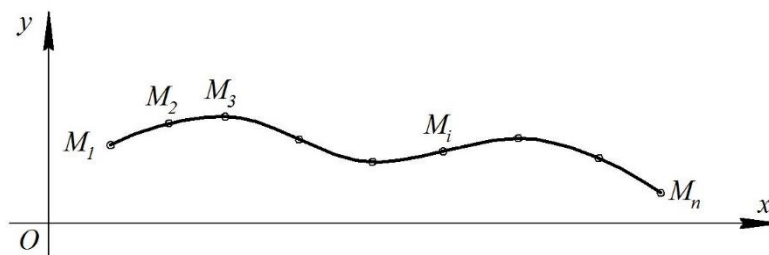


Рисунок 1 – Геометрическая модель однофакторного процесса

Причём обязательным является выполнение условия принадлежности геометрического объекта к определённому пространству. В данном случае принадлежность дуги кривой к плоскости OE_1E_2 декартовой системы координат: $\sum_{i=1}^n p_i(u) = 1$ [1-2]. Следует отметить, что точечные уравнения, полученные на инвариантах параллельного проецирования, являются справедливыми как для декартовой, так и для аффинной системы координат.

Точечное уравнение (1) представляет собой символическую запись. Выполнив по координатный расчёт, для плоской кривой получим систему двух параметрических уравнений:

$$\begin{cases} x_M = \sum_{i=1}^n x_{M_i} p_i(u), \\ y_M = \sum_{i=1}^n y_{M_i} p_i(u), \end{cases} \quad (2)$$

где x_M и y_M – координаты текущей точки дуги кривой, проходящей через наперёд заданные точки;

x_{M_i} и y_{M_i} – координаты исходных точек, соответствующих исходным экспериментальным данным.

Эта же кривая в трёхмерном пространстве будет определяться аналогичной системой трёх уравнений. Подобным образом любое точечное уравнение можно представить в виде системы n параметрических уравнений.

Система параметрических уравнений (2) представляет собой аналитическое описание проекций дуги плоской кривой на оси глобальной системы координат. При этом оси Ox соответствует фактор, влияющий на процесс или явление, а оси Oy – искомая функция отклика. Связь между функцией отклика и фактором влияния аналитически осуществляется с помощью параметра u . Графически она представляет собой линию проекционной связи, если представить проекции кривой на комплексном чертеже Монжа.

Следует отметить, что каждой точке, через которую проходит геометрический объект, соответствует конкретное значение экспериментально-статистической информации исходных данных искомой модели процесса или явления.

В свою очередь, двухфакторный процесс определяется двухпараметрическим множеством точек – поверхностью трёхмерного пространства, проходящей через наперёд заданные точки (рис. 2):

$$\begin{cases} M_1 = \sum_{j=1}^n M_{1j} p_{1j}(u), \\ \dots\dots\dots \\ M_i = \sum_{j=1}^n M_{ij} p_{ij}(u), \\ \dots\dots\dots \\ M_m = \sum_{j=1}^n M_{mj} p_{mj}(u), \\ M = \sum_{i=1}^m M_i q_i(v), \end{cases} \quad (3)$$

где M – текущая точка отсека поверхности, проходящей через наперёд заданные точки в количестве $m \times n$. Она же текущая точка образующей дуги кривой;

M_i – текущая точка i -й опорной (направляющей) дуги кривой, проходящей через наперёд заданные точки. Движение всех опорных линий согласовывается с помощью единого параметра u ;

M_{ij} – исходные точки, через которые должен проходить искомый отсек поверхности. Координаты этих точек соответствуют исходным экспериментально-статистическим данным;

$p_{ij}(u)$ – функции от параметра u (в данном случае полиномиальные, но в общем случае могут быть и другие, в том числе трансцендентные), определяющие вид опорных линий;

$q_i(v)$ – функции от параметра v , определяющие вид образующей линии;

m – количество опорных линий;

n – количество исходных точек в каждой опорной линии (на рис. 2 n одинаково для всех опорных линий, но это не обязательно);

i – порядковый номер опорной линии;

j – порядковый номер исходной точки в каждой опорной линии;

u – текущий параметр опорных линий отсека поверхности, который изменяется от 0 до 1;

v – текущий параметр образующей линии отсека поверхности, который также изменяется от 0 до 1;

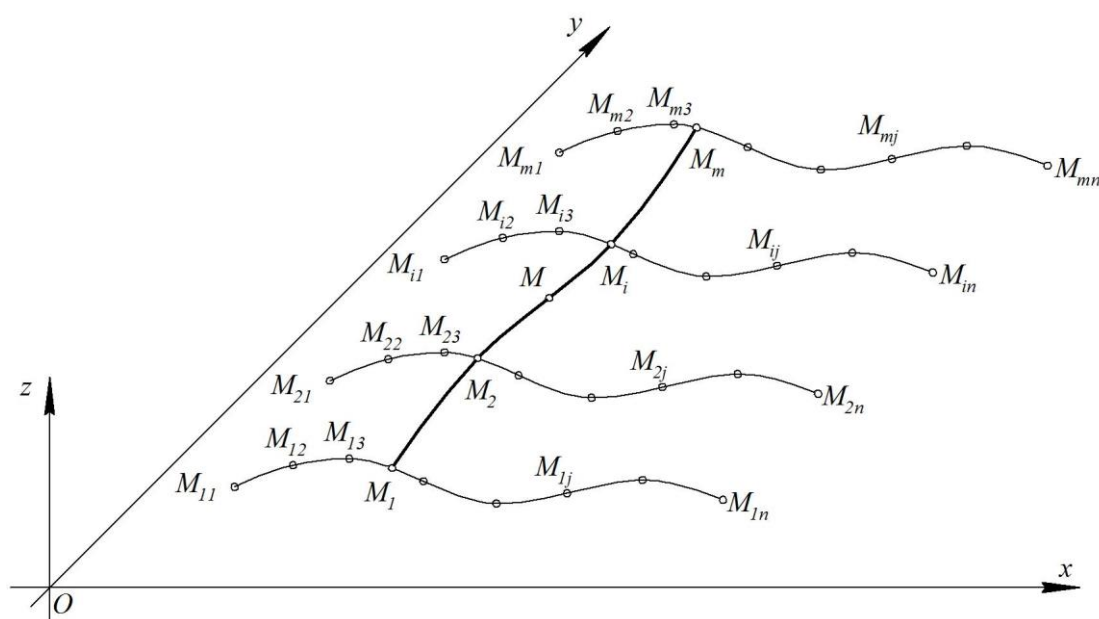


Рисунок 2 – Геометрическая модель двухфакторного процесса

Для описания геометрической модели двухфакторного процесса (рис. 2) используется трёхмерная, в частном случае Декартова, система координат. Причём осям Ox и Oy соответствуют факторы, влияющие на процесс, а ось Oz – искомая функция отклика.

Аналогичным образом, трёхфакторный процесс определяется трёхпараметрическим множеством точек – гиперповерхностью четырёхмерного пространства, проходящей через наперёд заданные точки, (рис. 3):

$$\left[\begin{array}{l} M_{ij} = \sum_{k=1}^l M_{ijk} p_{ijk}(u), \\ \dots\dots\dots \\ M_i = \sum_{j=1}^n M_{ij} q_{ij}(v), \\ \dots\dots\dots \\ M = \sum_{i=1}^m M_i r_i(w), \end{array} \right. \quad (4)$$

где M – текущая точка отсека гиперповерхности, проходящей через наперёд заданные точки в количестве $m \times n \times l$;

M_i – текущая точка образующей i -го отсека поверхности, проходящая через наперёд заданные точки,

которая является опорной для построения гиперповерхности;

M_{ij} – текущая точка j -й опорной дуги кривой, проходящей через наперёд заданные точки;

M_{ijk} – исходные точки, через которые должен проходить искомый отсек гиперповерхности. Координаты

этих точек соответствуют исходным экспериментально-статистическим данным;

$p_{ijk}(u)$ – функции от параметра u , определяющие вид опорных линий;

$q_{ij}(v)$ – функции от параметра v , определяющие вид образующих линий отсеков поверхностей;

$r_i(w)$ – функции от параметра w , определяющие вид образующей линии отсека гиперповерхности;

l – количество исходных точек в каждой опорной линии;

n – количество опорных линий для построения опорных поверхностей;

m – количество опорных поверхностей для построения гиперповерхности;

i – порядковый номер опорной поверхности;

j – порядковый номер опорной линии;

k – порядковый номер исходной точки в каждой опорной линии;

u , v и w – текущие параметры, которые изменяются от 0 до 1.

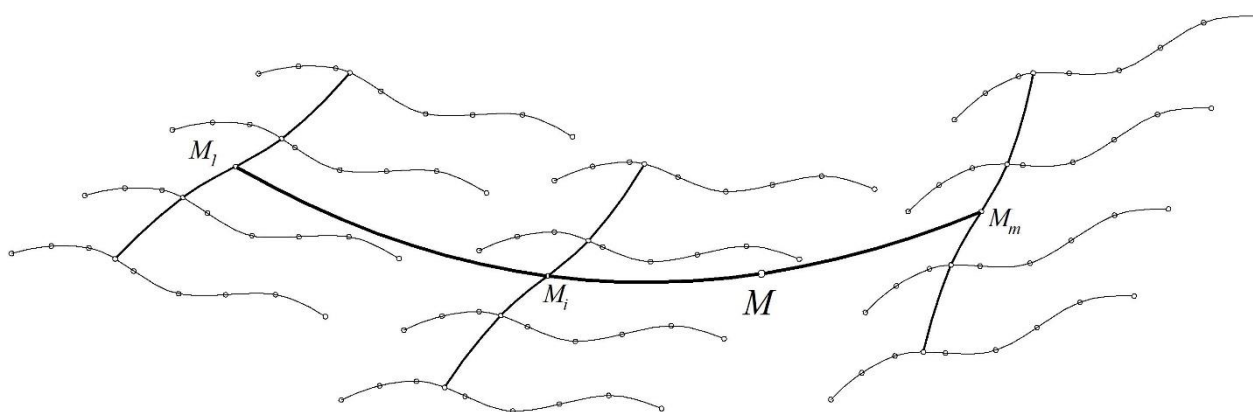


Рисунок 3 – Дерево геометрической модели трёхфакторного процесса

Выполнив покоординатный расчёт, получим проекции гиперповерхности на оси глобальной системы координат, для которой оси Ox , Oy и Oz будут соответствовать факторам влияния, а Oz – искомым функциям отклика.

Обобщая предложенный подход, получим геометрическую модель n -факторного процесса, как n -параметрическое множество точек или гиперповерхность $(n+1)$ -го пространства. Инцидентность искомого геометрического объекта заданным точкам обеспечивается прохождением всех точек через направляющие линии на каждом этапе формирования дерева модели. При этом любая линия представляется организованным определённым образом однопараметрическим множеством точек. Таким образом, для геометрического моделирования многофакторных процессов и явлений необходимо иметь библиотеку специальных дуг кривых, проходящих через наперёд заданные точки.

Принцип 3. Моделирование дуг алгебраических кривых инцидентных заданным точкам

Для моделирования дуг кривых, инцидентных заданным точкам, возможны два принципиальных подхода: моделирование единых кривых линий и составных. Чтобы сделать выбор необходимо знать конкретные условия задачи моделирования. Однако следует учесть, что сложность визуального восприятия многомерного пространства, приводит к необходимости использования не зрительной, а логической наглядности, основанной на методах обобщения и аналогии. Поэтому использование единой кривой является более предпочтительным и даёт возможность не только определить оптимальные значения факторов, влияющих на функцию отклика, но и в перспективе спрогнозировать её поведение. Это приводит к необходимости создания методики определения универсальных дуг кривых, которые должны быть носителями исходных точек вне зависимости от их взаимного положения.

Рассмотрим методику моделирования дуг алгебраических кривых, проходящих через наперёд заданные точки на основе полиномов Бернштейна [10]. Пусть задана ломаная линия $A_1A_2...A_n, A_{n+1}$. Дуга алгебраической кривой n -го порядка имеет следующее точечное уравнение:

$$M = A_1 \bar{u}^n + A_2 C_n^1 \bar{u}^{n-1} u + A_3 C_n^2 \bar{u}^{n-2} u^2 + \dots + A_n C_n^{n-1} \bar{u} u^{n-1} + A_{n+1} u^n = \sum_{i=0}^n A_{i+1} C_n^i \bar{u}^{n-i} u^i, \quad (5)$$

где M – текущая точка дуги кривой;

A_n^1, \dots, A_n^i – исходные точки, определяющие форму дуги кривой;

C_n^1, \dots, C_n^i – коэффициенты Бернштейна;

u – текущий параметр, который изменяется от 0 до 1;

$\bar{u} = 1 - u$ – дополнение параметра до 1.

Тогда исходное точечное уравнение кривых, полученных на основе полиномов Бернштейна, которые в общем случае находятся в пространстве размерности n , являются кривыми $(n-1)$ -й кривизны (например, пространственная дуга кривой 3-го порядка в общем случае является дугой двойкой кривизны) и могут быть записаны в следующем виде:

$$M = \sum_{i=0}^n A_{i+1} \frac{n!}{i!(n-i)!} \bar{t}^{n-i} t^i. \quad (6)$$

Используя равномерное распределение текущего параметра $t = \frac{j}{n}$, получим:

$$M_{j+1} = \sum_{i=0}^n A_{i+1} \frac{n!}{i!(n-i)!} \left(\frac{n-j}{n}\right)^{n-i} \left(\frac{j}{n}\right)^i. \quad (7)$$

Переопределим точки ломаной линии $A_1 A_2 \dots A_n, A_{n+1}$ через точки $M_1 M_2 \dots M_n, M_{n+1}$, которые принадлежат дуге кривой, определенной с помощью уравнения (7). В результате получим систему линейных $n+1$ алгебраических уравнений, каждая строка которой определяется следующим уравнением:

$$\sum_{i=0}^n \frac{n!}{i!(n-i)!} (n-j)^{n-i} (j)^i A_{i+1} = n^n M_{j+1}. \quad (8)$$

Решив эту систему уравнений методом Крамера и подставив в исходное уравнение (7), получим уравнение дуги кривой, проходящей через наперёд заданные точки $M_1 M_2 \dots M_n, M_{n+1}$.

Полученная дуга кривой представлена в точечной форме (т.е. в символьной форме, для которой вместо координат используются непосредственно точки). Для практического использования полученных уравнений необходимо выполнить по координатный расчёт [1-2]. При этом одно точечное уравнение заменяется системой параметрических уравнений, которые представляют собой аналитическое описание проекций текущей точки на оси глобальной системы координат. Здесь следует отметить очень важную отличительную особенность полученного уравнения. Для точечных уравнений принадлежность геометрического объекта к пространству конкретной размерности определяется суммой функций от параметра u , которая обязательно должна быть равна 1. Поскольку функции от параметра u определяются биномом Ньютона, который раскладывается для параметра u и его дополнения до 1, то условие принадлежности дуги кривой конкретному пространству будет выполняться вне зависимости от размерности пространства. Иными словами полученные параметрические уравнения дуги кривой могут быть использованы для пространства любой размерности.

Другой важной особенностью полученной дуги кривой является равномерное распределение параметра, изначально заложенное в методику определения дуги кривой, проходящей через наперёд заданные точки. При этом для каждой конкретной координатной оси, имеющей равномерное распределение координат исходных точек справедлива линейная зависимость между натуральным значением фактора, принадлежащего i -й оси проекций, и текущим параметром:

$$x_i = n l_i u + b_i, \quad (9)$$

где x_i – i -я ось проекций глобальной системы координат;

n – порядок дуги кривой;

b_i – начальное значение фактора влияния, соответствующее i -й оси проекций;

l_i – шаг равномерного распределения проекции исходных точек на i -ю ось.

Эта особенность в значительной мере сокращает объём необходимых вычислений при моделировании геометрических объектов на регулярной сети точек. При этом сама методика носит универсальный характер и без внесения каких-либо модификаций, может в полном объёме использоваться как для регулярной, так и для нерегулярной сети точек.

Принцип 4. Проверка достоверности результатов моделирования

Для оценки точности аппроксимации и проверки достоверности полученных моделей существует

достаточно много различных критериев, к которым относятся коэффициент детерминации, стандартная ошибка модели, средняя относительная ошибка аппроксимации и т.п. Однако работ, посвящённых оценке достоверности геометрических моделей с учётом их специфики, автору найти не удалось. Наиболее часто в литературе встречается точность аппроксимации применительно к методам математической статистики (например, к регрессионному анализу), которые основаны на сравнении результатов моделирования с исходными данными. Этот же принцип может быть заложен в основу оценки точности аппроксимации геометрической модели.

В общем, точность любой модели характеризуется разностью между фактическими и расчетными значениями исследуемого показателя. Такой мерой точности является стандартная ошибка модели $S_{\text{мод}}$, которую можно вычислить по формуле:

$$S_{\text{мод}} = \sqrt{\frac{\sum_{i=1}^n (y_i - \hat{y}_i)^2}{n-2}},$$

где $\sum_{i=1}^n (y_i - \hat{y}_i)^2$ – сумма квадратов регрессионных остатков, которая включает фактические y_i и расчётные \hat{y}_i значения исследуемой переменной.

Однако наиболее удобно оценивать точность модели с помощью средней относительной ошибки аппроксимации $E_{\text{отн}}$, которая показывает, на сколько процентов в среднем модельные значения \hat{y}_i отличаются от фактических y_i :

$$E_{\text{отн}} \approx 0,8 \frac{S_{\text{мод}}}{\bar{y}} 100\%,$$

где \bar{y} – выборочное среднее.

Если $E_{\text{отн}} \leq 5\%$, то считается, что модель имеет достаточно высокую точность.

Наиболее же удобным инструментом для оценки адекватности полученной модели, с точки зрения автора, является коэффициент детерминации R^2 , определяемый по формуле:

$$R^2 = 1 - \frac{\sum_{i=1}^n (y_i - \hat{y}_i)^2}{\sum_{i=1}^n (y_i - \bar{y})^2},$$

где $\sum_{i=1}^n (y_i - \bar{y})^2$ – общая дисперсия.

Коэффициент детерминации для модели принимает значения от 0 до 1. Чем ближе значение коэффициента к 1, тем сильнее зависимость. При оценке геометрических моделей будем это интерпретировать как соответствие полученной модели исходным экспериментальным (или статистическим) данным. Для приемлемых моделей предполагается, что коэффициент детерминации должен быть хотя бы не меньше 50%. Модели с коэффициентом детерминации выше 80% можно признать достаточно хорошими. Равенство коэффициента детерминации 1 означает, что было достигнуто полное соответствие между геометрической моделью и исходными данными.

В общем случае, вне зависимости от выбора критериев адекватности геометрической модели, заимствованных из математической статистики, для моделей с высоким уровнем достоверности любой из перечисленных критериев будет показывать хороший результат. Другими словами, достоверность модели является инвариантной по отношению к способу её проверки. Исходя из этого, для анализа результатов моделирования будем использовать коэффициент детерминации. Однако здесь следует учесть важную особенность. В случае многомерной интерполяции итоговый геометрический объект является носителем исходной экспериментально-статистической информации, а потому коэффициент детерминации такой модели будет равен 1. Поэтому для более точной проверки достоверности и устойчивости результатов необходимо использование дополнительных экспериментально-статистических данных.

Принцип 5. Аппроксимации модели более простым геометрическим объектом

Общая концепция компьютерного моделирования геометрических объектов методом многомерной интерполяции, применительно к решению задач моделирования и оптимизации многофакторных процессов и явления, включает четыре принципа, изложенные выше. Однако итоговый результат, включающий в себя большое количество исходных данных, может иметь значительный размер, что с одной стороны не является неразрешимой проблемой, учитывая мощность современной вычислительной техники, но с другой стороны это значительно затрудняет анализ и дальнейшее использование результатов моделирования. Одним из выходов из

этой ситуации является использование аппроксимации полученных геометрических объектов более простым геометрическим объектом. Кроме того использование многомерной аппроксимации позволяет аналитически описать общую тенденцию протекания процесса и исключить ошибочную информацию, возникающую в ходе проведения эксперимента. При этом важно, чтобы аппроксимированная модель оказывала незначительное влияние на достоверность результатов моделирования, которые должны быть подтверждены соответствующим значением коэффициента детерминации, рассмотренными выше.

В соответствии с математической энциклопедией [14], под аппроксимацией следует понимать замену одних математических объектов другими, в том или ином смысле близкими к исходным. В данном случае под аппроксимацией будем понимать замену одних геометрических объектов другими, имеющими одинаковое количество текущих параметров с исходными объектами (т.е. одинаковое количество факторов, влияющих на функцию отклика). Тогда дуга кривой, как однопараметрическое множество точек, имеет один текущий параметр и аппроксимирующий её геометрический объект также должен быть однопараметрическим множеством точек. Аналогичным образом, отсек поверхности, как двухпараметрическое множество точек, аппроксимируется другим отсеком поверхности и т.д. для отсека гиперповерхности многомерного пространства.

Для реализации многомерной аппроксимации геометрической модели процесса эффективно используются дуги алгебраических кривых, проходящие через наперёд заданные точки (принцип 3). Тогда задача сводится к определению узловых точек аппроксимации. В некоторых случаях их можно выделить из имеющихся узлов интерполяции. Методика выбора узловых точек аппроксимации включает в себя вычислительный алгоритм, который реализует анализ нескольких аппроксимированных моделей с оценкой результатов моделирования с помощью коэффициента детерминации и выбором наиболее оптимального соотношения между точностью построения модели процесса и сложностью его аналитического описания.

Принцип 6. Определение экстремумов компьютерной модели геометрического объекта

Определение экстремумов компьютерной модели геометрического объекта имеет особое значение для исследования многофакторных процессов и явлений, поскольку, в большинстве случаев, позволяет оптимизировать результат моделирования. При этом возможны два принципиальных подхода для определения экстремальных точек, первый из которых подразумевает использование методов дифференциального исчисления функции многих переменных (по числу факторов), второй – численное определение минимальных и максимальных значений путём дискретизации значений варьируемых параметров, каждый из которых соответствует фактору, влияющему на процесс. Реализация обоих подходов, учитывая объём и сложность вычислений, практически невозможна без использования современной вычислительной техники. Следует отметить, что выбор подхода к определению экстремумов напрямую зависит от конкретной задачи. В случае использования методов дифференциального исчисления следует учитывать, что экстремальные точки могут быть не только внутри исследуемой области, но и на её границах. Для проверки границ исследуемой области на экстремумы необходимо чтобы текущие параметры, соответствующие факторам влияния, поочередно принимали значения 0 и 1.

Заключение

Изложенные принципы построения компьютерных моделей многофакторных процессов и явлений позволяют строить, анализировать и оптимизировать компьютерные модели на основе дискретного массива точек в виде исходной экспериментально-статистической информации. Такой подход особенно эффективен при моделировании процессов и явлений, основанных на экспериментально-статистических данных и имеющих большое количество взаимосвязанных факторов, и нашли успешное отражение при решении научных и прикладных задач компьютерного моделирования и оптимизации физико-механических свойств и составов композиционных строительных материалов, а также социально-экономических, термодинамических, светотехнических и других процессов и явлений.

Литература

1. Балюба, И.Г. Конструктивная геометрия многообразий в точечном исчислении: дис. ... доктора техн. наук: 05.01.01 // И.Г. Балюба. – Макеевка, 1995. – 227 с.
2. Балюба, И.Г. Точечное исчисление: учебное пособие // И.Г. Балюба, В.М. Найдыш. – Мелитополь: МГПУ им. Б. Хмельницкого, 2015. – 236 с.
3. Введение в математический аппарат БН-исчисление [Текст] / Бумага А.И., Конопацкий Е.В., Крысько А.А., Чернышева О.А. // Проблемы качества графической подготовки студентов в техническом ВУЗе: традиции и инновации. Материалы VII Международной научно-практической интернет-конференции. – Пермь: ПНИПУ, 2017. – Вып. 4. – С. 76-82.
4. Калиткин, Н.Н. Численные методы // Н.Н. Калиткин. – М. Наука, 1978. – 512 с.

5. Пахнатов, И.А. Многомерная интерполяция // И.А. Пахнатов. – Интерактивная наука. – 2017. – №15. – Точка доступа: <https://cyberleninka.ru/article/n/mnogomernaya-interpolyatsiya> (дата обращения: 26.08.2018).
6. Многомерная теоретико-числовая Фурье интерполяция // Н.М. Добровольский, А.Р. Есаян, О.В. Андреева, Н.В. Зайцева. – Чебышевский сборник, 2004. – Т.5. – Вып. 1. – С.122-143.
7. Шустов, В.В. Многомерная интерполяция сеточной вектор-функции // В.В. Шустов. – Молодой учёный. – Чита: Издательство Молодой учёный, 2010. – №8(19)/2010 – Т.1. – С.17-20.
8. Бежаев, А.Ю. Сплайн-интерполяция многомерных данных большого объема // А.Ю. Бежаев. – Сибирский журнал вычислительной математики. – Новосибирск, 2003. – Т. 6. – №3. – С.249-261.
9. Бахвалов, Ю.Н. Метод многомерной интерполяции и аппроксимации и его приложения // Ю.Н. Бахвалов. – М.: Спутник+, 2007. – 108 с.
10. Конопацкий, Е.В. Геометрическое моделирование и оптимизация многофакторных процессов и явлений методом многомерной интерполяции // Е.В. Конопацкий. – Труды Международной научной конференции по физико-технической информатике СРТ2018, 28-31 мая 2018 г. – Москва-Протвино, 2018. – С.299-306.
11. Короев, Ю.И. Начертательная геометрия: учебник // Ю.И. Короев. – 3-е изд., стер. – М.: КНОРУС, 2013. – 424 с.
12. Метод подвижного симплекса при конструировании 2-поверхностей многомерного пространства / [Балюба И.Г., Полищук В.И., Горягин Б.Ф. и др.] // Моделювання та інформаційні технології: Зб. наук. праць. – К.: Інститут проблем моделювання в енергетиці ім. Г.Є. Пухова НАН України. – Т.1. – С.310-318.
13. Бумага, А.И. Геометрическое моделирование физико-механических свойств композиционных строительных материалов в БН-исчислении: дис. ... канд. техн. наук: 05.23.05, 05.01.01. // А.И. Бумага. – Макеевка, 2016. – 164 с.
14. Математика: Энциклопедия / под ред. Ю.В. Прохорова. – М.: Большая Российская энциклопедия, 2003. 845 с. ил.

Конопацкий Е.В. Принципы построения компьютерных моделей многофакторных процессов и явлений методом многомерной интерполяции. В статье изложены шесть основных принципов составляющих методику компьютерного моделирования многофакторных процессов и явлений методами многомерной интерполяции и аппроксимации, в основу которых заложено конструирование геометрических объектов, инцидентных заданным точкам. Предложенная методика позволяет строить, анализировать и оптимизировать модели многофакторных процессов и явлений с любым, наперед заданным, количеством факторов, влияющих на функцию отклика.

Ключевые слова: компьютерная модель, многофакторный процесс, многомерная интерполяция, функция отклика, факторы влияния, точечное уравнение, дерево геометрической модели, дуга кривой, отсек поверхности, отсек гиперповерхности.

Konopatskiy E.V. Principles the construction of computer models of multifactor processes and phenomena's by multidimensional interpolation. The article presents six basic principles of the method computer modeling of multifactor processes and phenomena's by methods of multidimensional interpolation and approximation, which are based on the design of geometric objects incident to a given point. The proposed method allows to build, analyze and optimize models of multifactor processes and phenomena's with any predetermined number of factors affecting the response function.

Key words: computer model, multifactor process, multidimensional interpolation, response function, factors of influence, point, equation, geometric tree models, the arc of the curve, the compartment surface, the compartment hypersurface.

Динамическое распределение файлов по узлам компьютерной сети

Лазебная Л.А., Едемская Е.Н., Бельков Д.В.
 Донецкий национальный технический университет
 l_lazebnaya@mail.ru, 6otba@list.ru, belkov65@list.ru

Лазебная Л.А., Едемская Е.Н., Бельков Д.В. Динамическое распределение файлов по узлам компьютерной сети. В статье предложен метод динамического размещения файлов по узлам компьютерной сети, позволяющий уменьшить трудоемкость размещения за счет оптимальной периодичности мониторингов уровней запросов к файлам. С использованием метода фазового укрупнения сложных систем определен оптимальный период времени между мониторингами уровней локальных запросов к файлам. Критерием оптимальности является минимизация задержки обработки запросов к файлам, вызванной мониторингом.

Ключевые слова: узлы компьютерной сети, запросы к файлам, периодичность мониторингов, состояния системы.

Введение

Быстрое развитие информационных технологий в современном мире привело к широкому распространению распределенных систем на основе компьютерных сетей. Анализ работ по размещению файлов в компьютерных сетях свидетельствует, что для обеспечения эффективного функционирования распределенной системы необходимо рациональное размещение файлов по узлам компьютерной сети в зависимости от интенсивности запросов к ним. Наиболее рационально размещать файл в тот узел, который чаще других выдает запрос на его использование. Это позволяет существенно повысить эффективность функционирования сети за счет уменьшения нагрузки на каналы связи. Для статического размещения файлов по узлам компьютерной сети необходимо при фиксированных значениях интенсивностей запросов к файлам так распределить файлы по узлам компьютерной сети, чтобы время отклика сети было минимальным [1,2]. В процессе длительной работы распределенной системы интенсивность потока запросов к файлам со временем изменяется. Исходное размещение файлов может стать не оптимальным. Поэтому размещение файлов должно корректироваться. Непрерывный мониторинг уровней локальных запросов к файлам требует больших вычислительных и временных затрат. Мониторинг целесообразно проводить периодически [3].

Целью статьи является разработка метода динамического размещения файлов по узлам компьютерной сети. Задача работы: определение оптимального интервала между мониторингами уровней локальных запросов к файлам для уменьшения трудоемкости размещения файлов.

Метод динамического распределения файлов

Введем обозначения:

N_j - суммарная интенсивность сетевых запросов к файлам узла j :
$$N_j = \sum_{i=1}^m F_{ij}(1 - X_{ij}),$$

L_j - суммарная интенсивность локальных запросов к файлам узла j :
$$L_j = \sum_{i=1}^m F_{ij} X_{ij},$$

Y_j - функция полезности узла j : $Y_j=0$, если $L_j > N_j$, иначе $Y_j=1$.

Во время мониторинга узел j выполняет следующие шаги:

Шаг 1. Измерить текущий уровень локальных запросов L_j .

Шаг 2. Если $L_j > N_j$, то $Y_j=0$, иначе $Y_j=1$.

Шаг 3. Перейти к шагу 1.

Узел, в котором хранится наибольшее число файлов, является сервером. Он выполняет следующий метод динамического размещения файлов:

Шаг 1. В текущий момент времени t распределить файлы по узлам сети и ожидать Ω_{opt} единиц времени.

Шаг 2. В момент времени $(t + \Omega_{opt})$ выдать узлам сети сигнал начать мониторинг за уровнем локальных запросов, мониторинг длится t_{mon} единиц времени.

Шаг 3. В момент времени $(t + \Omega_{opt}) + t_{mon}$ выдать узлам сети сигнал завершить мониторинг и переменной

t присвоить значение $(t + \Omega_{opt}) + t_{mon}$.

Шаг 4. Если число узлов с низкой локальной нагрузкой превысило допустимый порог Δ : $\sum_{j=1}^n Y_j \geq \Delta$, то

установить сигнал RD . Требования перераспределить файлы системы: $RD=1$ и перейти к шагу 1, иначе установить сигнал $RD=0$ и перейти к шагу 2.

Величина Δ характеризует чувствительность метода к уменьшению уровней локальных запросов в узлах. Она может принимать значения в диапазоне от 1 до n с шагом 1. При $\Delta=\Delta_{min}=1$ метод имеет наибольшую чувствительность к уменьшению локальной нагрузки узлов. При $\Delta=\Delta_{max}=n$ метод имеет наименьшую чувствительность к уменьшению локальной нагрузки узлов.

Обозначим: k – номер мониторинга, $k=1,2,\dots$, Ω_{opt}^k - период, в течение которого сеть функционирует от k -ого мониторинга до $(k+1)$ -го мониторинга.

Процесс динамического размещения файлов состоит в следующем. Пусть в момент времени t_k завершен k -й мониторинг и установлен счетчик шагов: $s:=k$. Осуществляется шаг по времени длиной Ω_{opt}^k из точки с

координатами $(0, t_k)$ в точку с координатами $(0, t_{k+1})$. Если в момент t_{k+1} выполняется условие $\sum_{j=1}^n Y_j \geq \Delta$, то

устанавливается сигнал требования перераспределить файлы системы: $RD=1$, иначе осуществляются следующие действия: $RD:=0$, $s:=s + 1$ и процесс повторяется с момента t_{k+1} . Таким образом, интервал между

перераспределениями файлов равен сумме $\sum_{k=1}^s \Omega_{opt}^k$.

Процесс функционирования сервера показан на рисунке 1. Пусть $\Delta=n$ и в момент t_1 сделано первоначальное распределение файлов. В момент $t_2 = t_1 + \Omega_{opt}^1$ проведен первый мониторинг. Так как в момент

t_2 условие $\sum_{j=1}^n Y_j \geq \Delta$ выполняется, то в этот момент необходимо перераспределить файлы второй раз. В

момент $t_3 = t_2 + \Omega_{opt}^2$ проведен второй мониторинг. Так как в момент t_3 условие $\sum_{j=1}^n Y_j \geq \Delta$ не выполняется,

то в момент t_3 перераспределение файлов не нужно. Поэтому функционирование распределенной системы

продолжается до момента t_4 . Так как в момент $t_4 = t_3 + \Omega_{opt}^3$ условие $\sum_{j=1}^n Y_j \geq \Delta$ выполняется, то в этот

момент времени необходимо перераспределить файлы в третий раз.

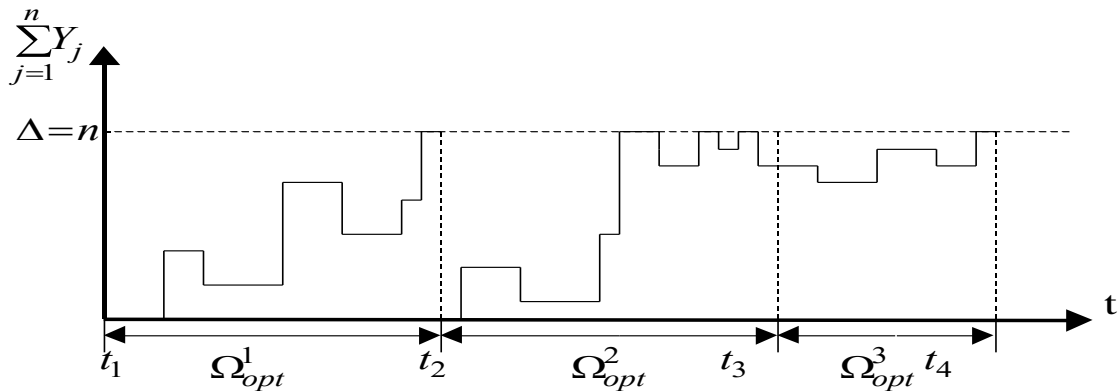


Рисунок 1 - Процесс функционирования сервера

Процесс функционирования узла j показан на рисунке 2.

Пусть $S_j = L_j - N_j$ в момент t_1 сделано первоначальное распределение файлов. В момент $t_2 = t_1 + \Omega_{opt}^1$ проведен первый мониторинг. Так как в момент t_2 условие $S_j \leq 0$ выполняется, то в этот момент формируется сигнал Y_j : $Y_j=1$. В момент $t_3 = t_2 + \Omega_{opt}^2$ проведен второй мониторинг. Так как в момент t_3 условие $S_j \leq 0$ не

выполняется, то в момент t_3 сигнал Y_j не формируется: $Y_j = 0$. В момент $t_4 = t_3 + \Omega_{opt}^3$ проведен третий мониторинг. Так как в момент t_4 условие $S_j \leq 0$ выполняется, то в этот момент формируется сигнал Y_j : $Y_j = 1$.

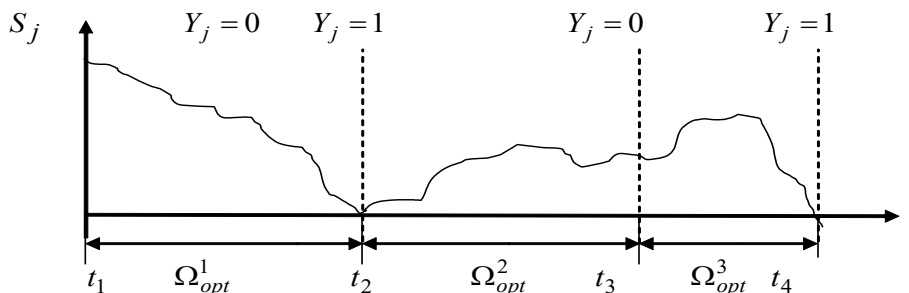


Рисунок 2 - Процесс функционирования узла j

Определение оптимальной периодичности мониторингов

Обозначим: τ_k - среднее время выполнения запроса к файлу, s - число запросов к файлам при выполнении вычислительного задания.

Состояния, в которых может находиться система, разделяются на два класса. Один класс содержит единственное состояние перераспределения файлов, другой содержит множество состояний функционирования. В состоянии перераспределения файлов выполняется процесс размещения файлов по узлам сети, запросов к файлам не возникает. После размещения файлов по узлам сети система начинает функционировать: в каждом узле выполняются запросы к файлам. В одном состоянии функционирования система выполняет один запрос к файлу, поэтому количество состояний равно s . Момент завершения выполнения запроса, совпадающий с моментом перехода системы к выполнению другого запроса, есть момент перехода системы в другое состояние функционирования. Из любого состояния функционирования система может перейти в другое состояние функционирования или в состояние перераспределения файлов. Граф состояний системы имеет вид, показанный на рисунке 3.

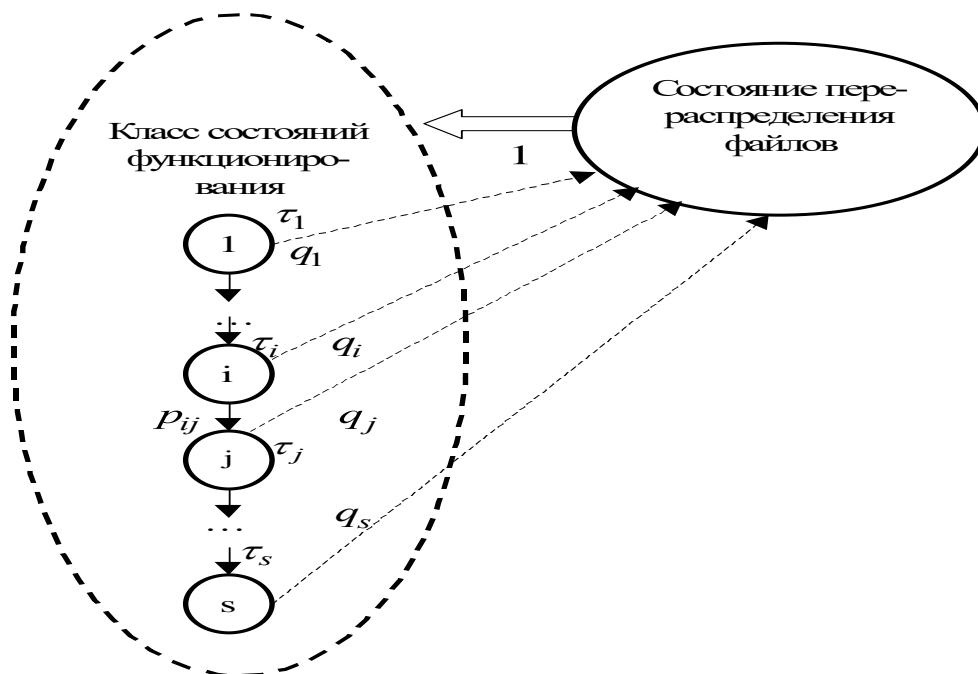


Рисунок 3 - Граф состояний системы

Пусть в i -м состоянии функционирования система пребывает случайное время с произвольной функцией распределения и средним значением τ_i , q_i - вероятность перехода из i -го состояния функционирования в состояние перераспределения файлов, p_{ij} - вероятность перехода из i -го состояния функционирования в j -е

состояние функционирования, $\sum_{j=1}^s p_{ij} < 1$, вектор начального распределения – $p_1 = \{p_{11}, p_{12}, \dots, p_{1s}\}$.

Такой процесс функционирования системы является полумарковским. Отличие от марковского процесса состоит в том, что матрица переходов является полустохастической $\left(\sum_{j=1}^s p_{ij} < 1 \right)$ и случайное время пребывания системы в состояниях функционирования имеет произвольную функцию распределения.

Для определения среднего времени пребывания системы в классе состояний функционирования до перехода в состояние перераспределения файлов используем методику [4].

Пусть α_i – среднее время пребывания системы в классе состояний функционирования до перехода в состояние перераспределения файлов при условии, что начальным состоянием является состояние i . Время α_i состоит из времени τ_i и суммы времен, накопленных в состояниях j , $j \neq i$:

$$\alpha_i = \tau_i + \sum_{j=1}^s p_{ij} \alpha_j \quad (1)$$

Для реальных компьютерных сетей сложно определить точные значения переходных вероятностей p_{ij} . Число s в реальных сетях достигает значения 10^5 и решение системы уравнений (1) трудоемко. Поэтому среднее время T пребывания системы в классе состояний функционирования до перехода в состояние перераспределения файлов целесообразно определить приближенно.

В реальных условиях работы распределенная система редко находится в состоянии перераспределения файлов. Поэтому вероятности переходов между состояниями функционирования близки к единице ($p_{ij} \approx 1$) и вероятность перехода системы из класса состояний функционирования в состояние перераспределения файлов близка к нулю ($q_i \approx 0$). Следовательно, выполняются условия укрупнения полумарковских систем и для определения значения T можно использовать метод фазового укрупнения полумарковских систем.

Для оценки погрешности, возникающей при укрупнении состояний системы, введем идеальную систему, которая отличается от реальной системы тем, что для нее никогда не требуется перераспределять файлы. Идеальная система не имеет состояния перераспределения файлов, все ее состояния являются состояниями функционирования.

Обозначим: p_{ij}^* – вероятность перехода из i -го состояния идеальной системы в j -е состояние идеальной системы, $\sum_{j=1}^s p_{ij}^* = 1$, $p_1^* = \{p_{11}^*, p_{12}^*, \dots, p_{1s}^*\}$ – вектор начального распределения идеальной системы.

Переходные вероятности идеальной системы связаны с переходными вероятностями реальной системы по формуле: $p_{ij}^* = p_{ij} + q_{ij}$, где q_{ij} – вероятность перехода из i -го состояния функционирования в состояние перераспределения файлов при условии, что следующим состоянием функционирования является j -е состояние. Матрица Q характеризует параметр близости идеальной системы реальной системе γ : $\gamma = \max(\gamma_{ij})$, где $\gamma_{ij} = p_{ij} / p_{ij}^* = (p_{ij} - q_{ij}) / p_{ij}^* = 1 - q_{ij} / p_{ij}^*$. Для реальной распределенной системы значение γ близко к единице, следовательно, в системе уравнений (1) возможно заменить значения p_{ij} значениями p_{ij}^* :

$$\alpha_i = \tau_i + \sum_{j=1}^s p_{ij}^* \alpha_j \quad (2)$$

Таким образом, погрешность, возникающая при укрупнении состояний системы, определяется значением $(1 - \gamma)$.

Обозначим: q – вероятность перехода системы из класса состояний функционирования в состояние перераспределения файлов, p – вероятность пребывания системы в классе состояний функционирования: $p = 1 - q$. Все состояния функционирования системы склеиваются в одно состояние функционирования.

Граф состояний полученной укрупненной системы имеет вид, показанный на рисунке 4.

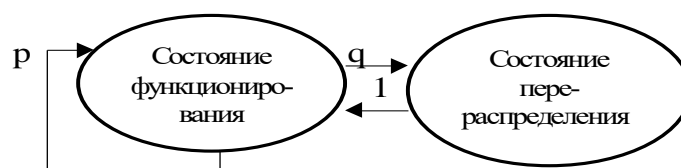


Рисунок 4 - Граф состояний укрупненной системы

Процесс переходов укрупненной системы является марковским с двумя состояниями. Переключения состояний управляются дискретной цепью Маркова с матрицей переходов:

$$\pi = \begin{vmatrix} p & q \\ 1 & 0 \end{vmatrix}$$

Основные характеристики укрупненной системы приближенно, с погрешностью $(1 - \gamma)$, могут быть приняты в качестве характеристик реальной системы. Поэтому время пребывания реальной системы в классе состояний функционирования приблизительно равно времени пребывания укрупненной системы в состоянии функционирования. Это время распределено показательно с параметром λ , где λ - интенсивность переходов системы из состояния функционирования в состояние перераспределения файлов. Пусть событием является переход системы из состояния функционирования в состояние перераспределения файлов.

На укрупненную систему, находящуюся в состоянии функционирования действует пуассоновский поток событий с интенсивностью λ , переводящий ее в состояние перераспределения файлов. Из состояния перераспределения файлов система переходит в состояние функционирования. Пусть значение t_a равно среднему времени перераспределения файлов и μ - интенсивность переходов системы из состояния перераспределения файлов в состояние функционирования, $\mu = 1/t_a$. Граф переходов укрупненной системы показан на рисунке 5.

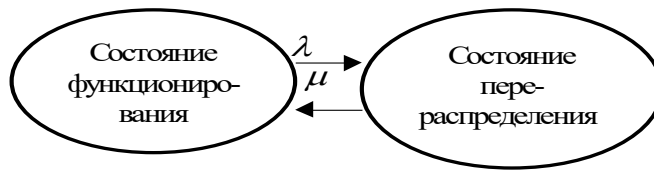


Рисунок 5 - Граф переходов укрупненной системы

Поведение вероятностей p_1 и p_2 состояний укрупненной системы описывается уравнениями Колмогорова [4]:

$$dp_1(t)/dt = p_2(t)\mu - p_1(t)\lambda; dp_2(t)/dt = p_1(t)\lambda - p_2(t)\mu; p_2(t) = p_1(t) \quad (3)$$

Пусть в начальный момент времени укрупненная система находилась в состоянии функционирования: $p_1(0)=1$ и $p_2(0)=0$. В таком случае решениями уравнений являются следующие выражения:

$$p_1(t) = \mu/(\lambda + \mu) + (\lambda + \mu)e^{-(\lambda + \mu)t}; p_2(t) = \lambda/(\lambda + \mu) - \lambda e^{-(\lambda + \mu)t}/(\lambda + \mu) \quad (4)$$

При большой величине t в укрупненной системе устанавливается стационарный режим: $p_1 = \mu/(\lambda + \mu)$, $p_2 = \lambda/(\lambda + \mu)$.

Обозначим: τ_i - среднее время пребывания системы в i -м состоянии функционирования, ρ_i - вероятность пребывания системы в i -м состоянии функционирования, q_i - вероятность перехода системы из i -го состояния функционирования в состояние перераспределения файлов.

Для определения значений T и D_T используются известные [4] формулы:

$$q = \sum_{i=1}^s \rho_i \cdot q_i; \lambda = q / \sum_{i=1}^s \rho_i \tau_i; T = 1/\lambda; D_T = 1/\lambda^2 \quad (5)$$

Пусть состояния функционирования системы равновероятны: $\rho_i = 1/s$, переходы системы из каждого состояния функционирования в состояние перераспределения файлов равновероятны: $q_i = 1/s$. В таком случае:

$$q = 1/s; T = s \sum_{i=1}^s \tau_i / s = \sum_{i=1}^s \tau_i; D_T = T^2. \quad (6)$$

Таким образом, среднее время T пребывания системы в классе состояний функционирования до перехода в состояние перераспределения файлов определяется по формуле:

$$T = \sum_{i=1}^s \tau_i \quad (7)$$

Обозначим: θ - среднее значение интервала между событиями с применением мониторингов. Длительность мониторинга много меньше, чем время T , поэтому для определения θ можно использовать известную [4] формулу: $\theta = T + D_T/T$. В таком случае:

$$\theta = 2T = 2 \sum_{i=1}^s \tau_i = 2/\lambda \quad (8)$$

Таким образом, если мониторинги применяются, то поток событий является потоком Эрланга 2-го порядка с интенсивностью Λ : $\Lambda = 1/\theta = \lambda/2$.

Определим оптимальную периодичность мониторингов. Обозначим: Ω - пауза между мониторингами, t_{mon} - время проведения мониторинга. Целевая функция при периодическом мониторинге состоит из двух частей. Во-первых, число мониторингов и время t_{mon} должны быть минимальными, чтобы задержка обработки

запросов к файлам, вызванная мониторингом, была минимальной. Для выполнения этого условия необходимо минимизировать величину t_{mon}/Ω . Во-вторых, на интервале θ нужно с минимальной задержкой опознать событие. Задержка опознания события будет минимальной, если оно возникло во время мониторинга. Величина этой задержки не превышает t_{mon} . Для минимизации этой задержки примем, что длительность мониторинга равна времени вычисления функции полезности сервера. Максимальная задержка опознания события будет, если событие возникло сразу по окончании мониторинга. Она больше значения t_{mon} на величину Ω . Для минимизации задержки опознания события время между мониторингами на интервале θ должно быть минимальным. Поэтому необходимо условие: $\Omega / \theta \rightarrow \min$. Таким образом, целевая функция задачи определения оптимального времени Ω между мониторингами: $E = t_{mon}/\Omega + \Omega / \theta \rightarrow \min$. Решив уравнение $dE/d\Omega=0$, найдем значение Ω , при котором величина E достигает минимума:

$$\Omega_{opt} = \sqrt{t_{mon} \cdot \theta} \quad (9)$$

Таким образом, с использованием метода фазового укрупнения сложных систем, по критерию минимизации задержки обработки запросов к файлам, вызванной мониторингом, определяется оптимальный период времени между мониторингами уровней локальных запросов к файлам.

Выводы

В процессе длительной работы компьютерной сети интенсивность запросов к файлам со временем изменяется, поэтому необходим мониторинг уровней запросов с целью определения момента перераспределения файлов. В статье предложен метод динамического размещения файлов по узлам компьютерной сети, позволяющий уменьшить трудоемкость размещения за счет оптимальной периодичности мониторингов уровней запросов к файлам. С использованием метода фазового укрупнения сложных систем определен оптимальный период времени между мониторингами уровней локальных запросов к файлам. Критерием оптимальности является минимизация задержки обработки запросов к файлам, вызванной мониторингом.

Литература

1. Бельков Д.В. Методы и вычислительные структуры для оптимизации размещения файлов в компьютерных сетях. Автореферат диссертации. Донецк: 2004. – 20 с.
2. Жуков И.А. Модель распределения информационных ресурсов в компьютерных сетях / И.А. Жуков, Ю.П. Зайченко, Н.К. Печурин // Проблемы информатизації тауправління. – К.: НАУ. - 2005. – Вип. 3 (14). – С. 9-14.
3. Демиденко О.М. Средства и технология организации мониторинга параметров вычислительного процесса и рабочей нагрузки на локальную вычислительную сеть // Проблемы управления и информатики. - 2002. - № 4. - С. 119 – 127.
4. Королук В.С. Фазовое укрупнение сложных систем / В.С. Королук, А.Ф. Турбин. – К : Вища школа, 1978. – 108 с.

Лазебная Л.А., Едемская Е.Н., Бельков Д.В. Динамическое распределение файлов по узлам компьютерной сети. В статье предложен метод динамического размещения файлов по узлам компьютерной сети, позволяющий уменьшить трудоемкость размещения за счет оптимальной периодичности мониторингов уровней запросов к файлам. С использованием метода фазового укрупнения сложных систем определен оптимальный период времени между мониторингами уровней локальных запросов к файлам. Критерием оптимальности является минимизация задержки обработки запросов к файлам, вызванной мониторингом.

Ключевые слова: узлы компьютерной сети, запросы к файлам, периодичность мониторингов, состояния системы.

Lazebnaya Ludmila, Edemskaya Eugeniya, Belkov Dmitriy Dynamic allocation of files on nodes of computer network. The article suggests the dynamic file allocation method for computer network nodes to reduce labor input properties due to the optimal frequency of monitoring levels of queries to files. Using the phase of consolidation of complex systems determined the optimum period of time between monitoring the levels of local requests to files. The optimality criterion is the minimization of delay in processing requests to files caused by monitoring.

Keywords: computer network nodes, file requests, monitoring frequency, system status.