

УДК 004.65

АНАЛИЗ ДАННЫХ НА ОСНОВЕ ПЛАТФОРМЫ SQL-MAPREDUCE

С.А. Алексеев

Научный руководитель – к.т.н., доцент А.К. Дорожкин

В работе рассматривается проблема построения современных аналитических платформ на основе новых способов интеграции существующих технологических решений в области хранения и анализа данных, позволяющих удовлетворять постоянно растущие требования как со стороны разработчиков программного обеспечения, ориентированного на анализ данных, так и со стороны организаций, заинтересованных в его использовании на практике.

Цель работы – определение эффективности подхода к расширению парадигмы, заложенной в основе реляционных систем управления базами данных (РСУБД), моделью распределенных вычислений MapReduce, используемой для параллельных вычислений над большими объемами данных в компьютерных кластерах.

Основой деятельности большинства организаций из различных сфер деятельности человека является анализ данных посредством сложных аналитических инструментов. От качества проводимого анализа во многом зависит эффективность деятельности всей организации в целом. На сегодняшний день рынок подобных инструментов широко представлен в виде набора готовых решений. Однако решения такого рода обычно сложно поддаются настройке под конкретные требования организации. В связи с этим в последнее время возрос интерес к построению систем анализа данных на основе готовых аналитических платформ, позволяющих на основе интегрированных в них решений быстро создавать новые аналитические инструменты, учитывающие все потребности конкретной организации. Такие аналитические системы представляют собой сложные программно-аппаратные комплексы, основным элементом которых являются средства хранения анализируемых данных. Согласно статистике, в качестве таких средств на сегодняшний день наибольшее распространение получили системы управления базами данных (СУБД), среди которых лидирующие позиции занимают реляционные системы управления базами данных (РСУБД). Поэтому в работе основное внимание было решено уделить рассмотрению вопроса построения современной аналитической платформы на основе развития концепции РСУБД.

На развитие таких средств оказывают влияние такие факторы как: стремительный рост объемов данных, повышенный спрос к качеству и скорости проведения анализа данных внутри организации, расширяющиеся требования пользователей и разработчиков, а также архитектурные изменения в области использования компьютеров [1].

В качестве основного вектора рассмотрения в работе приводится подробный анализ, направленный на определение вопроса применимости уже существующих проверенных решений для преодоления проблем рассмотренных выше. Поясняются причины, по которым структурированный язык запросов SQL и параллельные архитектуры СУБД не удовлетворяют требуемым показателям производительности по анализу данных. После этого формулируются основные направления, которые могли бы решить существующие ограничения.

Одним из наиболее эффективных решений в данной области является подход, предложенный компанией Teradata [2, 3]. Он заключается в расширении возможностей языка SQL посредством хорошо распараллеливаемых по модели MapReduce, предложенной компанией Google, самоопределяемых и полиморфных табличных функций, которые можно вызывать прямо из операторов выборки. Для обеспечения работы данного механизма

компанией была разработана инфраструктура под названием SQL-MapReduce, реализованная в SQL-ориентированной массивно-параллельной СУБД nCluster. Основная часть альтернативных решений, рассмотренных в работе, уступает решению, предложенному компанией Teradata по нескольким причинам: во-первых, из-за новизны данного направления эти проекты находятся лишь на стадии развития, во-вторых, некоторые решения содержат значительные ограничения синтаксиса языка SQL и обладают очень сложными механизмами интеграции и разработки решений на их основе.

Для выполнения практической части исследования была построена специальная модель, позволяющая поставить ряд экспериментов по применимости данной технологии в различных областях, требующих проведения анализа данных.

Детальное изучение данной области показало, что область анализа данных широко открыта для появления новых, более эффективных решений, одно из которых подробно рассматривается в данной работе. При этом немаловажно, чтобы такой «поиск» основывался на хорошей теоретической и практической базе с привлечением лучших специалистов из области информационных технологий.

Литература

1. Agrawal Rakesh, O'Reilly Tim, Stonebraker Michael. The Claremont Report on Database Research// Sigmod Record. – 2008. – V. 37. – № 3. – P. 9–19.
2. Eric Friedman, Peter Pawlowski, John Cieslewicz. SQL/MapReduce: A practical approach to self-describing, polymorphic, and parallelizable userdefined functions // Proceedings of the 35th VLDB Conference, Lyon, France. – 2009.
3. Rick F. van der Lans. Using SQL-MapReduce@for Advanced Analytical Queries. 2011. [Электронный ресурс]. – Режим доступа: http://www.asterdata.com/resources/assets/ar_SQL-MapReduce_for_Advanced_Analytics.pdf, своб.

УДК 004.934.8'1

АВТОМАТИЗАЦИЯ ПРОЕКТИРОВАНИЯ РЕЧЕВЫХ БАЗ ДАННЫХ

А.А. Алсуфьев

Научный руководитель – к.и.н, доцент П.В. Игнатов

Хорошо структурированные речевые базы данных (иначе – речевые корпуса) являются важным инструментом для развития речевых технологий, в частности для настройки и тестирования алгоритмов как в задачах распознавания речи, так и в задачах идентификации и верификации диктора, а также в фонетических исследованиях. Речевые базы данных для каждой из этих задач обладают своими особенностями, которые необходимо учитывать при проектировании каждого конкретного корпуса.

Создание качественного речевого корпуса является сложной и затратной задачей. Для упрощения данного процесса встает вопрос об автоматизации проектирования и создания речевых баз данных. На рубеже XX–XXI веков в фокусе внимания разработчиков и других заинтересованных специалистов оказались вопросы стандартизации методов представления данных, аннотаций и инструментария корпусных ресурсов [1]. Мы ставим своей задачей обобщить опыт проектирования речевых корпусов и предложить элементы автоматизации данного процесса, заключающиеся в первую очередь в разработке методики создания речевых баз данных.

При проектировании речевой базы данных необходимо решить ряд вопросов. Необходимо определить задачи, для решения которых планируется использовать каждый

конкретный речевой корпус. Требования, предъявляемые к составу текстового материала, могут иметь различный характер и обычно определяются конкретными задачами, для решения которых формируется речевая база данных [2]. Так, для задач текстозависимой верификации нужны корпуса, содержащие многократное произнесение небольшого количества парольных фраз относительно небольшим количеством дикторов (обычно не более 100) [3], для целей распознавания речи, напротив, необходимы записи заведомо большего количества дикторов, при этом каждый из них обычно должен иметь только по одной записи [2, 4], для речевых баз для создания систем синтеза речи обычно необходимо 1–5 дикторов [3].

После определения количества дикторов необходимо определить состав дикторов по таким параметрам, как пол, возраст, родной язык, диалект, так и по социальным критериям – образование, профессия, социальная группа [3, 5], а также уровень речевой компетенции. Также важно решить, сколько сессий записи будет иметь каждый диктор [6].

Следующим этапом должен стать подбор произносимого материала. В первую очередь определяют, будет ли записываться чтение или спонтанная речь [2, 7]. Для целей текстозависимой верификации, распознавания и синтеза речи необходимы записи чтения, для целей текстонезависимой идентификации/верификации – спонтанная речь. Необходимо осуществить подбор текстового материала (специализированный/репрезентативный, тип произносимых речевых образцов – слова, отдельные предложения, тексты, образцы спонтанной речи; фонетически сбалансированный/несбалансированный, тип балансировки, статистическая представительность звуковых единиц и т.п.) [2]. На практике также часто используются записи речи, собранной из теле- и радиозэфира [7].

Также необходимо решить, какая информация будет ассоциирована с речевым корпусом. Это может быть текстовая расшифровка, транскрипция, различные уровни разметки, информация о дикторе, информация об условиях записи, просодическая аннотация, другие [5]. Также отдельно делаются специальные пометки, указывающие на наличие возможных шумов, оговорок, обрывов записи [8]. Разметка проводится главным образом по 3 типам: фонемическая, фонетическая, орфографическая [7]. Возможно проведение разметки не на всем собранном материале: так, в речевой базе CORP-ORAL из 53 часов речи на орфографическом уровне было размечено 32 часа, а на фонетическом – лишь 1 час [9].

Подводя итог вышесказанному, мы можем заключить, что разработка подробной методики создания речевых баз данных является важнейшим шагом на пути к автоматизации процесса создания речевых баз данных.

Литература

1. Кривнова О.Ф. Речевые корпуса на новом технологическом витке // Речевые технологии. – 2008. – №2. – С. 13–24.
2. Богданов Д.С., Брухтий А.В., Кривнова О.Ф., Подрабинович А.Я., Строкин Г.С. Технология формирования речевых баз данных // Организационное управление и искусственный интеллект. Сборник трудов Института системного анализа РАН. – М.: Эдиториал УРСС, 2003/2004. – С. 239–259.
3. Кривнова О.Ф., Захаров Л.М., Строкин Г.С. Речевые корпуса (опыт разработки и использование) // Труды Международного семинара «Диалог-2001» по компьютерной лингвистике и ее приложениям (в двух томах). – М., 2001. – Т. 2.
4. Jonathan Harrington. The Phonetic Analysis of Speech Corpora. – New York: Wiley-Blackwell, 2010. – 424 p.
5. Богданов Д.С., Кривнова О.Ф., Подрабинович А.Я. Современный инструментарий для разработки речевых технологий // Информационные технологии и вычислительные системы. – М.: Эдиториал УРСС. – 2004. – № 2.

6. F. Alvin Martin. Encyclopedia of biometrics // National Institute of Standards and Technology. Gaithersburg, Maryland. – USA, 2009. – P. 1245–1253.
7. Pétur Helgason. Speech databases and speech corpora // Computer-Based Tools, 7,5hp, HT. – Uppsala University. – 2010.
8. Галунов В.И., Кочанина Ю.Л., Остроухов А.В. Речевые базы данных русского языка в рамках европейских программ SPEECHDAT – Auditech. Ltd [Электронный ресурс]. – Режим доступа: <http://auditech.ru/page/galunov.html>. – Загл. с экрана.
9. Fabíola Santos, Tiago Freitas. CORP-ORAL: a spontaneous European Portuguese speech resource // Propor: Applications of Portuguese Speech and Language Technologies. – 2008. – P. 42–45.

УДК 004.048

ПОДГОТОВКА ИСХОДНЫХ ДАННЫХ К ПРОВЕДЕНИЮ КЛАСТЕРИЗАЦИИ ДОКУМЕНТОВ

В.В. Белозёров

Научный руководитель – к.т.н., доцент А.К. Дорожкин

В работе рассматриваются трудности, возникающие при автоматической обработке текстовых документов на естественном языке, варианты их решения. Обобщается практический опыт ведущих университетов и научных центров мира, занимающихся компьютерной лингвистикой. Проводится сравнение английского и русского языков, следующие из этого изменения, которые необходимо вносить при использовании алгоритмов, разработанных для англоязычных документов.

Цель работы – выделение главных этапов автоматической обработки текстовых документов на естественном языке.

Лексический анализ – процесс преобразования массива символов (текст документа) в массив слов. Полученные слова будут использоваться в дальнейшем процессе обработки, например, для создания алфавитного указателя коллекции документов. В простейшем случае лексический анализ определяет символы пробела, как разделители между словами. При работе с реальными текстами такого подхода недостаточно, необходимо отдельно обрабатывать специальные символы, знаки препинания, цифры. Эти символы являются бесполезными терминами, поскольку сами по себе не имеют смысла, а зависят от контекста. С другой стороны эти символы, а чаще всего числа, могут использоваться в названиях и иностранных аббревиатурах (например, J2SE или GT2835R). Соответственно в таких случаях термины разделить нельзя.

В каждом документе только небольшой процент слов действительно имеет значение, остальные 80–90% слов практически бесполезны для проведения анализа над текстом. Такие слова называются стоп-словами и исключаются из алфавитного указателя. По умолчанию к этой категории относятся предлоги, союзы, частицы. В то же время знаменитая фраза «Быть или не быть?» полностью состоит из стоп-слов, поскольку «быть» является наиболее часто используемым глаголом в русском языке, точно также обстоят дела и с оригиналом на английском. Существует большое количество стоп-листов для английского языка, однако поиск подобных списков на других языках является более сложной задачей, требующей самостоятельного поиска стоп-слов.

При формировании своих поисковых запросов не только пользователи могут ошибиться, но еще и возможно различное написание тех или иных слов. Но все же самая частая причина использования проверки орфографии – незнание того, как правильно пишутся некоторые слова.

Как правило, в Интернете встречаются и неверные варианты, поэтому задачей системы является подсказать пользователю корректное написание.

Исключение стоп-слов, стемматизация и проверка орфографии предполагают, что нам уже известен язык, на котором написан обрабатываемый документ. Такое предположение справедливо для небольших коллекций документов, но в общем случае язык неизвестен, более того, в одном документе могут встречаться вставки на различных языках. Если же говорить о крупных коллекциях документов, например, о всемирной паутине, то это – сочетание десятков различных языков. Очевидно, что необходимо использовать алгоритмы для автоматического определения языков.

Некоторые пользователи поисковых систем вместо использования поисковых запросов, использования соответствующего синтаксиса, задают обычные вопросы. Этот тип поисковых запросов направлен на документы, содержащие ответ на конкретный вопрос, а не на страницы с несколькими ключевыми словами. Оптимальным решением было бы выводить вместо целого документа только ответ на вопрос.

Этот процесс очень сложный, вопросы зависимы от контекста, вкладываемого смысла. Обработка естественного языка (NaturalLanguageProcessing, NLP) – направление искусственного интеллекта.

На данный момент из крупных поисковых систем, использующих NLP, работает лишь система Start, увидевшая свет в 1993 году.

В работе выделены следующие проблемы автоматической обработки документов:

- неверное написание запросов пользователем, а также неверное написание слов в тексте;
- различные вариации написания одного и того же слова;
- развитая морфология некоторых языков;
- зависимость от контекста;
- наличие омонимов.

Выделены следующие этапы автоматической обработки документов:

- лексический анализ;
- стемматизация и лемматизация;
- проверка орфографии;
- создание стоп-листов;
- определение языка;
- обработка естественного языка.

УДК 65.011, 519.256, 004.043

ОПТИМАЛЬНАЯ ДИСЦИПЛИНА УПРАВЛЕНИЯ ОЧЕРЕДЬЮ В СМО С ОГРАНИЧЕННОЙ ПРОИЗВОДИТЕЛЬНОСТЬЮ

В.Ю. Боркунов

Научный руководитель – к.ф.-м.н., доцент Д.А. Зубок

Задача оптимального управления системой без потерь с ограниченными ресурсами по производительности и ограниченной очередью решается методами ТМО, связанными с определением пороговой дисциплины обслуживания [1] для критерия минимизации времени обслуживания требования. Такие задачи тяжело поддаются аналитическому исследованию и часто исследуются средствами имитационного моделирования и численными методами [2]. Применение методов управляемых систем массового обслуживания [3] позволяет утверждать существование оптимальных стратегий управления для широкого класса систем. Достаточно полные результаты для систем с фиксированным числом приборов получены в [4–6].

Цель работы – определение оптимального управления очередью в системе с групповой

обработкой требований и ограниченной производительностью.

Дадим описание анализируемой системы. На вход в систему подается пуассоновский поток требований с определенной интенсивностью. В системе предусмотрена очередь ограниченной длины. Дисциплина обслуживания: требования из очереди принимаются на выполнение в том случае, если система свободна или очередь полностью заполнена в момент поступления очередного требования в систему. Производительность системы зависит от числа обрабатываемых требований.

Требуется определить влияние величины очереди на среднее время выполнения поступающих требований и определить минимально возможное пороговое значение очереди, минимизирующее среднее время пребывания требования в системе. Для выходящего потока – моментов времени окончания обслуживаний – проанализировать распределение случайной величины – времени обслуживания.

При пуассоновском характере входящих требований, моменты времени окончания обслуживания требований – «выходящий поток» – уже не образуют стационарный процесс. Параметры выходящего потока зависят от интенсивности входящего потока и закона падения производительности при групповой обработке требований.

В данной работе в среде имитационного моделирования AnyLogic построена модель системы и проведены имитационные эксперименты для некоторой группы параметров, описывающих систему. Модель построена с использованием стандартных объектов и классов библиотеки EnterpriseLibrary пакета AnyLogicProfessional 6.4.1. Реализованы параметрические статистические оптимизационные эксперименты. Проводилось усреднение по серии более 100 экспериментов для каждого набора параметров. Вычислялось минимальное значение очереди по критерию минимального среднего времени выполнения требований в системе с учетом пребывания требований в очереди.

Основной результат работы: при заданном законе падения производительности системы в зависимости от интенсивности входящего потока определены пороговые значения очереди, минимизирующее функцию критерия – среднее время выполнения требований. Проведена статистическая оценка дисперсии процесса в зависимости от длины очереди – дисперсия случайной величины также минимизируется, начиная с того же порогового значения очереди. При длине очереди больше пороговой среднее число одновременно выполняемых требований в системе не зависит от длины очереди. Статистическая оценка ковариационной функции дает основание предполагать, что выходящий поток является нестационарным. Фактически рассматриваемая система коррелирует события входящего потока на этапе их выполнения.

Установлено, что при некоторых значениях интенсивности входящего потока, введение очереди является фактором, стабилизирующим фактическую производительность системы.

Литература

1. Рыков В.В., Ефросинин Д.В. К анализу характеристик производительности СМО с неоднородными потоками // Автоматика и телемеханика. – 2008. – № 1. – С. 64–82.
2. Рыков В.В., Вербицкий С.Н. Численное исследование оптимальных политик управления скоростью обслуживания // Автоматика и телемеханика. – 1998. – № 11. – С. 59–70.
3. Висков О.В., Ширяев А.Н. Об управлениях, приводящих к оптимальным стационарным режимам // Сборник работ по теории вероятностей, Тр. МИАН СССР, 71. – М.: Наука, 1964. – С. 35–45.
4. Миллер Б.М., Миллер Г.Б., Семенихин К.В. Методы синтеза оптимального управления марковским процессом с конечным множеством состояний при наличии ограничений // Автоматика и телемеханика. 2011. – № 2. – С. 111–130.
5. Rykov V. Monotone control of queueing systems with heterogeneous servers // Queueing systems. – 2001. – № 37. – P. 391–403.

6. Rykov V., Efrosinin D. Optimal control of queueing systems with heterogeneous servers // Queueing systems. – 2004. – № 46. – P. 389–407.

УДК 004.383.8.032.26

РАЗРАБОТКА ВЫСОКОПАРАЛЛЕЛЬНОЙ ОБУЧАЮЩЕЙСЯ НЕЙРОННОЙ СЕТИ НА ПЛИС

К.А. Булыгин

Научный руководитель – д.т.н., профессор А.Е. Платунов

Постановка проблемы. Многие теоретические и прикладные задачи можно свести к определению функции (в той мере, которая достаточна для ее использования) вектора характеристик системы от вектора ее параметров по известным точкам этой функции и сведениям о ее поведении (например, выведение научной теории из множества фактов, прогнозирование погоды, задачи оптимизации и т.д.). Обычно такие функции настолько сложны, что не поддаются формализации не только на алгоритмическом уровне, но и на уровне словесного описания: в этих случаях говорят об интуиции эксперта как о механизме реализации такой функции. Однако формирование знаний эксперта можно промоделировать с помощью нейронных сетей.

Большинство существующих нейросетевых систем имеют программную реализацию на вычислительных системах с фоннейманновской архитектурой, что существенно ограничивает скорость их работы и сокращает возможности использования во встроенных вычислительных системах. Опубликованные аппаратные реализации нейросетей в определенной мере следовали общей тенденции к повторению последовательного характера работы программных реализаций, что упрощает их создание и понимание, но приводит к простому используемых аппаратных ресурсов.

Цель работы: создание обучающейся нейронной сети на ПЛИС с минимальным простым аппаратных ресурсов.

Базовые положения исследования. Указанный простой можно сократить, во-первых, за счет высокой степени параллелизации вычислительного процесса и, во-вторых, за счет детальной параметризации нейросети для конкретной задачи (с помощью задания числа слоев нейросети, числа нейронов в каждом слое, разрядности всех типов данных, влияющей прежде всего на разрядность АЛУ и шин данных и т.д.). В качестве альтернативы подходу прямого отображения многослойной нейросети на аппаратные блоки (слишком затратного с точки зрения аппаратных ресурсов, в особенности на ПЛИС, обеспечивающих ограниченную связность между блоками) используется подход с одним аппаратным слоем нейронов, что обеспечивает линейную (теоретически наилучшую в данном случае) зависимость длительности эпохи от числа слоев и линейную зависимость числа аппаратных элементов от числа нейронов в одном слое.

Промежуточные результаты. Спроектирована однослойная и многослойная нейронные сети. Для обоих типов сетей пользователем задаются бинарные векторы, являющиеся произвольными строками непротиворечивой таблицы истинности (т.е. конкатенациями вектора слов x и слов $y_1(x), \dots, y_M(x)$), и при сходимости процесса возвращается вектор синаптических весов каждого нейрона $w_i(y_i(x))=w_i x$. Строго говоря, задачи не обязательно должны быть линейно разделимыми: например, задача XOR решается одним нейроном с помощью переполнения (хотя в качестве основной используется

арифметика с насыщением).

Для реализации однослойной нейросети предлагается следующая организация. Основные блоки (управляющее устройство, слой нейронов, память и адаптивный тактогенератор) работают в одном из общесистемных режимах работы («перезагрузка», «выборка», «вычисление ответа», «проверка», «корректировка» по дельта-правилу), разделенных на два подрежима, соответствующих тактам («инициация» – с минимальной длительностью, и «выполнение» – с длительностью, зависящей от текущего режима и задаваемой на основе временного симулирования).

В многослойной нейросети для обучения используется алгоритм обратного распространения ошибки. Состав основных аппаратных блоков такой же, но вследствие того что при обратном проходе каждый нейрон использует веса и значения ошибок других нейронов, блок «слой нейронов» концентрирует информацию о нейронах и распределяет ее по отдельным блокам нейронов, являющихся идентичными АЛУ для прямого и обратного прохода. Значения сигма-функции для параллельного доступа хранятся в идентичных блоках памяти, связанных с нейронами (используется арифметика с фиксированной точкой).

В настоящий момент реализована однослойная и реализуется многослойная нейросеть на языке Verilog.

Практические результаты. В результате была спроектирована и проверена идея работы высокопараллельной аппаратной нейросети с линейной зависимостью длительности эпохи от числа слоев и с линейной зависимостью числа элементов от числа нейронов в одном слое. Нейронную сеть планируется использовать в качестве IP-блока, в том числе в проекте по распознаванию протоколов с помощью логического анализатора.

УДК 004.051

СРАВНИТЕЛЬНЫЙ АНАЛИЗ МЕТОДОВ ПРОГНОЗИРОВАНИЯ НАГРУЗКИ НА ИТ-СЕРВИСЫ

Е.В. Быстрова

Научный руководитель – к.т.н., доцент Н.Ф. Гусарова

В настоящее время информационные технологии стали неотъемлемой частью практически любого предприятия – как коммерческих, так и некоммерческих организаций. Поддержание ИТ-инфраструктуры в актуальном состоянии является одним из важнейших критериев для успешного осуществления деятельности предприятия. Для организаций, основной род деятельности которых связан с предоставлением ИТ-услуг, данный фактор является критическим, так как влияет на качество услуг, и, как следствие, на имидж и финансовое состояние организации.

В ситуации быстро меняющихся условий рынка и постоянно развивающихся технологий руководителям важно вовремя получать информацию о том, стоит ли приобретать дополнительное оборудование или же состояние инфраструктуры удовлетворяет текущим потребностям.

Поэтому особое внимание стоит уделять вопросам производительности и нагрузки на ИТ-сервисы, применяемые как для внутренних нужд организации, так и с целью предоставления ИТ-услуг потребителям.

Для решения подобных задач в различных сферах деятельности (медицина, экономика, электрические сети) применяется широкий спектр методов прогнозирования.

Все методы прогнозирования можно разделить на две группы: интуитивные и формализованные.

Интуитивные методы прогнозирования не предполагают разработку моделей прогнозирования и отражают индивидуальные суждения специалистов (экспертов) относительно перспектив развития процесса.

Для исследования проблемы прогнозирования нагрузки на ИТ-сервисы данная группа методов не рассматривается, так как предполагается, что сервисы введены в эксплуатацию и по результатам функционирования доступны статистические данные для последующего анализа. За основу берется вторая группа методов – формализованные методы.

Формализованные методы рассматривают модели прогнозирования.

Модели прогнозирования разделяются на статистические модели и структурные модели.

В статистических моделях функциональная зависимость между будущими и фактическими значениями временного ряда, а также внешними факторами задана аналитически. К статистическим моделям относятся следующие группы:

- регрессионные модели;
- авторегрессионные модели;
- модели экспоненциального сглаживания.

В структурных моделях функциональная зависимость между будущими и фактическими значениями временного ряда, а также внешними факторами задана структурно. К структурным моделям относятся следующие группы:

- нейросетевые модели;
- модели на базе цепей Маркова;
- модели на базе классификационно-регрессионных деревьев.

Проблему прогнозирования нагрузки на ИТ-сервисы предлагается рассмотреть с нескольких точек зрения, построив различные модели (статистические и структурные). Также планируется на основе полученных результатов провести сравнительный анализ методов прогнозирования, исследование качества выбранных моделей, адекватность их реальному процессу и выбор лучшей из моделей.

Для составления прогноза планируется провести анализ производительности, взяв за основу следующие показатели:

- время обработки запроса сервером;
- время отклика сервера.

В качестве факторов, в зависимости от которых меняется значение показателей, рассматриваются следующие пункты:

- количество активных пользователей сервиса;
- количество запросов в единицу времени (например, в секунду);
- объем информации в единицу времени (например, КВ/сек).

Полученные результаты прогнозов необходимо сравнить с данными, которые будут получены при проведении эксперимента – нагрузочного тестирования рассматриваемых сервисов. В качестве инструмента для проведения тестирования предполагается использовать ApacheMeter. Один из примеров графиков, получаемых в результате проведения тестирования, представлен на рисунке – он отражает зависимость времени отклика от числа активных пользователей для двух различных запросов.

В качестве полигона для эксперимента и источника данных для составления прогноза предполагается рассматривать Web-сервисы. Результаты прогноза и тестирования помогут определить значения, при которых система перестанет соответствовать ожидаемым значениям определенных показателей нагрузки. Подобная информация позволит вовремя принимать управленческие решения – в частности решения, связанные с модификацией текущей ИТ-инфраструктуры.

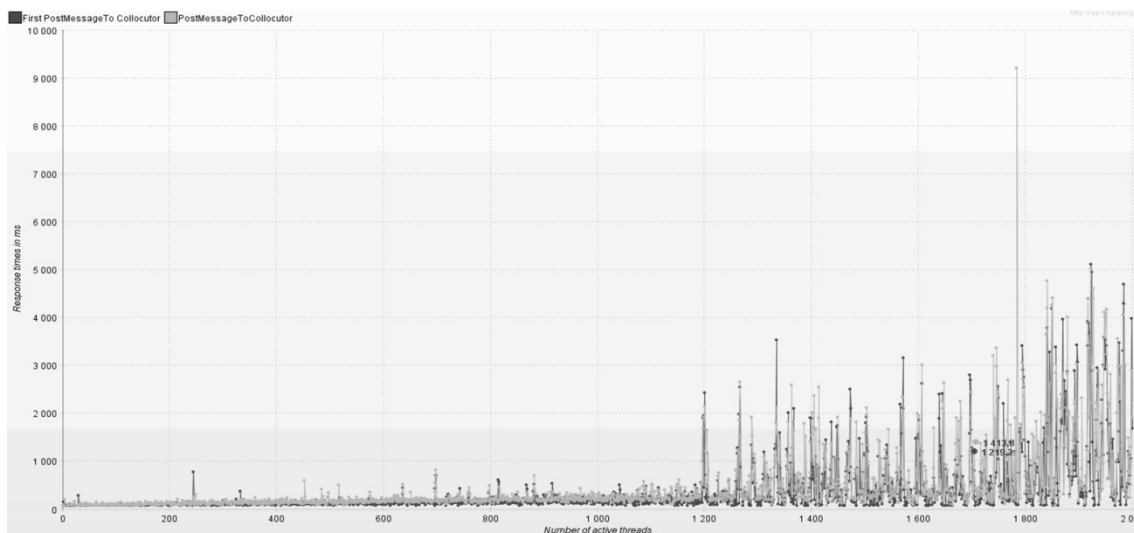


Рисунок. Время отклика (в миллисекундах) относительно количества активных пользователей (2000 пользователей)

УДК 004.91

АВТОМАТИЗАЦИЯ РАЗРАБОТКИ ПРОДУКТОВОЙ ДОКУМЕНТАЦИИ НА ИННОВАЦИОННОМ ПРЕДПРИЯТИИ

М.А. Вишнякова-Садовская

Научный руководитель – к.т.н. М.В. Хитров

Решаемая проблема. Разработка технической документации предполагает своевременное и качественное создание документации в условиях постоянного развития компании. При разработке технической документации неавтоматизированными, а стандартными офисными средствами этот процесс становится наиболее дорогостоящей и трудоемкой частью проекта. Автоматизированное документирование способно избавить технического писателя от части трудоемкой, единообразной работы: среда разработки – MicrosoftOffice – создает шаблон будущей документации на основе реализуемого проекта, что упрощает процесс ее создания. Язык разработки – VisualBasicforApplications.

Основной **целью работы** является алгоритм (программа) автоматизации процесса создания технической документации. Это позволит: повысить управляемость жизненного цикла технической документации; минимизировать рутинность, трудоемкость и ресурсоемкость в процессе создания документации; сократить сроки разработки и поддержания в актуальном состоянии технической документации; повысить качество документации; снизить влияние «человеческого фактора» на конечный результат.

Нормативные документы, применяемые при разработке технической документации, а также большинство создаваемых документов взаимосвязаны, повторяются по структуре и содержанию, имеют типовое наполнение, взаимоувязанное на уровне структурных единиц – пунктов, подпунктов, абзацев, разделов, подразделов. Изученные особенности являются предпосылками использования автоматизированного алгоритма по созданию технической документации, который основан на принципе «единого шаблона».

Функциональная схема программы. При автоматизации процесса создания единой технической документации выполняются следующие задачи: выделение типовой информации (в дальнейшем данная информация будет добавлена в текст); временное хранение полученной информации; подстановка выделенной информации в

документированные файлы в формате .docx; предоставление конечной документации пользователю. Такой процесс назовем шаблонным: в основу документации ложится шаблон, который в дальнейшем предоставляет разработчику (техническому писателю) применять его в процессе создания типового комплекта документации.

В результате работы было составлено техническое задание на программу AutoDoc по ГОСТ 19.201-78. Настоящий стандарт устанавливает порядок построения и оформления технического задания на разработку программы AutoDoc для вычислительных машин, комплексов и систем независимо от их назначения и области применения.

Требования к составу выполняемых функций. Промежуточные результаты.

Программа должна обеспечивать возможность выполнения перечисленных ниже функций:

- разработка комплекта технической документации;
- создание нового (пустого) файла (документа);
- открытие (загрузка) существующего файла из текстового процессора MicrosoftWord;
- редактирование открытого (далее – текущего) файла путем ввода, замены, удаления содержимого файла с применением стандартных устройств ввода;
- редактирование текущего документа с применением буфера обмена операционной системы;
- автоматическое формирование имени файла ЦВАУ.00XXX-0X_0X_0X в формате .docx;
- сохранение документа с исходным именем;
- сохранение документа с именем, отличным от исходного;
- автоматическое оформление выходных документов в соответствии с требованиями ГОСТ 19 и 34;
- автоматическое формирование структуры (состава) выходных документов в соответствии с требованиями ГОСТ 19.101-77;
- профилирование (отбор) текста (общей части текста) для формирования выходного документа по различным признакам, включая аппаратные платформы, операционные системы, язык программирования, технические средства и т.д.
- автоматизированное формирование оглавлений и перекрестных ссылок как внутри документа, так и между документами;
- автоматическое формирование титульных листов и листов утверждений документов по ГОСТ 19 и 34;
- формирование выходных документов в формате .docx;
- конвертация документа в формате .pdf (AdobeAcrobat);
- публикация технической документации на (бумажных носителях (печать документа), в электронном виде (сохранение документа));
- модификация (внесения изменений в текущий документ), отслеживание изменений технической документации – сопровождение;
- обмен технической документацией между подразделениями компании (SharePoint) (из состава функций MicrosoftWord);
- передача технической документации (pdf, docx) заказчику (или конечному пользователю) электронной почтой с помощью клиентской почтовой программы (MicrosoftOutlook – из состава функций MicrosoftWord);
- вывод оперативных справок (подсказок) в строковом формате;
- интерактивная справочная система;
- отображение названия программы, версии программы, копирайта и комментариев разработчика.

Требования к организации входных и выходных данных. Входные данные

программы планируется организовать в виде технического задания формата .docx, соответствующего спецификации. Файлы указанного формата должны размещаться (храниться) на локальных или съемных носителях, отформатированных согласно требованиям операционной системы.

Выходные данные программы должны быть организованы в виде комплекта документов (заготовки, шаблоны по ГОСТ 19 и 34, которые потом должны заполняться; автоматически создается спецификация, титульный лист, лист утверждения). В шаблонах автоматически формируются заголовки, подписи к рисункам, названия, имя файла – ЦВАУ.00XXX-0X_0X_0X, литера (Э, О и т.д.).

Форматы выходных документов для работы, хранения и печати:

- – pdf (AdobeAcrobat);
- – docx (MicrosoftWord 2007, 2010).

Основной результат. В результате работы был частично реализован алгоритм автоматизации процесса создания технической документации. Результаты тестов его производительности и эффективности будут представлены в форме графика. Работа над проектом продолжается.

УДК 004.912

СИСТЕМА ИЗВЛЕЧЕНИЯ ВЗВЕШЕННОГО СОДЕРЖИМОГО ИЗ ТЕКСТОВЫХ ДОКУМЕНТОВ

В.А. Власов

(Тульский государственный университет)

Научный руководитель – д.т.н., профессор Ф.А. Данилкин

(Тульский государственный университет)

В современном мире, в условиях ежедневного обращения существенных больших объемов информации, становится важной задача автоматического структурирования этой информации в соответствии с ее семантическим содержанием, включающим выделение ключевых слов для анализируемых текстов, и последующего распределения текстов по степени их схожести.

Извлечение текстового содержимого из документов офисных форматов может основываться на автоматизации внешних программных решений, соответствующих целевым форматам файлов, однако подобный подход не обеспечивает достаточной эффективности в силу разветвления зависимостей СОМ-объектов и отсутствия переносимости между платформами. Исходя из этого, решение задачи извлечения взвешенного содержимого из текстовых документов основывается на осуществлении взаимодействия с целевыми файлами на уровне двоичного определения информационных полей согласно внутреннему формату хранения данных.

Наиболее распространенными форматами хранения электронных документов являются WCBFF, OOXML, OpenOffice и Postscript, которые ассоциированы с файловыми расширениями doc, docx, odt и pdf соответственно. Электронные документы указанных форматов осуществляют хранение наибольшего процентного объема текстовых информационных полей на общий объем хранимых данных, соответственно извлечение из них текстового содержимого позволяет осуществить качественный семантический анализ на основе обработки естественного языка.

Извлечение простых групп текстовых информационных полей из оригинальных файловых объектов рассмотренных форматов является недостаточным для выполнения

полного семантического анализа, поскольку оригинальный электронный документ обеспечивает хранение такой информации о маркировании отдельных слов и параграфов, как кегль и стиль гарнитуры шрифта, которая указывает степень их значимости. Использование информации о маркировании и размещении текстовых информационных полей позволяет осуществить получение взвешенного содержимого текстового документа.

Внутренним форматом хранения электронных документов на время процесса извлечения текстовых информационных полей выбран OOXMLDocument, позволяющий, с одной стороны, включить все необходимые данные о форматировании и маркировании элементов, совместимые с остальными форматами электронных документов, и с другой стороны, являющийся классическим XML-документом.

Организация извлечения взвешенного содержимого из текстовых документов форматов WCBFF, OpenOffice и Postscriptdocument обеспечивается на основании предварительного преобразования содержательных групп информационных полей, в соответствии с правилами разметки OOXMLDocument, осуществляемого в виде последовательного процесса формирования ассоциированных элементов текстовых и интегрируемых сущностей.

Процесс конвертации электронных документов в базовый формат предполагает осуществление сопоставления кодовых таблиц групп информационных полей в соответствии с первичным повышенным уровнем базовой элементной трактовки. Производится сопоставление уровней хранения информационных полей для кодовых таблиц US-ASCII, UTF-8, CP1251, CP1252, UTF-16LE/Unicode, базирующейся на создании массивов соответствия байтов и средств интернационализации и локализации целевой среды.

Результатом функционирования системы извлечения взвешенного содержимого из текстовых документов является список слов естественного языка, для которых осуществлено взвешивание относительно содержащего их документа, основанного на визуальном и позиционном выделении ассоциированных групп информационных полей. Исходя из полученного списка взвешенных слов, возможно осуществление дальнейшего анализа оригинальных документов средствами обработки естественного языка.

С использованием приведенной в докладе методики и алгоритмов, авторами осуществляется проект по созданию системы конфигурирования текстовой информации на персональном компьютере FileComfort.

Литература

1. Microsoft Corporation // Windows Compound Binary File Format Specification [Электронный ресурс]. – Режим доступа: <http://download.microsoft.com/download/0/B/E/0BE8BDD7-E5E8-422A-ABFD-4342ED7AD886/WindowsCompoundBinaryFileFormatSpecification.pdf> (дата обращения 22.02.13).
2. Microsoft Corporation // MICROSOFT OFFICE WORD 97-2007 BINARY FILE FORMAT SPECIFICATION [* .doc] [Электронный ресурс]. – Режим доступа: [http://download.microsoft.com/download/0/B/E/0BE8BDD7-E5E8-422A-ABFD-4342ED7AD886/Word97-2007BinaryFileFormat\(doc\)Specification.pdf](http://download.microsoft.com/download/0/B/E/0BE8BDD7-E5E8-422A-ABFD-4342ED7AD886/Word97-2007BinaryFileFormat(doc)Specification.pdf) (дата обращения 22.02.13).
3. Microsoft Corporation // [MS-DOC]: Word Binary File Format (.doc) Structure Specification [Электронный ресурс]. – Режим доступа: <http://download.microsoft.com/download/2/4/8/24862317-78F0-4C4B-B355-C7B2C1D997DB/%5BMS-DOC%5D.pdf> (дата обращения 22.02.13).

ИССЛЕДОВАНИЕ ПОГРЕШНОСТИ АЛГОРИТМА СИГМА-ТОЧЕЧНОГО ФИЛЬТРА КАЛМАНА ПРИ ВОССТАНОВЛЕНИИ ОГИБАЮЩЕЙ КВАЗИГАРМОНИЧЕСКОГО СИГНАЛА

Е.Л. Дмитриева, М.А. Волынский

Научный руководитель – д.т.н., профессор И.П. Гуров

Бесконтактные методы исследования и контроля объектов применяются во многих отраслях науки и техники. Интерферометрические методы являются наиболее точными методами бесконтактных измерений геометрических характеристик объектов [1, 2].

Квазигармонические сигналы, получаемые в интерферометрических системах, содержат полезную информацию о свойствах исследуемых объектов. Одной из характеристик объекта является положение максимума огибающей сигнала. Значение огибающей в максимуме характеризует степень отражения от границы слоя среды. При сканировании по глубине многослойной среды формируются пики огибающей интерференционных полос малой когерентности, положение которых соответствует границам слоев. Для извлечения полезной информации об исследуемых объектах требуются быстродействующие помехоустойчивые алгоритмы обработки сигналов [3]. Особый интерес представляет класс алгоритмов на основе фильтра Калмана [1–4]. Одним из таких алгоритмов является сигма-точечный фильтр Калмана (UnscentedKalmanfilter) [5], свойства которого, применительно к восстановлению параметров квазигармонического сигнала, исследованы в данной работе.

Сигма-точечный фильтр Калмана использует подход, состоящий в замене вектора параметров на набор сигма-точек, покрывающий истинные математическое ожидание и ковариацию случайной величины, подвергшейся нелинейному преобразованию. Усреднение значений сигма-точек позволяет производить устойчивую оценку параметров системы [5].

В результате исследований, проведенных в работе, реализован алгоритм сигма-точечного фильтра Калмана для восстановления параметров квазигармонических сигналов. Исследована погрешность алгоритма сигма-точечного фильтра Калмана при восстановлении огибающей квазигармонического сигнала.

Исследованный в работе алгоритм обработки данных применим для восстановления информации о структуре сред в динамическом режиме. В работе показано, что рассмотренный метод обеспечивает высокую разрешающую способность и позволяет разрешать пики огибающей, расстояние между которыми значительно меньше длины когерентности. Недостаток этого метода состоит в сильной зависимости погрешности предсказания параметров сигнала от начальных условий.

Литература

1. Волынский М.А., Захаров А.С. Анализ интерференционных сигналов малой когерентности на основе моделей линейной и нелинейной стохастической фильтрации // Научно-технический вестник СПбГУ ИТМО. – 2007. – №37. – С. 4–9.
2. Васильев В.Н., Гуров И.П. Компьютерная обработка сигналов в приложении к интерферометрическим системам. – СПб: БХВ Санкт-Петербург, 1998. – 240 с.
3. Волынский М.А., Гуров И.П., Захаров А.С. Нелинейная стохастическая фильтрация сигналов в интерферометрах с частично когерентным освещением // Научно-технический вестник СПбГУ ИТМО. – 2007. – № 43. – С. 276–283.
4. Kalman R.E. A New Approach to Linear Filtering and Prediction Problems // Transactions of the ASME – Journal of Basic Engineering. – 1960. – № 82. – С. 35–45.

5. Julier S.J., Uhlmann J.K. A New Extension of the Kalman Filter to Nonlinear Systems // Defense Sensing, Simulation and Controls. – 1997. – V. 3068. – P. 182–193.

УДК 65.011, 519.256, 004.043

ОПТИМАЛЬНАЯ ДИСЦИПЛИНА УПРАВЛЕНИЯ ОЧЕРЕДЬЮ В СМО С ВЫСОКОЙ НАДЕЖНОСТЬЮ

А.Ю. Ефимов

Научный руководитель – к.п.н., доцент А.В. Маятин

Обеспечение надежности хранения данных является приоритетной задачей управления информационно-технологической инфраструктурой. Один из распространенных методов повышения надежности хранения данных заключается в их резервном копировании из основной системы хранения данных (СХД) в другие хранилища. В связи с большими объемами и постоянным ростом объемов данных, как правило, применяется инкрементное резервное копирование данных через равные промежутки времени. Объем копируемой информации в этом случае определяется интенсивностью требований на запись в период, прошедший с предыдущего резервного копирования.

Высокая стоимость СХД и телекоммуникационной инфраструктуры требует обеспечения высокой эффективности использования ресурсов, поэтому при проектировании центра обработки данных конфигурация СХД подбирается таким образом, чтобы одновременно минимизировать простой оборудования и обеспечить минимальное среднее время выполнения требований на чтение и запись данных при ожидаемой интенсивности этих требований. Производительность СХД существенно зависит от количества одновременно выполняющихся операций чтения и записи данных и определяется ее аппаратной конфигурацией. Операции резервного копирования в этом случае приводят к значительному снижению производительности СХД в связи с одновременным появлением группы требований на резервирование данных, выполнение которых приводит к превышению оптимального количества одновременно выполняющихся операций в СХД.

Для сохранения оптимального режима работы СХД в случае резкого возрастания интенсивности требований во время выполнения резервного копирования используется их буферизация. В терминах ТМО такой буфер является очередью. **Цель работы** состоит в определении оптимального управления очередью требований на чтение, запись и резервирование данных в условиях ограниченной производительности прибора (СХД).

Дадим описание анализируемой системы. Система обладает двумя состояниями: операционное состояние и нагруженное состояние. Для операционного состояния характерны два потока требований, поступающих в систему: поток требований на чтение и поток требований на запись данных. СХД сконфигурирована таким образом, что в операционном состоянии среднее время выполнения требования стремится к номинальному времени выполнения соответствующей операции в СХД. Для нагруженного состояния добавляется третий поток требований на резервирование. Изменение состояния с операционного на нагруженное происходит с заданным временным интервалом. Изменение состояния в обратном направлении происходит по окончании выполнения требований на резервирование и накопившихся в очереди требований на чтение и запись и возвращения возросшего в процессе резервного копирования среднего времени выполнения требований к времени, характеризующему операционное состояние.

Система обладает очередью определенной длины. Дисциплина обслуживания требований следующая: требования от трех возможных источников поступают в очередь. Если СХД не занята выполнением других требований, то требование берется на выполнение.

В случае заполнения очереди требование из очереди поступает на выполнение одновременно с уже выполняющимися требованиями. Производительность системы снижается при росте числа одновременно обрабатываемых требований.

Требуется определить влияние величины очереди на время возвращения системы к операционному состоянию после выполнения резервного копирования и определить минимально возможное пороговое значение очереди, минимизирующее среднее время нахождения системы в нагруженном состоянии.

В работе в среде имитационного моделирования AnyLogic построена модель системы и проведены имитационные эксперименты для некоторой группы параметров, описывающих систему. Модель построена с использованием стандартных объектов и классов библиотеки EnterpriseLibrary пакета AnyLogicProfessional 6.4.1. Реализованы параметрические статистические оптимизационные эксперименты. Проводилось усреднение по серии более 100 экспериментов для каждого набора параметров. Вычислялось минимальное значение длины очереди по критерию минимального среднего времени между переходом системы из операционного состояния в нагруженное и возвращения в операционное состояние с учетом пребывания требований трех типов, при этом интенсивность пребывания требований на резервирование выше, чем двух других типов.

Основной результат работы: при заданном законе падения производительности системы в зависимости количества одновременно выполняемых требований и заданных интервалах между операциями резервного копирования определены пороговые значения очереди, минимизирующие функцию критерия – среднее время между изменением состояний системы (с операционного на нагруженный и обратно). При длине очереди больше пороговой среднее время между изменением состояний системы не зависит от длины очереди. Таким образом введение очереди (буфера) заданной длины позволяет уменьшить негативное влияние инкрементного резервного копирования на производительность СХД.

УДК 004.414

АВТОМАТИЗАЦИЯ ПРОЦЕССА УПРАВЛЕНИЯ ТРЕБОВАНИЯМИ ПРИ РАЗРАБОТКЕ ПО

М.А. Залилов

Научный руководитель – П.Г. Гордеев

Процесс формирования требований и управления ими уже давно признается необходимым действием на пути к успешному созданию сложных систем и успешному завершению проектов, связанных с программными разработками.

Создание и управление требованиями является достаточно сложной задачей при выполнении ИТ-проектов – координировать совместные действия с заказчиками, подрядчиками и другими заинтересованными лицами. Организации, занимающиеся разработкой ПО, нуждаются в правильном формировании и эффективном управлении требованиями, чтобы быть уверенными в том, что конечный результат отвечает потребностям заказчика, соответствует техническому заданию и что работа над проектом ведется в рамках выделенного бюджета и запланированных сроков. Их особенностью является большое количество средних и малых проектов, требующих формализации требований в сжатые сроки и управления ими с учетом балансировки нагрузки на исполнителей. Отсутствие практического опыта автоматизации процесса управления требованиями в Санкт-Петербурге делает тему актуальной для рассмотрения. При этом в данной проблеме можно выделить несколько существенных моментов:

– требования динамично меняются на протяжении всего периода разработки и эксплуатации

программных продуктов. Дело тут не только в изменении намерений заказчика, уточнении набора функций и т.д. Довольно часто оказывается, что необходимость коррекции исходного задания определяется технологическими проблемами, которые возникают уже в ходе самой разработки;

- в процессе разработки занято много специалистов, и всех нужно своевременно информировать об изменениях;
- прямого взаимодействия «заказчик – программисты», как правило, уже давно нет на практике. Связь осуществляется через довольно длинную цепочку «заказчик – руководитель проекта – руководитель группы – программист», и нужно обеспечить, чтобы исходные требования заказчиков были адекватно переданы исполнителями.

Целью работы является создание прототипа модели на основе IBM RationalDoors с учетом анализа особенностей бизнес-модели организации.

Результатом работы является обоснованное предложение по оптимизации процесса «серийной» разработки ПО за счет тиражирования собранных требований и программных решений по их исполнению с использованием методологии Rational.

УДК 004.65

ТЕХНОЛОГИИ РЕПЛИКАЦИИ БАЗ ДАННЫХ В РАЗНОРОДНОЙ ИНФОРМАЦИОННОЙ СИСТЕМЕ

А.А. Иванычев

Научный руководитель – к.т.н., профессор В.В. Кириллов

Проблемы организации разнородных информационных систем, построенных на системах мульти-баз данных или распределенных баз данных, связаны со сложностью предметных областей, неоднозначностью взаимодействия между различными информационными средами, трудностью представления метаданных из всех баз данных, входящих в разнородную информационную систему, отсутствием прозрачности преобразования данных, сложными механизмами управления и одновременной обработки запросов [1]. Помимо этого, слабо развита методология распределения, размещения и обработки данных [2].

Цель работы – определение наиболее эффективной технологии репликации данных на основании минимизации временных затрат на инициализацию и выполнение репликации на различных наборах данных, отличающихся по типу, связанности, распределенности, объему в разнородной информационной системе.

Для достижения цели были рассмотрены средства репликации между разнородными СУБД, а также представлены собственные механизмы репликации баз данных средствами самих СУБД с использованием декларативных языков программирования.

В работе рассмотрены следующие технологические решения ведущих производителей СУБД:

- Oracle Streams;
- Oracle GoldenGate;
- Sybase Replication Server.

Помимо этого, разработаны алгоритмы репликации данных, реализуемые без учета представленных выше технологических решений, основанные на использовании словарей данных, механизмов DBLINK и «заданий по расписанию». Следует отметить, что все исследуемые технологические решения могут реализовывать репликацию слиянием, транзакционную репликацию и репликацию снимками [3].

Все исследования проводились на объектно-реляционных СУБД, а именно:

- MySQL 5.5.29;
- Oracle Enterprise Edition 11.2.0.3;
- Microsoft SQL Server 2012.

Основным критерием эффективности технологического решения репликации данных является время, необходимое на инициализацию и выполнение репликации на различных наборах данных, отличающихся по типу, связанности, распределенности и объему [4].

На сегодняшний день любая технология репликации данных в гетерогенной системе сталкивается с проблемами согласованности копий данных, что соотносится с теоремой Брюера и нарастающим потоком данных [5–7].

В результате проведенного исследования было определено, что при объемах реплицируемых связанных данных менее 20 ГБ эффективна технология OracleStreams, однако при увеличении количества реплицируемых данных свыше 20 ГБ данная технология не рекомендуется к использованию. Главным преимуществом технологии OracleStreams являются механизмы автоматического обнаружения и разрешения коллизий реплицируемых данных, но нестабильная работа OracleStreams является существенным недостатком. Технология OracleGoldenGate эффективна при передаче больших объемов несвязанных, статичных данных и может использоваться в совокупности с технологией OracleStreams. Несомненным преимуществом данной технологии является простота организации репликации между разнородными СУБД. Оптимальным решением представляется технология SybaseReplicationServer, предоставляющая все инструменты для быстрого внедрения и развертывания сервера репликации между разнородными СУБД. Помимо этого, алгоритмы, основанные на использовании словарей данных, механизмов DBLINK и «заданий по расписанию», показывают лучший результат при объеме данных менее 5 ГБ.

Литература

1. Дейт К.Дж. Введение в системы баз данных. 8-е издание: Пер. с англ. – М.: Вильямс, 2005. – 1316 с.
2. Белоусов В.Е. Алгоритмы репликации данных в распределенных системах обработки информации: Диссертация на соискание ученой степени кандидата технических наук. – Пенза: ПГУ, 2005. – 184 с.
3. Thomas M. Connolly, Carolyn E. Begg. Database Systems: A Practical Approach to Design, Implementation, and Management 4nd Revised edition. – Addison Wesley, 2009. – 112 p.
4. Ларичев О.И. Теория и методы принятия решений, а также хроника событий в Волшебных странах: Учебник. Изд. 2-е, перераб. и доп. – М.: Логос, 2002. – 392 с.
5. Ceri S. and Pelagatti G. Distributed Databases: Principles and Systems. – New York: McGraw-Hill, 1984. – 393 p.
6. Ceri S., Pernici B. and Wiederhold G. Distributed Database Design Methodologies // Proc. IEEE. – 1987. – № 75(5). – P. 533–546.
7. Brewer Eric A. Towards robust distributed systems (англ.) // Proceedings of the XIX annual ACM symposium on Principles of distributed computing. – Portland, OR: ACM, 2000. – Т. 19. – № 7.

ИДЕНТИФИКАЦИЯ НАГРУЗКИ В ВИРТУАЛИЗОВАННЫХ ЦЕНТРАХ ОБРАБОТКИ ДАННЫХ

Д.С. Иванов, С.М. Кляус, Н.М. Лукьянов

Научный руководитель – к.т.н., доцент Б.Д. Тимченко

Нагрузка, определяемая как совокупность исполнимых программных компонентов, является определяющим фактором функционирования любых компьютерных систем, в том числе и центров обработки данных (ЦОД) различных типов. Обычным в современных условиях является широкое использование виртуализации, и в дальнейшем в контексте данной работы ЦОД по умолчанию считается виртуализованным, а среда работы – корпоративной. Следующие признаки следует считать специфическими для таких ЦОД: аппаратная и программная гетерогенность, неоднородность нагрузки – наличие в ее составе работ различных классов, сложная динамика нагрузки – нестационарность в широком смысле, изменчивость ЦОД в аппаратно-программном отношении.

Понимание нагрузки, наличие достоверных сведений о ее составе, ресурсных свойствах динамике является необходимым для решения важнейших задач системной эксплуатации и ИТ-управления [1]. Данная работа имеет целью получение представления (модели) нагрузки ЦОД не только для управления, но и для обеспечения экспериментальных натурных и модельных исследований, нагрузочных испытаний и параметрической настройки систем обработки данных. Совокупность системных исследований, необходимых для получения модели нагрузки определяется как идентификация нагрузки, включает в себя структурный и параметрический аспекты.

Структурно модель нагрузки ЦОД предлагается определять совокупностью классов, каждый из которых в терминах теории массового обслуживания в общем случае задается потоком и некоторым набором законов обслуживания. Структурная идентификация, при соответствующем составе первичных данных формально выполняемая путем статистической классификации (кластеризации) компонентов нагрузки, практически требует знания программно-технической организации ЦОД, состава приложений и их роли в обеспечении бизнес задач, а также некоторых абстракций в представлении этого знания. В работе для этих предлагается потоково-серверное представление ЦОД, в котором потоки разделяются на входные, выходные и внутренние, а сервер определяется как любая сущность, для которой может быть заданы дисциплина и законы обслуживания классов нагрузки.

Параметрическая идентификация состоит в количественном определении значений параметров классов нагрузки. При вероятностных представлениях о среде функционирования, процессах и нагрузке большинство параметров оцениваются статистически на уровне, соответствующем составу первичных данных и требованиям к идентификации: одним или более моментом, либо эмпирическим распределением. В работе предлагается важное расширение традиционного подхода к параметрическому определению нагрузки, в котором основой служит статистика распределений [2]. Исходная предпосылка состоит в присущей корпоративным системам цикличности функционирования, первичное проявление которой выражается форматом 24×7 и последующим агрегировании недели, месяца, года. Многие параметры нагрузки в связи с этим необходимо оценивать методами анализа временных рядов, а не методами статистики распределений. Показываются и обсуждаются следствия использования в идентификации модели временных рядов, важнейшее из которых касается проблемы установившегося режима как необходимого условия при использовании статистики распределений.

Осуществимость и качество идентификации нагрузки определяющим образом зависят

от организации и проведения мониторинга, что по сути дела является отдельной сложной задачей системного характера. В работе на примере конкретного ЦОД, имеющего представляющей общий интерес архитектуру, ресурсы и нагрузку, обсуждаются детали конфигурирования мониторинга, сохранения и первичной обработки данных и их последующего анализа. Показывается, что целесообразно комбинирование использования универсальных мониторов, в нашем случае Zabbix, с инструментами вендоров, в нашем случае средства мигрирующего гипервизора VMware vSphere 5.1, операционной системы Microsoft Windows Server 2008 и системы хранения EMC Clarion [3]. Обсуждается проблема мониторинга в части получения первичных данных в форме трасс, полученный опыт оценивается и представляется в формате предварительного варианта методики идентификации нагрузки ЦОД. Обсуждаются перспективы использования полученной модели для проведения управляемых и воспроизводимых экспериментов по исследованию производительности, а также перспективы повторного использования полученных результатов как ссылочных моделей нагрузки, так и в качестве первичных данных при их представлении в открытом доступе.

Литература

1. Ингланд Роб. Введение в реальный ITSM / Пер. с англ. – М.: Лайвбук, 2010. – 132 с.
2. Dror G. Feitelson. Workload Characterization and Modeling Book [Электронный ресурс]. – Режим доступа: <http://www.cs.huji.ac.il/~feit/wlmod/wlmod.pdf> (последняя дата обращения 22.02.2012).
3. vSphere Monitoring and Performance: vSphere 5.1 [Электронный ресурс]. – Режим доступа: <http://pubs.vmware.com/vsphere-51/topic/com.vmware.ICbase/PDF/vsphere-esxi-vcenter-server-51-monitoring-performance-guide.pdf>, своб.

УДК 623.611

ОСОБЕННОСТИ ПРИМЕНЕНИЯ ТЕХНОЛОГИИ ОПЕРАТИВНОГО ПЛАНИРОВАНИЯ ДЛЯ РЕШЕНИЯ ЗАДАЧ ОПТИМАЛЬНОГО РАСПРЕДЕЛЕНИЯ РЕСУРСА В СИСТЕМАХ СПУТНИКОВОЙ СВЯЗИ

А.А. Ковальский

(Военно-космическая академия им. А.Ф. Можайского)

Научный руководитель – д.т.н., профессор К.Ю. Цветков

(Военно-космическая академия им. А.Ф. Можайского)

В современных условиях эффективность боевого применения войск в значительной степени определяется эффективностью функционирования системы управления. В настоящее время основу системы боевого управления Вооруженными силами (ВС) Российской Федерации (РФ) составляет объединенная система связи, которая включает в себя системы спутниковой связи (ССС). Проведенный анализ показал, что потребности ВС РФ в спутниковой связи в период с 2010 по 2020 год вырастут в 12 раз, однако используемый частотный ресурс увеличится незначительно, что окажет влияние на суммарную пропускную способность СССР. Таким образом, явно прослеживается противоречивая ситуация, связанная с наличием перспективных потребностей в обеспечении спутниковой связью и ограниченным ресурсом орбитальной группировки спутников-ретрансляторов (СР). Возможными вариантами выхода из сложившейся ситуации являются:

1. совершенствование технологической структуры процессов оперативного планирования и распределения ресурса СССР в части разработки и внедрения перспективных элементов математического (моделей, методов и алгоритмов управления), методического (подходов

и методик решения задач планирования) и специального программного обеспечения, что позволит более эффективное распределение частотно-временного ресурса при планировании и управлении ССС;

2. расширение частотного ресурса ССС, что предполагает вывод на орбиту новых космических аппаратов и аренду дополнительных стволов СР;
3. аппаратное повышение эффективного использования частотного ресурса, которое включает в себя:
 - модернизацию каналообразующего оборудования (доработка модемов для получения новых тактико-технических характеристик);
 - разработку дополнительного оборудования (разработка устройств для эффективного использования частотного ресурса бортового ретрансляционного комплекса);
 - разработку нового оборудования (применение принципиально новых технологий и оборудования).

Перечисленные направления совершенствования ССС являются взаимодополняющими. Данная статья посвящена особенностям применения первого из перечисленных направлений, поскольку предлагаемые структурные изменения и материально-технические затраты на реализацию первого из направлений совершенствования ССС представляются существенно меньшими по сравнению со вторым и третьим направлениями.

Решение указанной практической задачи требует применения нетривиальных методов в связи с существенной сложностью их формализации, высокими требованиями по оперативности и качеству формируемых решений. При этом в основу принимаемых решений должна быть положена известная концепция ситуационного управления, опирающиеся на нее принципы финитного управления, отвечающие идеологии программно-целевого подхода, и реализующая их технология гибких стратегий.

В реальных условиях сформированные оперативные планы работы ССС в силу целого ряда объективных и субъективных факторов могут выполняться со значительными погрешностями, которые, естественно, могут существенно отразиться на эффективности распределения ресурса ССС. Для снижения потерь, обусловленных действиями этих факторов целесообразно отказаться от технологии жесткого планирования и перейти к технологии так называемых гибких стратегий управления, обеспечивающих оперативную коррекцию оперативного плана работы ССС. При этом коррекция плана осуществляется за счет организации оперативной корректирующей обратной связи в контуре управления, обеспечивающей комплексирование этапов оперативного планирования и оперативного управления с доминированием технологии оперативного планирования. Коррекция же распределения ресурса ССС позволяет оперативную структурную коррекцию с учетом текущей важности сеансов связи, а также наличия у земной станции спутниковой связи (ЗССС) временных, энергетических, организационно-технических и других видов ресурсов.

Центр управления ССС с учетом корректирующей обратной связи в контуре управления целесообразно объединить в систему оперативного планирования и управления, оснастить ее необходимым алгоритмическим и программным обеспечением, реализованным в виде системы поддержки принятия решений (СППР), которая позволит не только автоматизировать наиболее рутинные операции сбора, обработки и анализа информации о состоянии СР и ЗССС, но и обеспечит оперативную генерацию оптимальных вариантов распределения ресурса с соответствующей коррекцией сеансов связи ЗССС. При создании такой СППР представляется возможным в качестве основных ее элементов использовать унифицированные проблемно-ориентированные модули из состава специального программно-математического обеспечения созданного на основе автоматизированной системы управления центра управления ССС. Основные усилия при этом можно сосредоточить на разработке непосредственно модуля генерации гибких стратегий управления ЗССС и орбитальной группировкой СР, входящих в ССС.

Алгоритм функционирования этого модуля должен реализовывать оптимальные или локально-оптимальные стратегии управления в реальном масштабе времени. Для синтеза этих стратегий задача оперативного планирования и распределения ресурса ССС должна быть сформулирована в оптимизационной постановке.

Литература

1. Мануйлов Ю.С., Птушкин А.И., Стародубов В.А. Методологические основы применения гибких стратегий управления космическими аппаратами. – СПб: МО РФ, 2002. – С. 102.
2. Мануйлов Ю.С., Новиков Е.А. и др. Алгоритм локально-оптимального управления комплексом операций обслуживания однотипных объектов // Сборник алгоритмов и программ типовых задач. – СПб: МО РФ, 2004. – Вып. 22.– С. 3–17.
3. Мануйлов Ю.С., Новиков Е.А. Метод и алгоритм оперативного планирования работы наземных средств на основе принципа гибких стратегий // Сборник алгоритмов и программ типовых задач. – СПб: МО РФ, 2005. – Вып. 23.– С. 3–17.

УДК 004.415.2

РАЗРАБОТКА ПРОГРАММНОГО КОМПЛЕКСА ДЛЯ СОПРОВОЖДЕНИЯ УДАЛЕННОГО МОДУЛЯ «ТИС-ПРОЦЕСС»

И.В. Колбик

Научный руководитель – д.т.н., профессор Д.Д. Куликов

В современных условиях для эффективного решения задач ТПП необходимо повышать уровень автоматизации. САПР ТП «ТИС-Процесс» предназначена для формирования модели технологического процесса и печати комплекта технологических карт.

Целью работы является переработка САПР ТП «ТИС-Процесс» под web-сервис.

В конструкторской и технологической подготовке производства важную роль играет модель технологического процесса создаваемого изделия. Эта модель представляет собой описание процесса, содержащее информацию о размерах, номер чертежа, тип процесса и др. В ходе работы с системой выполняется редактирование модели технологического процесса, либо поиск процесса аналога по каталогу технологических процессов с последующей записью результатов поиска в буфер для дальнейшего использования в других системах.

Но на данный момент существует проблема отсутствия системы «ТИС-Процесс» в удаленном доступе.

Используя все преимущества удаленного приложения, мы можем применить данную систему для проектирования ТП в других САПР ТП.

На данный момент реализована работа с каталогом системы «ТИС-Процесс», в котором реализован поиск технологического процесса аналога по базе данных технологических процессов.

В этом докладе описывается процесс разработки программного комплекса для сопровождения удаленной системы «ТИС-Процесс». Показана структура модулей системы, их возможности по параметризации, графический интерфейс приложения.

В конечном счете, будет создан программный комплекс, позволяющий удаленно осуществлять полноценную работу с системой «ТИС-Процесс» и использовать результаты работы для проектирования ТП в других системах.

ИСПОЛЬЗОВАНИЕ ПАРАЛЛЕЛЬНЫХ ВЫЧИСЛЕНИЙ НА НОВОМ СЕМЕЙСТВЕ ГРАФИЧЕСКИХ ПРОЦЕССОРОВ KEPLER ДЛЯ РЕШЕНИЯ МАТРИЧНЫХ СЕТЕВЫХ УРАВНЕНИЙ БОЛЬШИХ РАЗМЕРНОСТЕЙ

А.Ю. Коровяковская

Научный руководитель – к.т.н., профессор Т.А. Павловская

В работе рассмотрена архитектура семейства графических процессоров Kepler фирмы nVidia. Приведен алгоритм оптимизации обработки сетевых моделей сложных многомерных задач и доказаны его преимущества при решении матричных уравнений большой размерности с использованием параллельных вычислений на новом семействе графических процессоров Kepler.

В работе [1] было показано, что производительность штатных библиотек типа CUBLAS, используемых для манипуляций с матрицами, недостаточна для расчета больших сетевых моделей. При моделировании сложных газовых и электрических инфраструктур сетевые матрицы могут достигать достаточно больших размеров (10000×10000 элементов и более). Для решения указанных уравнений, моделирующих поведение этих систем, необходимо использовать более производительный алгоритм.

В работе [1] был описан алгоритм, позволяющий распараллелить вычисления и тем самым увеличить производительность при обработке большого массива данных. Согласно оценкам, указанным в работе, производительность может быть увеличена в 10–15 раз.

В последнее время фирма NVIDIA выпустила новое семейство графических процессоров GK104 под кодовым названием Kepler. Данное семейство графических процессоров, как анонсировано разработчиком, позволяет значительно ускорить работу с параллельными потоками данных. В статье [2] приведены данные изменения производительности графических процессоров фирм NVIDIA и Intel различных поколений. Теоретическая производительность графического процессора GK104 может достигать порядка 3TFlops в секунду. Поэтому было бы интересно исследовать указанную разработку для решения нашей задачи моделирования сложной сетевой структуры.

Рост производительности графического процессора GK104Kepler обусловлен следующим главными факторами:

1. Количество одновременно работающих ядер увеличено до 1536 единиц. Объем памяти устройства может достигать 4 Гб (в зависимости от конкретного изготовителя), а тактовая частота видеопамати составляет рекордные для настоящего времени 6 000 МГц. Частота видеопамати вкупе с шиной памяти обеспечивают пропускную способность порядка 192,3 Гб/с. Эти факторы позволяют одновременно манипулировать данными больших объемов и обеспечивать быструю загрузку данных из Host устройства (центральный CPU) в Device устройство (графический процессор GPU) и обратно.
2. Общее количество потоковых мультипроцессоров равно 8 (четыре кластера по 2 SMX), а в каждом SMX встроено 192 CUDA-ядра (ранее только 32), а суммарное количество ядер, как уже отмечалось, равно $192 \times 8 = 1536$.
3. Упрощен процесс программирования GPU, позволяя с легкостью ускорять все параллельные вложенные циклы. Это приводит к тому, что GPU динамически порождает новые потоки без возврата к CPU и может совмещать загрузку данных с расчетами.
4. Реализован механизм Hurd-Q, который сокращает время ожидания современного мультиядерного CPU, позволяя всем ядрам CPU одновременно использовать один GPU на базе архитектуры Kepler, значительно увеличивая возможности программирования [2].

Одна из целей настоящего исследования заключалась в том, чтобы испытать предложенный в работе [1] алгоритм в рамках новой архитектуры CUDA и получить

численные оценки быстродействия тестов при манипуляции с матрицами больших размеров. Для этого была создана программа, написанная на языке высокого уровня C++ с использованием API CUDA, ориентированного на архитектуру Kepler.

В таблице приведены результаты испытаний алгоритма умножения матриц различной размерности (все элементы матриц являются ненулевыми и значения их заполнялись с помощью датчика случайных чисел). В первом столбце таблицы приведено время, затраченное графическим ядром видеокарты GeForce 260, во втором – ядром видеокарты GeForce 680), в третьем – число элементов в матрице (произведение $m \times n$), а в последнем столбце – выполненное количество операций.

Таблица 1. Результаты испытаний

GeForce 260, τ_1	GeForce 680, τ_2	$m \times n$	Non
$0,97 \cdot 10^{-3}$	$0,254 \cdot 10^{-3}$	$1,54 \cdot 10^5$	$98 \cdot 10^6$
$1,62 \cdot 10^{-3}$	$0,424 \cdot 10^{-3}$	$2,21 \cdot 10^5$	$169 \cdot 10^6$
$2,6 \cdot 10^{-3}$	$0,675 \cdot 10^{-3}$	$3,01 \cdot 10^5$	$269 \cdot 10^6$
$3,77 \cdot 10^{-3}$	$0,987 \cdot 10^{-3}$	$3,93 \cdot 10^5$	$403 \cdot 10^6$

Из таблицы видно, что реальная производительность для новой архитектуры возрастает в 4–5 раз по сравнению с более ранними разработками фирмы NVidia и может достигать 400–600 Gflops. Следует отметить, что при испытаниях не использовалось важное свойство сетевых матриц – их слабая заполненность. Следует ожидать еще большего выигрыша производительности новой архитектуры при манипуляции со слабозаполненными матрицами.

Литература

1. Коровяковская А.Ю. Опыт применения современных программно-аппаратных методов для решения многомерных сетевых задач // Мат. докл. науч.-теорет. конф., Информационные системы и технологии. – Красноярск, май 2012. – С. 159–173.
2. C Programming Guide CUDA Toolkit Documentation [Электронный ресурс]. – Режим доступа: <http://docs.nvidia.com/cuda/cuda-c-programming-guide/>, своб.

УДК 004.056.3

ОБЕСПЕЧЕНИЕ НЕПРЕРЫВНОСТИ РАБОТЫ WEB-ДОСТУПНОГО КОНТЕНТ-ХРАНИЛИЩА

А.Г. Красиков

Научный руководитель – к.т.н., доцент Б.Д. Тимченко

В работе адресуется проблема обеспечения непрерывности функционирования приложений в виртуализированных средах.

Цель работы – обобщить лучшие практики обеспечения непрерывности работы приложений и применить их в реализации контент-системы СЗ (система файлового хранения, предназначенная для группового взаимодействия, работа с которой осуществляется через web-интерфейс).

Непрерывность функционирования (continuous availability, непрерывная доступность) согласно [1] определяется как задача по проектированию ИТ-услуги, приложения или системы, направленная на достижение 100% доступности.

Задача обеспечения непрерывности работы сетевых и web-приложений является одной из наиболее актуальных и сложных задач из области разработки программного обеспечения.

Известно, что простои в работе приводят к значительным потерям прибыли: недоступность сервиса в течение одного часа стоит медиа-компаниям порядка \$90 тысяч, тогда как потери крупной брокерской конторы за тот же период времени могут составлять миллионы долларов [2].

Технологии виртуализации открывают определенные возможности решения задачи обеспечения непрерывности на основе миграции приложений, обеспечиваемой средствами самих платформ виртуализации [3].

В связи с повсеместным распространением технологий виртуализации актуальным становится рассмотрение проблемы отдельно для систем (приложений), работающих на виртуальных машинах. Примером такой системы является и контент-система СЗ, работающая в среде под управлением VMwarevSphere. В работе анализируются как общие подходы к обеспечению непрерывности работы приложений в целом, так и специфичные для виртуализированных сред.

Основным документом, регламентирующим действия по обеспечению непрерывности работы конкретных систем, считается ContingencyPlan («План [обеспечения] непрерывности»). В качестве стандарта для создания такого плана принято использовать руководство [4].

Представляется анализ различных типов планов обеспечения непрерывности и предлагается его вариант для виртуализированных систем. Первичной задачей является обеспечение возможности резервного копирования и восстановления данных. При этом, должно быть доступно создание как копии всей виртуальной машины средствами гипервизора, так и копии данных, используемых системой (программой). Второй шаг для обеспечения непрерывности – реализация поддержки синхронной и/или асинхронной репликации данных системы. Создание процедуры миграции системы на альтернативный хост для запуска – обязательный пункт плана. Последняя (по порядку, но не по значению) задача – назначение лиц, ответственных за исполнение разработанного плана.

В рамках работы создан план непрерывности для контент-системы СЗ. Обеспечено локальное и удаленное резервирование данных системы (с использованием сервиса RedHatOpenShift по протоколу SFTP). Целостность бэкапов при восстановлении проверяется с помощью вычисления контрольной суммы с использованием алгоритма MD5 [5]. Опционально доступно добавление электронной цифровой подписи к резервной копии. Реализована синхронная и асинхронная локальная репликация отдельных хранилищ данных между собой. Кроме того, предложен способ резервного копирования образов виртуальных машин с использованием средств VMwarevSphere, в частности vSphere «livemigration» [3]. Для миграции и репликации был предложен общий алгоритм, позволяющий синхронизировать данные разных экземпляров системы СЗ, которые могут быть запущены как на одном, так и на разных серверах. Репликация (и миграция) может производиться как в пределах локальной сети, так и по сети Интернет, все передаваемые данные при этом шифруются. Развертывание системы на удаленном хосте автоматизировано.

В работе обсуждаются детали реализации, а также процесс тестирования и внедрения плана обеспечения непрерывности.

Литература

1. ITIL V3 Glossary Russian Translation [Electronic resource]. – Режим доступа: http://www.itexpert.ru/rus/biblio/itil_v3/ITILV3_Glossary_Russian_v092_2009.pdf, своб.
2. How Much Does Downtime Really Cost? [Electronic resource] – Режим доступа: http://www.information-management.com/infodirect/2009_133/downtime_cost-10015855-1.html, своб.
3. vSpherevMotion: Live migration of virtual Machines [Electronic resource]. – Режим доступа: <http://www.vmware.com/ru/products/datacenter-virtualization/vsphere/vmotion.html>, своб.

4. Marianne Swanson, Pauline Bowen, Amy Wohl Phillips, Dean Gallup, David Lynes. Contingency Planning Guide for Federal Information Systems // National Institute of Standards and Technology Special Publication 800-34. USA, May 2010. – 59 p.
5. The MD5 Message-Digest Algorithm [Electronic resource]. – Режим доступа: <http://tools.ietf.org/html/rfc1321>, своб.

УДК 65.011.56, 004.62

УПРАВЛЕНИЕ ИНТЕГРАЦИЕЙ MICROSOFT DYNAMICS CRM И СИСТЕМЫ ЗАПИСИ ЗВОНКОВ DIGITEL

А.А. Крупин

Научный руководитель – к.ф.-м.н., доцент С.Е. Иванов

Краткое введение, постановка проблемы. Часто в компаниях функционирует несколько информационных систем. Иногда информация в этих системах дублируется. Интеграция таких систем необходима для автоматического переноса информации из одной информационной системы в другую, что оптимизирует рабочий процесс и повышает его эффективность. Исходя из этого, появляется потребность в разработке соответствующего метода интеграции. Примером таких систем являются MicrosoftDynamics CRM и система записи звонков Digitel. В обеих системах ведется история совершенных телефонных звонков, причем в первой системе она ведется вручную, а во второй – в автоматическом режиме. Организовав перенос информации о звонках из Digitel в MicrosoftDynamics CRM, мы избавим работников от ручного ввода информации в систему, убрав человеческий фактор ошибки, присутствующий при ручном вводе.

Цель работы. Необходимо спроектировать и реализовать метод интеграции MicrosoftDynamics CRM и системы записи звонков Digitel. Метод должен обеспечивать следующую функциональность:

- при появлении информации о звонке в системе Digitel передавать ее в MicrosoftDynamics CRM (в том числе, ссылку на mp3-файл с записью разговора);
- на основе телефонных номеров по базе данных MicrosoftDynamics CRM определять бизнес-партнера и сотрудника компании, автоматически внося их в создаваемый объект «Звонок»;
- отображать у целевого сотрудника графическое окно с только что созданным в MicrosoftDynamics CRM звонком.

Базовые положения исследования. Исследуются две системы, MicrosoftDynamics CRM и Digitel, на наличие интеграционных возможностей. Основной сложностью в данной ситуации является то, что MicrosoftDynamics CRM 4.0 не имеет встроенной возможности интеграции с продуктами, не построенными на платформе .NET и не являющимися продуктами Microsoft, кроме как методом экспорта-импорта файлов и SOAP-запросов.

Промежуточные результаты. Сформулированы нефункциональные требования к методу интеграции MicrosoftDynamics CRM и Digitel:

- мгновенность: данные должны передаваться с минимальными задержками;
- универсальность: создаваемый метод должен иметь возможность впоследствии использоваться для интеграции MicrosoftDynamics CRM с другими не-Microsoft продуктами.

На основе требований проведен информационно-аналитический обзор возможных

способов интеграции данных информационных систем, выявлены их достоинства и недостатки. Был выбран метод интеграции на основе веб-сервисов. В качестве промежуточного звена будет использоваться ASP .NET страница, которой Digitel будет передавать информацию через GET-запросы при возникновении событий.

Основной результат. Реализован и протестирован метод интеграции MicrosoftDynamics CRM и Digitel, полностью соответствующий поставленным требованиям. Предлагаемый метод можно в дальнейшем модифицировать для интеграции MicrosoftDynamics CRM с другими не-Microsoft продуктами.

УДК 004.93+57.087.1

СИСТЕМА РАЗДЕЛЕНИЯ ДИКТОРОВ В ПОЛИЛОГЕ

О.Ю. Кудашев

Научный руководитель – к.ф.-м.н., доцент Т.С. Пеховский

Задача разделения дикторов состоит в выделении речевых сегментов фонограммы и кластеризации (объединении) выделенных сегментов по принадлежности к одному диктору. Эта задача является важной и неотъемлемой частью систем автоматической обработки речи. К таким системам можно отнести автоматическую голосовую идентификацию дикторов, индексацию аудиоданных.

Основным отличием задачи разделения дикторов в полилоге является отсутствие априорной информации о числе дикторов. В случае наличия такой информации задача разделения дикторов решается, к примеру, применением вариационного байесовского анализа в совокупности с факторным анализом, дающего в настоящее время наименьшие значения ошибки сегментации фонограммы.

Для разработки системы разделения дикторов в полилоге было решено использовать методы, хорошо зарекомендовавшие себя в задачах идентификации человека по голосу.

В первую очередь к таким методам следует отнести создание голосовой модели диктора на основе факторного анализа в пространстве «полной изменчивости». В соответствии с современными исследованиями, наиболее эффективным типом генеративной модели диктора является смесь гауссовых распределений (GaussianMixtureModel, GMM), аппроксимирующая распределение акустических признаков. При этом априорным распределением для каждого диктора является так называемая универсальная фоновая модель (UniversalBackgroundModel, UBM) –GMM, полученная на большой базе обучения. Модель каждого диктора получается путем адаптации средних значений гауссойд UBM без изменения матриц ковариаций. Объединение полученных таким образом векторов средних значений называют супервектором средних диктора. В качестве априорного распределения супервектора средних диктора (s) принимается сумма супервектора средних УФМ (μ) и произведения матрицы «полной изменчивости» (T) на низкоразмерный вектор скрытых параметров (факторов) (w): $s = \mu + Tw$.

Следующим наиболее важным методом является метод компенсации канальных искажений. Таким методом является проекция вектора скрытых параметров на новое пространство, заданное матрицей, полученной при помощи линейного дискриминантного анализа (ЛДА). Результат этой проекции является максимизация отношения меж-дикторской и внутри-дикторской вариативностей.

Необходимо отметить, что все параметры UBM, а также матрицы «полной изменчивости» и матрицы ЛДА были получены на следующих речевых базах данных: NISTSRE 1996-2006, NISTSRE 2008, NISTSRE 2010, RuSTeN. Суммарная длительность

чистой речи на указанных базах составила 1900 часов, а количество дикторов – 5400. Таким образом, полученная проекция вектора скрытых параметров на плоскость ЛДА отвечает наиболее информативным признакам, пригодных для дискриминации речевых сегментов по принадлежности к разным дикторам. При этом результат дискриминации остается приемлемым даже для коротких речевых сегментов длительностью от 1–1,5 секунды.

Разработанная система разделения дикторов в полилоге состоит из следующих этапов:

- выделение речевых сегментов фонограммы;
- расчет акустических признаков;
- построение проекции векторов скрытых параметров на плоскость ЛДА (l -векторов) для каждого окна, содержащего 1 секунду чистой речи, с шагом 0,5 секунд;
- попарное сравнений полученных l -векторов;
- поиск точек смены дикторов на фонограмме на основе полученных матрицей сравнений с заранее заданным порогом λ_1 ;
- агломеративное объединение множества речевых сегментов, ограниченных найденными точками смены дикторов, путем сравнения соответствующих l -векторов, с заранее заданным порогом λ_2 ;
- пересегментация точек смены дикторов на основе алгоритма Витерби;
- окончательное агломеративное объединение множества речевых сегментов с заранее заданным порогом λ_3 .

Поскольку данная система имеет три дополнительных порога ($\lambda_1, \lambda_2, \lambda_3$), определяемых на базе развития, ключевым вопросом при сравнении l -векторов является робастность метрики. Было рассмотрено три основные метрики: косинусное ядро; CLR (Cross-LikelihoodRatio); дистанция на основе SVM (SupportVectorsMachine) с косинусным ядром. В результате численных экспериментов было показано, что наиболее робастной метрикой является использование дистанции на основе метода опорных векторов.

Описанная система разделения дикторов была реализована на кафедре речевых информационных систем НИУ ИТМО, базовой кафедре ООО «ЦРТ-инновации». Она показала хорошие результаты, как на русскоязычных, так и англоязычных базах тестирования. Указанная система успешно внедрена и применяется в системах автоматической голосовой идентификации, а также системе индексирования записей новостных передач компании ООО «ЦРТ-инновации».

УДК 004.056

МЕТОДИКА ДОКУМЕНТИРОВАНИЯ АРХИТЕКТУРЫ ПРОГРАММНОГО ОБЕСПЕЧЕНИЯ

В.А. Лежнин

Научный руководитель – к.т.н., доцент Б.Д. Тимченко

Задачей данной работы является разработка методики документирования архитектуры, основанной на последних формулировках, определениях и рекомендациях международных организаций по стандартизации. Основными стандартами, охватывающими область документирования архитектуры программного обеспечения, являются IEEE 1471:2000 и разработанный на его основе стандарт ISO/IEC 42010.

Особое внимание в работе уделяется определению архитектуры программного обеспечения. Предметная область архитектуры ПО является относительно молодой, поэтому определение архитектуры нельзя считать установившимся. Согласно данным Института Программной Инженерии Карнеги Меллон, насчитывается более 150 версий определения

архитектуры ПО [1]. Приводится обоснование выбора определения архитектуры, наиболее соответствующего целям ее документирования.

Стандарт ISO/IEC 42010 дает определение методики документирования архитектуры или архитектурного фреймворка – «архитектурный фреймворк определяет набор указаний по созданию, интерпретированию, анализу и использованию архитектурных описаний в рамках определенной предметной области или сообщества заинтересованных лиц» [2].

В работе анализируются и оцениваются существующие методики документирования. Базовым элементом большинства из них является архитектурное представление. Еще в 70х годах прошлого столетия различные авторы пришли к выводу, что архитектура имеет множество аспектов, которые должны быть описаны с использованием разных представлений [1]. Широкое распространение получила методика «4+1 представление», предложенная в 1995 году Филиппом Кратченом из компании Rational [3]. На данной методике основано документирование архитектуры ПО в RationalUnifiedProcess [4]. Эта методика, как и некоторые другие, применяемые в отрасли, основана на фиксированном наборе представлений. С другой стороны, в 2000 году был выпущен стандарт IEEE 1471, рекомендуемый использование набора представлений, исходя из потребностей заинтересованных лиц конкретного проекта [2]. При разработке методики учтены рекомендации комитетов по стандартизации по использованию динамического набора представлений. Дана классификация возможных архитектурных представлений, включая модульные представления, представления компонентов и коннекторов, а также представления распределения.

Проводится анализ влияния потребностей заинтересованных лиц (stakeholders) на выбор архитектурных представлений. Приводится состав методики и обоснование включаемых в нее разделов. Описываются структура комплекта архитектурной документации и приводятся шаблоны документов. Демонстрируется применение разработанной методики для документирования архитектуры системы хранения и групповой работы СЗ, разработка которой ведется на кафедре вычислительной техники НИУ ИТМО. Приводятся достоинства и недостатки разработанной методики, выявленные на практике.

Литература

1. Clements P., Bachmann F., Bass L., Garlan D., Ivers J., Little R., Merson P., Nord R., Stafford J. Documenting Software Architectures: Views And Beyond Second Edition, Addison-Wesley. – 2011.
2. International Standard ISO/IEC/IEEE 42010 – Systems and software engineering – Architecture description [Электронный ресурс]. – Режим доступа: <http://cabibbo.dia.uniroma3.it/asw/altrui/iso-iec-ieee-42010-2011.pdf>, своб.
3. Philippe Kruchten Architectural Blueprints– The «4+1» View Model of Software Architecture, Rational Software Corp [Электронный ресурс]. – Режим доступа:
 1. <http://www.cs.ubc.ca/~gregor/teaching/papers/4+1view-architecture.pdf>, своб.
4. Арлоу Д., Нейштадт А. UML 2 и унифицированный процесс / Пер. с англ. Н. Шатохиной. – СПб: Издательство Символ-Плюс, 2007. – 624 с.

НЕСТАЦИОНАРНОСТЬ И СТРУКТУРНАЯ СЛОЖНОСТЬ ИНФОРМАЦИОННЫХ ПОТОКОВ В ПРОЦЕССЕ МОДЕЛИРОВАНИЯ МУЛЬТИСЕРВИСНЫХ ТЕЛЕКОММУНИКАЦИОННЫХ СИСТЕМ

С.И. Макаренко, К.В. Ушанев

(Военно-космическая академия им. А.Ф. Можайского)

Научный руководитель – д.т.н., профессор К.Ю. Цветков

(Военно-космическая академия им. А.Ф. Можайского)

В современном мире широкое развитие и применение получили высокоскоростные сети связи. Активность развития таких сетей является стремительной во многом благодаря востребованности на массовом уровне услуг сети Интернет, появлению новых приложений (IP-телефония, IPTV), работающих в режиме реального времени, мультимедийных приложений, а также развитию сотовых сетей связи. Все это привело к необходимости передачи по сети различных видов трафика, в том числе, чувствительного к задержкам. Активность развития сетей проявляется в развитии их инфраструктуры, повышении производительности каналов передачи и серверов обработки данных. Однако, даже значительно возросшие в последнее время пропускные способности каналов и серверов могут оказаться недостаточными из-за продолжающегося роста интенсивности трафика, повышения его структурной сложности, а также повышения требований к качеству обслуживания абонентов. Поэтому немаловажное значение имеют характеристики технической системы связи, определяющие эффективность ее функционирования которые закладываются на этапе проектирования сети. В связи с этим всю большую роль играют методы проектирования, адекватность математических моделей, используемых при этом, в том числе с учетом характера трафика. От них зависят свойства и эффективность будущей системы. Одним из этапов проектирования мультисервисных телекоммуникационных систем (ТС) является выбор параметров структурных элементов, таких как пропускные способности каналов связи, объемы буферов, используемых для обслуживания потоков в узлах сети, и другие. Решение данных задач осуществляется с применением методов параметрического синтеза.

Используемые в настоящее время методы параметрического синтеза базируются на применении классических математических моделей потоков, которые хорошо себя зарекомендовали при проектировании сетей с коммутацией каналов, таких как телефонные сети. Современные исследования трафика, передаваемого в ТС, показывают, что его статистические характеристики отличаются от тех, которые приняты в классической теории телетрафика. Использование представления о том, что объединение большого числа потоков от независимых источников информации приводит к получению процесса, описываемого пуассоновским потоком, не соответствует истине. Это приводит к тому, что использование традиционных методов расчета параметров ТС и их вероятностно-временных характеристик (ВВХ), неверные результаты, приводящие к недооценке нагрузки.

Последние исследования свойств информационных потоков в мультисервисных ТС показали, что использование моделей самоподобных (фрактальных) процессов (потоков) позволяет более точно описывать трафик, передаваемый в данных системах. С практической точки зрения это можно объяснить высокой изменчивостью интенсивности трафика и, как следствие, высокой пачечностью поступления пакетов в узел сети при высокой скорости передачи данных, что приводит, из-за ограниченности буфера, к потерям пакетов.

В отличие от пуассоновских процессов самоподобные характеризуются наличием последствия: вероятность поступления следующего события зависит не только от времени,

но и от предыдущих событий. Это означает, что число текущих событий может зависеть от числа предыдущих событий в отдаленные промежутки времени.

Исследование влияния самоподобных свойств информационного трафика на качество обслуживания абонентов представляется важным, поскольку при наличии свойства самоподобия, качество обслуживания (QoS – Quality of Service), как правило, ухудшается по сравнению с тем, что наблюдалось бы, в случае пуассоновского трафика. Учет самоподобных (фрактальных) свойств трафика позволит более точно описать и воспроизвести информационный трафик, что, в свою очередь, обеспечит возможность получения показателей QoS, соответствующих реально наблюдаемому. Поэтому актуальными представляются исследования свойств самоподобия трафика информации, их влияния на характеристики QoS в ТС и оптимизация входных параметров ТС с целью обеспечения заданного QoS.

В этом случае необходимо установить взаимосвязь в виде аналитических выражений между параметрами передаваемых потоков и параметрами качества обслуживания. Полученные в результате зависимости необходимо использовать в составе ограничений и целевой функции при решении задач параметрического синтеза.

Актуальными при решении данных задач в настоящее время являются методы управления интенсивностью сетевого трафика. Выделяют два основных направления.

1. Как известно самоподобные процессы обладают долгосрочной зависимостью, выражающейся в бесконечном интервале корреляции. Это в свою очередь позволяет предсказать поведение трафика на длительных временных интервалах. Используя прогнозируемые данные возможно создавать эффективные алгоритмы управления трафиком.

С практической точки зрения предсказание величины возможной пиковой нагрузки и времени ее появления является очень важным атрибутом для принятия соответствующих мер по обеспечению качества обслуживания QoS.

2. Использование механизмов Traffic Shaping (TS) и Traffic Policing (TP). Суть данных алгоритмов заключается в следующем:

- с помощью TS сглаживается трафик и пересылается с постоянной интенсивностью путем постановки в очередь (буферизации) пакетов, интенсивность передачи которых превысила среднее значение;
- механизмом TP просто отбрасываются пакеты, интенсивность которых выше согласованной скорости передачи. С одной стороны, тот факт, что в TS не допускаются отбрасывания пакетов, является положительным для управления передачей информации реального времени (речь, видео), с другой стороны – вносятся задержки, связанные с буферизацией, что отрицательно сказывается на характеристиках передаваемого трафика.

Таким образом, описанные методы параметрического синтеза могут быть использованы в процессе проектирования и являются обоснованием для выбора значений конфигурационных параметров телекоммуникационного оборудования в узлах сети при планировании и эксплуатации мультисервисных ТС, а также составлении спецификаций устанавливаемого оборудования, при развертывании новых и модернизации существующих мультисервисных ТС.

ИССЛЕДОВАНИЕ ЭФФЕКТИВНОСТИ СИСТЕМ САРТСНА

И.А. Меженин

Научный руководитель – к.т.н., доцент В.И. Поляков

Постановка задачи. Как известно, одним из механизмов борьбы со спамом и защиты от автоматизированных атак является САРТСНА (C Completely Automated Public Turing Test to Tell Computers and Humans Apart) – тест для различия компьютеров и людей. Самым распространенным вариантом САРТСНА являются сложные искажения символов, которые позволяют сделать их неразличимыми для алгоритмов распознавания образов, сохраняя их узнаваемость для человека. Одно из направлений при проектировании тестов – это использование цветовых решений, позволяющих, с одной стороны повысить читаемость для пользователей и с другой стороны увеличивать надежность теста.

Цель исследования. Изучение эффективности различных систем САРТСНА и разработка методов, позволяющих значительно повысить их защищенность. Исследование механизмов устойчивости к сегментации различных цветовых решений.

Базовые положения исследования. Обеспечение надежности и читаемости непростая задача. При проектировании САРТСНА необходимо использовать как традиционные подходы – искажение символов и добавление помех, так и новые, в частности использование различных цветовых схем:

- на сайтах Microsoft отдано предпочтение простой цветовой схеме, в которой сам текст темно-синего цвета, а фон – светло-серый;
- в сервисах Google используется только один цвет (зеленый, красный или синий) для символьных последовательностей и белый фон;
- на сайтах Yahoo применяются только черно-белые САРТСНА;
- в системе reСАРТСНА тоже используется черно-белая схема (reСАРТСНА– новая разработка изобретателей Gimpy-r и EZ-gimpy).

Любые системы автоматического распознавания преобразуют цветное изображение в оттенки серого. Для получения полутонового черно-белого изображения чаще всего используется следующая классическая формула:

$$I(C) = 0,3 \cdot R(C) + 0,59 \cdot G(C) + 0,11 \cdot B(C),$$

где I – интенсивность в точке полутонового изображения; R , G и B (значения 0..255) – красная, зеленая и синяя компонента цвета C . Этот способ подходит для разных задач и обладает своими преимуществами, однако имеет и недостаток – полученные изображения часто теряют четкость в сравнении с оригиналом. Особенно часто это случается при использовании однотонных цветов. Именно этот недостаток можно использовать для повышения устойчивости теста.

Проблема преобразования цветного изображения в оттенки серого аналогична квантованию цвета и сжатию размерности палитры. Однако, как квантование цвета, так и сжатие цветовой палитры, обычно сохраняют различия в яркости и не учитывают вклад цветности каждого значения цвета в изображении.

Основной результат. Можно предложить следующие рекомендации по использованию цвета:

- неоднородность символов и фона позволяют повысить сопротивляемость сегментационным атакам, поэтому рекомендуется при проектировании теста обеспечивать максимальную неоднородность переднего плана и фона, сохраняя читаемость;

- нежелательно использование цветов, усиливающих контраст между передним планом и фоном или контраст между разными символами, так как снижает устойчивость к сегментации;
- применять простые цветовые схемы, обеспечивающие удобство использования и не нарушающие уязвимость;
- разработка палитр учитывающих вклад цветности каждого значения цвета в тестовых изображениях.

УДК 004.415.53

ИССЛЕДОВАНИЕ ПРОБЛЕМ ТЕСТИРОВАНИЯ ПРОГРАММНОГО ОБЕСПЕЧЕНИЯ ТОЧКИ ДОСТУПА СЕТИ LTE

А.С. Митрофанов

Научный руководитель – ассистент А.В. Гаврилов

Тестирование – неотъемлемая часть современного процесса по разработке программного обеспечения. Данная область активно совершенствуется и регулярно предоставляет новые технологии и подходы, которые можно применять на практике. Но, тем не менее, очень часто возникают ситуации, когда тестирование замедляет весь ход разработки. В таких случаях обычно имеет место неправильная организация процесса проверки качества программного продукта либо использование не самых удачных технологий. Чтобы избежать подобных проблем, необходимо хорошо разбираться в существующих способах тестирования и постоянно стремиться к улучшению текущих решений.

В области беспроводных технологий к качеству программного обеспечения точек доступа традиционно предъявляются жесткие требования. Также необходимо отметить и специфику структуры и функционирования данного оборудования, которая обуславливает дополнительные сложности проверки качества соответствующего продукта. Исходя из приведенных фактов, можно сделать вывод о ключевой роли тестирования в процессе разработки программного обеспечения для точек доступа беспроводной сети. **Целью работы** является обзор текущих видов тестирования программного обеспечения (ПО) точки доступа сети стандарта LTE, выявление существующих в них проблем и рассмотрение способа их решения.

Для исследования были выбраны наиболее сложные и часто применяемые в ходе разработки ПО точки доступа сети LTE виды тестирования: интеграционное и системное. В интеграционном тестировании рассмотрены различия в подходах к проверке программных компонентов жесткого и мягкого реального времени. В системном, в свою очередь, предоставлен обзор используемых способов нагрузочного и функционального тестирования. Далее в ходе работы был выявлен список проблем, характерных для каждого из рассмотренных видов проверок качества программного обеспечения, и произведена выборка общих недостатков с целью систематизации и дальнейшего анализа. Затем, опираясь на результаты предыдущего шага, в статье было предложено решение, которое способно оптимизировать приведенные в работе виды тестирования и минимизировать выявленные недостатки или убрать их вовсе.

В результате проделанной работы выявлен ряд проблем, негативно влияющих на производительность тестирования ПО точки доступа сети LTE. В ходе их анализа обнаружены общие черты и отмечена одинаковая природа происхождения. На основе полученной информации сделан вывод о целесообразности внедрения облачных технологий для решения описанных проблем и улучшения производительности тестирования в целом.

АНАЛИЗ ФУНКЦИОНИРОВАНИЯ ПРОТОКОЛА SPANNINGTREEPROTOCOL

Р.Л. Михайлов

(Военно-космическая академия им. А.Ф. Можайского)

Научный руководитель – д.т.н., профессор К.Ю. Цветков

(Военно-космическая академия им. А.Ф. Можайского)

Алгоритм STA (SpanningTreeAlgorithm) был разработан Р. Перлман в 1983 г. [1] с целью расширения сферы применения прозрачных мостов, которые активно использовались в то время. Протоколом, реализующим алгоритм STA, является STP (SpanningTreeProtocol). STP – сетевой протокол, работающий на канальном уровне модели взаимодействия открытых систем. Основной задачей STP является приведение сети с множественными связями к древовидной топологии, исключающей циклы. Это достигается путем автоматического блокирования ненужных в данный момент для полной связности портов [1].

Обнаружение отказа элемента сети обеспечивается посредством использования таймеров возраста служебных сообщений. При получении этого сообщения таймер устанавливается в исходное состояние. Если какой-либо назначенный узел, его порт или линия связи отказывают, то служебные сообщения перестают поступать во все узлы, которые находятся в поддеревьях отказавшего элемента. Вследствие этого, в этих узлах истекает интервал таймера возраста сообщения и вызывается процедура реконфигурации сети.

Модель функционирования коммутатора в этом случае можно представить в виде процесса перехода между различными состояниями, параметры которого определяются вероятностью отказа каналов связи $P_{\text{отк}}(t)$ и вероятностью их восстановления $P_{\text{вост}}(t)$ за время t . Если сеть отказала с вероятностью $P_{\text{отк}}(t)$ то она переходит из работоспособного состояния S_0 в состояние S_1 . Состояние S_1 характеризуется непринятым решением о реконфигурации, когда недоставка служебного сообщения считается сбоем и информация продолжает направляться в соответствии со старой конфигурацией STP, что приводит к ее потере. Таким образом, состояние S_1 является худшим вариантом состояния системы. В случае, когда канал связи восстановил свою работоспособность (произошло событие восстановления) за время $T_{\text{Max age}}$ система возвращается в работоспособное состояние $S_1 \rightarrow S_0$ или если за время $T_{\text{Max age}}$ восстановление канала связи не произошло система производит реконфигурацию протокола $STP_{S_1} \rightarrow S_2$ и вновь переходит в работоспособное состояние $S_2 \rightarrow S_0$. Таким образом, состояния системы соответствуют следующим процессам:

S_0 – сеть работоспособна, коммутаторы работоспособны;

S_1 – необнаруженный отказ канала связи в сети, коммутатор сохраняет свою конфигурацию;

S_2 – реконфигурация коммутатора в соответствии с новой топологией сети.

Процессы переходов в различные состояния, введя допущения о пуассоновском процессе отказов каналов (например, в результате дестабилизирующего воздействия) и процессе их восстановления (в результате функционирования адаптивной системы связи) возможно преобразовать к вложенному Марковскому процессу, который описывается известным процессом отказа – восстановления из теории надежности.

Для аппроксимации времени восстановления работоспособности сети $\tau_{\text{вост}} = 10\text{--}70$ с, в процессе моделирования использовалось экспоненциальное распределение плотности вероятности времени восстановления с интенсивностью $\mu_{\text{вост}} = m(\tau_{\text{вост}}) = 0,04$, что соответствует $\tau_{\text{вост}} = 25$ с.

Практический интерес при моделировании представляет вероятность нахождения системы в работоспособном состоянии $P_0(S_0)$ на общем времени функционирования, а также изменение значения необнаруженного отказа $P_1(S_1)$.

Как показывает анализ результатов моделирования максимальное влияние на поведение системы оказывает параметры интенсивности отказов $\lambda_{\text{отк}}$ и таймера сохранения текущей конфигурации $T_{\text{Max age}}$.

Уже при $\lambda_{\text{отк}} = \mu_{\text{вост}}$ система находится в работоспособном состоянии с вероятностью $P_0(S_0) = 0,5$, а дальнейшее увеличение интенсивности отказов $\lambda_{\text{отк}}$ выше данного значения ведет к фактическому неработоспособному состоянию системы (так при $\lambda_{\text{отк}} = 2 \mu_{\text{вост}}$, вероятность работоспособного состояния составляет $P_0(S_0) \approx 0,3$). Необходимо отметить, что самым неблагоприятным состоянием системы является состояние S_1 «необнаруженного отказа». Значение таймера сохранения текущей конфигурации $T_{\text{Max age}}$ напрямую влияет на предельное значение вероятности состояния S_1 и скорость его достижения. Установлено также, что даже в случае гипотетической ситуации мгновенного обнаружения отказа канала связи ($T_{\text{Max age}} = 0$) в сети связи не произойдет потери пакетов, но процессы передачи данных в сети будут остановлены на время реконфигурации $2T_{\text{forward delay}}$.

Литература

1. Perlman R.J. An algorithm for distributed computation of a spanning tree in an extended LAN. // SIGCOMM. – 1985. – P. 44–53.

УДК 004

УПРАВЛЕНИЕ ТРЕБОВАНИЯМИ В ПРОЕКТИРОВАНИИ СПЕЦИАЛИЗИРОВАННОГО МОДУЛЯ ГОЛОСОВОЙ БИОМЕТРИИ

Е.Б. Мишкарева

Научный руководитель – к.т.н. М.В. Хитров

Современные тенденции развития индустрии информационных технологий позволяют разрабатывать, создавать решения практически для любой сферы применения. Но в ходе разработки любого программного продукта, от настольного приложения до крупномасштабной корпоративной системы, подразумевается сбор, анализ и управление требованиями, что является процессом длительным и сложным, в который вовлечен большой круг заинтересованных лиц. Работа с требованиями предполагает определение круга заинтересованных лиц, принятие во внимание всех их потребностей и гарантия того, что они понимают значение разрабатываемых систем.

Опыт индустрии информационных технологий однозначно показывает, что вопросы, связанные с управлением требованиями, оказывают критически важное влияние на программные проекты, в определенной степени – на сам факт возможности успешного завершения проектов.

Одной из тенденций последнего времени является создание конечного программного продукта при помощи готовых («off-the-shelf») модулей, приобретаемых вместе с технологиями. Такой подход позволяет в значительной мере сократить издержки и рабочий цикл на создание конечного программного продукта. В условиях современной экономики это один из важнейших факторов, диктуемый необходимостью конкурировать и предоставлять, в конце концов, продукт, отвечающий требованиям заказчика.

Указанная тенденция дает возможность монополизировать все выгоды от разработки новой технологии или продукта за счет быстрого выполнения проекта без привлечения больших производственных мощностей.

Разработка специализированного модуля голосовой биометрии также является нетривиальным вопросом, требующим детальной проработки и особого подхода. Специализированный модуль голосовой биометрии – это, что важно понимать, не

законченный продукт, т.е. не готовая информационная система, а элемент, который может применяться системным интегратором в качестве модуля решения для конечного заказчика.

Модуль голосовой биометрии предназначен для встраивания в программное обеспечение с целью построения моделей по фонограммам и проведения процедур идентификации и верификации диктора посредством сравнения построенных моделей.

Модуль может применяться для создания, как полностью автоматизированных многопоточных решений высокой производительности, так и одноканальных экспертных продуктов с высокими качественными характеристиками распознавания диктора. Таким образом, условия использования, или UseCases, могут быть абсолютно различными.

Работа с требованиями для таких продуктов имеет свои особенности и отличия от анализа требований готовых, «завершенных», продуктов.

Одной из особенностей является необходимость предусмотреть все возможные сценарии использования и дать «интегратору» всесторонние рекомендации по встраиванию модуля в решения для конечного заказчика, а также предусмотреть возможность встраивания модуля, как для работы в рамках десктоп-приложений, так и в масштабируемых корпоративных системах.

УДК 004.822

ПОДХОД К СОВМЕСТНОЙ РАЗРАБОТКЕ ОНТОЛОГИЙ

А.В. Невидимов

Научный руководитель – к.т.н., доцент И.А. Бессмертный

В настоящее время все более актуальными становятся семантические технологии. Созданий баз знаний в виде онтологий позволяет уменьшать временные затраты людей на обработку информации и принятие решений практически во всех предметных областях.

Для того чтобы создать качественную онтологию, разумным представляется задействовать не одного, а нескольких экспертов выбранной предметной области. Однако их совместная работа сопровождается следующими трудностями:

- онтологии одной и той же предметной области, созданные разными экспертами, могут значительно различаться. Отсюда возникает задача обеспечения слияния онтологий.
- часто разрабатываемая онтология является источником знаний для некоторого программного обеспечения. Процессы развития онтологии и использующего ее ПО могут идти параллельно. Таким образом, необходимо предоставить ПО возможность использовать то состояние используемой онтологии, с которой оно совместимо.

В работе предлагается подход, основанный на идеологии систем контроля версий с центральным репозиторием (таких, как git, mercurial, SVN). Приведем список ключевых правок онтологии в таком подходе с их описанием и особенностями:

1. добавление новых триплетов вида «Сущность-объект, Предикат, Сущность-субъект» не нарушает обратной совместимости онтологии с ее предыдущей версией, так как все множество фактов, которые могли быть получены из старой версии онтологии, могут быть получены и из новой. Важным является также необходимость определять случаи параллельного добавления разными экспертами схожих сущностей для исключения дублирования данных в онтологии;
2. удаление старых триплетов в некоторых случаях может привести к нарушению обратной совместимости. Факт наличия или отсутствия обратной совместимости предлагается определять путем попытки логического вывода всех удаленных триплетов из новой онтологии, и если хотя бы один триплет не может быть выведен – считаем выявленным

нарушение обратной совместимости. Также удаление триплетов может привести к невалидности правок, вносимым другим экспертом;

- удаление элементов из множества сущностей или множества отношений всегда приводит к нарушению обратной совместимости.

Таким образом, перечисленные задачи будут решены следующим образом:

- перед внесением в основную ветку онтологии набора правок они будут проверяться на совместимость с текущим состоянием онтологии, в случае несовместимости автору правок будет предложено внести соответствующие коррективы. Определять схожие сущности, параллельно добавленные разными экспертами, предлагается с помощью вычисления расстояния Левенштейна между их названиями, либо на основе данных ресурса WordNet о синонимичности понятий;
- при каждом нарушении обратной совместимости в основной ветке разработки онтологии будет публиковаться новый URL, являющийся SPARQL-endpoint, что обеспечит прозрачную для ПО пользоваться последней версией онтологии, с которой оно совместимо.

УДК 004.087.2, 004.056.3, 004.651

ФАЙЛОВАЯ СИСТЕМА С РАЗДЕЛЕНИЕМ ОБНАРУЖЕНИЯ И КОРРЕКЦИИ ОШИБОК ДЛЯ NAND FLASH НОВЫХ СЕМЕЙСТВ

М.Г. Нестеров

Научный руководитель – ст. преподаватель С.В. Бибииков

До недавнего времени чипам NAND FLASH памяти хватало встроенной коррекции ошибок 1 бит на 512 байт, однако с уменьшением производственного техпроцесса с сотен нанометров до близких к 10 нм стали возникать проблемы.

Как известно, чипы памяти NAND имеют ограниченное количество циклов перезаписи (Program/Erasecycles), и с уменьшением размера одного элемента (в данном случае транзистора) эта характеристика значительно ухудшается. Кроме того, такой элемент более подвержен внешним воздействиям, хотя бы потому, что хранимый заряд на плавающем затворе транзистора становится меньше, а значит, его легче изменить в ту или иную сторону. При этом разброс пороговых значений заряда (по которым, в простейшем случае, контроллером принимается решение о значении снятого бита информации: «0» или «1») также становится меньше. Точность измерения снятого с транзистора заряда контроллером, безусловно, повышается, однако она, во-первых, конечна, во-вторых, сам контроллер становится сложнее, а значит, повышается вероятность его сбоя.

Исходя из вышеизложенного, поставлена **цель работы**: компенсировать ухудшающуюся надежность NANDSLCFlash памяти путем написания и последующего использования специализированной файловой системы. Минимальные требования: обнаружение и коррекция до 8 ошибок на 512 байт, ошибки могут быть как одиночные, так и пакетные (серия из нескольких последовательных ошибок).

В первую очередь было рассмотрено текущее положение отрасли, как в техническом плане, так и в плане ПО. Из рассмотренного материала стоит выделить следующие факты:

- при техпроцессе менее 32 нм требуется коррекция более 4 бит на 512 байт, по данным компании Toshiba (являющейся создательницей Flash-памяти);
- многие компании провели свои исследования, которые также показывают, что количество ошибок будет тем больше, чем миниатюрнее будет транзистор (по данным MicrosoftResearch, Micron).

После этого мы рассмотрели хорошо известные и проверенные временем подходы к

обнаружению и коррекции ошибок. В первую очередь, это стандартный ЕСС, основанный на коде Хэмминга. Однако этот код хорош только для малого количества ошибок и имеет недетерминированную вероятность не обнаружить ошибку, в случае если ошибок будет больше чем заложено в конкретную реализацию кода. Более того, в случае если необходимо обнаруживать 8 ошибок длина блока данных ограничивается всего ничего порядка 50 битами (при этом исправлять такой код будет менее восьми ошибок).

Далее были рассмотрены код БЧХ и его разновидность – код Рида-Соломона. Несомненно, у них есть свои плюсы и минусы, и на некоторое время они стали фаворитами. Из явных плюсов: код Рида-Соломона применим при большом количестве одиночных и пакетных ошибок (с некоторыми ограничениями последних), есть готовые хорошие реализации кода Рида-Соломона, обладающие высокой производительностью, остается только применить «в нужных местах» архитектуры файловой системы.

В конце концов, была обнаружена работа, предлагающая альтернативный подход к решению поставленной задачи. Его название: «ReliableCompressedFlashFileSystem» (RCFFS, «Надежная со Сжатием Файловая для Флэш-памяти Система»). В чем его суть: обеспечение целостности данных состоит из двух процессов – обнаружения ошибок и коррекции ошибок. В устоявшемся подходе (код Хэмминга) эти процессы обеспечивает один и тот же код, т.е. он берет на себя две обязанности. Новый подход предлагает отдельный аппарат для каждого процесса: алгебраические сигнатуры для обнаружения ошибок и код Рида-Соломона для их коррекции.

Кроме того, стоит отметить одно важное решение, принятое в этом подходе, применительно к NANDFlash памяти: обычно для хранения кода Хэмминга используется небольшая специальная область памяти в каждом блоке Flash памяти, так называемая sparearea. Ни алгебраические сигнатуры, ни код Рида-Соломона не привязываются к этой области: сигнатуры расположены в конце блока, а код Рида-Соломона занимает целый отдельный блок.

Это самые важные принципы. В целом, предлагаемая их реализация учитывает особенности NANDFlash памяти: запись данных блоками и ограниченное количество циклов перезаписи.

На данный момент создается реализация предлагаемой файловой системы на языке C++ чтобы затем проверить заложенные характеристики на тестовом стенде. Уже готов и в какой-то мере протестирован класс создания алгебраических сигнатур, оперирующий байтами. Вероятно, в будущем будет использоваться также реализация, вычисляющая сигнатуру по словам (два байта) – от длины базового элемента зависит значение максимального количества этих элементов (255 для байтов и 65535 для слов). В процессе находится реализация самого базового уровня, который имеет смысл в данной ФС – страница Flash памяти.

РАЗРАБОТКА ПРОГРАММНОГО КОМПЛЕКСА ДЛЯ СОПРОВОЖДЕНИЯ УДАЛЕННОГО МОДУЛЯ «ТАБЛИЧНЫЙ ПРОЦЕССОР»

А.В. Никитин

Научный руководитель – д.т.н., профессор Д.Д. Куликов

В настоящее время автоматизированные системы технологической подготовки производства (АСТПП) подразумевают под собой комплекс программного обеспечения (ПО), который нуждается в соответствующем обслуживании и поддержке: покупка лицензий, установка ПО, последующие обновления и так далее. Предприятия вынуждены тратить большое количество времени, сил и материальных средств. Такие затраты могут себе позволить далеко не все предприятия и компании.

Решением этой проблемы могут стать удаленные web-приложения. Необходимо исследовать способы создания веб-сервисов и их сопровождения. Для системы «ТИС-ТАП», входящей в состав технологической интегрированной системы (ТИС), необходимо создать удаленное приложение, которое позволит работать с АСТПП, имея только доступ в интернет.

SaaS-технология – бизнес-модель продажи и использования программного обеспечения, при которой поставщик разрабатывает веб-приложение и самостоятельно управляет им, предоставляя заказчикам доступ к программному обеспечению через интернет. Основное преимущество модели SaaS для потребителя состоит в отсутствии затрат, связанных с установкой, обновлением и поддержкой работоспособности оборудования и работающего на нем программного обеспечения.

Программный модуль «Табличный процессор ТИС-ТАП» предназначен для формирования и сопровождения базы данных и знаний для расчета режимов резания и нормирования, а также принятия решений на основе таблиц соответствия.

На данном этапе ведется работа над каталогом создаваемой системы, а также определение способов переработки системы «ТИС-ТАП» под веб-сервис.

В этом докладе описывается процесс разработки каталога системы, организация поиска по каталогу и редактированию записей. Показана структура системы и предполагаемый графический интерфейс удаленного приложения.

В конечном счете, будет создана система, позволяющая работать с «ТИС-ТАП» через интернет. Будет организован авторизованный доступ к системе.

ИССЛЕДОВАНИЕ МЕТОДОВ СТЕГОАНАЛИЗА ЦИФРОВЫХ ИЗОБРАЖЕНИЙ

А.С. Николаева

Научный руководитель – д.т.н., профессор А.Ю. Тропченко

В работе рассматривается проблема выявления стеганографического скрытия информации в файлах формата JPEG. Исследуются распространенные методы обнаружения сторонней информации в цифровых изображениях, основанные как на использовании статистических характеристик, так и на особенностях формата JPEG.

Цель работы – определение эффективного метода стегоанализа цифровых изображений формата JPEG.

Формат JPEG является наиболее распространенным среди остальных форматов изображений и широко используется для решения задач стеганографии. Это делает актуальной проблему разработки методов стегоанализа данных в рассматриваемом формате.

Большинство исследований в области стегоанализа направлено на решение основной задачи: определение факта наличия скрытой информации. Для этого применяется многообразный инструментарий: от визуального анализа до различных статистических подходов.

В случае с JPEG, визуально определить присутствие скрытой информации невозможно. Даже если в результате скрытия возникают незначительные искажения изображения, их можно объяснить изменением степени его сжатия. При использовании файла формата JPEG в качестве контейнера, встраивание информации, как правило, производится в наименее значащие биты (НЗБ) дискретных косинусных коэффициентов (ДКП коэффициентов).

Задача стегоаналитика состоит в выявлении различий между теоретической моделью и исследуемым контейнером, на основе которых определяется вероятность наличия секретного сообщения.

В работе исследовались различные методы выявления скрытой информации:

- на основе критерия хи-квадрат;
- на основе «блочного» хи-квадрат;
- на основе оценки числа переходов значений младших бит в соседних квантованных ДКП коэффициентах;
- на основе анализа гистограмм, построенных по частотам квантованных ДКП коэффициентов;
- на основе анализа частот переходов битовых значений ДКП коэффициентов.

Метод обнаружения сторонней информации на основе критерия хи-квадрат [1] подходит для анализа изображений со встроенной информацией различными стеганографическими методами, т.е. является универсальным. Однако результаты стегоанализа в значительной мере зависят от способа скрытия данных. Приемлемые результаты достигались данным методом при последовательной записи в НЗБ коэффициентов ДКП, а при псевдослучайном выборе младших бит и распределении сообщения по всему контейнеру, метод хи-квадрат не срабатывал.

Предложенный в работе [2] вариант «блочного» хи-квадрат, анализирующего JPEG-файл не целиком, как в предыдущем методе, а по частям, показал хорошие результаты при анализе изображений, использующих псевдослучайные алгоритмы выбора коэффициентов ДКП для скрытия в них информации.

Метод, основанный на анализе числа переходов значений младших бит в соседних квантованных ДКП коэффициентах и описанный в [3], эффективен при реализации последовательного скрытия в НЗБ ДКП коэффициентов, при этом метод способен надежно выявлять факт стеганографического скрытия даже при незначительном заполнении контейнера. Данный метод следует применять в качестве дополнения к другим методам стегоанализа.

Метод, анализирующий гистограммы частот ДКП коэффициентов в JPEG-файлах [4], позволяет достаточно надежно выявлять присутствие информации, в то время, когда методы по критерию хи-квадрат и «блочному» хи-квадрат данную информацию не выявляют.

Результаты работы метода на основе анализа частот переходов битовых значений ДКП коэффициентов зависят от метода скрытия данных, а также от их объема. Выявление факта скрытия осуществимо при заполнении изображения-контейнера более чем на 60 процентов. Данный метод имеет право на существование при выявлении значительных объемов скрываемой информации.

Комплексное применение различных методов стегоанализа позволяет более гибко подходить к вопросу о возможном существовании встроенной информации, свести к минимуму вероятности ошибок первого и второго рода (случаи, когда пустой контейнер принимается за заполненный и заполненный контейнер за пустой, соответственно) при обнаружении стеганографического скрытия информации, обеспечить наглядность

результатов анализа (графики, гистограммы).

Литература

1. Грибунин В.Г., Оков И.Н., Туринцев В.И. Цифровая стеганография. – М.: СОЛОН-Пресс, 2002. – 272 с.
2. Голуб В.А., Дрюченко М.А. Комплексный подход для выявления стенографического скрытия в JPEG-файлах // Инфокоммуникационные технологии. – 2009. – № 1. – Т. 7. – С. 44–50.
3. Барсуков В.С., Романцов А.П. Оценка уровня скрытности мультимедийных стеганографических каналов хранения и передачи информации // Специальная Техника. – 2000. – № 1.
4. Fridrich J., Goljan M., Hogeя D. Steganalysis of JPEG Images: breaking the F5 algorithm // 5th Information Hiding Workshop, Noordwijkerhout, The Netherlands. – 2002. – P. 310–323.

УДК 351/354: 004.01

СРАВНИТЕЛЬНЫЙ АНАЛИЗ СИСТЕМ ЭЛЕКТРОННОГО ДОКУМЕНТООБОРОТА

А.С. Новиков

Научный руководитель – к.ф.-м.н., доцент С.Е. Иванов

В наше время владение, оперативность доступа к достоверной актуальной и объективной информации – это основа успеха в любой сфере. Эффективность организации процессов обмена оперативной информацией на предприятии становится актуальной для его развития. С ростом документации возникает необходимость большей формализации бизнес-процессов и перевода документооборота в электронный вид. Так формируется понятие системы электронного документооборота.

Развитие рынка СЭД в России началось позже, чем в развитых странах. Во многом это обусловлено поздним на фоне общемировой картины переходом к рыночной модели экономики. Однако проблема создания таких систем и их внедрения на предприятиях с ростом объемов информации оказалась актуальной. Объем рынка СЭД в России в 2008 году составил около \$210 млн. На долю зарубежных разработчиков платформ и решений приходится более 50% от рынка СЭД, однако по количеству внедрений соотношение явно не в пользу иностранных разработок.

Ввиду высокой стоимости лицензии зарубежных программ, а также недешевой и трудоемкой адаптации под требования российских стандартов они используются в основном в крупных компаниях и госучреждениях, для которых вопрос цены стоит не столь остро. Предприятия среднего и малого бизнеса больше склонны применять отечественные разработки из-за их гибкости и дешевизны. Однако и крупные компании все чаще обращаются к российским СЭД. Чем дальше, тем большую часть рынка, вероятно, будут завоевывать отечественные системы: с каждым годом они становятся функциональнее, существенно менее затратны во внедрении и, что очень важно, разрабатываются специально под российские реалии.

В настоящее время существует множество систем электронного документооборота, способных обеспечить повышение эффективности управления предприятием, добиться более полной прозрачности данных, повысить уровень управления предприятием

В работе рассмотрены основополагающие понятия электронного документооборота, задачи, которые преследует СЭД, предпосылки и цели для внедрения СЭД, этапы внедрения, платформы и СУБД. Будет представлен сравнительный анализ существующих систем электронного документооборота для предприятий различных видов деятельности, а также

предпосылки развития требований к системам электронного документооборота.

УДК 004.414.23

ПРИМЕНЕНИЕ ГЕНЕРАТОРОВ ТРАФИКА В ИМИТАЦИОННЫХ МОДЕЛЯХ КОМПЬЮТЕРНЫХ СЕТЕЙ

П.А. Нужина

Научный руководитель – д.т.н., профессор Т.И. Алиев

В настоящее время важной задачей является контроль работоспособности компьютерных сетей. Одним из средств диагностики сетевых проблем являются генераторы трафика (ГТ), т.е. программы, позволяющие создавать сформированные заданным образом пакеты различных уровней модели OSI, за исключением физического. Данный тип программ используется системными администраторами, сетевыми инженерами и другими специалистами в области компьютерных сетей. Необходимость использования данного типа программ ставит задачу оптимального выбора ГТ из большого количества различных вариантов. Однако, несмотря на широкую распространенность, ГТ недостаточно документированы, поэтому на сегодняшний день эта задача не имеет четкого решения.

Использование генераторов целесообразно как в уже функционирующих сетях, так и в моделируемых сетях. Имитационное моделирование в применении к компьютерным сетям может использоваться на этапе предварительного проектирования для анализа различных вариантов построения сети и выбора оптимального варианта. Именно на этом этапе применение генераторов трафика особенно эффективно для проведения оценки построенной модели.

В данной работе рассматриваются возможности существующих ГТ, в том числе возможности использования генераторов в имитационных моделях компьютерных сетей и особенности их использования. **Целью работы** является разработка метода выбора оптимального ГТ. Для достижения указанной цели нужно определить набор показателей эффективности, используемых для сравнения генераторов. Проведенный сравнительный анализ ГТ показал, что наиболее существенными являются следующие показатели эффективности:

- количество операционных систем (ОС), в которых может использоваться генератор;
- тип лицензии, предоставляемый пользователям генератора;
- интерфейс (графический или интерфейс командной строки);
- язык программирования, на котором написан генератор (язык может влиять на способ описания трафика; например, *Scapy* написан на Python – и это позволяет использовать конструкции данного языка при создании пакетов);
- количество уровней OSI-модели, пакеты которых может генерировать программа;
- поддерживаемые форматы дампов трафика (например: .pcap, .fws, .xml);
- используемые протоколы (например: ARP, Ethernet, FTP, HTTP, ICMP, IGMP, IP, OSPF, PPP, RIP, SMTP, TCP, UDP);
- количество параметров, используемых для описания генерируемого трафика (закон распределения времени между пакетами в трафике, размер пакетов, скорость передачи, адреса отправителя, назначения, порт источника, назначения, MAC-адрес, фрагментация).

В работе рассматривались следующие генераторы пакетов: Anettest, Bittwist, CatKaratPacketBuilder, ColasoftPacketBuilder, Mauzezahn, Nemesis, Ostinato, PackETH, Pierf, Pktgen, Scapy, UMPA, hping. Большинство из них можно использовать как в ОС Microsoft Windows, так и в *nix-подобных системах. Некоторые ориентированы на использование только в ОС Windows (CatKaratPacketBuilder, ColasoftPacketBuilder) или

только в *nix (Mauzezahn, Pktgen). Другие являются кросс-платформенными (Nemesis, Ostinato, UMPA, hping).

Генераторы CatKaratPacketBuilder, ColasoftPacketBuilder и UMPA имеют графический интерфейс, остальные – интерфейс командной строки, а PackETH комбинирует оба варианта интерфейса.

В основном все генераторы поддерживают большое количество разных протоколов с возможностью создания пользовательских заголовков. Так, в CatKaratPacketBuilder есть готовые шаблоны для 20 протоколов, в то время как ColasoftPacketBuilder по умолчанию поддерживает только Ethernet, ARP, IP, TCP, UDP, на основе которых с помощью встроенного редактора можно создавать пользовательские пакеты. Однозначно все рассматриваемые генераторы поддерживают Ethernet, ARP, IP, ICMP, TCP, UDP.

В работе продемонстрирована возможность проведения особого типа тестирования под названием Fuzzing с помощью генератора Scapy, а также показано, как этот генератор можно подключать в качестве модуля к другим утилитам, в том числе к собственноручно созданным. Проведение такого тестирования позволяет обнаруживать дефекты в программах при передаче ей случайных данных. Установлено, что из всех рассматриваемых генераторов данной возможностью обладает только генератор Scapy.

Проведенный сравнительный анализ ГТ позволил выявить важнейшие показатели эффективности ГТ и предложить метод выбора оптимального генератора в зависимости от области его применения, которая определяет требования, предъявляемые к ГТ.

УДК 004.422.83:004.6:004.72

РАЗРАБОТКА И ВНЕДРЕНИЕ МОДУЛЕЙ ДЛЯ СИСТЕМЫ МОНИТОРИНГА СОСТОЯНИЯ СЕТИ И СЕРВИСОВ

Е.С. Образцов

Научный руководитель – к.т.н., доцент А.С. Бузинов

Мониторинг является одним из важных критериев при обеспечении информационной безопасности и защиты пользовательских данных. В современном мире информационных технологий одним из основных критериев пользователей к системам является критерий стабильности работоспособности систем, которыми они пользуются. Компьютер, сеть между компьютерами или же какая-либо программа, написанная на одном из множества языков программирования, подвержены периодическим сбоям, перегрузкам или влияниям извне. Для того чтобы оперативно исправить ту или иную проблему, возникшую в процессе работы какой-либо системы, необходим непрерывный надзор за этой системой – мониторинг. Помимо этого в компьютерном мире существует несчетное количество разнообразных сервисов. Эти сервисы играют значительную роль в функционировании компьютеров. Сервисы операционной системы обеспечивают правильную работу компьютера. Сетевые сервисы помогают компьютерам общаться между собой. Мониторинг этих сервисов и своевременное восстановление их работоспособности – важная задача для компании любых размеров. Помимо сервисов также существует необходимость вести мониторинг различного рода программ, имеющих доступ к пользовательским данным или данным, являющимся конфиденциальными.

В качестве защиты информации мониторинг подойдет как нельзя, кстати, ведь он может стать средством непрерывного наблюдения за необходимыми данными и сигнализировать в случае их кардинальных изменений.

Актуальность реализации задачи мониторинга состоит в том, что:

1. повышается качество обслуживания сети и сервисов;

2. предотвращаются катастрофические ошибки;
3. производится количественный анализ качества работы сети и сервисов.

Предлагаемая система мониторинга в совокупности с дополнительными модулями позволяют заметно улучшить качество обслуживаемой локальной сети и повысить время реагирования на появившуюся проблему. Они могут использоваться на различных предприятиях с локальными сетями различного масштаба. Минимальные требования к аппаратной части делают данное решение доступным, а лицензионность и открытость программного кода операционной системы, системы мониторинга и дополнительных модулей позволяют не только снизить затраты на программные продукты, но и дают возможность увеличения функциональных возможностей.

УДК 004.891

ПРОЕКТИРОВАНИЕ ГИБРИДНОЙ ЭКСПЕРТНОЙ СИСТЕМЫ ПОДДЕРЖКИ ПРИНЯТИЯ РЕШЕНИЙ

Н.А. Полковникова

(Таганрогский кампус Южного федерального университета)

Научный руководитель – д.т.н., профессор В.М. Курейчик

(Таганрогский кампус Южного федерального университета)

Краткое вступление, постановка проблемы. По оценкам специалистов, причинами 80% аварий в море являются неправильные действия судового персонала, неправильные решения или ошибки при реализации принятого решения. На современном автоматизированном судне механик анализирует режимы работы и техническое состояние энергетического оборудования, выполняя функции оператора. При проведении контроля и постановки диагноза обслуживающий персонал вырабатывает решения на основе поступающей к нему информации. Чрезмерное увеличение потока сведений затрудняет работу оператора. Например, на танкере «NSLeader» СКФ «Новошип» контролируется до 300 параметров. Это порождает проблему выбора оптимального объема информации, поступающей к обслуживающему персоналу в процессе эксплуатации судовых энергетических установок. Экспериментально установлено, что пропускная способность (ПС) человека, т.е. максимальная скорость, с которой он может воспринимать и передавать информацию (бит/сек), равна 3–5 бит/сек. При ПС = 5–9 бит/сек информационная нагрузка несколько завышена, но поскольку она соответствует объему оперативной памяти (7 +/- 2 бит/сек), то вполне допустима. Зоной перегрузки является 10...100 бит/сек, в то время как скорость принятия сложных решений даже меньше скорости приема и запоминания информации, равную 1–9 бит/сек. На современных энергетических установках потоки информации могут достигать 250 бит/с. Приведенные данные вызывают опасения, что в процессе эксплуатации могут возникнуть ситуации, когда обслуживающий персонал не будет в состоянии перерабатывать всю поступающую к нему информацию. Если скорость поступления информации к оператору превышает ПС, то из-за перегрузки человек не сможет воспринимать всю информацию, будет допускать ошибки, что может привести к принятию неверных решений.

Экспертная система может полностью взять на себя функции, выполнение которых обычно требует привлечения опыта человека-специалиста, или играть роль ассистента для человека, принимающего решение. Человек, работающий в сотрудничестве с такой программой, может добиться с ее помощью результатов более высокого качества. Правильное распределение функций между человеком и машиной является одним из ключевых условий высокой эффективности внедрения экспертных систем.

Целью работы является разработка гибридной экспертной системы, в которой учтены следующие эргономические требования:

- информация ограничена только тем, что необходимо механику для принятия решений и соответствует пропускной способности человека;
- информация отображается только с такой точностью и степенью детализированности, какая требуется механику;
- информация отображается в форме, непосредственно пригодной к использованию.

Промежуточными результатами выполненной работы являются разработка:

1. базы знаний, предназначенной для хранения экспертных знаний, используемых при решении задач экспертной системой;
2. базы данных, предназначенной для хранения фактов или гипотез, являющихся результатом общения системы с человеком, ведущим диалог с экспертной системой;
3. подсистемы общения для ведения диалога с пользователем, в ходе которого экспертная система запрашивает у пользователя необходимые данные;
4. подсистемы объяснений, необходимую для обеспечения возможности пользователю контролировать ход рассуждений и, может быть, учиться у экспертной системы;
5. подсистемы приобретения знаний для корректировки и пополнения базы знаний.

При разработке экспертной системы была использована методология разработки программного обеспечения Rational Unified Process фирмы Rational Software Corporation. Проектирования экспертной системы производилось с помощью CASE-средства Rational Rose. Первым действием пользователя является ввод данных по результатам индентификации двигателя, затем экспертной системой производится расчет, построение графиков; вывод неисправностей с рекомендациями по их устранению, создание отчета и добавление новых неисправностей в базу данных. Следующим этапом проектирования в среде RationalRose было создание диаграммы взаимодействия модулей приложения (mainlogicalview). Минимальным элементом данной схемы является пакет (package). Каждый пакет соответствует программному модулю экспертной системы. На этом этапе было определено, что из главного окна программы будет доступен весь функционал экспертной системы и возможности каждого модуля. Далее в RationalRose была произведена разработка прототипов модулей экспертной системы в виде диаграмм классов.

В результате была получена объектная модель системы. Все классы и диаграммы, описывающие деятельность системы были спроектированы в Rational Rose с помощью программы Rose Delphi Link. Таким образом, при периодическом согласовании на основе технологии прямого и обратного проектирования обеспечивается соответствие модели Rose и кода Delphi, что позволяет разработчику иметь актуальные документированные модели, отражающие действительное состояние разработки проекта.

Основным результатом работы является разработанная экспертная система, которая обеспечивает математическую обработку результатов измерений, ретроспективный анализ информации, а разработанный интерфейс обеспечивает интерактивный режим работы с компьютером. Оператор (механик) получает информационную поддержку и рекомендации при решении задач технической эксплуатации главных и вспомогательных судовых дизелей, что позволяет обосновать и правильно принимать решения. Особенностью архитектуры экспертной системы является баланс быстродействия, гибкости и понятности.

Литература

1. Полковникова Н.А., Курейчик В.М., Полковников А.К. Экспертная система для диагностики судовых дизелей «Дизель эксперт». Свидетельство о гос. регистрации № 2012615935 от 28.06.2012 // Официальный бюллетень Федеральной службы по

интеллектуальной собственности, патентам и товарным знакам «Программы для ЭВМ, базы данных, топологии интегральных микросхем» RU ОБПБТ. – 2012. – №3(80). – С. 683–684.

2. Гаврилов А.В. Гибридные интеллектуальные системы: Монография. – Новосибирск: Изд-во НГТУ, 2002. – 142 с.

УДК 004.051+004.054

ПРИНЦИПЫ ПОСТРОЕНИЯ JAVA БЕНЧМАРКОВ

М.В. Пономарев

Научный руководитель – к.т.н., доцент М.С. Косяков

Высокая производительность программ всегда является целью разработчиков программного обеспечения, даже в наш век многоядерных процессоров и гигабайтов оперативной памяти. В этой связи измерение показателей производительности (бенчмаркинг) с целью выявления наиболее эффективных решений, является важной задачей программистов. Наиболее часто измеряемым показателем является скорость выполнения программ.

Написание бенчмарков (тестовых программ для измерения производительности) на современных языках программирования, выполняющихся на виртуальной машине, таких как Java, является нетривиальной задачей. Измерение скорости выполнения отдельных участков кода связано с большим количеством «подводных камней», приводящих к получению неадекватных и невозпроизводимых результатов. Чаще всего источником ошибок является неучитывание сложного поведения виртуальной машины при разработке бенчмарков. Данный вопрос мало освещен в литературе. В ходе исследования были рассмотрены причины получения неудовлетворительных результатов и принципы построения бенчмарков на Java, позволяющие получить адекватные и прогнозируемые результаты. Большая часть исследуемых принципов является общим для всех языков, исполняемых на виртуальной машине.

Первой и основной причиной получения неудовлетворительных результатов при написании бенчмарков является непринятие во внимание разработчиками такой особенности виртуальной машины Java (англ. JavaVirtualMachine, JVM), как смешанный режим. В этом режиме JVM использует технологию компиляции «на лету» (англ. Just-in-time compilation, JIT). JIT – технология увеличения производительности программ, использующих байт-код, заключающаяся в компиляции наиболее критичных к производительности участков из байт-кода в машинный код непосредственно во время выполнения и интерпретации остальной части программы. Это позволяет JVM снизить накладные расходы на компиляцию при старте программы и потратить больше времени на компиляцию и оптимизацию наиболее чувствительных к производительности мест. Выбор участка байт-кода, который будет скомпилирован, осуществляется на основе различных эвристических методов, зависящих от конкретной реализации виртуальной машины, но в целом в их основе лежит статистический (в машинный код компилируются наиболее часто выполняющиеся участки) и вероятностный (в машинный код компилируются участки, выполнение которых наиболее вероятно) принципы.

Второй причиной потери производительности является загрузка классов во время выполнения программы при первом обращении к ним.

В связи с вышесказанным, в случае если бенчмарк с целью измерения скорости выполнения выполняет интересующий программиста с точки зрения производительности код только один раз, будут получены заведомо плохие результаты, поскольку код будет

интерпретироваться, в то время как в реальной системе участок программы, наиболее вероятно, будет выполняться часто и, как следствие, полностью или частично оптимизирован JIT.

Во избежание вышеописанной проблемы достаточно «прогреть» JIT-механизм, т.е. выполнить код достаточно большое число раз «в холостую» (без замеров времени выполнения) и тем самым гарантировать с большой долей вероятности, что код будет оптимизирован JIT. Прогрев также поможет избежать потери производительности во время загрузки классов при первом обращении к ним.

При построении бенчмарка необходимо избегать «неожиданных» для виртуальной машины путей выполнения, таких как редко выполняющиеся ветки условного оператора или выбросы исключений. Это может спровоцировать деоптимизацию кода, т.е. удаление скомпилированного машинного кода и возврат к интерпретации байт-кода, что в свою очередь неизбежно повлияет на качество результатов.

Другой частой причиной ошибочных результатов при бенчмаркинге является выбор несоответствующих средств измерения времени выполнения. Очень часто программисты используют стандартный вызов `System.currentTimeMillis`, зависящий от внешних событий и предназначенный для получения текущего времени, вместо более подходящего `System.nanoTime`, специально разработанного для измерения количества времени между двумя событиями и обладающего большей точностью.

В работе представлены результаты экспериментов, демонстрирующие эффективность описанных методов.

УДК 004.8

ОСОБЕННОСТИ ПОСТРОЕНИЯ НЕЙРОСЕТЕВЫХ АРХИТЕКТУР ДЛЯ РАСПОЗНАВАНИЯ БИОМЕТРИЧЕСКИХ МОДАЛЬНОСТЕЙ

Е.А. Ржевский

Научный руководитель – д.т.н., профессор А.Ю. Тропченко

Краткое вступление, постановка проблемы. Биометрика в настоящее время является очень бурно-развивающимся направлением информационных технологий. Широкое распространение получают биометрические системы идентификации человека. Они позволяют однозначно идентифицировать личность человека по биометрическим особенностям. К числу таких особенностей относятся отпечатки пальцев, форма ладони, узор радужной оболочки, изображение сетчатки глаза, индивидуальные характеристики лица. На данный момент существует множество подходов к решению данных задач, а также наличие коммерческих решений в той или иной степени имеющих отношение к данной области. Но все эти методы в той или иной мере связаны с распознаванием образов. Одним из таких подходов является нейросетевой подход. О нем и пойдет речь в данной работе.

Целью работы является рассмотрение особенностей построения биометрических систем идентификации личности по голосу и лицу с применением нейросетевого подхода. А также выявление наиболее перспективных нейросетевых архитектур, позволяющих получать высокий процент распознавания в коррелированных, а также зашумленных данных. В частности, большой интерес представляют методы объединения (фузирования) различных модальностей.

Базовые положения исследования. Нейросетевым методам распознавания лиц и голоса посвящено множество научных публикаций. Как правило, за основу берется некая

нейросетевая архитектура, вокруг которой проводится исследования. Оцениваются качество распознавания, производительность, ограничения того или иного решения. В области биометрии нейронные сети в чистом виде используются очень редко. Это обусловлено достаточно высокой глобальной корреляцией биометрических особенностей отдельно взятой модальности. Поэтому исследователям приходится придумывать более сложные подходы, связанные различной предобработкой распознаваемых модальностей. Этот так называемый этап «featureextraction» позволяет улучшить пространственную делимость образов, и позволяет получать более высокие результаты распознавания.

Другим направлением улучшения качества результатов является выбор нейросетевых архитектур, более адаптированных к специфике задачи. Это направление связано с уменьшением структурной неопределенности НС, т.е. внесение априорной информации в структуру НС в виде добавления или удаления нейронных связей, выбора специальной функции активации и т.п.

В данной работе рассмотрены две нейросетевых архитектуры, наиболее подходящих под специфику задачи.

1. Сверточные нейронные сети. На их основе разработан подход Multimodaldeeplearning. Данный подход является одним из наиболее перспективных и показывает высокие показатели распознавания, и высокую способность сети к обобщению.
2. Параметрические нейронные сети (ПНС). Это рекуррентные сети Хопфилдовского типа. В их основе лежит механизм ассоциативной памяти. Исследования показывают, что при правильном выборе структуры сети она обладает значительно большей емкостью памяти и высокой помехоустойчивостью. Также рассмотрены методы по уменьшению корреляции образов в таких НС.

С точки зрения методологии обучения в данной работе рассмотрены одномодальные и межмодальные методы. Межмодальный метод представляет наибольший интерес. Он позволяет производить фузирование различных модальностей, например голоса и лица, повышая этим качество распознавания.

Основной результат. В результатах работы описаны асимптотические оценки и процентные соотношения улучшения качества распознавания при использовании описанных выше подходов.

УДК 004.942

ПРОЕКТИРОВАНИЕ ПРОГРАММНОГО ОБЕСПЕЧЕНИЯ ДЛЯ МОДЕЛИРОВАНИЯ СЕТЕЙ ПЕТРИ-МАРКОВА

И.А. Страхов

(Тульский государственный университет)

Научный руководитель – к.т.н., доцент А.Н. Ивутин

(Тульский государственный университет)

Использование сетей Петри-Маркова – один из наиболее подходящих способов моделирования систем с параллельным ходом выполнения компонентов. К системам такого вида можно отнести:

- выполнение независимых частей программы на ЭВМ;
- строительство объектов;
- аппаратное обеспечение ЭВМ.

Для анализа систем с использованием аппарата сетей Петри-Маркова удобно проводить моделирование с применением средств современных ЭВМ, что позволяет снизить стоимость

и сократить время исследования, по сравнению с тестированием на реальной системе.

Как показал анализ, в настоящее время не существует программного обеспечения моделирования сетей Петри-Маркова, которое позволяло бы производить его быстро и наглядно с сохранением приемлемой точности получаемого результата.

Для лучшего понимания метода сетей Петри-Маркова существует возможность представлять теоретические определения в виде графа. Граф сети Петри-Маркова – это ориентированный взвешенный двудольный граф. Позиции в нем обозначаются кругами, переходы – жирной чертой. Существование возможности выполнения полушага отмечается стрелкой, над которой указывается плотность распределения времени и вероятность для выполнения полушага из позиции в переход или логические условия для выполнения полушага из перехода в позицию.

Целью работы является создание программного инструмента, который бы позволял строить произвольные сети Петри-Маркова и удовлетворял следующим обязательным требованиям:

- поддерживать изменения или переконфигурирования сети;
- предоставлять функции загрузки и сохранения результатов;
- реализовывать инструменты для проведения процесса моделирования сети;
- позволять получить полную справочную информацию;
- выводить информацию о результате моделирования;
- гарантировать высокое быстродействие;
- обеспечивать удобный графический интерфейс;
- давать возможность легкой модификации проекта.

Анализируя исходные данные, были сделаны следующие выводы:

- необходимо наличие разных типов вершин – позиций и переходов;
- необходимо наличие связей между данными вершинами;
- количество связей между вершинами может быть больше 1;
- вершины и связи должны обладать свойствами, характеризующими сеть и позволяющими проводить моделирование;

При рассмотрении языка написания программы, выбор стоял между тремя наиболее распространенными: C++, C#, Java. После анализа, было решено использовать C#, так как C++ и Java не обладают встроенными возможностями для создания приложений со сложной графической составляющей. C# обладает обширными библиотеками и большим количеством технологий, позволяющим создавать быстродействующие и удобные графические приложения. Проводя сравнение скорости работы приложений, написанных на этих языках, было выявлено, что C#-приложения проигрывают в скорости работы C++ и выигрывают у Java. Но проигрыш C# не слишком значителен, и не заставляет отказываться от данного языка при решении поставленной задачи.

При написании программы было принято решение использовать интегрированную среду разработки MSVisualStudio 2010 совместно с технологией WindowsPresentationFoundation (WPF), в основе которой лежит графическая технология DirectX. За счет использования аппаратного ускорения графики через DirectX отклик графического интерфейса WPF-приложений выше, чем у приложений, созданных с использованием других графических технологий, например, таких как WindowsForms.

Во время проектирования программного обеспечения были решены три основные задачи:

1. каким способом задать сеть;
2. как расположить элементы сети при задании ее структуры из внешнего источника. При этом была решена задача построения графа с непересекающимися дугами;
3. как сохранить заданную структуру сети.

В результате проделанной работы было получено программное обеспечение, которое

полностью удовлетворяет поставленной цели. Дальнейшее развитие и модернизация программы будут направлены на оптимизацию используемых математических алгоритмов и улучшение пользовательского интерфейса приложения.

УДК 004.056

СПОСОБЫ КОНТРОЛЯ ЦЕЛОСТНОСТИ ОБРАЗА ОПЕРАЦИОННОЙ СИСТЕМЫ ПРИ УДАЛЕННОЙ ЗАГРУЗКЕ НА ТОНКИЕ КЛИЕНТЫ В СИСТЕМАХ ТЕРМИНАЛЬНОГО ДОСТУПА

О.А. Теплоухова

Научный руководитель – д.т.н., профессор Ю.А. Гатчин

Введение. За последние несколько лет проектирование распределенных автоматизированных систем все чаще осуществляется с учетом концепции облачных вычислений. Данная технология позволяет предоставить повсеместный и удобный сетевой доступ к общему пулу конфигурируемых вычислительных ресурсов, которые могут быть быстро предоставлены и освобождены с минимальными усилиями по управлению и необходимостью взаимодействия с провайдером услуг. Построенные таким образом системы называют системами терминального доступа (СТД).

Благодаря централизованному хранению и обработке информации значительно упрощается построение системы защиты информации, циркулирующей в СТД. Необходимо отметить, что при этом традиционно основное внимание уделяют безопасности терминальных серверов, осуществляющих непосредственно хранение и процесс обработки данных. Но терминальные клиенты также являются неотъемлемой частью системы, а для построения защищенной СТД необходимо обеспечить безопасность всех ее элементов. При этом задача защиты терминальных клиентов распадается на две подзадачи:

1. контроль доступа пользователя к портам ввода-вывода терминала с целью запретить подключение неразрешенных внешних устройств;
2. контроль целостности программных средств, с помощью которых пользователь взаимодействует с терминальным сервером.

Целью работы является изучение особенностей построения системы защиты СТД, основанной на тонких клиентах, и проведение анализа существующих способов контроля целостности образа операционной системы (ОС), загружаемого на тонкий клиент с удаленного сервера.

Базовые положения исследования. При построении защиты терминальных клиентов важно учитывать особенности их работы: небольшой размер загружаемой ОС, использование одного образа многими клиентами, разные способы хранения образа (на локальном жестком диске, на жестком диске сервера, на мобильном носителе) и соответствующие способы загрузки.

При создании распределенной СТД задача выбора аппаратной платформы клиентских терминалов для установки на рабочие места почти всегда решается в пользу бездисковых рабочих станций – так называемых, «тонких клиентов». При общении с данным видом терминалов пользователю предоставляются только экран, мышь, клавиатура и USB-порты, использование которых полностью контролируется сервером. С точки зрения администрирования построенных таким образом СТД наиболее удобной является загрузка образа ОС на терминальные станции по сети. Но распределенный характер данного способа загрузки несет в себе дополнительные угрозы несанкционированного доступа к информации

(НСД). Для защиты такой системы требуется контролировать состав оборудования каждого терминала, а также всех серверов, с которых может производиться загрузка. Кроме того, необходимо гарантировать целостность загружаемой ОС.

Обычно загрузка образа по сети осуществляется согласно стандарту PXE (PrebootExecutionEnvironment). Для этого на терминале должна быть установлена сетевая карта с поддержкой этой возможности. PXE-код, прописанный в сетевой карте, получает с сервера специальный загрузчик, который осуществляет загрузку образа ОС. При этом не существует способа гарантировать, что запрос на загрузку обслуживается настоящим сервером, соответственно, откуда ему придет ОС и что придет под видом ОС – неизвестно.

На данный момент существует два основных способа проверки целостности и аутентичности образов загружаемой ОС:

- вычисление контрольной суммы либо хэш-значения загружаемого образа;
- вычисление электронной цифровой подписи (ЭЦП), используя асимметричные алгоритмы шифрования.

Существенный недостаток первого способа состоит в том, что при каком-либо изменении загрузочного образа, меняется значение контрольной суммы (хэш-значения), которое нужно донести до терминальных клиентов, чтобы СЗИ могли распознать измененный образ. В связи с этим, предпочитают использовать ЭЦП для проверки целостности и аутентичности полученного образа. Таким образом, на тонком клиенте должен быть реализован механизм защищенного вычисления ЭЦП получаемого с сервера образа ОС.

Существующие на рынке аппаратно-программные модули доверенной загрузки (АПМДЗ), позволяющие выполнять криптографические вычисления, не предназначены для работы с бездисковыми машинами, а применение персональных средств криптографической защиты информации (ПСКЗИ) на многопользовательских терминалах не всегда возможно.

Таким образом, средства защиты, позволяющие гарантировать целостность операционной системы, загружаемой с сервера на тонкий клиент, на сегодняшний день отсутствуют как класс. Без такого средства невозможно гарантировать целостность функционирующих на рабочих местах программных средств защиты информации, а, следовательно, и обеспечить защиту информации в общем случае.

Заключение. В работе были получены следующие результаты:

- проведен анализ использования тонких клиентов при построении распределенных систем терминального доступа;
- рассмотрены особенности построения системы защиты тонких клиентов в СТД;
- изучены способы контроля целостности образа операционной системы, загружаемой на тонкие клиенты с удаленного сервера.

В результате проделанной работы были получены следующие выводы:

- представляется актуальной задача разработки средства защиты, реализующего доверенную загрузку операционной системы на бездисковые терминальные клиенты;
- для программно-аппаратной реализации данной задачи предлагается использовать сетевую карту с поддержкой протокола PXE;
- представляется целесообразным разработать защищенную версию протокола PXE для обеспечения основных функций безопасности разрабатываемого СЗИ.

**ПРОГРАММНОЕ ОБЕСПЕЧЕНИЕ ДЛЯ ИССЛЕДОВАНИЯ
ТОПОЛОГИИ ПОВЕДЕНИЯ И КЛАССИФИКАЦИИ ЭЛЕМЕНТАРНЫХ СЕТЕЙ
ПЕТРИ С ПОМОЩЬЮ ВЫЧИСЛЕНИЯ ИХ ГРУПП ГОМОЛОГИЙ**

Т.А. Тришина

(Комсомольский-на-Амуре государственный технический университет)

Научный руководитель – д.ф.-м.н., профессор А.А. Хусаинов

(Комсомольский-на-Амуре государственный технический университет)

Сеть Петри представляет собой двудольный ориентированный граф, состоящий из вершин двух типов – мест и переходов, соединенных между собой стрелками. В позициях могут размещаться фишки, способные перемещаться по сети.

Сеть Петри называется элементарной, если для каждой ее маркировки число фишек в каждом месте не больше единицы. Хорошо известно, что каждая элементарная сеть Петри эквивалентна безопасной.

В настоящее время теория сетей Петри содержит большое количество моделей, методов и средств анализа, имеющих обширное количество приложений практически во всех отраслях вычислительной техники и даже вне ее. Для изучения работы элементарной сети Петри необходимо исследовать ее группы гомологий.

Целью работы является разработка программного обеспечения для вычисления групп гомологий и направленных групп гомологий элементарных сетей Петри.

Алгоритм основан на методе вычисления групп асинхронной системы, соответствующей элементарной сети Петри. Состояниями этой асинхронной системы служат достижимые маркировки. Маркировка соединяется переходом с другой маркировкой, если последняя получена из нее с помощью срабатывания перехода (события).

Это позволяет строить и визуально проверять коэффициенты матрицы дифференциалов комплекса для вычисления групп гомологий. Группы гомологий вычисляются с помощью приведения этих матриц к нормальной форме Смита.

Разрабатываемое программное обеспечение призвано визуализировать процесс построения сети Петри и автоматизировать процесс вычисления групп гомологий сети Петри. Рассмотрение, использование и визуализация сетей Петри и их групп гомологий предполагает соответственно их непосредственное использование при программировании. Таким образом, основу математической модели разрабатываемого программного обеспечения составляют законы, формулы и соотношения из теории асинхронных систем.

На данный момент разработана программа визуализации движения меток сети Петри.

Алгоритм теоретически обоснован в работе «Husainov A.A. The Homology Groups of Partial Trace Monoid Actions and Petri Nets, Appl. Categor. Struct. 2012, DOI 10.1007/s 10485-012-9280-9». Все действия по созданию программы расписаны. Программа находится в стадии отладки. Алгоритмы проверены. Результаты вычисления согласованы с теоретическими, полученными в препринте «Homology groups of pipeline Petri nets», опубликованном по адресу: <http://arxiv.org/abs/1301.7521> и в статье, принятой к печати в МАИС. Данная работа выполнена в рамках программы стратегического развития государственных образовательных учреждений высшего профессионального образования, №2011-ПР-054.

АДАПТИВНЫЙ МЕТОД ПОВЫШЕНИЯ РАЗРЕШЕНИЯ ПОЛУТОНОВЫХ ИЗОБРАЖЕНИЙ НА ОСНОВЕ ДИСКРЕТНОГО ВЕЙВЛЕТ-ПРЕОБРАЗОВАНИЯ

Д.А. Фёдоров

(Санкт-Петербургский национальный исследовательский университет информационных технологий механики и оптики)

Научный руководитель – к.т.н., доцент Т.Ю. Фисенко

(Филиал ОАО «Корпорация «Комета» ОЭКН)

Задача повышения разрешения изображений является актуальной в связи с появлением дисплеев с высоким разрешением, развитием телевидения высокой четкости. В настоящей работе разработан метод адаптивного повышения разрешения на основе дискретного вейвлет-преобразования с использованием карт направлений. Выполнено экспериментальное исследование метода на серии тестовых изображений, показавшее, что разработанный метод позволяет получить высокое пиковое отношение сигнал/шум. Применение метода уменьшает блочность и размытие границ.

Введение. Одной из тенденций в задаче масштабирования изображений является применение дискретного вейвлет-преобразования (ДВП). Классический метод интерполяции изображений на основе ДВП, как и большинство других методов масштабирования (билинейная интерполяция или глобальный кубический сплайн), не зависит от вида сигнала и не использует априорную информацию о сигнале, поэтому его можно назвать изотропным. Выбор подходящего материнского вейвлета позволяет избавиться от большинства артефактов, за исключением размытости границ.

Модель представления изображения. Чтобы увеличить изображение с видимым повышением разрешения, необходимо учитывать его локальные особенности. Информация о контурах и границах объектов несут высокочастотные (ВЧ) коэффициенты ДВП. Поэтому для успешного решения задачи повышения разрешения, сначала строится на основе ВЧ коэффициентов ДВП модель, позволяющую решить задачу аппроксимации двумерного сигнала изображения. Выполним разложение сигнала по некоторому базису

$$y = y_B + y_r. \quad (1)$$

Чем меньше остаток y_r от разложения тем более точно базис характеризует рассматриваемый сигнал. Соответственно при декомпозиции сигнала необходимо выбирать такое разложение из всего множества разложений, которое минимизирует энергию остатков.

Можно минимизировать глобальную энергию остатков по всем векторам из «словаря» и по всему изображению – алгоритм преследования базиса [1, 2]. Алгоритм согласованного преследования, предложенный Малом и Зангом [2–4], требует меньше операций, но он также имеет высокую вычислительную сложность. Для уменьшения количества вычислений применим «жадный» алгоритм, который позволит выбрать преобладающие направления контуров, максимизируя y_B .

Для построения модели используется избыточная фильтрация ВЧ вейвлет-фильтром. Избыточность состоит в том, что не применяется децимация, в отличие от классической схемы. Фильтрация выполняется вдоль направлений. Это позволяет выделить контуры, ориентированные перпендикулярно направлению фильтрации. В данной работе используется ВЧ фильтр соответствующий вейвлету Хаара.

Для более точной оценки вес канала («энергия») вычисляется как сумма модулей сумм коэффициентов вейвлет-преобразования по одноэлементным направлениям в канале (полосам):

$$E(i, j) = \left| \sum_{k1} c_{k1} \right| + \left| \sum_{k2} c_{k2} \right|. \quad (2)$$

Из весов формируется общий массив. Из массива получаются карты направлений следующим образом. Выполняется сортировка по убыванию «энергии» в массиве весов коэффициентов ДВП. Далее последовательно наносятся каналы в соответствующие карты направлений и в общую карту. При этом если канал имеет пересечение с общей картой, то он исключается, не вносится в карты. В результате получается модель изображения, состоящая из карт направлений.

Далее, чтобы использовать карты для получения увеличенной версии изображения, необходимо произвести их интерполяцию методом ближайшего соседа и получить модель увеличенного изображения – M_B .

Метод адаптивного повышения разрешения изображений. На основании построенной модели увеличенное изображение можно представить как:

$$y_{\text{увелич.}} = \sum_B M_B U_B y + (1 - \sum_B M_B) U y, \quad (3)$$

где U_B – интерполятор вдоль направлений; U – изотропный интерполятор; $\mathbf{1}$ – матрица размера увеличенного изображения (X на Y), заполненная единицами.

Сначала следует получить низкочастотную (НЧ) версию изображения из исходного при помощи обратного ДВП (ОДВП) [5]. На следующем этапе несколько раз (по числу направлений) и отличными способами интерполируются коэффициенты исходного изображения. Вычисляемые значения соответствуют позициям в интерполированных картах направлений.

На заключительном этапе производится улучшение НЧ версии увеличенного изображения. В соответствии с масштабированными картами направлений значения, полученные изотропным методом, заменяются коэффициентами, вычисленными при помощи интерполяции вдоль направлений.

Экспериментальное исследование. Для исследования используется база тестовых изображений, состоящая из тридцати полутоновых восьмиразрядных изображений, имеющих различный размер, мелкие и крупные детали. На первом этапе изображения уменьшались. Далее производилось увеличение при помощи разработанного метода и ОДВП [5].

Эффективность метода оценивалась по объективному критерию: пиковому отношению сигнал/шум (ПОСШ):

$$\text{ПОСШ} = 20 \log_{10} (255 / \text{СКО}), \quad (4)$$

где СКО – среднее квадратическое отклонение, оцениваемое по формуле:

$$\text{СКО} = \sqrt{\frac{1}{XY} \sum_{j=1}^X \sum_{i=1}^Y (im1_{i,j} - im2_{i,j})^2}, \quad (5)$$

где $im1$, $im2$ – значения яркости в отсчетах оригинального и дважды масштабированного изображений. Результаты экспериментальных исследований позволили сделать вывод о преимуществе разработанного метода, как по объективному критерию, так и по качественной оценке масштабирования для всех изображений.

Заключение. Разработана модель представления изображения, на основании которой построен адаптивный метод повышения разрешения изображений с применением ДВП. Разработаны модели каналов на основе анализа контекстной информации о ВЧ составляющих по направлениям, полученных по методу ДВП с использованием вейвлетов Хаара. Разработан алгоритм формирования карт направлений. Разработан алгоритм увеличения изображения на основе ОДВП с использованием биортогональных вейвлетов для

формирования НЧ версии изображения на масштабе более высокого разрешения. Выполнено экспериментальное исследование алгоритма на серии тестовых изображений. Исследование показало, что разработанный метод позволяет получить более высокое ПОСШ в сравнении с базовой версией – увеличение при помощи ОДВП. Применение метода позволяет уменьшить блочность и размытие границ.

Литература

1. Малл С. Вейвлеты в обработке сигналов /Пер.с.англ. – М.: Мир, 2005. – 671 с.
2. Mallat S. Zhang Z. Matching pursuit with time-frequency dictionaries // IEEE Trans. Signal Process. – 1993. – V. 41. – № 12. – P. 3397–3415.
3. Chen S. Donoho D. Saunders M. Atomic decomposition by basis pursuit // SIAM Journal. – 2001. – V. 43. – P. 129–159.
4. Mallat S., Guoshen Yu Super-resolution with sparse mixing estimators // IEEE Trans. On Image Proc. – 2010. – V. 19. – №11. – P. 2889–2900.
5. Фёдоров Д.А. Метод интерполяции цифровых изображений на основе дискретного вейвлет-преобразования // Научно-технический вестник информационных технологий, механики и оптики. – 2012. – №3(79). – С. 37–40.

УДК 004.932.2

РАЗРАБОТКА ПОЛУАВТОМАТИЗИРОВАННОЙ СИСТЕМЫ ВИДЕОНАБЛЮДЕНИЯ ЗА ВОДНЫМИ ОБЪЕКТАМИ

В.И. Филатов

Научный руководитель – д.т.н., доцент А.С. Потапов

Невозможно оценить всю полноту и значимость области компьютерного зрения при решении задач. Несмотря на то, что область компьютерного зрения можно охарактеризовать, как молодую, динамично развивающуюся и появившуюся сравнительно недавно, на протяжении нескольких последних десятков лет эта область активно применяется, в том числе, и для автоматизированного (автоматического) распознавания объектов, слежения за движущимися целями.

Наиболее популярным применением компьютерного зрения при решении задач обнаружения, распознавания, является использование их в системах слежения, видеонаблюдения, охраны секретных и военных объектов [1].

Практически в каждой отрасли современной науки применяются алгоритмы и комплексы, построенные на знаниях, формирующих область компьютерного зрения, например в [2, 3]. Но, несмотря на все это, многие из компьютерных автоматизированных систем обладают ограниченностью, применяются для решения отдельных научных вопросов при определенных условиях и не подходят для общего случая. Такие ключевые факторы (на примере систем видеонаблюдения) как, изменение освещенности сцены, сложные динамически быстроменяющиеся сцены (задний план) [4, 5], тени как от искусственных источников излучения, так и от природных, скорость движения объектов, различные перекрытия объектов и т.д. не позволяют свести решение задачи к общему случаю, но даже при возможности такого сведения, решение получилось бы слишком ограниченным и малоприменимым на практике.

Из всего этого, можно сделать вывод, что в каждом отдельном случае получение новых решений, новых отдельных систем, решающих узкие специализированные проблемы необходимый в повседневности процесс.

Целью работы является создание морской полуавтоматизированной компьютерной

системы видеонаблюдения. Использование существующих знаний об области компьютерного зрения дает возможность разработки автоматизированных систем слежения без участия человеческого фактора.

В ходе исследования была разработана полуавтоматизированная система контроля безопасности морских объектов.

Задачи, выполняемые системой:

- обнаружение водных средств в области видимости камер слежения;
- слежение за обнаруженным объектом в пределах окна видимости камер по определенному идентификационному номеру;
- возможность подавления ложных выбросов (как автоматически, так и с помощью оператора связи);
- получение данных о местонахождении корабля, его скорости и направлении движения.

Литература

1. Лукьяница А.А., Шишкин А.Г. Цифровая обработка видеоизображений. – М.: Ай-Эс-Эс-Пресс, 2009. – 518 с.
2. Haralick R.M., Sternberg S.R., Zhuang X. Image Analysis using Mathematical Morphology // IEEE Transactions on Pattern Analysis and Machine Intelligence. –1987. – №9(4). – P. 532–550.
3. Garg S., Thapar S. Feature extraction using Morphological Operations on finger print images // International Journal of Computing and Business Research. – 2012.
4. Lucas B.D., Kanade T. An iterative image registration technique with an application to stereo vision // DARPA Image Understanding Workshop. – 1981. – P. 674–679.
5. Z Zivkovic. Improved Adaptive Gaussian Mixture Model for Background Subtraction // Pattern Recognition, ICPR. – 2004. – V.2. – P. 28–31.

УДК 681.5

АЛГОРИТМЫ ЭФФЕКТИВНОГО ОБУЧЕНИЯ В ИНФОРМАЦИОННЫХ СИСТЕМАХ

Д.А. Хвостов, К.А. Пантюхина, С.А. Чепинский

Научный руководитель – к.т.н., доцент С.А. Чепинский

Постановка задачи. При создании информационных систем зачастую возникает потребность в автоматической генерации базы знаний на основе входных данных. При нахождении системы в естественной среде процесс обучения такого рода должен обладать высокой скоростью, эффективной коррекцией и оптимальным форматом сохранения результатов, а также возможностью анализа результатов обучения и их ручной корректировки оператором.

Цель работы. В работе рассматривается задача синтеза эффективного алгоритма самообучения, позволяющего разнообразным информационным системам с помощью анализа и сопоставления данных корректировать собственные алгоритмы управления с целью сохранения или достижения оптимального состояния при изменении внешних условий.

Базовые положения исследования. Подход, используемый в работе, предусматривает использование поверхностей для создания базы знаний и представления входных данных с помощью функций принадлежности. Это определяет эффективность алгоритма, а также дает

возможность упростить представление базы для восприятия оператором и обеспечить быструю реакцию на изменение внешних условий.

Основные результаты. В результате работы был разработан алгоритм самообучения, который успешно решает поставленные перед ним задачи. Для иллюстрации работоспособности алгоритма было выполнено компьютерное моделирование.

УДК 004.93+57.087.1

МАШИНЫ БОЛЬЦМАНА ДЛЯ ВЕРИФИКАЦИИ ДИКТОРА

А.В. Шолохов

Научный руководитель – к.ф.-м.н., доцент Т.С. Пеховский

Распознавание человека по голосу занимает одну из наиболее исследуемых областей биометрии. Голос является такой биометрической характеристикой человека, для получения которой не требуется наличие дорогостоящего оборудования, в отличие от, например, отпечатков пальцев или изображений радужной оболочки глаза, но в то же время позволяющей достаточно надежно идентифицировать личность.

На сегодняшний день лучшие системы распознавания по голосу основаны на различных модификациях Вероятностного Линейного Дискриминантного Анализа (PLDA) [1], признаками для которого являются i -вектора [2]. Здесь рассматривается модель для верификации диктора, основой которой является Машина Больцмана [3], моделирующая совместную плотность пар i -векторов [4], что позволяет определить вероятностную оценку сходства. Модель обучается на данных методом максимального правдоподобия.

Рассматриваемая вероятностная модель пытается определить отношение между двумя вещественными векторами x и y . Также модель содержит слой бинарных скрытых переменных $h_1 - h_K$, и оценка схожести определяется как

$$E(x, y, h) = - \sum_{f=1}^F \left(\sum_{i=1}^I v_{if} x_i \right) \left(\sum_{j=1}^J w_{jf} y_j \right) \left(\sum_{k=1}^K u_{kf} h_k \right) - \sum_{k=1}^K u_k h_k + \frac{1}{2} \sum_{i=1}^I (x_i - v_i)^2 + \frac{1}{2} \sum_{j=1}^J (y_j - w_j)^2,$$

где F – количество базисных функций («фильтров» v_f, w_f, u_f), на которые проецируются x, y и h . И тогда совместное распределение будет выглядеть так:

$$p(x, y, h) \propto \exp(-E(x, y, h)).$$

Далее, когда параметры модели настроены, ненормированная маргинальная плотность наблюдаемых переменных используется для оценивания степени сходства.

Предложенная модель показала производительность, сопоставимую с baseline-системой верификации.

Литература

1. Prince S.J.D. and Elder J.H. Probabilistic linear discriminant analysis for inferences about identity // 11th International Conference on Computer Vision. – 2007. – P. 1–8.
2. Dehak N., Dehak R., Kenny P., Brummer N., Ouellet P. and Dumouchel P. Support vector machines versus fast scoring in the low-dimensional total variability space for speaker verification // INTERSPEECH. Brighton, UK. – 2009. – P. 1559–1562.

3. [Электронный ресурс]. – Режим доступа: http://www.scholarpedia.org/article/Boltzmann_machine, свобод.
4. Roland Memisevic and Geoffrey E Hinton. Learning to represent spatial transformations with factored higher-order Boltzmann machines // Neural Computation. – 2010. – №22(6). – P. 1473–1492.

УДК 658.511

МЕТОДИКА ВЫБОРА ОПТИМАЛЬНОЙ СТРУКТУРЫ БИЗНЕС-ПРОЦЕССА ПРЕДПРИЯТИЯ С ИСПОЛЬЗОВАНИЕМ МЕТРИК СВЯЗНОСТИ И СЦЕПЛЕНИЯ

А.Д. Аркуша

Научный руководитель – к.т.н., доцент Н.Ф. Гусарова

Постановка задачи. Бизнес-архитектура и информационная архитектура большинства современных предприятий образуют тесно взаимосвязанный комплекс. Необходимы информационные технологии, позволяющие с единых системных позиций формировать и оптимизировать структуру таких комплексов. Особенно актуальна эта проблема для предприятий, на которых часто вносятся изменения в бизнес-процессы. В этом случае к соответствующим технологиям предъявляются следующие требования:

- с ними должны работать рядовые сотрудники предприятия, хорошо ориентирующиеся в специфике бизнес-процесса, а не приглашенные специалисты-программисты, услуги которых требуют больших затрат ресурсов;
- внесение изменений должно выполняться быстро и не затрагивать уже запущенные процессы.

Перспективной в этом плане является технология Workflow, которая ориентирована на моделирование и координацию бизнес-процессов, описание взаимосвязей и последовательности их выполнения. Технология Workflow позволяет задать правила описания и взаимодействия бизнес-процессов и выполнить моделирования схемы потоков работ. Однако вопросы количественной оценки эффективности структуры построенных схем до сих пор остаются открытыми и являются актуальными.

Цель работы разработать методику построения альтернативных вариантов организации бизнес-процессов и оценки их эффективности, что позволит определить пути оптимизации существующего процесса. В качестве базовой предметной области предполагается рассмотреть деятельность ресурсоснабжающей организации, связанная с формированием разрешительной документации.

Полученные результаты. Методика представляется в виде последовательности шагов:

1. проанализируем существующие бизнес-процессы и на основании анализа построим диаграмму «как есть». В связи с отсутствием стандартизованного языка для реализации технологии Workflow для построения диаграммы выбрана произвольная нотация, где выделены действия (операции) и переходы, образующие множество D ;
2. выделим основные подразделения и других участников, задействованных в процессе. Для этого выделим на диаграмме группы бизнес-процессов, которые относятся к одному подразделению. Стоит обратить внимание, что в качестве участника процесса также могут выступать другие программы или оргтехника (например, принтеры и др.);
3. определим условия перехода для каждой операции. Условия могут быть различных типов:
 - логические условия AND, OR и др., которые используются для связи нескольких операций;

- условия перехода между двумя операциями, имеющие, в свою очередь, ряд типов:
 - для выполнения операции необходимо наличие другой завершенной операции;
 - запрещено выполнение операции, если есть связанная с ней операция;
 - одна операция закрывает связанную с ней операцию;
- 4. определим все ресурсы W , которые будут участвовать в выделенном бизнес-процессе, и привяжем их к каждой операции, в которой они используются. В понятие ресурсов может входить информация из других систем, формируемые документы или документы, полученные от других отделов, необходимые для выполнения перехода;
- 5. выделим повторяющиеся наборы операций и назовем их паттернами. Паттерны можно использовать в любом бизнес-процессе в рамках предприятия как готовый шаблон. Важно, что при внесении изменений в запущенный процесс, изменения в паттерны вноситься не могут, поскольку они являются единой структурой, а не набором операций. Процесс внесения изменений будет ожидать завершения обработки паттерна, и только после его выполнения «пускать» выполнение операций по новому пути;
- 6. после выполнения пп. 1–5 построим первую альтернативную модель путем группировки операций в действия и рассчитаем среднее сцепление бизнес-процессов, коэффициент их связности и соотношение между полученными величинами. После чего построим несколько других альтернативных вариантов бизнес-процесса и произведем сравнительную оценку их эффективности, используя метрики связности и сцепления;
- 7. проверяем на достижимость все информационные элементы, ресурсы и операции.

Структуру информационных процессов формализуем при помощи подхода, который состоит в описании структуры операций информационных процессов при помощи набора (D, W, O) , где

D – набор информационных элементов, которые получаются или используются в ходе процесса,

W – набор ресурсов или ролей, которые доступны для процесса,

$O \in D \times W \times P(D)$ – набор достижимых элементов P для W . Для определения достижимости элементов в системе построим матрицу смежности, а на основании матрицы смежности построим матрицу достижимости.

В работе рассчитываются следующие показатели:

- сцепление показывает, насколько сильно разные операции влияют друг на друга и связаны между собой;
- сцепление информационных элементов основано на информационных элементах, которые используются в качестве входа/выхода операция соответствующей деятельности. Метрика показывает, как много элементов используется более одного раза по отношению ко всем используемым элементам;
- общее сцепление действий показывает наилучший показатель сцепления по обеим метрикам, если наилучший результат получен, то он и будет общим сцеплением;
- связность процессов показывает, насколько сильно связаны действия в процессе. Два действия связаны между собой только когда оба используют одни и те же информационные элементы. Метрика показывает, сколько с активностью связано других активностей. Все пары действий учитываются дважды;
- соотношение сцепления и связности процесса = связность процессов/ сцепление.

После построения нескольких альтернативных вариантов и расчета их показателей выбирается оптимальный вариант построения схемы бизнес-процессов.

**ВЫБОР ВАРИАНТА РАЗВЕРТЫВАНИЯ ИНФРАСТРУКТУРНОГО КОМПОНЕНТА,
ОБЕСПЕЧИВАЮЩЕГО РАБОТУ СИСТЕМЫ ВИДЕОНАБЛЮДЕНИЯ ПО
КОНЦЕПЦИИ VSAAS НА КРУПНОМ ПРЕДПРИЯТИИ С РАЗВИТОЙ
ФИЛИАЛЬНОЙ СЕТЬЮ**

М.А. Тимофеев

(Санкт-Петербургский национальный исследовательский университет информационных технологий, механики и оптики)

Научный руководитель – А.Д. Береснев
(ООО «РОТЭК»)

Краткое вступление постановка проблемы. Объем продаж на мировом рынке хостинговых услуг по видеонаблюдению растет очень быстро: в 2011 г. он составил 500 млн.долл., что на 25% больше, чем годом ранее. К 2014 г. данный показатель возрастет вдвое. Такая оценка содержится в отчете британской компании IMS Research, которая исследовала перспективы развития этой новой, быстро растущей сферы, именуемой «Видеонаблюдение как услуга» (VideoSurveillanceas a Service, VSaaS).

Согласно классификации Национального института по стандартам и технологиям США (NIST) выделяется несколько способов развертывания облачных систем:

- как частное облако (private), развернутое непосредственно в самой организации – заказчике;
- как групповые облака (community), распределенные между несколькими организациями, объединенными общими интересами (по обслуживанию и расположению не отличаются от частных облаков);
- как публичное облако (public), развернутое на базе инфраструктуры провайдера.

Обзор VSaaS-систем представленных на российском рынке показал, что большинство из них предлагают свои услуги на базе публичных облаков. Такие сервисы предлагают использовать свой продукт в домашних целях или на мелких предприятиях, где количество камер наблюдения невелико и пропускная способность канала связи является достаточной для трансляции сразу всех камер в режиме реального времени. Такой способ развертывания облака дает им возможность заявлять о доступности подключения к своему сервису по всему миру. Но, ни один из подобных сервисов на данный момент не предлагает свои услуги крупным предприятиям с развитой системой филиалов и удаленных объектов собственности. Такие предприятия нуждаются в стабильной работе своих систем в режиме «24/7», при этом, качество трансляции должно соответствовать самым высоким требованиям.

В работе рассмотрены факторы, которые сдерживают внедрение облачного видеонаблюдения на крупных предприятиях, рассмотрены возможные пути устранения некоторых из них.

Цель работы заключается в обосновании выбора варианта развертывания инфраструктурного компонента, обеспечивающего работу системы видеонаблюдения по концепции VSaaS на крупном предприятии с развитой филиальной сетью.

Практические результаты:

- определены основные компоненты инфраструктуры необходимые для функционирования систем облачного видео наблюдения;
- выявлены основные стандарты на разработку инфраструктурного компонента, обеспечивающего работу системы видеонаблюдения на крупном предприятии;
- определены основные риски использования VSaaS на крупных предприятиях;

- дана первичная оценка экономической выгоды от использования VSaaS на крупном предприятии с развитой сетью филиалов и удаленных объектов наблюдения;
- проведено обоснование выбора варианта развертывания инфраструктурного компонента, обеспечивающего работу системы видеонаблюдения по концепции VSaaS на крупном предприятии с развитой филиальной сетью.

УДК 65.011.56

**АНАЛИЗ ТРЕБОВАНИЙ К РАЗВИТИЮ ИНТЕГРИРОВАННОГО РЕШЕНИЯ
«1С: УПРАВЛЕНИЕ ПРОЕКТНОЙ ОРГАНИЗАЦИЕЙ» ДЛЯ ЭФФЕКТИВНОГО
ИСПОЛЬЗОВАНИЯ В КОМПАНИИ ОКАЗЫВАЮЩЕЙ УСЛУГИ В СФЕРЕ
ПРОЕКТИРОВАНИЯ, ИНЖИНИРИНГА, СТРОИТЕЛЬСТВА И ЭКСПЛУАТАЦИИ
ПРОМЫШЛЕННЫХ, КОММЕРЧЕСКИХ И ГРАЖДАНСКИХ ОБЪЕКТОВ**

А.В. Филимонова

Научный руководитель – к.п.н., доцент А.В. Маятин

Для того чтобы компания оказывающая полный комплекс услуг в сфере проектирования, инжиниринга, строительства и эксплуатации промышленных, коммерческих и гражданских объектов была конкурентоспособной на рынке и успешно осуществляла свою деятельность, а также принимала быстрые и верные управленческие решения, основанные на оперативной и достоверной информации, необходимо, чтобы использующаяся автоматизированная система управления предприятием в полной мере соответствовала специфике организации деятельности этого предприятия.

Для автоматизации управления рассматриваемой в работе компании была внедрена система «1С: Управление проектной организацией». В процессе эксплуатации интегрированного решения от 1С, неоднократно выявлялись дополнительные требования к функционалу программного обеспечения, приводящие к необходимости развития системы, сопровождаемые значительными затратами. Это показало недостаточность базового функционала интегрированного решения и сделало актуальной проблему анализа и оптимизации процессов развития решения 1С: УПО для соответствия нуждам конкретного предприятия. Одной из первых задач является сравнительный анализ требований компании к информационной системе и возможностей базовой конфигурации «1С: Управление проектной организацией». В результате анализа будут сформулированы требования к изменению базовой конфигурации 1С.

В результате анализа деятельности компании был выделен основной бизнес процесс компании ориентированный на проектный инжиниринг и проектирование объектов энергетики, а также на строительство промышленных, коммерческих объектов. А также определены вспомогательные бизнес-процессы: управление работой с поставщиками и заказчиками, управление финансовой деятельностью предприятия, управление юридической деятельностью компании, управление кадровым составом компании, мониторинг и анализ деятельности компании.

Для выявления требований был проведен функциональный анализ, в результате были выделены отдельно процессы управления и проведена их декомпозиция до уровня отдельных операций. Проанализировав базовую конфигурацию «1С: Управление проектной организацией» и сопоставив с процессами организации, было выявлено соответствие подсистем ИС и процессов управления (табл. 1).

Таблица 1. Соответствие подсистем ИС и процессов управления

Подсистемы ИС	Процессы компании
PMO. Управление проектами и портфелями проектов	Управление портфелями проектов компании
Подсистема отсутствует	Проектирование объектов
ERP. Управление ресурсами предприятия	Управление запасами
	Управление финансами компании
	Работа с сотрудниками компании
CRM/SRM. Управление отношениями с заказчиками, поставщиками и подрядчиками	Взаимодействие компании с поставщиками
	Взаимодействие компании с заказчиками
Подсистема отсутствует	Управление автотранспортом

Как видно из приведенной таблицы базовая конфигурация поддерживает почти все процессы организации. Но при этом если детализироваться до конкретных операций, то выясняется, что часть из них не имеет требуемого уровня поддержки со стороны АСУ.

Пример сопоставления операций в рамках процесса «Управление портфелями проектов компании» и функциональных возможностей подсистемы «PMO. Управление проектами и портфелями проектов» приведены в табл. 2 – указаны только операции, которые не поддерживаются или частично поддерживаются интегрированной системой.

Таблица 2. Пример сопоставления операций

Операции процесса «Управление портфелями проектов компании»	Поддержка операции в подсистеме «PMO. Управление проектами и портфелями проектов»
Рентабельность проектов (план-фактный анализ БДР)	Отсутствует
Рентабельность проектов (план-фактный анализ ДДС)	Отсутствует
План-фактный анализ БДР (с данными на начало периода с нарастающим итогом)	Частично поддерживается, с помощью функции «Управление финансами проектов и портфелей проектов»
План-фактный анализ ДДС (с данными на начало периода)	Частично поддерживается, с помощью функции «Управление финансами проектов и портфелей проектов»
Автоматическое распределение движения денежных средств по проектам	Отсутствует

При анализе других процессов компании и возможностей базовой конфигурации, также были выявлены операции, не поддерживаемые АСУ.

Непосредственно проектирование объектов осуществляется с использованием программы AutoCAD, но ее интеграция с АСУ не осуществлена.

Процесс управления автотранспортом, необходимый для контроля использования имеющегося автотранспорта в компании в принципе не реализован в базовой конфигурации ИС: УПО и требует разработки отдельной подсистемы для автоматизации следующих задач:

- формирование заявки на автотранспорт;
- формирование списка заявок на автотранспорт для утверждения начальником отдела логистики;

- создание маршрутного листа по конкретным водителям;
- управление жизненным циклом заявки на автотранспорт.

В результате этапа исследования были получены требования увеличения функционала базовой конфигурации 1С: УПО. На основе этих результатов планируется сформировать требования к модернизации компонент базовой конфигурации на уровне изменений моделей данных, к изменению и разработке новых функций приложения.

УДК 007.51, 004.42

АНАЛИЗ ЛОГИСТИЧЕСКИХ МЕТОДИК ПОСТРОЕНИЯ АВТОМАТИЗИРОВАННОЙ СИСТЕМЫ ПЛАНИРОВАНИЯ ПОПОЛНЕНИЯ ЗАПАСОВ ДЛЯ КРУПНОЙ СЕТИ РОЗНИЧНЫХ ПРОДАЖ

П.С. Шерстобитова

Научный руководитель – О.А. Йылмаз

Краткое вступление. По утверждению П.Г. Афанасьева «Анализ современных тенденций развития интеграционных процессов на потребительском рынке показывает наличие ряда проблем, связанных с эффективностью функционирования торговых сетей: проблемы рационального размещения и строительства новых объектов торговой недвижимости, использование нерациональных подходов к управлению товарными запасами» [1]. Сюда же можно отнести многочисленные проблемы с товародвижением, в частности низким организационно-технологическим уровнем работы управленческого персонала в плане принятия эффективных управленческих решений.

Высокая практическая значимость и недостаточно изученная теория данных проблем, а также неэффективное применение ее на практике, например, в построении информационных систем поддержки данного направления деятельности предопределяет актуальность и важность выбранной темы.

Цель работы. В рамках данной области планируется произвести моделирование ключевых бизнес-процессов контура распределения товаров по точкам реализации на основе сравнительного анализа и выбора оптимального набора ключевых методик, как за основополагающие для построения единой модели процессов планирования распределения.

За объект исследования планируется взять лучшие практики применения информационных технологий для поддержки деятельности торгово-промышленных компаний, предприятия и бизнес-структур в сфере распределения продукции разных отраслей.

В рамках выбранной области данной работы планируется произвести сравнительный анализ ведущих современных методик планирования и распределения готовой продукции, которые в свою очередь можно разделить на две группы:

1. методики распределения по точкам заказа;
2. методики распределение с учетом прогноза продаж.

Данные направления необходимо рассмотреть с учетом анализа степени удовлетворения бизнес-требований организаций в сфере розничных сетевых продаж со своей спецификой построения процессов деятельности.

Предполагаемые исследования. Для решения поставленных задач необходимо произвести:

- обзор лучших практик внедрения SAP в области Retail (в части распределения продукции);

- формирование критериев оценки методов распределения, используемых в существующих информационных системах на основании требований компании-заказчика;
- разбор современных логистических концепций на основе сформированных критериев;
- на основании вышеперечисленных пунктов проведение сравнительного анализа существующих подходов и методов на предмет возможности и целесообразности применения к общей стратегии деятельности аппарата управления;
- оценку возможности применения выработанной методики в рамках построения системы класса ERP (SAPRetail).

Вывод. В качестве основного результата данного исследования планируется предоставить концептуальную методику максимально равномерного распределения продукции по точкам реализации, для возможности применения в последующем моделировании и проектировании прикладных и информационных процессов логистики. Результаты могут использоваться руководством проектов по реинжинирингу бизнеса или разработки и внедрения автоматизированных систем в целях оптимального покрытия требований бизнеса в автоматизированном управлении перемещениями товаров, исполнение которого позволит удовлетворить потребность магазинов розничной сети продаж в необходимом количестве запасов. Данный критерий обеспечит возможность обеспечения возможного повышения уровня обслуживания целевых покупателей.

Литература

1. Афанасьев П.Г. Формирование эффективных бизнес-процессов в системе распределения продукции лесопильно-деревообрабатывающей промышленности // дис. на соиск. учен. степ. канд. экон. наук – Специальность 08.00.05 «Экономика и упр. нар. хоз-вом». – 2008.
2. Дыбская В.В., Зайцев Е.И., Сергеев В.И., Стерлигова А.Н. Логистика: интеграция и оптимизация логистических бизнес-процессов в цепях поставок / Учебник под ред. проф. В.И. Сергеева. – М.: Эксмо, 2008. – 944 с.

УДК 004

РЕАЛИЗАЦИЯ МЕТОДИКИ JIT В ИНФОРМАЦИОННОМ ОБЕСПЕЧЕНИИ ДЛЯ ПРЕДПРИЯТИЙ МАЛОГО БИЗНЕСА, ИМЕЮЩИХ ИНДИВИДУАЛЬНОЕ И СЕРИЙНОЕ ПРОИЗВОДСТВО

Л.А. Беляева

Научный руководитель – к.п.н., доцент А.В. Маягин

В настоящий момент большое количество предприятий в нашей стране используют подход «точно-в-срок» (just-in-time — JIT). Для сферы малого бизнеса с индивидуальным и серийным производством в условиях общей нестабильности, присутствующей в России, эффективность подхода JIT сделала его внедрение почти обязательным средством конкурентной борьбы. Для эффективного использования подхода «JIT» компании должны иметь постоянное информационное обеспечение процессов управления.

Суть подхода JIT сводится к отказу от производства продукции крупными партиями. Взамен этого создается непрерывно-поточное предметное производство. При этом снабжение производственных цехов и участков осуществляется столь малыми партиями, что по существу превращается в поштучное. В нашей стране наиболее конкурентоспособным способом производства для предприятий малого бизнеса является мелкосерийное производство продукции, происходящее параллельно с индивидуальным (или штучным)

производством. Для использования в этом случае подхода JIT необходима точная информация о текущем состоянии производства, точные прогнозы на ближайшее будущее. Появляется задача оперативного получения, хранения и обработки соответствующей информации. Внедрение систем ERP класса в этом случае требует значительной доработки информационной архитектуры для каждого из предприятий, основывающейся на особенностях производственного процесса, организационной структуры и стиля управления. Существует возможность упростить этот процесс, введя абстрактную модель информационной архитектуры, учитывающую изменения в традиционной модели управления предприятием, связанные с использованием подхода JIT.

Модель предприятия, использующего подход JIT, состоит из четырех основных формирующих блоков, на которые должны быть направлены управленческие воздействия:

- производственное планирование. Включает процессы: планирование сроков и методов производственного процесса, назначение ответственных лиц, составление плана закупки ресурсов;
- ресурсы и производство. Включает процессы: контроль процесса закупки ресурсов, хранения и перемещения ресурсов, контроль производственного процесса;
- разработка изделий. Включает процессы: проектное проектирование, проектирование и моделирование изделий, разработка инноваций;
- контроль, анализ и изменение бизнес-процессов. Включает процессы: сбор данных о работе предприятия, анализ полученных данных, непрерывный контроль качества на основании данных о работе предприятия, изменение бизнес-процессов предприятия в рамках процесса управления качеством и в связи с изменением ситуации на рынке и внутренней политики предприятия;
- кадровые/организационные элементы. Включает процессы: управление людскими ресурсами.

В рамках данных контуров управления JIT приводит к необходимости пересмотра способа автоматизации управления следующими элементами бизнес-процессов предприятия.

- Запасы. При этом у предприятий, имеющих серийное производство в нашей стране так или иначе должен присутствовать небольшой запас постоянно используемых ресурсов, это позволит снизить логистические затраты и риск резкого изменения предложения на рынке поставщиков. Необходимо управление складскими запасами, причем кроме информации о количестве ресурса на складе очень важно наличие информации о том когда, для какого процесса он был закуплен и точные данные о нем самом. Достоверность этих данных является критичной для процесса планирования. При планировании учитываются как ресурсы, заказанные у внешних поставщиков, так и ресурсы, использующиеся несколькими процессами внутри предприятия. Для производства бездефектных изделий необходима постоянная отбраковка материальных ресурсов. Необходим отдельный учет ресурсов, закупленных у поставщика и поступивших сразу на производство (артикул, количество, цена, сроки гарантии), переходящих от одного к другому процессу внутри предприятия (также важны данные об используемом процессе, ответственном лице, о процессе, использующем ресурс в дальнейшем и др.), о браке (критично для процесса планирования).
- Поставщики. При возникновении потребности в материальных ресурсах необходимо сразу направлять запрос поставщику и осуществлять доставку ресурсов точно в срок, в нужное место в нужном объеме. Обмен информации с поставщиками должен происходить при использовании данных, понятных обеим сторонам. Необходимо наладить обмен данными между информационными системами компании и постоянных поставщиков. Необходимо иметь данные о качестве продукции отдельных поставщиков, сроках и способе поставок, контактная информация агентов и актуальные данные об ассортименте.

- Объем партий. Необходимо искать способы снижения объемов производственных партий и добиваться коротких производственных циклов. Это приводит к необходимости иметь очень точные данные о ресурсах, ошибки в данных могут быть критичными для процесса планирования, так как отсутствуют стратегические запасы ресурсов и времени.
- Качество. Необходимо стремиться к отсутствию брака и постоянному улучшению качества изделий и бизнес-процессов на основе комплексного управления качеством. Необходимо постоянно получать сведения о поломках, простаивании оборудования, нереализованных ресурсах, производственном браке и его причинах и т.д. Для управления качеством существует потребность в системе анализа полученных данных.
- Время выполнения заказов. Необходимо снижать время выполнения заказов, чтобы снизить вероятность вмешательства факторов неопределенности. Существует потребность в точном учете времени и планировании временных ресурсов.
- Надежность. Все операции должны выполняться без сбоев и быть непрерывно управляемыми, т.е. не должно быть поломок оборудования, брака, невыхода персонала на работу и т.п. Подобные случаи должны быть учтены и доведены до сведения руководства.
- Работники. На предприятиях малого бизнеса большинство работники задействованы сразу на нескольких ролях и в различных ситуациях вынуждены брать на себя часть функций других сотрудников. Нужно не только иметь персональные данные каждого работника и сведения об объеме проделанной им работы и выполняемых в данный момент времени функциях, но и возможность анализа имеющихся у каждого работника временных ресурсов, его функциональной и эмоциональной совместимости с другими работниками, проявленных инициативах, вести строгий учет отгулов и внеурочной работы.
- Информационная поддержка должна позволять оперативно обмениваться информацией и синхронизировать все процессы поставки ресурсов, производства и сборки, поставки готовой продукции.

Опираясь на выделенные особенности, возможно построение модели информационной архитектуры для широкого класса информационных систем, позволяющей повысить эффективность внедрения информационной системы на предприятиях рассмотренного типа.

УДК 004.04

ИССЛЕДОВАНИЕ ОСОБЕННОСТЕЙ ИНФОРМАЦИОННОЙ МОДЕЛИ ДАННЫХ ERP-СИСТЕМЫ ДЛЯ ВНЕДРЕНИЯ НА ПРЕДПРИЯТИЯХ С ПРОЦЕССНЫМ ПРОИЗВОДСТВОМ

Е.С. Калапышин

Научный руководитель – О.А. Йылмаз

Краткое вступление, актуальность темы. Для обоснования актуальности данной тематики необходимо рассмотреть типы производств с точки зрения их непрерывности и классификации методов учета затрат

Согласно классификации, предложенной Gartner Group, все многообразие производств можно свести к трем основным типам: проектное, дискретное и процессное. При этом именно для процессного производства сложнее всего объединить в единую информационную среду логистический и финансовый контур ввиду сложных систем учета затрат на этапах производства продукции и во время конвертации единиц измерения. А так как объединение этих контуров обеспечивают только ERP-системы, и именно они лучше всего подходят для ведения управленческого учета, то за основу будет взята информационная модель системы MS Dynamics AX.

Целью работы является анализ бизнес-процессов предприятий с процессным производством и выявление общих особенностей для построения информационной модели данных ERP-системы. На данный момент существует несколько вариантов отражения основных бизнес и технологических процессов на промышленных предприятиях с помощью ERP систем, к примеру, решения Microsoft Dynamics AX. На каждом крупном промышленном предприятии так же существует система исполнения производства, так называемая MES-система, которая отражает ход технологических процессов и этапы получения того, или иного продукта. Но, к сожалению, не позволяет получать нужную информацию в целях ведения управленческого учета, т.е. информацию о бизнес-процессах планирования и управления производством. И поэтому, зачастую требуется объединить функциональность ERP и MES в одном информационном пространстве.

Необходимо проанализировать существующую модель данных MS Dynamics AX с точки зрения применимости разных методов учета затрат и формирования себестоимости, при этом выявить процессы производства, которые унифицированы для предприятий в любой области, занимающихся процессным производством. При детальном рассмотрении проблемы распадается на следующие составляющие:

- анализ бизнес-процессов предприятий с процессным типом производства и выявления общих, унифицированных процессов для таких предприятий с учетом существующих MES и АСУТП-систем;
- анализ существующей информационной модели данных MS Dynamics AX и выявление недостатков модели для применения системы в целях предприятий с процессным типом производства и с учетом применяемых на предприятиях методов учета затрат и ведения управленческого учета
- доработка информационной модели с учетом выявленных недостатков и требований к системам ведения управленческого и операционного учета на предприятиях с процессным типом производства.

Вывод. Главной задачей после рассмотрения проблемы в целом является выделение основных процессов, характерных предприятиям с непрерывным типом производства и разработка на их основе информационной модели данных для применения ERP-системы на подобных предприятиях в целях оперативного ведения управленческого учета.

УДК 004.05/005

ОБЗОР И СРАВНЕНИЕ СИСТЕМ АВТОМАТИЧЕСКОГО ТЕСТИРОВАНИЯ ЗНАНИЙ

В.А. Редюк

Научный руководитель – к.т.н., доцент В.И. Поляков

Вся система нашего образования, особенно среднего и высшего, построена на проверке знаний учащихся. Способов проверки существует большое количество. В последние несколько лет большую популярность набирает проверка знаний в форме тестирования. Особенно актуально это стало после введения обязательного ЕГЭ. Тестирование в своей первоначальной форме более просто для проверки знаний, чем остальные виды проверки. Но само проведение тестирования и проверка его результатов могут занять весьма продолжительное время. С этой точки зрения более предпочтительным оказывается компьютерное тестирование.

Проведение тестирования имеет ряд существенных недостатков. Большая часть этих недостатков связана с разработкой тестовых заданий. При объективном анализе полученных

результатов в поле зрения попадает не просто совокупность ответов на вопросы, но и относительная сложность вопросов. Несмотря на постоянные попытки создавать более менее равные вопросы, все равно, какие-то будут сложнее, а какие-то легче. Кроме того, не исключена возможность списывания теста учащимся. Это искажает результаты теста и приводит к необходимости учета вероятностной составляющей при их анализе или может привести к ситуации, когда учащийся может ошибочно попасть под подозрение на предмет взлома или обмана системы. К сожалению, в настоящее время нет единого и надежного способа проверки адекватности результатов тестирования. Поэтому одним из главных вопросов является изучение способов оценки адекватности результатов тестирования.

Целью работы является обзор имеющихся методик тестирования, определение каким образом они борются с обозначенной проблемой и выявление ключевых особенностей каждой их методик.

Для анализа были выбраны интегративные, адаптивные и критериально-оценочные методики тестирования. Для каждой из методик было рассмотрено некоторое количество используемых алгоритмов. Далее в процессе работы были определены достоинства и недостатки каждого из алгоритмов, и произведена обобщенная оценка каждой из методик с целью дальнейшего сравнения. На следующем шаге полученные результаты были использованы для выделения общих недостатков в системах автоматического тестирования знаний и предложены несколько вариантов решения обозначенных проблем, а именно создание нового алгоритма с использованием преимуществ алгоритмов, рассмотренных на предыдущих шагах.

В результате работы был выделен ряд недостатков, влияющих на точность способов проверки адекватности результатов тестирования. На основе полученных данных был сделан вывод о рациональности создания нового алгоритма с целью повышения точности анализа результатов тестирования.

УДК 004.021 + 004.62

МЕТОДЫ ИНТЕГРАЦИИ БИБЛИОТЕЧНОГО РЕСУРСА С ИНФОРМАЦИОННЫМИ СИСТЕМАМИ НИУ ИТМО

А.К. Барабанов

Научный руководитель – к.т.н. А.В. Куркин

Краткое введение, постановка проблемы. Информационная система библиотеки традиционно направлена в первую очередь на автоматизацию внутренних библиотечных процессов. Такой подход ограничивает возможности использования существующих информационных сервисов библиотеки конечными пользователями – студентами, научными сотрудниками, преподавателями. Курс на создание информационно образовательной среды (ИОС) вуза, продиктованный требованиями федеральных государственных образовательных стандартов высшего профессионального образования, подразумевает рассмотрение библиотечного ресурса в качестве ведущего информационно-образовательного центра. Однако помимо библиотечного ресурса в вузе существуют и другие центры информационно образовательной среды.

В связи с этим необходимо рассмотреть функционирование библиотечного ресурса в рамках информационно образовательной инфраструктуры вуза и возможности интеграции с функционирующими в НИУ ИТМО информационными системами. В свою очередь, информационные системы вуза обладают своими особенностями, поэтому необходимо спроектировать собственные механизмы интеграции библиотечного ресурса с ними.

Цель работы. Необходимо определить особенности информационных систем НИУ ИТМО, для которых целесообразна интеграция с библиотечным ресурсом вуза, и разработать соответствующие методы интеграции.

Базовые положения исследования. Библиотечный ресурс построен на основе системы «Руслан». Среди информационных систем ВУЗа, интересных с точки зрения интеграции с ними, можно выделить систему «Библиотекарь 5» Института холода и биотехнологий (ИХиБТ), центр информационных систем (ЦИС) НИУ ИТМО, систему дистанционного обучения (СДО) НИУ ИТМО. Стандартные механизмы интеграции, представленные в ИС «Руслан», не срабатывают на системе, имеющейся в ИХиБТ. Необходимо исследовать специфику интеграционных возможностей ЦИС и СДО.

Промежуточные результаты. Проведен анализ информационной системы библиотечного ресурса НИУ ИТМО и информационных систем вуза.

Исходя из его результатов, выявлены возможности модернизации ИОС вуза:

- определены интеграционные возможности ЦИС НИУ ИТМО и СДО НИУ ИТМО;
- разработан план обновления программно-аппаратных средств библиотечного ресурса;
- выявлены особенности системы «Библиотекарь 5» ИХиБТ, определена необходимость миграции данных из этой системы в систему «Руслан».

Основной результат. Разработаны методы интеграции библиотечного ресурса с выбранными информационными системами:

- с центром информационных систем НИУ ИТМО;
- с системой дистанционного обучения НИУ ИТМО;
- с информационной системой «Библиотекарь 5» ИХиБТ путем разработки модели односторонней миграции данных в систему «Руслан».

В дальнейшем, описанные методы, возможно, модифицировать и применять для интеграции библиотечного ресурса с другими системами.