

№ 1 (81) январь-февраль 2014

Издается с 2002 года. Выходит 6 раз в год

Учредитель – федеральное государственное бюджетное образовательное учреждение
высшего профессионального образования
«Государственный университет —
учебно-научно-производственный комплекс» (Госуниверситет – УНПК)

Редакционный совет

Голенков В.А., председатель
Радченко С.Ю., заместитель председателя
Борзенков М.И., секретарь

Астафичев П.А., Иванова Т.Н., Киричек А.В.,
Колчунов В.И., Константинов И.С.,
Новиков А.Н., Попова Л.В., Степанов Ю.С.

Главный редактор

Константинов И.С.

Редколлегия

Архипов О.П. (Орел, Россия)
Аверченков В.И. (Брянск, Россия)
Бок Т. (Мюнхен, Федеративная Республика Германия)
Гайндрик К. (Кишинев, Молдова)
Долгий А. (Сент-Этьен, Франция)
Еременко В.Т. (Орел, Россия)
Иванников А.Д. (Москва, Россия)
Ипатов О.С. (Санкт-Петербург, Россия)
Колоколов Ю.В. (Ханты-Мансийск, Россия)
Коськин А.В. (Орел, Россия)
Маркарян Г. (Ланкастер, Великобритания)
Подмастерьев К.В. (Орел, Россия)
Поляков А.А. (Москва, Россия)
Распопов В.Я. (Тула, Россия)

Рубрики номера

1. Математическое и компьютерное моделирование.....5-29
2. Информационные технологии в социально-экономических и организационно-технических системах30-58
3. Автоматизация и управление технологическими процессами и производствами.....59-66
4. Математическое и программное обеспечение вычислительной техники и автоматизированных систем.....67-74
5. Телекоммуникационные системы и компьютерные сети.....75-121

Редакция

О.И. Константинова
К.Д. Оболенская
А.А. Митин

*Сдано в набор 15.12.2013 г.
Подписано в печать 26.12.2013 г.
Формат 60x88 1/8.*

*Усл. печ. л. 7,5. Тираж 300 экз.
Заказ № 01/14П1*

*Отпечатано с готового оригинал-макета
на полиграфической базе
ФГБОУ ВПО «Госуниверситет - УНПК»
302030, г. Орел, ул. Московская, 65*

*Подписной индекс 15998
по объединенному каталогу*

«Пресса России»

**Материалы статей печатаются в авторской редакции.
Право использования произведений предоставлено
авторами на основании п. 2 ст. 1286 Четвертой части
ГК РФ.**

Журнал входит в **Перечень ведущих рецензируемых научных журналов и изданий**, определенных ВАК для публикации трудов на соискание ученых степеней кандидатов и докторов наук.

Адрес учредителя журнала

302020, г. Орел, Наугорское шоссе, 29
(4862) 42-00-24; www.gu-unpk.ru;
E-mail: unpk@ostu.ru

Адрес редакции

302020, г. Орел, Наугорское шоссе, 40
(4862) 43-40-39; www.gu-unpk.ru;
E-mail: isit@ostu.ru

Зарег. в Федеральной службе по надзору в сфере связи, информационных технологий и массовых коммуникаций.
Св-во о регистрации средства массовой информации ПИ № ФС77-47350 от 03.11.2011 г.

№ 1 (81) January-February 2014

The journal is published since 2002, leaves six times a year

The founder – State University – Education-Science-Production Complex

Editorial council

Golenkov V.A., president
Radchenko S.Y., vice-president
Borzenkov M.I., secretary

Astafichev P.A., Ivanova T.N., Kirichek A.V.,
Kolchunov V.I., Konstantinov I.S.,
Novikov A.N., Popova L.V., Ctepanov Y.S.

Editor-in-chief

Konstantinov I.S.

Editorial board

Arhipov O.P. (Orel, Russia)
Averchenkov V.I. (Bryansk, Russia)
Bok T. (Munich, Federal Republic of Germany)
Gaidrik K. (Kishinev, Moldova)
Dolgij A. (Saint-Etienne, France)
Eremenko V.T. (Orel, Russia)
Ivannikov A.D. (Moscow, Russia)
Ipatov O.S. (St. Petersburg, Russia)
Kolokolov J.V. (Khanty-Mansiysk, Russia)
Koskin A.V. (Orel, Russia)
Markaryan G. (Lancaster, Great Britain)
Podmasteriev K.V. (Orel, Russia)
Polyakov A.A. (Moscow, Russia)
Raspopov V.Ya. (Tula, Russia)

*It is sent to the printer's on 15.12.2013,
26.12.2013 is put to bed
Format 60x88 1/8.*

*Convent. printer's sheets 7,5. Circulation 300 copies
The order № 01/14П1*

*It is printed from a ready dummy layout
on polygraphic base of State University – ESPC
302030, Orel, Moskovskaya street, 65*

*Index on the catalogue
«Pressa Rossii» 15998*

In this number

1. Mathematical and computer simulation.....5-29
2. Information technologies in social and economic and organizational-technical systems.....30-58
3. Automation and control of technological processes and manufactures.....59-66
4. Software of the computer facilities and the automated systems.....67-74
5. Telecommunication systems and computer networks.....75-121

The editors

Konstantinova O.I.
Obolenskaya K.D.
Mitin A.A.

The address of the founder of journal

302020, Orel, Highway Naugorskoe, 29
(4862) 42-00-24; www.gu-unpk.ru;
E-mail: unpk@ostu.ru

The address of the editorial office

302020, Orel, Highway Naugorskoe, 40
(4862) 43-40-39; www.gu-unpk.ru;
E-mail: isit@ostu.ru

*Journal is registered in Federal Service for
Supervision in the Sphere of Telecom, Information
Technologies and Mass Communications.
The certificate of registration
ПИ № ФС77-47350 from 03.11.2011.*

© State University – ESPC, 2014

Journal is included into the list of the Higher Attestation Commission for publishing the results of theses for competition the academic degrees.

СОДЕРЖАНИЕ

МАТЕМАТИЧЕСКОЕ И КОМПЬЮТЕРНОЕ МОДЕЛИРОВАНИЕ

АБАКАРОВА О.Г.

Обобщенная модель OLAP-системы поддержки аналитической деятельности правоохранительных органов.....5-11

АФАНАСЬЕВ В.В., ЛЕБЕДЕНКО Е.В.

Графоаналитическая модель процесса формирования глобальной схемы мультибазы данных с учетом этапов ее реструктуризации..... 12-18

ДЕМКИН В.И., СО ЛИН МАУНГ

Старт-стопное управление шаговым двигателем на одном шаге без переключения фаз..... 19-23

ПЛЕТНЁВ К.В., ПРОЗОРОВ Д.Е.

Анализ метода марковской параметризации речевых сигналов..... 24-29

ИНФОРМАЦИОННЫЕ ТЕХНОЛОГИИ В СОЦИАЛЬНО-ЭКОНОМИЧЕСКИХ И ОРГАНИЗАЦИОННО-ТЕХНИЧЕСКИХ СИСТЕМАХ

КАРМИНСКАЯ Т.Д., ТЕЙ Д.О., ТАТЬЯНИН В.М., РУСАНОВ М.А.

Формирование контрольных цифр приема с учетом прогноза потребностей региональной экономики.....30-38

МАКАРОВА М.Ю.

Построение модели управления профориентационными процессами.....39-46

ТОЛКУНОВ А.А.

Алгоритмы формирования выборок типичных и полезных комментариев в ходе обработки результатов общественного обсуждения законопроектов.....47-58

АВТОМАТИЗАЦИЯ И УПРАВЛЕНИЕ ТЕХНОЛОГИЧЕСКИМИ ПРОЦЕССАМИ И ПРОИЗВОДСТВАМИ

ЕРЕМЕНКО А.В., ОСИПОВ А.Н., МИШИН Д.С., ПЕНЬКОВ Н.И., ПОЛЯКОВА А.Г., ИВАНОВА С.Н.

Проблемы обмена данными в интегрированной среде АСУ промышленных предприятий.....59-66

МАТЕМАТИЧЕСКОЕ И ПРОГРАММНОЕ ОБЕСПЕЧЕНИЕ ВЫЧИСЛИТЕЛЬНОЙ ТЕХНИКИ И АВТОМАТИЗИРОВАННЫХ СИСТЕМ

НОВИКОВ Е.И.

Методика оценки оперативности автоматизированных информационных систем..... 67-74

ТЕЛЕКОММУНИКАЦИОННЫЕ СИСТЕМЫ И КОМПЬЮТЕРНЫЕ СЕТИ

ОЛАДЬКО А.Ю., АТКИНА В.С.

Многоагентная система защиты информации в операционных системах, построенная с использованием технологий иммунокомпьютинга..... 75-83

БОРИСОВ О.М., РЯЗАНЦЕВ П.Н., КУЗЬМИН В.В., ФОМИН Н.И.

Алгоритмы обеспечения живучести виртуальной частной сети автоматизированной системы управления промышленного предприятия (ВЧС АСУ ПП) на основе оптимального распределения сетевых ресурсов....84-91

ГРИШАКОВ В.Г.

Представление самоорганизующейся ИТ-инфраструктуры в виде модели жизнеспособной системы.....92-102

ЛЯКИШЕВ А.А.

Способы и приемы оптимизации процессов информационного обмена в модульных структурах систем сбора и обработки данных АСУ газотранспортного предприятия..... 103-109

ФИСЕНКО В.Е., ФИСУН А.П.

Формирование требований к показателям надежности элементов телекоммуникационных систем.....110-113

ШУЛЬГИН Р.Н.

Математическое моделирование процессов сбора и обработки данных о нарушениях функционирования в вычислительных сетях газотранспортного предприятия..... 114-121

CONTENT

MATHEMATICAL AND COMPUTER SIMULATION

O.G. ABAKAROVA

The generalized OLAP-systems model of support of analytical activity of law enforcement agencies.....5-11

V.V. AFANAS'EV, E.V. LEBEDENKO

Graphic-analytical process model of global scheme multidatabase generation considering its stages of restructuring.....12-18

V.I. DYOMKIN, SO LIN MAUNG

Start-stop management of step motor on one step without phase switching.....19-23

K.V. PLETNYOV, D.E. PROZOROV

Analysis of the method of speech signals parameterization by simple Markov's chains.....24-29

INFORMATION TECHNOLOGIES IN SOCIAL AND ECONOMIC AND ORGANIZATIONAL-TECHNICAL SYSTEMS

T.D. KARMINSKAYA, D.O. TEJ, V.M. TAT'YANKIN, M.A. RUSANOV

Definition of education order subject to requirement forecast of regional economy.....30-38

M.Yu. MAKAROVA

The construction of model of vocational guidance processes control.....39-46

A.A. TOLKUNOV

Algorithms for generating typical and utility comments samples during the processing of the results of public discussion of bill.....47-58

AUTOMATION AND CONTROL OF TECHNOLOGICAL PROCESSES AND MANUFACTURES

A.V. ERYOMENKO, A.N. OSIPOV, D.S. MISHIN, N.I. PEN'KOV, A.G. POLYAKOVA, S.N. IVANOVA

Problems of data exchange in the integrated environment of the automated control systems of the industrial enterprises.....59-66

SOFTWARE OF THE COMPUTER FACILITIES AND THE AUTOMATED SYSTEMS

E.I. NOVIKOV

Evaluation methods of efficient property of automated information system.....67-74

TELECOMMUNICATION SYSTEMS AND COMPUTER NETWORKS

A.Yu. OLAD'KO, V.S. ATKINA

Multi-agent systems of information protection in the operating system built using the immunocomputing technology.....75-83

O.M. BORISOV, P.N. RYAZANCEV, V.V. KUZ'MIN, N.I. FOMIN

Survivability algorithms of virtual private network of automated control systems industrial enterprise (VPN ACS IE) based on the optimal allocation of network resources.....84-91

V.G. GRISHAKOV

Modeling self-organization IT-infrastructure as a viable system model.....92-102

A.A. LYAKISHEV

Methods and techniques of optimization of the process information exchange in the modular structure of data collection and processing systems in the ACS of a gas transportation company.....103-109

V.E. FISENKO, A.P. FISUN

Formation of requirements to reliability of elements of telecommunication systems.....110-113

R.N. SHUL'GIN

Mathematical modeling of collecting and data processing about operational occurrences in computer networks of gas transmission company.....114-121

УДК 681.3.06:65.011.2

О.Г. АБАКАРОВА

ОБОБЩЕННАЯ МОДЕЛЬ OLAP-СИСТЕМЫ ПОДДЕРЖКИ АНАЛИТИЧЕСКОЙ ДЕЯТЕЛЬНОСТИ ПРАВООХРАНИТЕЛЬНЫХ ОРГАНОВ

Предложена обобщенная модель многомерного представления данных, основанная на использовании OLAP-технологии. Данные в модели связываются между собой при помощи ключевых признаков без организации отдельных таблиц. Рассмотрено практическое применение модели в аналитической деятельности региональных правоохранительных органов.

Ключевые слова: хранилище данных; модель OLAP-системы; гиперкуб; аналитическая деятельность; правоохранительные органы.

ВВЕДЕНИЕ

Для эффективного выполнения своих профессиональных функций, координации действий подразделений, анализа складывающейся ситуации по слабоструктурированным массивам данных правоохранительные органы территорий остро нуждаются в информационной поддержке своей деятельности. Возрастает роль аналитической работы, выражающейся в виде таких ключевых операций, как оперативный анализ, прогнозирование динамики, тенденций процессов, подготовка специфической отчетности, планирование реакций и долгосрочных действий. Накоплены большие массивы данных, на основе которых извлекается полезная информация, выявляются скрытые тенденции, строятся гипотезы и модели разрешения проблем. Основу информационных решений такого рода задач в региональных правоохранительных органах составляют сегодня предметно-специализированные информационные системы с реляционными системами управления базами данных (СУБД).

Однако специалисты признали, что реляционные и постреляционные (например, объектно-ориентированные) СУБД не могут эффективно использоваться для решения оперативно-аналитических задач в силу возникающих противоречий между обычными базами данных и системами анализа по степени детализации и формату хранимых данных, управлению данными, характеру запросов к ним, времени отклика системы, приоритетности характеристик системы и др. Назрела необходимость внедрения в правоохранительные органы регионов новой модели для эффективной поддержки аналитической деятельности и произвольных запросов аналитиков подразделений. Тем более, что современные технологии анализа информации и системы поддержки принятия решений интенсивно развиваются и используются в различных сферах деятельности для сбора, хранения и анализа многомерных данных в целях поддержки процессов принятия решений [1].

ПОСТАНОВКА ЗАДАЧИ

Основная сложность технологии анализа в правоохранительных органах заключается в том, что необходимая для анализа цифровая информация может быть представлена в виде отсканированных документов, электронных таблиц, реляционных и многомерных баз данных, изображений, речевой информации и т.д. Кроме того, от аналитиков правоохранительных органов зачастую требуют вести строгую отчетность разного уровня и степени сложности на основе накопленной за различные промежутки времени информации. Контроль непосредственной деятельности региональных правоохранительных органов в целях улучшения организации дальнейшей работы обуславливает формирование и расчет разнообразных статистических показателей.

Системы хранения и манипуляции данными используются в правоохранительных органах для выполнения разнообразных операций над информацией предметной области. Однако специфика такова, что данные приходится извлекать из множества отдельных таблиц реляционной или объектно-ориентированной базы данных (БД). А эта задача нетривиальна и технически сложно реализуема. Если в подготавливаемых отчетах отображается некоторый относящийся к конкретному вопросу комплекс данных, необходимо выполнить транзакции к распределенной БД, учитывая сложную структуру самих таблиц и многочисленных связей между ними. Чтобы устранить перечисленные недостатки, обеспечить аналитическую обработку на основе многомерного представления данных, создаются специализированные хранилища данных (ХД). Концепция ХД позволяет создать единую модель данных организации, удобный доступ к разнородной информации. Основные свойства ХД определены в известных правилах Кодда [1].

Высокая степень агрегации данных, собранных в единое хранилище, позволяет выполнять запросы и формировать необходимые отчеты за более короткое время при меньшей сложности самих запросов. По истечении некоторого промежутка времени данные в ХД устаревают и нуждаются в обновлении.

Хотя организация ХД по сравнению с оперативными БД с точки зрения анализа данных по многим измерениям является более продвинутой технологией, тем не менее, использование такого подхода нельзя признать максимально эффективным: данные по-прежнему хранятся в отдельных связанных между собой таблицах и для вывода зависимостей между ними необходимо каждый раз выполнять сложный запрос к базе данных. Необходима разработка эффективной модели многомерного представления данных.

РАЗРАБОТКА МОДЕЛИ МНОГОМЕРНОГО ПРЕДСТАВЛЕНИЯ ДАННЫХ ДЛЯ ПРАВООХРАНИТЕЛЬНЫХ ОРГАНОВ

Предлагается модель многомерного представления данных, основанная на использовании OLAP(On-Line Analytical Processing)-технологии. Ее основная особенность заключается в том, что в модели не предусмотрена организация данных в виде отдельных таблиц, а все данные связаны между собой при помощи ключевых признаков. Модель позволяет конечному пользователю находить скрытые связи между признаками, основываясь только на данных хранилища; анализ, вычисления и другие операции производятся по атрибутам признаков, тем самым обеспечивается необходимая детализация данных и поиск зависимостей.

При этом в разработанной модели структура данных содержит множество агрегированных данных. Это позволяет значительно повысить скорость обработки запросов и общую производительность системы. Оставаясь в пределах одного отчета, что особенно важно для пользователей в правоохранительных аналитических подразделениях, они могут выполнить анализ на разных иерархических уровнях – субъекта, района, перейти к рассмотрению детальной информации на уровне города, улицы.

В основе любой OLAP-системы лежит специально спроектированное хранилище данных, которое наследует свойства как навигационных, так и иерархических баз данных. Как правило, это хранилище строится на основе уже существующей базы данных или другого источника данных, что позволяет производить комплексный анализ собранной информации. В OLAP-системах данные представляются в виде гиперкуба, ячейки которого соответствуют некоторым исследуемым событиям (фактам) и содержат данные, количественно характеризующие анализируемые факты. Каждое событие может быть описано рядом измерений, соответствующих граням куба [2].

При проектировании ХД нашли широкое применение две схемы организации таблиц – «звезда» и «снежинка». Особенность схемы «звезда» – в наличии денормализованной таблицы фактов, содержащей уникальный составной ключ, объединяющий первичные ключи таблиц измерений. Поля таблиц соответствуют измерениям гиперкуба. Таблица фактов содержит числовые поля, на основании которых и получают далее агрегатные данные.

В сложных системах со многими иерархическими измерениями имеет смысл использовать схему «снежинка». В этом случае отдельные таблицы фактов создаются для возможных сочетаний уровней обобщения различных измерений. Хотя эта схема дает лучшую производительность, часто приводит к избыточности данных и усложнению структуры БД.

Обе модели представлены центральной таблицей фактов, связанной со множеством измерений. Она содержит понятия о событиях, которые, будучи выраженными при помощи численных данных, представляют собой сведения об определенном анализируемом процессе. Главной целью введения таблиц измерений является необходимость описания измерений некоторыми наборами атрибутов, каждая таблица измерений находится в отношении «один ко многим» с таблицей фактов [3].

Предлагаемая обобщенная модель OLAP-системы для правоохранительных органов, показанная на рисунке 1, содержит следующие компоненты: источники данных, хранилище данных, OLAP-сервер и информацию для конечных пользователей.

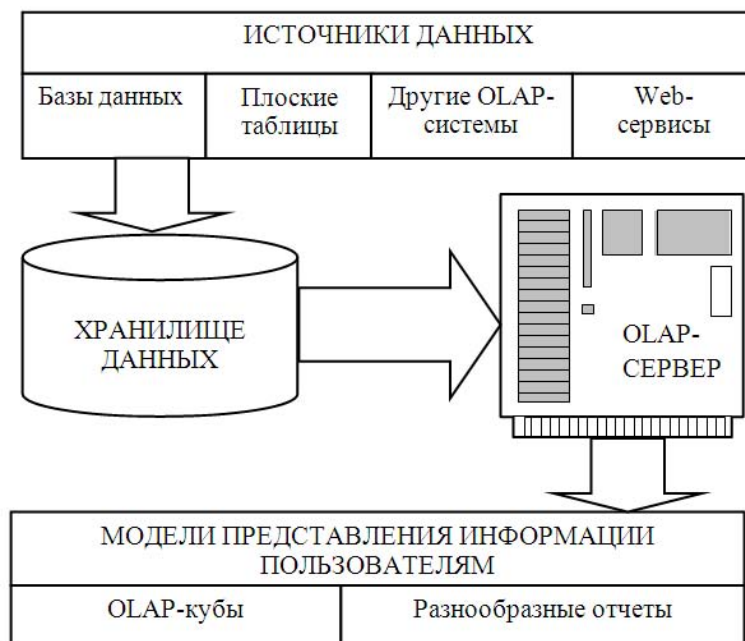


Рисунок 1 – Обобщенная модель OLAP-системы для правоохранительных органов

В качестве источников данных могут выступать источники различного типа – ограничений здесь нет, но наиболее распространены реляционные БД, таблицы Microsoft Excel, CSV-таблицы (их называют «плоскими» таблицами), другие OLAP-системы, web-сервисы (RSS, WSDL, CSDL) [4].

В OLAP-системах ХД используется как наиболее оптимальный источник информации для построения OLAP-куба, что обусловлено следующими факторами: наряду с необходимой для OLAP-системы информацией источники данных могут содержать и избыточную информацию, что может привести к усложнению модели и снижению скорости работы системы. К данным, на основе которых строится OLAP-куб, предъявляются особые требования, например, денормализация. Выполнение их для некоторых источников данных может быть затруднительным или даже невыполнимым; как правило, создание OLAP-куба требует больших вычислительных мощностей.

В случае использования ХД перенос информации из исходных источников в хранилище может быть организован с учетом минимизации нагрузки на источники данных. Одним из программных продуктов, позволяющих строить OLAP-кубы на основе хранилищ данных, является Microsoft Business Intelligence Development Studio. Для построения OLAP-

куба в данной среде разработчику вначале необходимо выбрать ХД как источник информации, а затем определить таблицы фактов, измерений и их иерархии, связи между разными сущностями в названной модели. Эти процессы выполняются в полуавтоматическом режиме, когда программное обеспечение самостоятельно выбирает наиболее подходящие опции для каждого шага, оставляя разработчику возможность корректировать данный набор параметров [5].

Для конечных пользователей Microsoft Business Intelligence Development Studio построенный куб используется и как средство для просмотра стандартных отчетов, предоставляемых системой, и как инструмент для формирования собственных отчетов с помощью установленного программного обеспечения, например, Microsoft Excel. Дополнительно существует возможность организации работы пользователей непосредственно с локальной копией OLAP-куба либо с использованием информационного портала как централизованного интерфейса доступа к данным отчетности и мониторинга правоохранительного органа.

ПРАКТИЧЕСКОЕ ПРИМЕНЕНИЕ МОДЕЛИ

Рассмотрим разработанную модель применительно к мониторингу деятельности правоохранительных органов. Согласно требованиям, регламентирующим деятельность правоохранительных органов, при совершении правонарушения заполняются регистрационные формы (карточки) о преступлении, лице, совершившем преступление, подозреваемом, обвиняемом, физическом лице, потерпевшем в результате преступления.

Принимая во внимание стандартные регистрационные формы, можно проследить основные виды возможных связей: преступление – сотрудник, расследующий дело; преступление – оружие и т.д., но отражение более сложных связей при такой организации данных является затруднительным. Применение реляционного хранилища данных позволяет решить эту проблему, связав, к примеру, персональные сведения о подозреваемом с конкретной статьей правонарушения (проследив зависимость применимости конкретной статьи от условного типа подозреваемых).

Задачу анализа данных можно решить поэтапно. Сначала строится абстрактная модель хранилища данных, содержащая следующие таблицы: «Преступление», «Подозреваемый», «Потерпевший». Таблицы представляют собой измерения в ХД. Фактом или процессом, которые в дальнейшем будут анализироваться, являются статус уголовного дела или сам факт совершения преступления. В любом отделении правоохранительных органов хранится общая БД по всем зарегистрированным лицам, аналогом которой может стать таблица «Лицо». Исходя из этого, ХД можно спроектировать так, чтобы наиболее близко соответствовать структурам БД, применяемых в государственных правоохранительных органах.

В качестве центральной можно выбрать таблицу фактов, содержащую информацию о совершенном преступлении, в частности: уголовное дело возбуждено или закрыто (рис. 2). Ее основной целью является построение связей между фактами и описывающими их измерениями. Записи по всем заведенным уголовным делам содержатся в соответствующей таблице, в которой также находится информация, напрямую связанная с преступлением: оружие, уровень алкоголя в крови подозреваемого и др.

Особенностью создаваемого ХД является использование единой картотеки, содержащей данные по всем лицам, которые зарегистрированы в общей базе правоохранительных органов. Этот факт позволяет рассматривать любого зарегистрированного человека в качестве возможного соучастника преступления. Единая картотека является аналогом реальной модели, используемой в органах правопорядка. Каждое зарегистрированное лицо содержит уникальный идентификационный номер, а также полную информацию, способную представлять интерес для следствия.

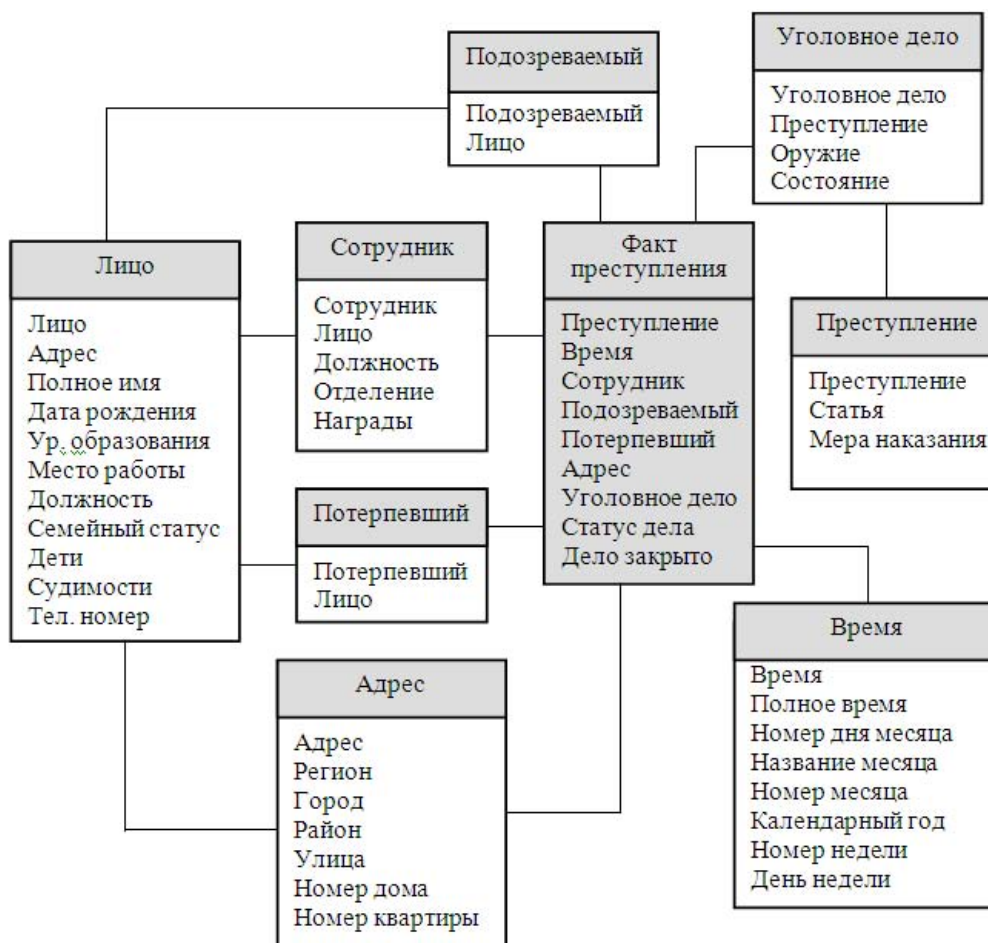


Рисунок 2 – Модель хранилища данных системы поддержки аналитической деятельности правоохранительных органов

Для моделирования перечня статей по различным правонарушениям можно ввести отдельную таблицу, содержащую основные статьи Уголовного кодекса. Ключевой информацией в данном контексте являются номер статьи и наказание за определенное преступление. Совершенные преступления и лица, зарегистрированные в картотеке, связаны с уникальным географическим адресом. Это позволяет проследить плотность распределения преступлений в каждом отдельно взятом регионе. Таблицы «Подозреваемый», «Потерпевший» и «Сотрудник» реализуют тип связи «многие ко многим» и связаны с единой картотекой лиц. Время как ключевое измерение в любой многомерной БД описано при помощи отдельной таблицы и представлено такими атрибутами, как день и номер недели, название месяца, номер месяца по порядку, а также полная дата.

Для организации данных в единую структуру можно использовать модель «снежинка». Измерения нормализовать для представления в виде нескольких таблиц, промежуточные таблицы ввести для реализации связей «многие ко многим». Полученное ХД заполнить при помощи случайных тестовых значений.

На втором этапе решения задачи анализа данных создается многомерная БД, фактическими записями которой является статистика преступлений со всеми необходимыми для исследования показателями. Итоговые диаграммы и сводные таблицы, представляющие собой готовые отчеты, можно построить при помощи программного продукта Microsoft Excel на основе данных, полученных из OLAP-куба в Business Intelligence Development Studio. Использование приведенных отчетов позволяет производить углубленный анализ данных, определять скрытые закономерности и находить пики различных событий.

Разработанная модель аналитической системы позволяет генерировать следующие виды отчетов:

- распределение количества заведенных уголовных дел по дням недели в контексте выбранного географического региона;
- преступления, совершенные в течение определенного сезона года, в развертке по дням недели;
- количество раскрытых преступлений, связанных с уровнем образования сотрудника правоохранительных органов;
- криминогенная обстановка в выбранном городе, рассмотренная относительно дней недели. Данная выборка позволяет определить пики преступной активности; количество совершенных преступлений определенным лицом относительно его семейного статуса, должности, уровня образования и др.; статистику по определенному виду оружия, связанному с выбранной статьей Уголовного кодекса.

Таким образом, основываясь только на информации, содержащейся в стандартных регистрационных формах, заполняемых на месте преступления, можно получить различные отчетные данные. Количество уголовных дел и их статус могут быть показаны в проекции на выбранный географический регион либо применительно к определенному отделению правоохранительных органов, сотруднику, подозреваемому и т.д.

Реляционные данные могут быть представлены в развертке на различные атрибуты измерений: уровень образования подозреваемого, его семейный статус, оружие и др.

ЗАКЛЮЧЕНИЕ

Применение разработанной модели системы поддержки аналитической деятельности правоохранительных органов позволяет достичь следующих преимуществ над традиционными методами анализа данных в подобных системах:

- возможность построения системы информационного обеспечения, являющейся достаточно гибкой для формирования отчетов любой сложности на основе огромного количества данных за короткий период времени;
- возможность анализа различных показателей по атрибутам признаков, отслеживания взаимосвязей и облегчения принятия решений на базе выявленных отношений;
- автоматизация аналитической системы позволяет уменьшить количество сотрудников, обслуживающих данное информационное решение;
- построение информационной системы, ориентированной на конечного пользователя, приводит к отсутствию необходимости ручного сбора информации из различных источников.

СПИСОК ЛИТЕРАТУРЫ

1. Барсегян А.А., Куприянов М.С., Степаненко В.В., Холод И.И. Технологии анализа данных: Data Mining, Visual Mining, Text Mining, OLAP. – СПб: БХВ-Петербург, 2008. – 384 с.
2. Kimpball R. The Data Warehouse Toolkit. – Wiley India Pvt. Ltd., 2009. – 460 p.
3. Wrembel R., Koncilia Chr. Data Warehouses and OLAP: Concepts, Architectures and Solutions. – Idea Group Inc., 2007. – 332 p.
4. Ponniah P. Data Warehousing Fundamentals for IT Professionals. – John Wiley&Sons Inc., 2010. – 571 p.
5. Бергер А.Б., Горбач И.В. Microsoft SQL Server 2005 Analysis Services. OLAP и многомерный анализ данных. – СПб: БХВ-Петербург, 2007. – 928 с.

Абакарова Оксана Гасановна

ФГБОУ ВПО Дагестанский государственный технический университет, г. Махачкала

Кандидат экономических наук, доцент, доцент кафедры «Прикладная информатика в юриспруденции»

Тел.: 8 (8722) 62-00-58

E-mail: oksgasan@gmail.com

O.G. ABAKAROVA (*Candidate of Economic Science, Associate Professor, Associate Professor of the Department «Applied Computer Science in Jurisprudence» Dagestan State Technical University, Makhachkala*)

THE GENERALIZED OLAP-SYSTEMS MODEL OF SUPPORT OF ANALYTICAL ACTIVITY OF LAW ENFORCEMENT AGENCIES

The generalized model of multidimensional data presentation based on using of OLAP-technology is offered. Data in model communicate among themselves by means of key signs, without the organization of separate tables. Practical application of model in analytical activity of regional law enforcement agencies is considered.

Keywords: *data warehouse; OLAP-system model; hypercube; analytical activity; law enforcement agencies.*

BIBLIOGRAPHY (TRANSLITERATED)

1. Bargesyan A.A., Kupriyanov M.S., Stepanenko V.V., Xolod I.I. *Texnologii analiza danny'x: Data Mining, Visual Mining, Text Mining, OLAP.* – SPb: BXV-Peterburg, 2008. – 384 s.
2. Kimpball R. *The Data Warehouse Toolkit.* – Wiley India Pvt. Ltd., 2009. – 460 p.
3. Wrembel R., *Koncilia Chr.Data Warehouses and OLAP: Concepts, Architectures and Solutions.* – Idea Group Inc., 2007. – 332 p.
4. Ponniah P. *Data Warehousing Fundamentals for IT Professionals.* – John Wiley&Sons Inc., 2010. – 571 p.
5. Berger A.B., Gorbach I.V. *Microsoft SQL Server 2005 Analysis Services. OLAP i mnogomerny'j analiz danny'x.* – SPb: BXV-Peterburg, 2007. – 928 s.

УДК 004.75

В.В. АФАНАСЬЕВ, Е.В. ЛЕБЕДЕНКО

ГРАФОАНАЛИТИЧЕСКАЯ МОДЕЛЬ ПРОЦЕССА ФОРМИРОВАНИЯ ГЛОБАЛЬНОЙ СХЕМЫ МУЛЬТИБАЗЫ ДАННЫХ С УЧЕТОМ ЭТАПОВ ЕЕ РЕСТРУКТУРИЗАЦИИ

В статье обосновывается выбор графового представления глобальной схемы, а также представляется графоаналитическая модель процесса ее формирования с учетом реструктуризации информационно-логических моделей в локальных базах данных узлов распределенной базы данных особого типа – мультибазе данных.

Ключевые слова: мультибаза данных; графовое представление глобальной схемы; графоаналитическая модель; реструктуризация.

ВВЕДЕНИЕ

Базовой концепцией хранения данных в распределенных информационных системах (РИС), функционирующих в составе автоматизированных систем управления производством (АСУП), является приближение к местам их обработки. Это означает, что в каждый момент времени в узлах РИС, реализующих те или иные функциональные подсистемы АСУП, в их базах данных (БД) должны быть актуальные для обработки данные – как локальные (формируемые самим узлом РИС), так и распределенные (получаемые от других узлов РИС).

В случае организации подсистемы хранения данных РИС в виде распределенной базы данных (РБД) проблема актуальности данных решается использованием методов реплицирования, встроенных в большинство современных СУБД [1].

Между тем практика формирования крупномасштабных корпоративных АСУП показывает, что инфраструктура РИС в них формируется путем интеграции разнородных информационных систем, функционирующих в составе предприятий, являющихся составной частью интегрированного производственного процесса [2].

В условиях динамично изменяющихся потребностей рынка такое интегрированное производство зачастую требует не только оперативного изменения хранящихся в узлах РИС данных, но и самой модели данных, связанной с динамической модификацией предметной области, на основе которой формируется БД.

Использование РБД, поддерживающей единую модель данных всей предметной области интегрированного производства, является нецелесообразным. Наиболее распространенным подходом организации системы хранения данных в подобных условиях является технология мультибаз данных (МБД) – совокупности локальных БД в узлах РИС, поддерживающих собственные модели данных предметной области, специфической для данного узла, и интегрирующих их в рамках единой структуры, именуемой глобальной схемой [3].

Применение в корпоративных АСУП технологии МБД с глобальной схемой позволяет отчасти добиться компромисса между независимым функционированием отдельных подсистем такой АСУП и обеспечением их интеграции для реализации динамически изменяемого производственного процесса. В частности, благодаря использованию МБД появляется возможность расширения функциональности локальной ИС или изменения модели предметной области без оказания существенного влияния на функционирование остальных узлов РИС.

Тем не менее, выполнение подобной локальной реструктуризации данных в узлах МБД требует постоянной корректировки и согласования глобальной схемы МБД – интеграционной основы РИС. В условиях, когда моменты локальной реструктуризации являются относительно редкими, корректировка глобальной схемы МБД может быть реализована организационно путем согласования структуры глобальной схемы

совокупностью администраторов локальных БД РИС. Однако с повышением частоты моментов реструктуризации, связанных, например, с динамически изменяемым производственным циклом, требующим существенного изменения моделей предметных областей локальных БД или даже включения в состав РИС новых функциональных узлов, качественное согласование глобальной схемы МБД, обеспечивающее поддержание ее актуального состояния, становится проблематичным.

В данной статье описывается подход к решению задачи повышения актуальности глобальной схемы МБД в условиях реструктуризации входящих в ее состав локальных БД, основанный на разработке графоаналитической модели процесса ее формирования.

ГРАФОВОЕ ПРЕДСТАВЛЕНИЕ ГЛОБАЛЬНОЙ СХЕМЫ МУЛЬТИБАЗЫ ДАННЫХ

Существует несколько подходов формального представления глобальной схемы (ГС), предназначенных для решения различных задач, связанных с оптимизацией процесса информационного обмена в рамках РИС, базирующихся на МБД [4, 5].

В [6] показано, что средствами теории графов можно моделировать различные структуры, в том числе глобальные схемы мультибаз данных. Пусть некоторая база данных DB_0 определена информационно-логической моделью (ИЛМ), заданной простым ориентированным графом $G(X, U)$, где $X = \{x_i \mid i = \overline{1, I}\}$ – множество вершин графа G (сущностей, входящих в состав ИЛМ), $U = \{u_m(x_i, x_j)\}$, $m = \overline{1, M}$, $i = \overline{1, I}$, $j = \overline{1, I}$ – множество ребер графа G (связей между сущностями).

Введем ограничения:

1. Граф не будет иметь петель (ребер, соединяющих вершину саму с собой), если выполняется ограничение

$$O_1 : \neg \exists u_m(x_k, x_k) \in U; \quad m = \overline{1, M}; \quad k = \overline{1, I}. \quad (1)$$

2. Граф не будет мультиграфом, если выполняется ограничение

$$O_2 : \forall u_m(x_i, x_j) \in U; \quad m = \overline{1, M}; \quad i \in \overline{1, I}; \quad j \in \overline{1, I}; \Rightarrow \\ \neg \exists u_l(x_i, x_j) \in U; \quad l \neq m; \quad l \in \overline{1, M}. \quad (2)$$

3. Вершина-приемник связана только с одной вершиной-источником (ограничения на число связей нет):

$$O_3 : \forall u_m(x_i, x_j) \in U; \quad m = \overline{1, M}; \quad i \in \overline{1, I}; \quad j \in \overline{1, I}; \Rightarrow \\ \neg \exists u_n(x_k, x_j) \in U; \quad m \neq n; \quad i \neq k; \quad n \in \overline{1, M}; \quad k \in \overline{1, I}. \quad (3)$$

Указанные предикатные ограничения позволяют ввести шаблоны ориентированных графов, моделирующих структуру ИЛМ БД узла мультибазы данных (рис. 1).

С другой стороны, существуют методы отображения в рамках БД реляционного типа графовых структур. Классификация методов хранения графовых структур в БД реляционного типа представлена на рисунке 2.

В работах [6, 7] показано, что в общем случае такие графовые структуры, как направленные деревья, мультиграфы без петель и замкнутые мультиграфы без петель, могут быть адекватно представлены эквивалентной схемой данных БД реляционного типа. На рисунке 3 представлена схема данных $DB_{uz} = \langle X, G \rangle$ – реляционная БД, отображающая G , где $X = \{ID_X, \text{Наименование_вершины}\}$, $G = \{ID_G, ID_XI \text{ (FK)}, ID_XJ \text{ (FK)}, \text{Наименование_ребра}\}$.

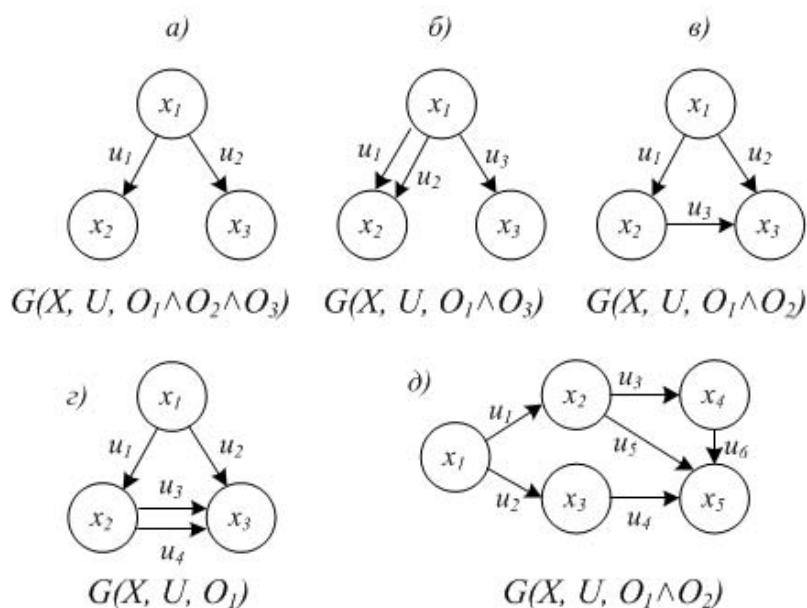


Рисунок 1 – Шаблоны ориентированных графов, моделирующих структуру ИЛМ БД узла мультибазы данных

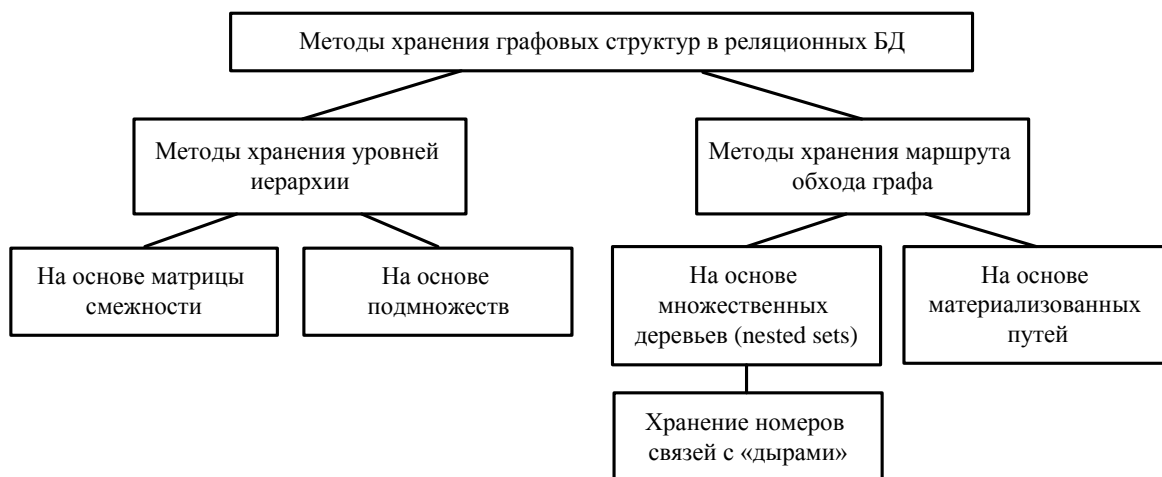


Рисунок 2 – Классификация методов хранения графовых структур в реляционных БД

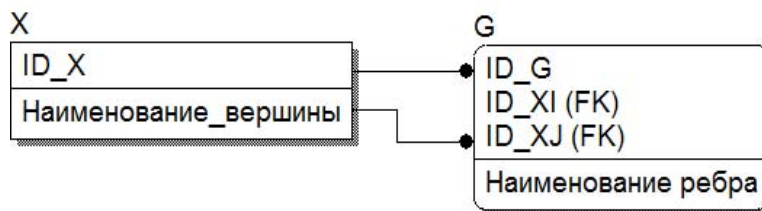


Рисунок 3 – Схема данных $BD_{уз}$ отображающая структуру простого ориентированного графа G_0

Поскольку в общем случае структура ГС МБД является вложенной иерархией информационно-логических моделей локальных баз данных (ИЛМ ЛБД), входящих в состав МБД и моделей физического уровня каждой ЛБД, представленных множествами таблиц в ЛБД и связей между ними [8], с учетом рассмотренной выше связи графовых структур и БД реляционного типа, выдвинем гипотезу о возможности задания ГС МБД, представляемой схемой данных реляционной БД, в виде ориентированного графа без петель.

В обобщенном виде такой граф представлен на рисунке 4.

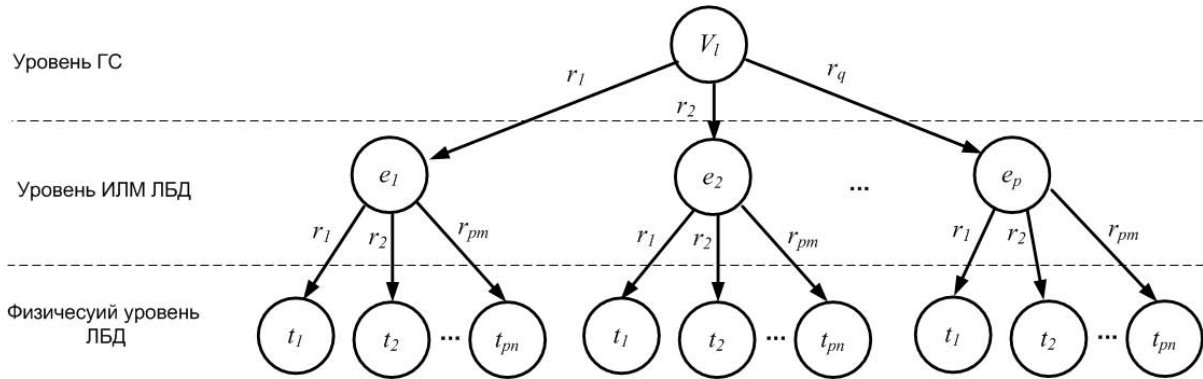


Рисунок 4 – Обобщенное представление ГС МБД в виде ориентированного графа без петель

Обоснуем данное представление, определив:

- подграфы e_1, e_2, \dots, e_p модели данных уровня ИЛМ ЛБД;
- подграфы t_1, t_2, \dots, t_{pn} модели данных физического уровня ЛБД;
- модели базовых отношений (1:М, М:1, М:М) сущностей.

Пусть $G_l(V_l, E_l)$ – граф ИЛМ ЛБД, где $V_l = (e_1, \dots, e_p), p \in N$ – множество сущностей;

$E_l = (r_1, \dots, r_q), q \in N$ – множество связей.

При этом с каждой $e \in V_l$ связано множество атрибутов сущностей

$Attr(e) = \{a_{e,0}, a_{e,1}, \dots, a_{e,m}\}$, где $a_{e,m} = \langle name, ref \rangle$ – составляющие атрибута.

На графе G_l можно определить следующие модели отношений сущностей.

Определим главную таблицу (master table – MT) сущности e_p как:

$$MT(e) = t, e \in V_l, t \in V_{ph},$$

а множество ее данных как

$$S(e), e \in V_l.$$

Тогда отношения 1 : М и М : 1 определяются как

$$(\forall e \in V_l)(\exists T_e \subset V_{ph}) : T_e = MT(e) \cup S(e). \tag{4}$$

Для определения отношений М : М в [8] обоснуется необходимость введения промежуточной (middle table):

$$Mid(r = (e_1, e_2)) = \begin{cases} t_m, (e_2, e_1) \in E_l \\ \emptyset, (e_2, e_1) \notin E_l \end{cases}, \tag{5}$$

а сами отношения определяются как

$$(\forall r = (e_1, e_2) \in E_l, e_1, e_2 \in V_l)(\exists E_r \subset E_{ph}) :$$

$$E_r = \{(t_1, t_2) \in E_{ph} \mid t_1, t_2 \in V_{ph}, t_1 = MT(e_1), t_2 = MT(e_2), Mid(r) = \emptyset\} \vee \tag{6}$$

$$\vee \{(t_1, t_m) \in E_{ph}, (t_2, t_m) \in E_{ph} \mid t_1, t_2, t_m \in V_{ph}, t_1 = MT(e_1), t_2 = MT(e_2), t_m = Mid(r) \neq \emptyset\}$$

На физическом уровне ЛБД может быть представлена графом реляционного представления данных $G_{ph}(V_{ph}, E_{ph})$, где $V_{ph} = \{t_1, t_2, \dots, t_{pn}\}, pn \in N$ – множество таблиц, представленных множеством полей $Fields(t) = \{f_{t,0}, f_{t,1}, \dots, f_{t,n_t}\}; E_{ph} = \{r_1, r_2, \dots, r_{pm}\}, pm \in N$ – множеством связей между таблицами.

При этом объект – местоположение конкретных данных – в каждой таблице определяется как

$$O(t, place), t \in V_{ph}, place \in Place, Place = \{db, pk\}, \tag{7}$$

где db – глобальная схема, pk – схема реструктуризации.

Тогда из определенных в выражениях (4-7) отношений сущностей множество всех вершин графа ЛБД на физическом уровне можно определить, как

$$V_{ph} = \left\{ \bigcup_{e \in V_l} MT(e) \cup S(e) \cup \left(\bigcup_{\substack{d \in V_l, \\ (e,d) \in E_l}} Mid(e,d) \right) \right\}. \quad (8)$$

Очевидно, что в каждый момент реструктуризации одной и более ИЛМ ЛБД узлов, входящих в состав МБД, структура графа ГС МБД, представленная выше, требует модификации путем изменения структуры подграфов ИЛМ ЛБД узлов на уровне ИЛМ или на физическом уровне, причем такой, чтобы время обработки распределенного SQL-запроса, в состав которого входят реструктурируемые ИЛМ ЛБД узлов, будет минимальным.

Благодаря переходу к графовому представлению модели данных ГС МБД решение этой задачи возможно численными методами, относящимися к классу задач динамического программирования (алгоритмы Флойда-Уоршалла, Беллмана-Форда) [9].

ГРАФОАНАЛИТИЧЕСКАЯ МОДЕЛЬ ПРОЦЕССА ФОРМИРОВАНИЯ ГЛОБАЛЬНОЙ СХЕМЫ МУЛЬТИБАЗЫ ДАННЫХ С УЧЕТОМ ЭТАПОВ ЕЕ РЕСТРУКТУРИЗАЦИИ

Введем понятие схемы реструктуризации (СР). СР – это подграф модели данных уровня ИЛМ, описывающей изменения ГС МБД в моменты реструктуризации моделей данных ИЛМ ее узлов.

Исходя их графового представления ГС МБД, СР можно задать графом $G_{sc}(V_{sc}, E_{sc})$, где $V_{sc} = (e_1, \dots, e_n) \subset V_l, n \in N, n \leq p$ – множество сущностей, входящих в состав

$$СР; E_{sc} = E_l \setminus \{r = (e_j, e_k) \in E_l \mid e_j \in V_l \setminus V_{sc} \vee e_k \in V_l \setminus V_{sc}\},$$

$$r_i = (e_j, e_k), i = 1..m; e_j, e_k \in V_{sc} - \text{множество связей между сущностями СР.}$$

Определим также понятие корневых (root) сущности $e_{start} \in V_{sc}$ и объекта $o_{start} \in O(e_{start})$.

Тогда для имеющейся в каждый момент реструктуризации СР в ЛБД узлов необходимо:

1. Создать копию данных – подграф модели данных узла ЛБД, подвергшегося реструктуризации.
2. Передать созданную копию данных на узел МБД, поддерживающий ГС.
3. Восстановить на узле МБД, поддерживающем ГС, модель данных СР, учитывая множество полученных копий данных.

Первую задачу предлагается решать с использованием алгоритма рекурсивного обхода графа СР в глубину [9] – алгоритм A_1 .

Исходными данными для его выполнения являются значения $e_{start} \in V_{sc}$ и $o_{start} \in O(e_{start})$, а результатом работы будет подграф копии данных (pk) $G_{pk}(V_{pk}, E_{pk})$, множество вершин в котором определяется как

$$V_{pk} = \left\{ \bigcup_{e \in V_{sc}} MT(e) \cup S(e) \cup \left(\bigcup_{\substack{d \in V_{sc}, \\ (e,d) \in E_{sc}}} Mid(e,d) \right) \right\}. \quad (9)$$

Исключим из рассмотрения особенности процесса передачи pk на узел МБД, поддерживающий ГС, считая их идеальными (задача 2). Тогда для решения задачи 3

возможно использование алгоритма попарного сравнения строк множеств V_{ph} и V_{pk} – алгоритм A_2 .

Очевидно, что в случае совпадения результатов сравнения (соответствует отсутствию изменений в данной ветви подграфа) достаточно выполнить обновление значений атрибутов, в противном случае (соответствует наличию изменений в данной ветви подграфа) требуется добавить объект $O(t, pk)$ в подмножество объектов $O(t, db)$.

Таким образом, в моменты реструктуризации моделей данных ЛБД некоторого подмножества узлов МБД над графом ГС, представленным подграфом СР, итерационно выполняются алгоритмы A_1 и A_2 . Число итераций при этом соответствует данному подмножеству узлов МБД (рис. 5).



Рисунок 5 – Наглядное представление графоаналитической модели процесса формирования глобальной схемы мультибазы данных с учетом этапов ее реструктуризации

ЗАКЛЮЧЕНИЕ

Процесс учета динамики реструктуризации моделей данных в локальных БД узлов мультибазы данных требует формализованного представления глобальной схемы, позволяющего производить над ней операции модификации, приводящие к формированию ее актуального (на момент выполнения реструктуризации) представления.

В статье обосновывается выбор графового представления глобальной схемы, а также представляется графоаналитическая модель процесса ее формирования с учетом реструктуризации моделей данных в ЛБД узлов, составляющих мультибазу данных.

СПИСОК ЛИТЕРАТУРЫ

1. Рябков Н.С. Аналитический обзор методов репликации и синхронизации баз данных // Качество. Инновации. Образование, 2006. – № 4. – С. 56-63.
2. Черняк Л. Интеграция данных: синтаксис и семантика // Открытые системы, 2009. – № 10.
3. Саймон А.Р. Стратегические технологии баз данных: менеджмент на 2000 год / пер. с англ.; под ред. М.Р. Когаловского. – Москва: Финансы и статистика, 1999. – 479 с.
4. Когаловский М.Р. Перспективные технологии информационных систем. – Москва: ДМК Пресс; Компания АйТи, 2003. – 288 с.
5. Шаша Д., Бонне Ф. Оптимизация баз данных: принципы, практика, решение проблем. – Москва: КУДИЦ-ОБРАЗ, 2004.
6. Мокрозуб В.Г. Графовые структуры и реляционные базы данных в автоматизированных интеллектуальных информационных системах. – Москва: ООО Издательский дом «Спектр», 2011. – 108 с.
7. Мокрозуб В.Г. и др. Применение N-ориентированных гиперграфов и реляционных баз данных для структурного и параметрического синтеза технических систем // Прикладная информатика, 2010. – № 4(28). – С. 115-122.
8. Лядова Л.Н. и др. Тиражирование данных в динамически настраиваемых распределенных информационных системах. В кн.: Natural and Artificial Intelligence. Sofia: Institute of Information Theories and Applications FOI ITHEA, 2010. – С. 165-172.

9. Кормен Т.Х. и др. Алгоритмы: построение и анализ, 2-е издание. – Москва: Издательский дом «Вильямс», 2005. – 1296 с.

Афанасьев Вадим Владимирович

Академия ФСО России, г.Орел

Преподаватель

Тел.: 8 920 805 48 61

E-mail: affa@mail.ru

Лебеденко Евгений Викторович

Академия ФСО России, г.Орел

Кандидат технических наук, преподаватель

Тел.: 8 915 501 76 20

E-mail: lebedenko_eugene@mail.ru

V.V. AFANAS'EV (*Teacher*)

E.V. LEBEDENKO (*Candidate of Engineering Science, Teacher*)
Academy of Federal Agency of Protection of the Russian Federation, Orel

**GRAPHIC-ANALYTICAL PROCESS MODEL OF GLOBAL SCHEME MULTIDATABASE
GENERATION CONSIDERING ITS STAGES OF RESTRUCTURING**

The article explains the choice of graph representation of a global schema and graphic-analytical generation process model taking into the restructuring of information and logical models in the local databases of a distributed special type database node – multidatabase.

Keywords: multidatabase; graph representation of a global schema; graphic-analytical model; restructuring.

BIBLIOGRAPHY (TRANSLITERATED)

1. Ryabkov N.S. Analiticheskoy obzor metodov replikacii i sinkronizacii baz danny'x // Kachestvo. Innovacii. Obrazovanie, 2006. – № 4. – S. 56-63.
2. Chernyak L. Integraciya danny'x: sintaksis i semantika // Otkry'ty'e sistemy', 2009. – № 10.
3. Sajmon A.R. Strategicheskie tehnologii baz danny'x: menedzhment na 2000 god / per. s angl.; pod red. M.R. Kogalovskogo. – Moskva: Finansy' i statistika, 1999. – 479 s.
4. Kogalovskij M.R. Perspektivny'e tehnologii informacionny'x sistem. – Moskva: DMK Press; Kompaniya AjTi, 2003. – 288 s.
5. Shasha D., Bonne F. Optimizaciya baz danny'x: principy', praktika, reshenie problem. – Moskva: KUDICz-OBRAZ, 2004.
6. Mokrozub V.G. Grafovy'e struktury' i relyacionny'e bazy' danny'x v avtmatizirovanny'x intellektual'ny'x informacionny'x sistemax. – Moskva: OOO Izdatel'skij dom «Spektr», 2011. – 108 s.
7. Mokrozub V.G. i dr. Primenenie N-orientirvanny'x gipergrafov i relyacionny'x baz danny'x dlya struktornogo i parametriceskogo sinteza texnicheskix sistem // Prikladnaya informatika, 2010. – № 4(28). – S. 115-122.
8. Lyadova L.N. i dr. Tirazhirovanie danny'x v dinamicheski nastroevaemy'x raspredelyonny'x infrmacionny'x sistemax. V kn.: Natural and Artificial Intelligence. Sofia: Institute of Information Theories and Applications FOI ITHEA, 2010. – S. 165-172.
9. Kormen T.X. i dr. Algoritmy': postroenie i analiz, 2-e izdanie. – Moskva: Izdatel'skij dom «Vil'yams», 2005. – 1296 s.

УДК 621.313.13

В.И. ДЕМКИН, СО ЛИН МАУНГ

СТАРТ-СТОПНОЕ УПРАВЛЕНИЕ ШАГОВЫМ ДВИГАТЕЛЕМ НА ОДНОМ ШАГЕ БЕЗ ПЕРЕКЛЮЧЕНИЯ ФАЗ

В данной статье описывается способ устранения колебаний ротора шагового двигателя (ШД) при отработке единичного шага без переключения фаз двигателя путем регулирования уровня возбуждения фаз в зависимости от сигнала датчика положения ротора ШД.

Ключевые слова: шаговый двигатель; старт-стопное управление; устранение колебаний ротора; уровень возбуждения фаз.

Шаговый двигатель находит широкое применение в системах управления технологическими процессами, а также в станках с ЧПУ и роботах [1].

Способы устранения колебаний ротора ШД в режиме единичных шагов при сравнительно низкой частоте тактовых импульсов называются «старт-стопным» управлением ШД. В конце отработки каждого шага колебания ротора ШД происходят из-за того, что в систему вводится при каждом переключении избыток электромагнитной энергии, который рассеивается при колебаниях ротора вокруг точки покоя. Колебания ротора неблагоприятно сказываются на точностных характеристиках приводного механизма и его долговечности. Основным недостатком ШД является колебательный характер отработки шагов.

Рассмотрим режим старт-стопного управления ШД без переключения фаз при движении ротора на одном шаге. В процессе отработки шага ротор ШД, проходя точку устойчивого равновесия за счет избытка кинетической энергии, попадает в зону действия тормозных моментов. Если уменьшить величину моментов торможения на этом участке, то ротор, двигаясь по инерции, дойдёт до точки устойчивого равновесия, соответствующей включённой следующей фазе или комбинации фаз. Включением этой фазы или комбинации фаз в данной точке можно зафиксировать ротор без колебаний на следующем шаге.

Все способы старт-стопного управления при отработке единичных шагов основаны на регулировании энергии, вводимой в приводную систему на участке разгона ротора и на участке торможения в процессе движения на шаге. Переход на участок торможения происходит или естественным путем после прохождения ротором точки устойчивого равновесия, или принудительно переключением фазных обмоток в сторону, противоположную направлению движения.

Способ управления ШД, который для четырехфазного ШД обеспечивает перемещение, равное единичному шагу, без колебаний в конце, иллюстрируется моментными кривыми, показанными на рисунке 1.

На рисунке 1 цифрами 1, 2, 3, 4 обозначены соответствующие фазы двигателя,

$M_H \text{sign}(\dot{\theta})$ – момент нагрузки типа «сухого трения» (не зависит от скорости), θ – угол поворота ротора, $\dot{\theta} = \frac{d\theta}{dt}$ – скорость вращения ротора, β – относительное уменьшение

амплитуды синхронизирующего тормозного момента, $\text{sign}(\dot{\theta})$ – знак скорости.

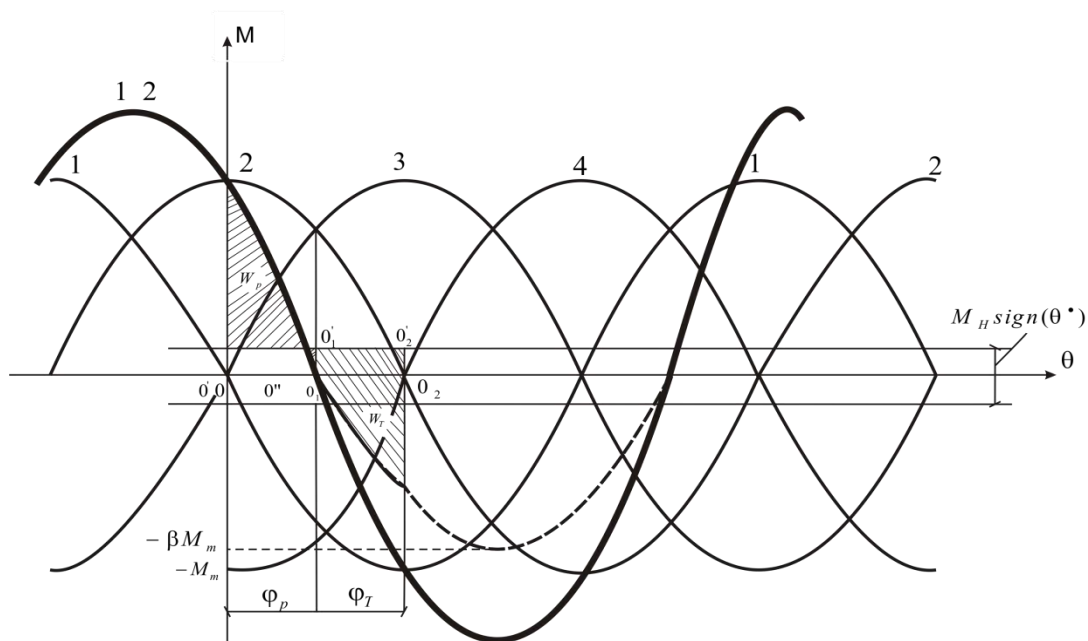


Рисунок 1 – Старт-стопный способ обработки единичного шага при естественном торможении без переключения фаз

Из исходного положения фиксации при включенной первой фазе (точка 0) ротор ШД включением второй фазы дополнительно к первой (суммарная кривая 12) разгоняется на участке $0 \div 0_1$. Торможение начинается уже на участке $0_1 \div 0_2$ при прохождении точки 0_1 , ротор окончательно тормозится на участке $0_2 \div 0$. По сигналу датчика шагов в точке 0_2 отключается 1 фаза, и ротор фиксируется в новой точке покоя. Баланс энергии разгона W_p и энергии торможения W_T обеспечивается уменьшением уровня возбуждения каждой из фаз, участвующих в работе на данном участке, в β раз. Уменьшение уровня возбуждения фаз осуществляется в точке 0_1 по сигналу датчика шагов.

Задачей анализа старт-стопного управления на одном шаге является нахождение коэффициента уменьшения амплитуды тормозного момента β и времени движения τ при обработке единичного шага α . Энергия разгона и торможения определяются выражениями [2]:

$$W_p = \int_0^{\frac{\pi}{n}} M_m [-\sin \theta + \sin(\frac{2\pi}{n} - \theta)] d\theta - \int_0^{\frac{\pi}{n}} M_H d\theta, \tag{1}$$

$$W_T = \int_{\frac{\pi}{n}}^{\frac{2\pi}{n}} \beta M_m [-\sin \theta + \sin(\frac{2\pi}{n} - \theta)] d\theta - \int_{\frac{\pi}{n}}^{\frac{2\pi}{n}} M_H d\theta. \tag{2}$$

где M_m – амплитуда кривой (статического) синхронизирующего момента; n – число тактов коммутации ШД в цикле; M_H – приведенный к ротору момент нагрузки; θ – угол поворота ротора.

Из условий баланса энергии W_p и W_T найдем выражение параметра настройки:

$$\beta = \frac{\frac{2\pi}{n} \mu_{H+1} - 2 \cos \frac{\pi}{n} + \cos \frac{2\pi}{n}}{1 - 2 \cos \frac{\pi}{n} + \cos \frac{2\pi}{n}}. \quad (3)$$

где $\mu_H = \frac{M_H}{M_m}$ – относительный момент нагрузки.

Относительные значения среднего момента разгона и торможения определяются, соответственно, формулами:

$$\mu_{ср.р} = \frac{n}{\pi} (2 \cos \frac{\pi}{n} - 1 + \cos \frac{2\pi}{n}) - \mu_H, \quad (4)$$

$$\mu_{ср.т} = \frac{n}{\pi} \beta (\cos \frac{2\pi}{n} + 1 - 2 \cos \frac{\pi}{n}) - \mu_H. \quad (5)$$

Для четырёхфазного ШД при $n = 4$ формулы 3 ÷ 5 преобразуются к виду:

$$\beta = 1 + \frac{\frac{\pi}{2} \mu_H}{1 - \sqrt{2}}, \quad (6)$$

$$\mu_{ср.р} = \frac{4(\sqrt{2}-1)}{\pi} - \mu_H, \quad (7)$$

$$\mu_{ср.т} = -\frac{4(\sqrt{2}-1)}{\pi} + \mu_H. \quad (8)$$

Зависимость коэффициента уменьшения амплитуды тормозного момента β от относительного момента нагрузки μ_H при отработке единичного шага приведена на рисунке 2.

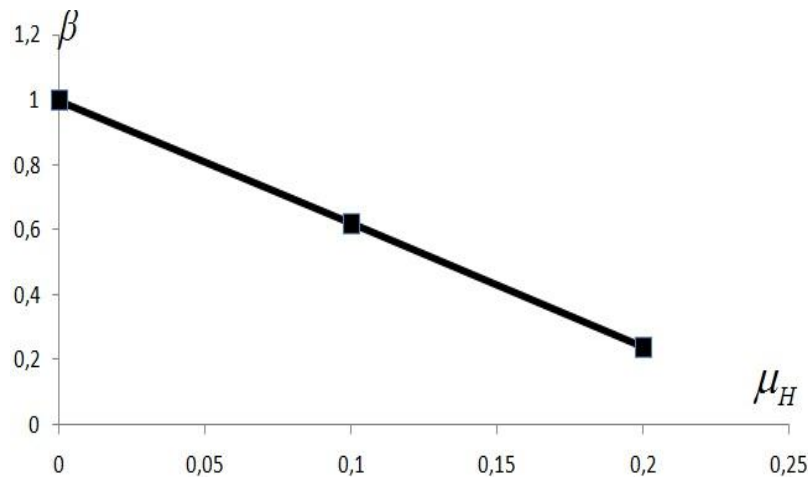


Рисунок 2 – Зависимость коэффициента уменьшения амплитуды тормозного момента β от относительного момента нагрузки μ_H при отработке единичного шага

Очевидно, что время разгона τ_p и торможения τ_t в данном случае будут одинаковыми. С учетом того, что при $n = 4$ углы разгона и торможения $\varphi_p = \varphi_t = \frac{\alpha}{2} = \frac{\pi}{4}$, из

формулы $\tau_\varphi = \sqrt{\frac{\varphi}{\mu_{cp}}}$ определим $\tau_p = \tau_T$ и $\tau_\alpha = 2\tau_p$. Эти зависимости графически приведены на рисунке 3.

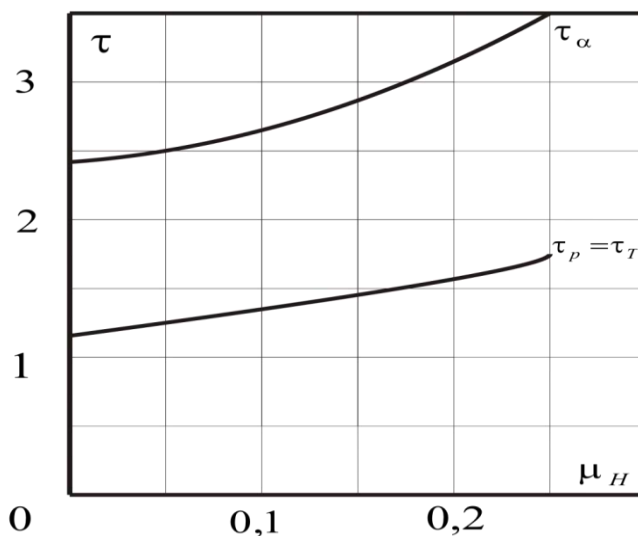


Рисунок 3 – Зависимость времени разгона τ_p , торможения τ_T и полного времени движения τ_α от относительного момента нагрузки μ_H

В данной статье приведен способ устранения колебаний ротора ШД при отработке единичного шага без переключения фаз двигателя за счет снижения уровня возбуждения фазы двигателя при его торможении.

СПИСОК ЛИТЕРАТУРЫ

1. Sagarika Pal, Niladri S. Tripathy. Remote Position Control System of Stepper Motor Using DTMF Technology, International Journal of Control and Automation, 2011. – Vol. 4. – № 2. – P. 35-41.
2. Ред М., Чиликин Г. Дискретный электропривод с шаговыми двигателями. – М.: Энергия, 1971.

Демкин Василий Иванович

Национальный исследовательский университет «МИЭТ», г. Москва, Зеленоград
 Доцент кафедры систем автоматического управления и контроля
 Тел.: 8 916 671 00 02
 E-mail: d_v_i@mail.ru

Со Лин Маунг

Национальный исследовательский университет «МИЭТ», г. Москва, Зеленоград
 Аспирант кафедры систем автоматического управления и контроля
 Тел.: 8 916 960 67 56
 E-mail: sawlinmg.miet@gmail.com

V.I. DYOMKIN (Associate Professor of the Department of Automatic Control and Monitoring Systems)

SO LIN MAUNG (Post-graduate Student of the Department of Automatic Control and Monitoring Systems)
 National Research University of Electronic Technology, Zelenograd

**START-STOP MANAGEMENT OF STEP MOTOR
ON ONE STEP WITHOUT PHASE SWITCHING**

The article describes the method of elimination of step motor rotor oscillations with unit step completion without phase switching by regulating the phase excitation's level depending on the signal of the rotor's step motor sensor.

Keywords: *step motors; start-stop control; vibration's eliminate rotor; phase energizing level.*

BIBLIOGRAPHY (TRANSLITERATED)

1. Sagarika Pal, Niladri S. Tripathy. Remote Position Control System of Stepper Motor Using DTMF Technology, International Journal of Control and Automation, 2011. – Vol. 4. – № 2. – P. 35-41.
2. Red M., Chilikin G. Diskretny'j e'lektroprivod s shagovy'mi dvigatelyami. – M.: E'nergiya, 1971.

УДК 004.934.1'1

К.В. ПЛЕТНЕВ, Д.Е. ПРОЗОРОВ

АНАЛИЗ МЕТОДА МАРКОВСКОЙ ПАРАМЕТРИЗАЦИИ РЕЧЕВЫХ СИГНАЛОВ

В статье рассмотрена марковская модель речевого сигнала. Предложен метод параметризации речевого сигнала, основанный на использовании модели простой цепи Маркова. Произведена оценка эффективности предлагаемого метода в системах распознавания изолированных команд по сравнению с методом параметризации мел-кепстральными коэффициентами.

Ключевые слова: марковские цепи; параметризация речи; распознавание речи.

ВВЕДЕНИЕ

Речевые технологии широко используются для построения справочных и поисковых систем, систем интерактивного обучения иностранным языкам, голосового управления сложными техническими системами и т.д. Одной из наиболее важных задач в области речевых технологий является задача автоматического распознавания речи (АРР).

Существенное влияние на эффективность АРР оказывают методы, используемые на этапе параметризации речевого сигнала [1]. В общем виде задача параметризации речевого сигнала состоит в выделении наиболее информативных признаков сигнала, позволяющих различать вокализованные участки речи. Рассматриваемая задача давно известна и имеет ряд решений, среди которых наиболее распространены методы, использующие кепстральные преобразования в частотной или временной области [1-4, 12, 13].

Методы параметризации, основанные на кепстральных преобразованиях, пригодны для обработки сигналов, записанных в шумных помещениях [5], однако сравнительно высокая вычислительная сложность этих методов и большое количество рассчитываемых параметров негативно влияют на производительность систем АРР [2]. Поэтому в условиях широкого распространения встраиваемых систем с малой энергоемкостью и мобильных приложений, реализующих речевые технологии, возникает необходимость в разработке и анализе методов параметризации речевых сигналов с относительно невысокими требованиями к объемам памяти и вычислительным ресурсам.

В данной работе исследуется метод параметризации, основанный на марковской модели речевого сигнала, позволяющий снизить требования системы автоматического распознавания команд (АРК) к вычислительным ресурсам при сохранении качества распознавания.

Первые работы, посвященные марковским моделям речевых сигналов, появились в 1970-х годах. Так, в монографии [6] проведен подробный анализ марковских моделей дельта-модулированных (ДМ) речевых сигналов. Однако дальнейшие исследования выявили ряд недостатков линейной дельта-модуляции, в связи с чем были разработаны более эффективные способы кодирования и форматы представления речевых сигналов [7]. Кроме того, в работе [8] было показано, что существенная зависимость параметров марковской модели речевых ДМ-сигналов от скрытых параметров ДМ-кодека значительно сужает область применимости данного метода.

Тем не менее, малое количество публикаций, посвященных исследованию возможности марковской параметризации речевых сигналов, и такие достоинства марковских моделей, как относительно небольшое количество параметров модели и линейная вычислительная сложность алгоритмов расчета параметров, вызывают интерес к указанному методу параметризации.

МОДЕЛЬ РЕЧЕВОГО СИГНАЛА НА ОСНОВЕ ПРОСТОЙ ЦЕПИ МАРКОВА

Наиболее общей формой представления речевого сигнала можно считать двоичную N -разрядную последовательность $s[k]$ ($k = 1, \dots, n$), полученную в результате дискретизации с частотой f_0 и равномерного квантования исходного сигнала. Двоичная последовательность $s[k]$ на выходе квантователя является комбинацией бинарных битовых последовательностей $s_b[k]$

$$s[k] = \sum_{b=0}^{N-1} (2^b \cdot s_b[k]), \quad (1)$$

где $s_b[k] = \overline{0,1}$ – бинарная последовательность, представляющая b -й бит (разряд) сигнала $s[k]$.

Как показывает проведенный авторами эксперимент, корреляционные связи в последовательности $s[k]$ на ограниченном интервале времени ослабевают практически монотонно (рис. 1, 2). Поэтому можно указать интервал (несколько тактов), за пределами которого корреляционные связи практически не распространяются. С учетом этого допущения для описания и анализа коротких фрагментов речевых сигналов используют математический аппарат простых или сложных цепей Маркова [9, 10].

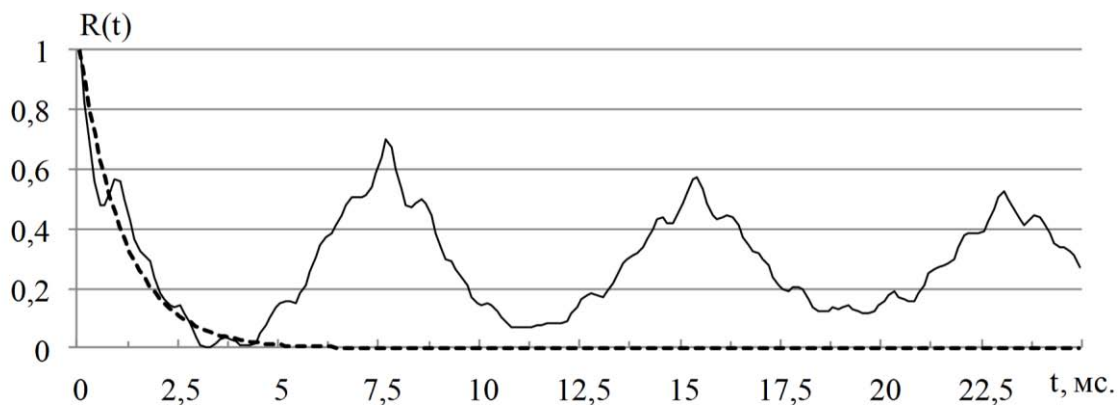


Рисунок 1 – Автокорреляционная функция ИКМ сигнала звука «а»

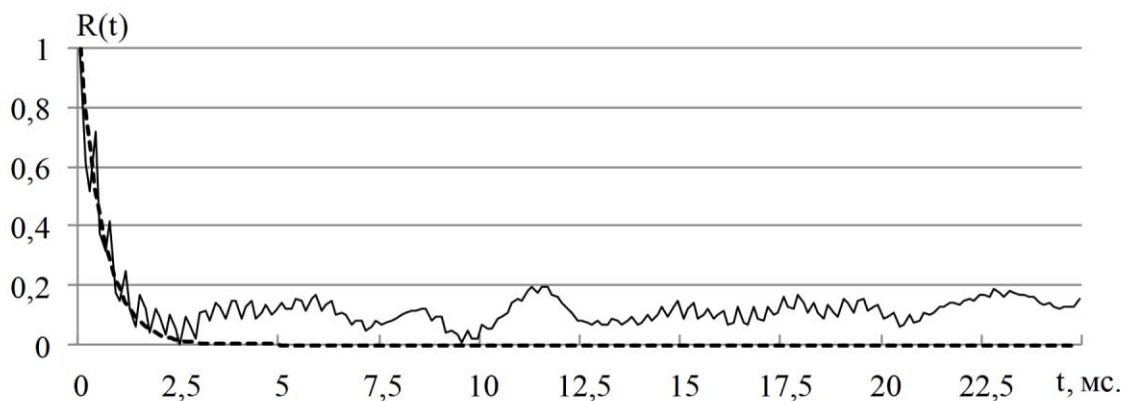


Рисунок 2 – Автокорреляционная функция ИКМ сигнала звука «ч»

Пунктиром на графиках (рис. 1, 2) обозначена экспоненциальная аппроксимация корреляционных функций $R(t)$ фрагментов $s[t]$ речевых сигналов.

Здесь и далее предполагается, что частота дискретизации речевого сигнала $f_{\partial} = 8$ кГц, количество уровней квантования $2^N = 256$ (8 бит).

Бинарные битовые последовательности $s_b[k]$, представленные моделью простой однородной цепи Маркова, описываются матрицей вероятностей перехода от i -го значения последовательности в k -м такте $s_{b,i}[k]$ к j -му значению последовательности в $(k+1)$ -м такте $s_{b,j}[k+1]$:

$$p(s_{b,j}[k+1]|s_{b,i}[k]) = \Pi_b = \begin{vmatrix} \pi_{b,11} & \pi_{b,12} \\ \pi_{b,21} & \pi_{b,22} \end{vmatrix}, \quad i, j = \overline{1,2} \quad (2)$$

и вектором вероятностей начальных состояний

$$P_b = \begin{bmatrix} p_{b,1} \\ p_{b,2} \end{bmatrix}.$$

Элементы $\pi_{b,ij}$ матрицы вероятностей перехода (2) удовлетворяют условиям нормировки

$$\sum_{j=1}^2 \pi_{b,ij} = 1, \quad i = \overline{1,2}. \quad (3)$$

МЕТОД МАРКОВСКОЙ ПАРАМЕТРИЗАЦИИ

Метод марковской параметризации содержит следующую последовательность шагов.

1. Из фрагмента речевого сигнала $s[k]$ выделяются бинарные последовательности $s_b[k]$.

2. Последовательности $s_b[k]$ разделяются на m пересекающихся сегментов с коэффициентом перекрытия h . В работе [11] показано, что наилучшие результаты распознавания достигаются при значении коэффициента перекрытия $h \approx 20\%$ от размера сегмента.

3. Для каждого l -го сегмента последовательностей $s_b^{(l)}[k]$ вычисляются значения матрицы вероятностей переходов $\pi_{b,ij}^{(l)}$ ($i, j = \overline{1,2}$). С учетом условия нормировки (3) для представления сегмента достаточно двух элементов матрицы, $\pi_{b,ij}^{(l)}$ ($i = \overline{1,2}$). В результате каждая битовая последовательность $s_b[k]$ характеризуется матрицей параметров вида

$$M_b = \begin{bmatrix} \pi_{b,11}^{(1)} & \pi_{b,11}^{(2)} & \pi_{b,11}^{(3)} & \dots & \pi_{b,11}^{(m)} \\ \pi_{b,22}^{(1)} & \pi_{b,22}^{(2)} & \pi_{b,22}^{(3)} & \dots & \pi_{b,22}^{(m)} \end{bmatrix}. \quad (4)$$

где m – количество сегментов.

Для исследования эффективности методов параметризации речевых сигналов авторами разработан программный комплекс на языке высокого уровня Scilab. Проведен эксперимент на коллекции речевых сигналов, содержащей варианты произношения цифр «0»...«9» одним диктором по пятнадцать вариаций произношения каждой. В качестве модели речевых сигналов использована простая однородная цепь Маркова. Параметры модели вычислялись для трех старших бинарных последовательностей $s_b[k]$, что достаточно для большинства практических применений [11]. Для распознавания параметризованных фрагментов речевых сигналов применен метод динамического искажения времени (Dynamic Time Warping, DTW) [4].

Таблица 1 – Вероятность правильного распознавания

Метод параметризации	Вероятность правильного распознавания, %
MFCC	96,5
SMC	98,7

В таблице 1 обобщены результаты оценки вероятности правильного распознавания изолированных фрагментов речевых сигналов при использовании метода марковской параметризации (SMC) и метода параметризации мел-кепстральными коэффициентами (MFCC).

Анализ результатов показывает, что марковская параметризация в условиях дикторозависимого изолированного распознавания фрагментов речевых сигналов позволяет достичь сопоставимого качества распознавания речи в сравнении с методом параметризации мел-кепстральными коэффициентами.

Предлагаемый метод параметризации обладает линейной вычислительной сложностью при сравнительно небольшом объеме получаемых параметров, характеризующих речевой сигнал. Минимальный размер вектора параметров для метода MFCC составляет $12m$ [2] (в эксперименте – $24m$), где m – количество сегментов, на которые был поделен сигнал. Поскольку для марковской параметризации использовались три старших битовых последовательности речевого фрагмента, в методе SMC матрица параметров содержит $6m$ значений. Оценка вычислительной сложности других распространенных методов параметризации приведена в таблице 2.

Таблица 2 – Вычислительная сложность методов параметризации

Метод параметризации	Вычислительная сложность	Относительная вычислительная сложность	Примечание
MFCC	$O(2N \cdot \log_2 N)$	~ 6,7	При использовании БПФ
PLP [12]	$O\left(MN + 3N + M^2 + \frac{N}{2} \log_2 N\right)$	~ 6,7	M – количество коэффициентов линейного предсказания
LPCC [2]	$O(3MN + 3N + M^2)$	~ 13,1	n – количество LPCC коэффициентов
STFT [2]	$O\left(\frac{N}{2} \log_2 N\right)$	~ 1,7	При использовании БПФ
Вейвлет анализ [13]	$O(n) > O\left(\frac{N}{2} \log_2 N\right)$	> 1,7	Нижняя оценка
SMC	$O(N \cdot C)$	1	C – количество используемых бит

При вычислении относительной вычислительной сложности методов параметризации использованы следующие коэффициенты: $N = 1024$, $M = 12$, $n = 20$, $C = 3$.

ЗАКЛЮЧЕНИЕ

Низкие требования к вычислительным ресурсам позволяют рекомендовать метод марковской параметризации речевых сигналов для применения в мобильных и встраиваемых системах автоматического дикторозависимого распознавания команд.

СПИСОК ЛИТЕРАТУРЫ

1. Oppenheim A.V., Schaffer R.W. Digital Signal Processing // Prentice-Hall. – Englewood Cliffs, New Jersey, USA, 1975.
2. Huang X. Spoken Language Processing: A Guide to Theory, Algorithm and System Development / X. Huang, A. Acero, H. Hon // Prentice Hall, 2001.
3. Mermelstein P. Distance Measures for Speech Recognition, Psychological and Instrumental // Chen C.H. (ed.) Pattern Recognition and Artificial Intelligence, Academic, New York, 1976. – P. 374-388.
4. Рабинер Л.Р. Шафер Р.В. Цифровая обработка речевых сигналов: пер. с англ.; под ред. Прохорова Ю.Н., Назарова М.В. – М.: Радио и связь, 1981. – 496 с.
5. Picone J.W. Signal Modeling Techniques in Speech Recognition / Proceedings of the IEEE, September 1993. – P. 1215-1247.
6. Венедиктов М.Д. Дельта-модуляция. Теория и применение / М.Д. Венедиктов, Ю.П. Женеvский, В.В. Марков. – М.: Связь, 1976. – С. 104-114.
7. Рихтер С.Г. Кодирование и передача речи в цифровых системах подвижной радиосвязи: учеб. пособие. – М.: Горячая линия-Телеком, 2010. – 302 с.
8. Плетнёв К.В., Прозоров Д.Е. Параметризация речевых сигналов цепями Маркова // Advanced Science, 2012. – №1. – С. 19-28.
9. Кемен Дж. Снелл Дж. Конечные цепи Маркова: пер. с англ.; под ред. А.А. Юшкевича. – М.: «Наука», 1970. – 272 с.
10. Тихонов В.И., Кульман Н.К. Нелинейная фильтрация и квазикогерентный прием сигналов. – М.: «Советское радио», 1975. – 704 с.
11. Плетнёв К.В. Метод параметризации речевых сигналов простыми цепями Маркова // Сб. трудов МНТК «Актуальные направления фундаментальных и прикладных исследований», 2013. – Том 2. – С. 86-88.
12. Hermansky N. Perceptual Linear Predictive (PLP) Analysis of Speech // The Journal of the Acoustical Society of America, 1990. – № 87(4). – P. 1738-1752.
13. Grossmann A., Morlet J. Decomposition of Hardy Functions into Square Integrable Wavelets of Constant Shape // SIAM J. Math., Anal, 1984. – № 15(4). – P. 723-736.

Плетнёв Константин Владимирович

ФГБОУ ВПО «Вятский государственный гуманитарный университет», г. Киров
Ассистент кафедры прикладной математики и информатики
Тел.: 8 909 140 24 71
E-mail: pletnev.kirov@gmail.com

Прозоров Дмитрий Евгеньевич

ФГБОУ ВПО «Вятский государственный университет», г. Киров
Профессор кафедры радиоэлектронных средств
Тел.: 8 (8332) 35-72-59, 8 912 711 21 11
E-mail: prozorov.de@gmail.com

K.V. PLETNYOV (*Assistant of the Department of Applied Mathematics and Computer Science*)
Vyatka State University of Humanities, Kirov

D.E. PROZOROV (*Professor of the Department of Radioelectronic Devices*)
Vyatka State University, Kirov

**ANALYSIS OF THE METHOD OF SPEECH SIGNALS
PARAMETERIZATION BY SIMPLE MARKOV'S CHAINS**

The paper describes a Markov model of the speech signal. Method of speech signal parameterization based on the using of a simple Markov chain model is proposed. The effectiveness of the proposed method has been evaluated in the command recognition systems compared to the method of parameterization of mel-frequency cepstral coefficients.

Keywords: markov's chain; parameterization of speech; speech recognition.

BIBLIOGRAPHY (TRANSLITERATED)

1. Oppenheim A.V., Schafer R.W. Digital Signal Processing // Prentice-Hall. – Englewood Cliffs, New Jersey, USA, 1975.
2. Huang X. Spoken Language Processing: A Guide to Theory, Algorithm and System Development / X. Huang, A. Acero, H. Hon // Prentice Hall, 2001.
3. Mermelstein P. Distance Measures for Speech Recognition, Psychological and Instrumental // Chen C.H. (ed.) Pattern Recognition and Artificial Intelligence, Academic, New York, 1976. – P. 374-388.
4. Rabiner L.R. Shafer R.V. Cifrovaya obrabotka rechevy'x signalov: per. s angl.; pod red. Proxorova Yu.N., Nazarva M.V. – M.: Radio i svyaz', 1981. – 496 s.
5. Picone J.W. Signal Modeling Techniques in Speech Recognition / Proceedings of the IEEE, September 1993. – P. 1215-1247.
6. Venediktov M.D. Del'ta-modulyaciya. Teoriya i primeneniye / M.D. Venediktov, Yu.P. Zhenevskij, V.V. Markov. – M.: Svyaz', 1976. – S. 104-114.
7. Rixter S.G. Kodirovaniye i peredacha rechi v cifrovyy'x sistemax podvizhnyj radisvyazi: ucheb. posobie. – M.: Goryachaya liniya-Telekom, 2010. – 302 s.
8. Pletnyov K.V., Prozorov D.E. Parametrizaciya rechevy'x signalov cepyami Markova // Advanced Science, 2012. – №1. – S. 19-28.
9. Kemen Dzh. Snell Dzh. Kontchny'e cepi Markova: per. s angl.; pod red. A.A. Yushkevicha. – M.: «Nauka», 1970. – 272 s.
10. Tixonov V.I., Kul'man N.K. Nelinejnaya fil'traciya i kvazikogerentny'j priyom signalov. – M.: «Sovetskoe radio», 1975. – 704 s.
11. Pletnyov K.V. Metod parametrizacii rechevy'x signalov prosty'mi cepyami Markova // Sb. trudov MNTK «Aktual'ny'e napravleniya fundamental'ny'x i prikladny'x issledovanij», 2013. – Tom 2. – S. 86-88.
12. Hermansky H. Perceptual Linear Predictive (PLP) Analysis of Speech // The Journal of the Acoustical Society of America, 1990. – № 87(4). – P. 1738-1752.
13. Grossmann A., Morlet J. Decomposition of Hardy Functions into Square Integrable Wavelets of Constant Shape // SIAM J. Math., Anal, 1984. – № 15(4). – P. 723-736.

УДК 303.732.4

Т.Д. КАРМИНСКАЯ, Д.О. ТЕЙ, В.М. ТАТЬЯНКИН, М.А. РУСАНОВ

ФОРМИРОВАНИЕ КОНТРОЛЬНЫХ ЦИФР ПРИЕМА С УЧЕТОМ ПРОГНОЗА ПОТРЕБНОСТЕЙ РЕГИОНАЛЬНОЙ ЭКОНОМИКИ

В работе рассматривается вопрос распределения регионального заказа на подготовку кадров среди учреждений профессионального образования по укреплённым группам специальностей в условиях ограниченных ресурсов. Определена структура дополнительных данных, предложен алгоритм распределения.

Ключевые слова: региональный заказ; рейтингование; оптимальное распределение; прогнозирование.

ВВЕДЕНИЕ

Развитие и внедрение экспертно-аналитических систем прогнозирования потребностей в профессиональном кадровом обеспечении является стратегической задачей в концепции долгосрочного социально-экономического развития Российской Федерации на период до 2020 года и в стратегии инновационного развития Российской Федерации на период до 2020 года [1].

В настоящее время сформированы различные подходы к решению задачи определения кадровых потребностей региональной экономики [2-6]. Современные методики определения кадровых потребностей региональной экономики позволяют определить будущие кадровые потребности региона. Однако актуальной является задача оптимального распределения заказа на подготовку кадров по сети региональных учреждений профессионального образования (УПО) в соответствии с требованиями региональной экономики. При этом в методиках [2-6] вопрос оптимального распределения регионального заказа остается слабо проработанным.

Решение задачи оптимального распределения регионального заказа на подготовку кадров по сети учреждений профессионального образования требует формирования значительного объема дополнительной информации, описывающей «производственную мощность», движение контингента студентов, бюджет региона, рейтинговые показатели УПО, данные рынка труда, количество школьников и т.д.

В статье рассматривается вариант решения задачи оптимального распределения регионального заказа по сети учреждений профессионального образования.

СИСТЕМА ФОРМИРОВАНИЯ КОНТРОЛЬНЫХ ЦИФР ПРИЕМА

В Югорском государственном университете разработана система формирования регионального заказа на подготовку кадров (рис. 1).

Элементами системы являются: база данных, подсистема прогнозирования социально-экономических параметров региона, подсистема формирования рейтинга учреждений профессионального образования и укрупнённых групп специальностей (УГС), подсистема формирования контрольных цифр приёма региональной системы профессионального образования. Входными данными подсистемы прогнозирования кадровой потребности региональной экономики являются: ВРП, инфляция, дефлятор, объём производства, численность занятых по видам экономической деятельности, эмиграция, иммиграция и т.д.

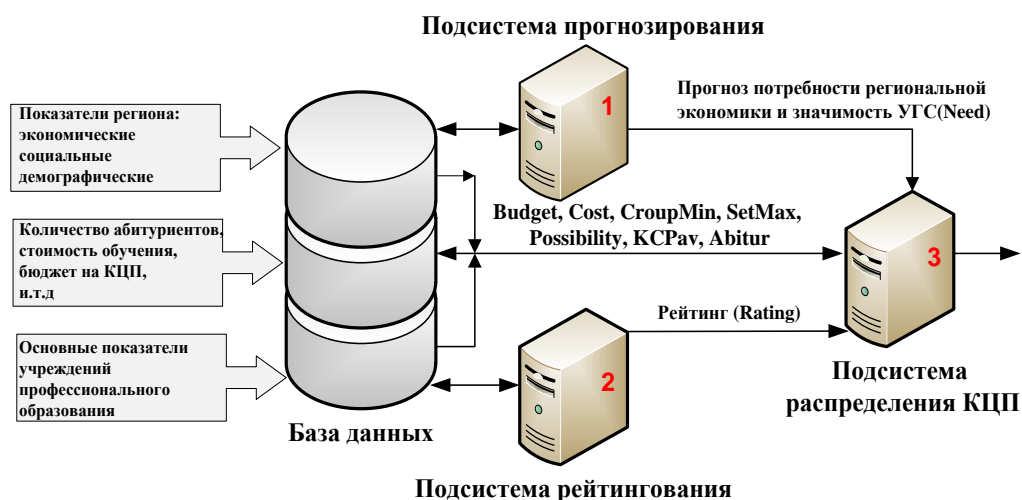


Рисунок 1 – Система формирования регионального заказа на подготовку кадров

Выходными данными подсистемы прогнозирования кадровой потребности региональной экономики являются: количество кадров по УГС с определённым уровнем образования, которое необходимо подготовить УПО, и коэффициент «значимости» УГС для региональной экономики. Подсистема рейтингования формирует рейтинг подготовки кадров по УГС в рамках региональной сети профессионального образования. В качестве входных параметров подсистемы рейтингования используются основные показатели учреждений профессионального образования: площадь, техническое оснащение, доля безработных выпускников и т.д. В работе будет рассмотрена подсистема определения контрольных цифр приёма по УПО (далее подсистема КЦП).

Основным назначением подсистемы КЦП является определение регионального заказа на подготовку кадров для учреждений профессионального образования Ханты-Мансийского автономного округа – Югры (ХМАО-Югры). Критерий и ограничения, используемые при определении регионального заказа на подготовку кадров, сформулированы департаментом образования и молодежной политики в рамках проекта разработки системы определения контрольных цифр приема для ХМАО-Югры.

Для определения регионального заказа на подготовку кадров необходимо выполнить следующие этапы:

1. Определение общего заказа на подготовку кадров по трем уровням образования. Для выполнения этого этапа необходимо:

- проанализировать ограничения, определяющие количество абитуриентов, которых можно обучить по 3 уровням профессионального образования;
- определить объем регионального заказа целевого заказа на подготовку кадров по 3 уровням профессионального образования.

2. Распределение общего регионального заказа по укрупненным группам специальностей, полученного на первом этапе, требует выполнения следующих шагов:

- определение коэффициента значимости укрупненных групп специальностей для задач региональной экономики, исходя из данных прогноза кадровых потребностей и данных регионального рынка труда;
- анализ ограничений подготовки кадров в рамках УГС и учреждений профессионального образования;

– определение объема регионального заказа целевого заказа на подготовку кадров по УГС в рамках 3 уровней образования.

3. Распределение регионального заказа по УГС, полученного на втором этапе по учреждениям профессионального образования региона. Для выполнения этого этапа необходимо выполнить:

– определение рейтинга учреждений профессионального образования и рейтинга специальностей в рамках УГС согласно методике, утвержденной Министерством образования и науки РФ;

– определение регионального заказа на подготовку кадров по УГС и уровням образования для учреждений профессионального образования на основе их рейтинга, показателей движения контингента студента (отчисление, болезнь, декретный отпуск и т.д.);

– анализ соответствия регионального заказа на подготовку кадров «производственной мощности» учреждений профессионального образования и корректировка регионального заказа на подготовку кадров по УГС и уровням образования для учреждений профессионального образования.

Параметр «производственная мощность» определяет максимальное количество абитуриентов, которое учреждение профессионального образования может принять на 1 курс по определенной УГС.

ГРАНИЧНЫЕ УСЛОВИЯ СПОСОБА ФОРМИРОВАНИЯ КОНТРОЛЬНЫХ ЦИФР ПРИЕМА

Целевая функция способа формирования контрольных цифр приема по УГС и УПО описывается следующим выражением:

$$\sum_{l=1}^k \sum_{i(l)=1}^{n(l)} \sum_{j(l)=1}^{m(l)} K_ves_{i(l)j(l)}^l * X_{i(l)j(l)}^l \rightarrow \max, \quad (1)$$

где $K_ves_{i(l)j(l)}^l$ – весовой коэффициент значимости i УГС; $l = 1...3$ – уровень образования (ВПО, СПО, НПО); $X_{i(l)j(l)}^l$ – заказ на подготовку кадров по $i(l)$ УГС $j(l)$ УПО; $m(l)$ – количество учреждений профессионального образования; $n(l)$ – количество УГС для l уровня образования.

Весовой коэффициент значимости УГС определяется через весовой коэффициент, полученный при прогнозировании кадровой потребности региональной экономикой, и весовой коэффициент, основанный на текущей ситуации на рынке труда:

$$K_ves_{i(l)}^l = \frac{K_p_{i(l)}^l + K_T_{i(l)}^l}{\sum_{i(l)=1}^{n(l)} K_p_{i(l)}^l + \sum_{i(l)=1}^{n(l)} K_T_{i(l)}^l}, \quad (2)$$

где $K_p_{i(l)}^l$ – отношение требуемого количества кадров в прогнозном году к имеющемуся в прогнозном году. Данные о весовом коэффициенте K_p определяются подсистемой прогнозирования; $K_T_{i(l)}^l$ – отношение текущих вакансий на рынке труда к количеству безработных.

Поиск решения (1) требует определения минимальной численности учебной группы и «производственной мощности» учреждения профессионального образования:

$$MinGroup_{i(l)j(l)}^l \leq X_{i(l)j(l)}^l \leq MaxNabor_{i(l)j(l)}^l \quad \text{или} \quad X_{i(l)j(l)}^l = 0, \quad (3)$$

где $MinGroup_{i(l)j(l)}^l$ – минимальное количество человек в группе по $i(l)$ УГС; $l = 1...3$ – уровень образования (ВПО, СПО, НПО); $MaxNabor_{i(l)j(l)}^l$ – максимальное количество человек, которое может принять УПО по $i(l)$ УГС. Финансовые ограничения целевой функции (1) описываются следующим выражением:

$$\sum_{l=1}^k \sum_{i(l)=1}^{n(l)} \sum_{j(l)=1}^{m(l)} Cost_{i(l)j(l)}^l * X_{i(l)j(l)}^l \leq Budget, \quad (4)$$

где $Cost_{i(l)j(l)}^l$ – стоимость обучения на $i(l)$ УГС; $Budget$ – региональный бюджет, направленный на подготовку кадров, который принимается 1 раз в год.

Региональный заказ на подготовку кадров (1) не должен превышать количества выпускников 9, 11 классов Ханты-Мансийского автономного округа Югры:

$$Zakaz(l) = \sum_{i(l)=1}^{n(l)} \sum_{j(l)=1}^{m(l)} X_{i(l)j(l)}^l \quad (5)$$

$$Zakaz(l) \leq Abitur(l) \quad ,$$

где $Zakaz(l)$ – суммарное количество кадров с уровнем образования l на подготовку; $Abitur(l)$ – количество абитуриентов, которые могут поступить в УПО с уровнем образования l .

Требования региональной целевой программы «Новая школа Югры» определяют желаемый объем подготовки кадров с СПО, что описывается следующим неравенством:

$$\frac{Zakaz(2)}{\sum_{l=1}^3 Zakaz(l)} \geq K_spo, \quad (6)$$

где K_spo – доля кадров, которые требуется подготовить со средним профессиональным образованием.

Распределение регионального заказа на подготовку кадров среди УГС определяется следующим выражением:

$$\sum_{j(l)=1}^{m(l)} X_{i(l)j(l)}^l \leq Zakaz(l) * K_ves_{i(l)}^l \quad i(l) = 1 \dots n(l) \quad , \quad (7)$$

$$ZakazUGS_{i(l)}(l) = \sum_{j(l)=1}^{m(l)} X_{i(l)j(l)}^l$$

где $ZakazUGS_{i(l)}(l)$ – региональный заказ по УГС.

Региональный заказ на подготовку кадров должен учитывать движение контингента студентов (отчисление, болезнь, декретный отпуск и т.д.). Корректировку регионального заказа на подготовку кадров, исходя из данных движения контингента студентов, можно описать следующим выражением:

$$\sum_{j(l)=1}^{m(l)} Possibility_{i(l)j(l)}^l * X_{i(l)j(l)}^l \leq Prognoz_{i(l)}(l) \quad i(l) = 1 \dots n(l), \quad (8)$$

где $Prognoz_{i(l)}(l)$ – потребность в кадрах с уровнем образования l по $i(l)$ УГС, $Possibility_{i(l)j(l)}^l$ – вероятность выпуска кадров по $i(l)$ УГС $j(l)$ УПО.

Распределение бюджетных средств по сети учреждений профессионального образования производится согласно рейтингу УПО в рамках УГС:

$$\left[\begin{array}{l} Rating_{i(l),2}^l * X_{i(l),1}^l - Rating_{i(l),1}^l * X_{i(l),2}^l = 0 \\ \vdots \\ Rating_{i(l),m(l)}^l * X_{i(l),m(l)-1}^l - Rating_{i(l),m(l)-1}^l * X_{i(l),m(l)}^l = 0 \end{array} \right], \quad (9)$$

где $Rating_{i(l)j(l)}^l$ – рейтинг $i(l)$ УГС $j(l)$ УПО.

АЛГОРИТМ ОПТИМАЛЬНОГО РАСПРЕДЕЛЕНИЯ РЕГИОНАЛЬНОГО ЗАКАЗА НА ПОДГОТОВКУ КАДРОВ

Для численной реализации предложенной методики разработан алгоритм оптимального распределения контрольных цифр приёма среди УПО. В алгоритме используются следующие начальные условия:

1. Рейтинг выпускающих УГС УПО составлен на основе приказа Минобразования РФ от 26 февраля 2001 г. № 631 (подсистема формирования рейтинга).
2. Равное количество человек в группе по уровням образования у всех УПО.

Для реализации алгоритма оптимального распределения регионального заказа на подготовку кадров требуется выполнение 6 этапов:

1. Определяем максимально допустимый региональный заказ на подготовку кадров по трём уровням образования:

$$Zakaz(l) = \min[Abitur(l), MaxNabor(l), \frac{Budget}{\min(Cost(l))}, \frac{Prognoz(l)}{\min(Possibility(l))}]. \quad (9)$$

2. Корректируем региональный заказ по уровням образования согласно условию (6).

3. Определяем региональный заказ на подготовку кадров по уровням образования и определённой УГС через коэффициент значимости УГС:

$$ZakazUGS_{i(l)}(l) = K_{ves_{i(l)}}(l) * Zakaz(l) \quad i(l) = 1 \dots n(l). \quad (10)$$

4. Выполняем проверку на ограничение по «производственной мощности» и минимальной группе для всех уровней образования и всех УГС.

5. Определяем заказ на подготовку кадров по требуемым УГС с учётом ограничений (8-9).

Для каждого уровня образования определяем УГС с максимальным коэффициентом значимости и выполняем следующие действия:

$$X_{i(l)j(l)}^l = \frac{Need_{i(l)}(l) * Rating_{i(l)j(l)}^l}{\sum_{j(l)=1}^{n(l)} Possibility_{i(l)j(l)}^l * Rating_{i(l)j(l)}^l}, \quad (12)$$

определяем нераспределенный бюджет на подготовку кадров:

$$Budget = Budget - \sum_{j(l)=1}^{n(l)} X_{i(l)j(l)}^l * Cost_{i(l)j(l)}^l, \quad (13)$$

если $Budget > 0$, возвращаемся в начало этапа 5 и выполняем всё заново, иначе 6.

1. Проверяем полученные контрольные цифры приёма на выполнение условий (3, 6, 7). Если условия (3, 6, 7) выполняются, значения контрольных цифр приема в УПО по УГС являются оптимальными согласно целевой функции (1), в противном случае необходимо пересчитать региональный заказ по уровням профессионального образования согласно следующего выражения:

$$Zakaz(l) = \sum_{i(l)=1}^{n(l)} \sum_{j(l)=1}^{m(l)} X_{i(l)j(l)}^l \quad (14)$$

и выполнить этапы 3-7.

АПРОБАЦИЯ АЛГОРИТМА

В таблице 1 представлена информация о потребности региона в кадрах по уровням образования и укрепленным группам специальностям. В таблице 2 содержится максимальное количество кадров, которые может подготовить учреждение профессионального образования, в таблице 3 – рейтинг УГС, в таблице 4 – вероятность выпуска кадров по УГС, в таблице 5 – стоимость обучения по УГС, а в таблице 6 – значимость УГС. Будем считать, что минимальное количество студентов в группе больше или равно 20.

Таблица 1 – Потребность в кадрах

УГС	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19
ВПО	220	291	260	25	109	110	205	179	236	110	61	25	231	61	116	165	68	192	145
СПО	33	235	87	181	289	129	208	227	129	196	32	280	56	79	239	146	230	118	81
НПО	37	39	27	2	126	196	216	159	32	189	37	40	29	42	50	58	95	94	65

Таблица 2 – «Производственная мощность» УПО

УГС	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19
ВПО1	0	60	0	0	40	30	70	70	60	0	0	50	30	90	90	20	80	80	50
ВПО2	0	90	0	80	30	50	80	90	60	20	60	0	40	60	90	80	40	70	40
СПО1	50	50	0	80	40	20	80	70	80	20	60	60	0	40	20	0	20	80	80
СПО2	80	70	50	60	40	40	0	80	30	20	30	30	50	50	30	0	50	60	90
СПО3	0	50	0	0	60	80	90	50	90	50	50	30	40	40	0	0	0	40	80
НПО1	50	60	80	50	20	0	40	90	60	60	70	30	50	50	0	50	60	40	0
НПО2	70	70	0	30	50	0	0	20	80	0	80	0	60	50	20	50	0	60	0

Таблица 3 – Рейтинг УПО по УГС

УГС	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19
ВПО1	0	69	0	48	56	70	52	68	67	59	55	53	62	40	45	42	79	59	76
ВПО2	0	58	0	58	76	78	62	39	46	55	73	62	61	75	39	41	58	64	54
СПО1	62	63	61	52	51	69	53	63	44	41	62	66	45	57	60	0	43	66	41
СПО2	47	54	41	79	49	61	76	46	47	78	58	68	42	40	54	0	49	46	65
СПО3	71	38	70	66	50	71	56	53	43	51	65	49	43	37	64	0	55	61	58
НПО1	66	58	65	72	63	0	52	46	49	68	69	64	40	44	52	68	65	58	0
НПО2	53	71	58	62	58	0	58	51	70	60	53	73	55	54	53	69	50	53	0

Таблица 4 – Вероятность выпуска кадров в процентах

УГС	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19
ВПО1	0	86	0	59	70	87	65	85	83	73	68	66	77	50	56	52	98	73	94
ВПО2	0	74	0	74	96	99	79	50	58	70	93	79	77	95	50	52	74	81	69
СПО1	89	90	87	75	73	98	76	90	63	58	89	95	65	82	86	0	61	95	58
СПО2	59	67	51	99	61	76	95	57	59	98	73	85	53	50	67	0	61	58	81
СПО3	99	53	97	92	69	98	77	73	59	70	90	68	59	51	89	0	76	85	81
НПО1	90	79	88	98	85	0	71	62	66	92	94	86	54	60	71	92	88	78	0
НПО2	71	95	77	83	77	0	77	68	94	80	71	97	73	72	71	92	67	71	0

Таблица 5 – Стоимость обучения в тысячах рублей

УГС	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19
ВПО	80	70	80	54	61	91	65	91	63	72	99	61	87	99	50	60	87	66	97
СПО	62	70	58	66	87	91	76	90	84	69	87	76	80	96	77	66	87	91	94
НПО	66	69	59	88	84	64	66	77	61	76	99	52	92	70	60	54	77	77	67

Таблица 6 – Значимость УГС

УГС	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19
ВПО	0,00	0,06	0,00	0,11	0,03	0,06	0,08	0,11	0,11	0,10	0,08	0,06	0,01	0,05	0,04	0,03	0,03	0,01	0,03
СПО	0,10	0,01	0,08	0,10	0,01	0,08	0,05	0,05	0,06	0,05	0,03	0,09	0,04	0,09	0,03	0,00	0,02	0,04	0,08
НПО	0,01	0,08	0,12	0,08	0,08	0,00	0,01	0,06	0,06	0,08	0,03	0,06	0,05	0,06	0,01	0,06	0,06	0,11	0,00

Предполагаем, что абитуриентов, готовых поступить в учреждения ВПО, не более 10 тысяч человек, в учреждения СПО – не более 4 тысяч человек, в НПО – не более 1 тысячи человек. Рассмотрим два различных региональных бюджета на подготовку кадров: 250 миллионов рублей, 2 миллиона рублей. Результат применения предложенного алгоритма формирования контрольных цифр приведен в таблицах 7-8.

Таблица 7 – Заказ на подготовку кадров при бюджете в 2 миллиона рублей

УГС	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19
ВПО1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
ВПО2	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
СПО1	38	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
СПО2	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
СПО3	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
НПО1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
НПО2	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0

Таблица 8 – Заказ на подготовку кадров при бюджете в 250 миллионов рублей

УГС	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19
ВПО1	0	60	0	0	39	30	69	70	60	0	0	33	0	26	90	20	0	0	0
ВПО2	0	89	0	36	30	49	80	89	59	0	59	0	0	50	89	79	0	0	0
СПО1	38	0	0	52	0	20	70	70	70	20	0	59	0	40	0	0	0	65	26
СПО2	0	0	25	60	0	40	0	78	30	20	0	30	30	48	0	0	0	38	42
СПО3	0	0	0	0	0	79	74	50	68	49	0	30	31	34	0	0	0	40	38
НПО1	0	0	32	0	20	0	0	89	0	59	0	20	0	28	0	31	59	40	0
НПО2	0	43	0	0	49	0	0	20	37	0	0	0	0	34	0	31	0	59	0

В таблице 7 представлен вариант ограниченного бюджета на подготовку кадров. В этом случае заказ на подготовку кадров получил УПО с самым высоким рейтингом по УГС с наибольшей для региональной экономики значимостью. В таблице 8 заказ распределён в рамках имеющегося бюджета на подготовку кадров в соответствии с рейтингом УПО, значимостью УГС, «производственной мощностью», движением контингента студентов.

ЗАКЛЮЧЕНИЕ

Решение задачи определения кадровых потребностей региональной экономики позволяет получить данные, необходимые для управления системой профессионального образования через инструмент формирования заказа на подготовку квалифицированных кадров. Однако задача распределения регионального заказа на подготовку кадров по сети учреждений профессионального образования требует сбора дополнительных данных, описывающих учреждения среднего и высшего профессионального образования, выпуск школ, региональный бюджет на подготовку кадров.

В статье описаны структура дополнительных данных, требуемая для распределения заказа на подготовку кадров по сети учреждений профессионального образования региона. Описан алгоритм, позволяющий распределить региональный заказ на подготовку кадров с учетом существующих ограничений, прогноза потребностей на подготовку кадров, данных рынка труда, рейтинга учреждений профессионального образования.

СПИСОК ЛИТЕРАТУРЫ

1. Проект по концепции долгосрочного социально-экономического развития Российской Федерации на период до 2020 года [Электронный ресурс]. – URL: <http://www.economy.gov.ru/minec/press/news/doc1224245909936>.
2. Прогнозирование потребности региональной экономики в подготовке квалифицированных кадров: монография / А.Г. Мокроносов и др. // Екатеринбург: Изд-во Рос. гос. проф-пед. университета, 2010. – 111 с.
3. Питухин Е.А., Гуртов В.А. Математическое моделирование динамических процессов в системе «экономика - рынок труда - профессиональное образование». – СПб.: Изд-во С.-Петербург. ун-та, 2006. – 350 с.
4. Система прогнозирования потребности в трудовых ресурсах [Электронный ресурс]. – URL: <http://ibs.ru/content/rus/629/6296-article.asp>.

5. Карминская Т.Д., Алексеев В.И. Гибридная модель прогнозирования потребностей в квалифицированных кадрах для социально-экономических систем регионального уровня управления // Известия ОрелГТУ, № 2/52 (563), 2009. – С. 3-11.
6. Тей Д.О, Татьянкин В.М. Недостатки статистических данных для систем прогнозирования кадровых потребностей [Электронный ресурс]. – URL: <http://irsit.ru/article172>.
7. Босс В. Лекции по математике: линейная алгебра. – Т. 3. – М.: КомКнига, 2005. – 224 с.

Карминская Татьяна Дмитриевна

Югорский государственный университет, г. Ханты-Мансийск
Кандидат технических наук
Тел.: 8 (3467) 35-75-04
E-mail: ktd@mail.ru

Тей Дмитрий Олегович

Югорский государственный университет, г. Ханты-Мансийск
Кандидат технических наук
Тел.: (3467) 35-78-54
E-mail: teyd@ugrasu.ru

Татьянкин Виталий Михайлович

Югорский государственный университет, г. Ханты-Мансийск
Аспирант
Тел.: 8 950 501 05 98
E-mail: bambar@bk.ru

Русанов Михаил Александрович

Югорский государственный университет, г. Ханты-Мансийск
Аспирант
Тел.: 8 951 980 48 91
E-mail: m_rusanov@ugrasu.ru

T.D. KARMINSKAYA (*Candidate of Engineering Science*)

D.O. TEJ (*Candidate of Engineering Science*)

V.M. TAT'YANKIN (*Post-graduate Student*)

M.A. RUSANOV (*Post-graduate Student*)
Ugra State University, Khanty-Mansiysk

**DEFINITION OF EDUCATION ORDER SUBJECT
TO REQUIREMENT FORECAST OF REGIONAL ECONOMY**

In this paper the question of education order grouping among establishments of professional education into extended specialty divisions is discussed in in some resource-limited settings. Structure of additional data and distribution algorithm are defined and approved.

Keywords: *regional order; rating; optimal assignment; forecasting.*

BIBLIOGRAPHY (TRANSLITERATED)

1. Proekt po koncepcii dolgosrochnogo social'no-e'konomicheskogo razvitiya Rossijskoj Federacii na period do 2020 goda [E'lektronny'j resurs]. – URL: <http://www.economy.gov.ru/minec/press/news/doc1224245909936>.
2. Prognozirovanie potrebnosti regional'noj e'konomiki v podgotovke kvalificirovanny'x kadrov: monografiya / A.G. Mokronosov i dr. // Ekaterinburg: Izd-vo Ros. gos. prof-ped. universiteta, 2010. – 111 s.
3. Pituxin E.A., Gurtov V.A. Matematicheskoe modelirovanie dinamicheskix processov v sisteme «e'konomika - ry'nok truda – professional'noe obrazovanie». – SPb.: Izd-vo S.-Peterb. un-ta, 2006. – 350 s.

4. Sistema prognozirovaniya potrebnosti v trudovy'x resursax [E'lektronny'j resurs]. – URL: <http://ibs.ru/content/rus/629/6296-article.asp>.
5. Karminskaya T.D., Alekseev V.I. Gibridnaya model' prognozirovaniya potrebnostej v kvalificirovanny'x kadrax dlya social'no-e'konomicheskix sistem regional'nogo urovnya upravleniya // Izvestiya OryolGTU, № 2/52 (563), 2009. – S. 3-11.
6. Tej D.O., Tat'yankin V.M. Nedostatki statisticheskix danny'x dlya sistem prognozirovaniya kadrovy'x potrebnostej [E'lektronny'j resurs]. – URL: <http://irsit.ru/article172>.
7. Boss V. Lekcii po matematike: linejnaya algebra. – T. 3. – M.: KomKniga, 2005. – 224 s.

УДК 004.942

М.Ю. МАКАРОВА

ПОСТРОЕНИЕ МОДЕЛИ УПРАВЛЕНИЯ ПРОФОРИЕНТАЦИОННЫМИ ПРОЦЕССАМИ

В данной статье рассмотрен пример использования латентно-семантического метода для установления соответствий между компетенциями, имеющимися в государственных образовательных стандартах, и должностными инструкциями к профессиям, а также группирования схожих по смыслу видов деятельности. На основе данного подхода построена модель управления профориентационными процессами, позволяющая устранить противоречия между требованиями работодателей и устремлениями соискателей.

Ключевые слова: управление профориентационными процессами; метод латентно-семантического анализа.

ВВЕДЕНИЕ

Профориентационная деятельность приобретает в последнее время особую актуальность, поскольку множество специалистов не востребованы из-за перенасыщенности рынка труда одинаковыми профессиями, а существующие вакансии продолжают оставаться открытыми. Это происходит потому, что большинство школьников, на плечи которых, в основном, ложится выбор будущей профессии, не имеют однозначного представления ни о профессиях, и о том, какими навыками и знаниями необходимо обладать для их успешного освоения, ни о собственных возможностях и предпочтениях. Существующие тесты по профориентации направлены на выявление склонностей, способностей и особенностей характера пользователей, в качестве результатов предлагают наиболее подходящие профессии, но многие из них либо вовсе исчезли с рынка труда, либо имеют другие (современные) названия. Это создает путаницу в понимании потребностей общества, а, соответственно, лишь затрудняет процесс профессионального самоопределения. Для решения данной проблемы и устранения противоречий между потребностями общества в кадрах и профессиональными устремлениями большей части этого общества была разработана информационная система поддержки профориентационных процессов, рассматриваемая в данной статье.

Анализ диссертаций [1-5] позволил сделать вывод, что профориентационная деятельность довольно подробно изучена и является очень актуальной для современного общества, а основные принципы профориентации могут быть успешно автоматизированы. Однако существующие программные комплексы по профориентации [6-10] имеют ряд недостатков:

- данные комплексы не бесплатны и профориентационная поддержка не осуществляется дистанционно;
- имеется ограниченный набор методик для диагностики пользователей;
- подбор предпочтительных профессий осуществляется без учета современных требований к образованию и квалификации;
- отсутствует возможность получения обобщенной рекомендации по выбору будущей профессии и направления обучения для освоения данной профессии.

Рассматриваемая в данной статье информационная система направлена на устранение описанных недостатков и имеет своей целью создание механизма управления профориентационными процессами на различных уровнях посредством автоматизации процедур профориентационной диагностики и консультирования по выбору будущей профессии. Результатом работы с системой является путь профессионального становления личности в соответствии с индивидуальными данными и требованиями рынка труда, а

именно, наилучшим образом подходящие профессии, способы овладения той или иной профессией и вакансии на рынке труда.

Научная новизна данной работы состоит в построении модели управления профориентационными процессами с использованием метода латентно-семантического анализа для установления соответствий между результатами освоения государственных образовательных программ и требованиями работодателей к соискателям. Применение латентно-семантического анализа в данной информационной системе позволило решить основную проблему профориентации – определение направлений обучения, которые подходят для освоения той или иной профессии и устранение противоречий в создании образа профессии за счет объединения схожих по смыслу видов деятельности в группы.

ИСПОЛЬЗОВАНИЕ ЛАТЕНТНО-СЕМАНТИЧЕСКОГО АНАЛИЗА ДЛЯ РЕШЕНИЯ ПРОБЛЕМЫ ПРОФОРИЕНТАЦИИ

Метод латентно-семантического анализа подробно описан в источнике [11]. В рамках данной статьи рассмотрим основные этапы, из которых состоит процедура установления соответствий между описаниями профессий и компетенциями, представленными в государственных образовательных стандартах (ГОС):

1) Предварительная обработка исходной текстовой информации: исключение стоп-слов (предлоги, союзы, частицы).

2) Проведение операции стемминга с помощью алгоритма Портера для выделения основы слова, то есть получения так называемых термов.

3) Исключение термов, встречающихся в блоке текста один раз.

4) Составление частотной матрицы, в которой строки – полученные на предыдущем этапе термы, нулевой столбец – блок с описанием профессии, следующие столбцы – блоки компетенций. В каждой ячейке указывается, сколько раз данный терм используется в этом текстовом блоке.

5) Нормализация частотной матрицы с использованием меры TF-IDF. Мера TF-IDF является произведением двух сомножителей TF и IDF:

$$TF = \frac{n_i}{\sum_k n_k}, \quad (1)$$

где n_i – число вхождений терма в текстовый блок; $\sum_k n_k$ – общее число термов в данном блоке.

$$IDF = \lg \frac{|D|}{\left| \left(d_i \supset t_i \right) \right|}, \quad (2)$$

где $|D|$ – общее количество текстовых блоков; $\left| \left(d_i \supset t_i \right) \right|$ – количество блоков, в которых встречается данный терм.

6) Реализация сингулярного разложения, позволяющего выделить ключевые составляющие матрицы, игнорируя шумы, путем разложения полученной матрицы M на три составляющие:

$$M = U \cdot W \cdot V^T, \quad (3)$$

где U и V^T – ортогональные матрицы; W – диагональная матрица, элементы которой упорядочены по убыванию (сингулярные числа).

7) Расчет коэффициента соответствия компетенций описанию профессии. Для каждого из блоков-компетенций (столбцы матрицы, кроме нулевого) определяется

коэффициент сходства с описанием профессии (нулевой столбец), используя значение косинусной меры по следующей формуле:

$$c_i = \frac{\sum_{k=0}^2 (V_{k,0} \cdot V_{k,i})}{\sqrt{\sum_{k=0}^2 (V_{k,0})^2} \cdot \sqrt{\sum_{k=0}^2 (V_{k,i})^2}}, \quad (4)$$

где c_i – коэффициенты сходства блоков-компетенций с описанием профессии; $V_{k,0}$ – значения нулевого столбца (описание профессии); $V_{k,i}$ – значения блоков-компетенций; i –

порядковый номер блока-компетенции; k – порядковый номер строки матрицы V .

Причем из формулы видно, что для расчета используются только первые три строки полученной на предыдущем этапе матрицы V^T . Иными словами, реализована так называемая трехмерная декомпозиция, поскольку в ходе испытаний данный подход позволил получить наилучшие результаты.

8) Расчет общей оценки соответствия (OCR – Overall Conformity Rating) образовательного стандарта описания профессии путем нахождения общего арифметического полученных на предыдущем этапе коэффициентов:

$$OCR = \frac{\sum_{i=0}^{|C|} (c_i)}{|C|} \cdot 100\%, \quad (5)$$

где $|C|$ – общее количество блоков-компетенций.

На основе описанного алгоритма были определены коэффициенты соответствия образовательных стандартов тем или иным профессиям, имеющимся в базе данных информационной системы, представленные на рисунке 1.

Код	Наименование	Врач	Менеджер	Программист	Учитель	Экономист	Юрист
010400	Прикладная математика и информатика	20.1	59.2	23.2	43.4	23.5	20.4
030900	Юриспруденция	41.1	81	4.3	15	58	86.6
050100	Педагогическое образование	20.2	41.1	26.5	70.5	11.5	13.3
060101	Лечебное дело	90.5	46.9	6	7.4	15.5	9.3
060103	Педиатрия	90.1	68.4	5	8.7	11.7	11.9
080100	Экономика	32.5	70.4	56	37.6	84	28.2
080200	Менеджмент	29.3	89.8	38.6	16.7	62.3	17.2
100700	Торговое дело	36	89.3	6.3	7.2	52.2	15.1
151000	Технологические машины и оборудование	20.6	62.7	7	18	46.4	18.4
151900	Конструкторско-технологическое обеспечение машиностроительных производств	7.4	47.9	45.4	12.6	60.5	29.4
230100	Информатика и вычислительная техника	30.5	36.8	78.8	38.3	33.4	50.9
230400	Информационные системы и технологии	48.3	67.3	77.9	31.9	72.7	28.7
230700	Прикладная информатика	26.8	57.7	60	41.4	47.3	23.6

Рисунок 1 – Коэффициенты соответствия образовательных стандартов профессиям

Как видно из рисунка 1, результаты испытаний вполне можно считать удачными, поскольку интуитивно близкие профессиям направления обучения получили высокие баллы, **№1(81)2014**

а по смыслу не подходящие той или иной профессии образовательные стандарты получили низкие оценки.

Еще одной задачей, которую можно решить с помощью латентно-семантического анализа, является устранение противоречий в составлении образа профессии путем объединения схожих по смыслу профессий в группы, что, в свою очередь, облегчит выбор будущей профессии. Например, для профессии «менеджер» являются по смыслу похожими профессии «менеджер по рекламе», «менеджер по связям с общественностью», «менеджер по персоналу», поэтому данные виды деятельности можно назвать разновидностями профессии «менеджер». Рассмотрим, каким образом данный процесс реализован в информационной системе.

При добавлении новой профессии производится латентно-семантический анализ по описанному ранее алгоритму с целью обнаружения похожих по описанию профессий, а результаты представляются пользователю для выбора лучшего варианта (рис. 2).

	Наименование профессии	Коэффициент сходства
<input type="radio"/>	Юрист	97.9%
<input type="radio"/>	Экономист	98.7%
<input checked="" type="radio"/>	Менеджер	99.1%

Добавить новую профессию Добавить разновидность профессии

Рисунок 2 – Результаты определения коэффициентов сходства для профессии «Менеджер по персоналу»

Как видно из рисунка 2, наибольший коэффициент схожести получила профессия «менеджер», которая и была выбрана в данном случае. Коэффициенты соответствия ГОС, которые были рассчитаны для профессии «менеджер», признаются аналогичными и для данной профессии «менеджер по персоналу». Если же схожие профессии отсутствуют в системе, происходит автоматическое добавление новой профессии и производится расчет коэффициентов соответствия ГОСов. На рисунке 3 представлены профессии и их разновидности, которые были определены с помощью латентно-семантического анализа.

Стоит отметить, что поскольку описания профессий представляют собой один блок текста, информация довольно сжата и однотипна, для получения наилучших результатов формула (1) для расчета меры TF была изменена и выглядит следующим образом:

$$TF = \frac{n}{\max_k(n_k)}, \quad (6)$$

где $\max_k(n_k)$ – максимальное количество термов в данном блоке.

Таким образом, использование латентно-семантического анализа позволило, во-первых, определить требования к образованию и квалификации соискателей, во-вторых, устранить неопределенности в создании образа профессии. Также следует отметить, что данный подход лежит в основе построения модели управления профориентационными процессами в рассматриваемой информационной системе.

Профессии		
+ Добавить - Удалить Выделить все Снять выделение		
Название	Разновидности профессии	Описание
Врач	врач-интерн, врач-патологоанатом, врач-специалист, врач-суд.мед.эксперт, врач скорой медицинской помощи	оказание своевременной плановой и экстренной медицинской просветительную работу среди больных и их родственников п
Менеджер	менеджер по персоналу, менеджер по рекламе, менеджер по связям с инвесторами, менеджер по связям с общественностью	работа с другими людьми в организациях для достижения орга стабильного функционирования, поддержания деловой репутаи внешних связей и обмена опытом. Осуществляет анализ спро
Программист	техник-программист	На основе анализа математических моделей и алгоритмов рео основе анализа выходных данных. Разрабатывает инструкции действиям, составление алгоритма или выбор оптимального д
Учитель	воспитатель, педагог-организатор, педагог дополнительного образования, социальный педагог	Осуществляет обучение и воспитание обучающихся с учетом разрабатывает рабочую программу по предмету, курсу на осн обучающихся, поддерживает учебную дисциплину, режим посе
Экономист	экономист по сбыту, экономист по финансовой работе	Выполняет работу по осуществлению экономической деятельн выпускаемой продукции, производительности труда, снижению правильностью осуществления расчетных операций. Ведет уч методы и средства проведения вычислительных работ; прави
Юрист	-	консультирование лиц, нуждающихся в юридической помощи (Разрабатывает или принимает участие в разработке документ рассмотрения претензий, судебных и арбитражных дел, практ

Рисунок 3 – Результаты определения разновидностей профессий с помощью латентно-семантического анализа

МОДЕЛЬ УПРАВЛЕНИЯ ПРОФОРИЕНТАЦИОННЫМИ ПРОЦЕССАМИ

Для реализации механизма управления профориентационными процессами с учетом выше указанных требований необходимо наполнить базу данных системы следующими данными:

- государственные образовательные стандарты с учетом разбиения на блоки общекультурных и профессиональных компетенций, которые разделены по областям деятельности;
- современные профессии или указание разновидностей профессий;
- тестовые методики, с помощью которых будет проводиться профориентационная диагностика.

Для успешной работы информационной системы эти действия необходимо выполнить один раз, но по мере необходимости дополнительная информация также может быть добавлена. На рисунке 4 представлена модель управления профориентационными процессами, реализованная в рассматриваемой системе.

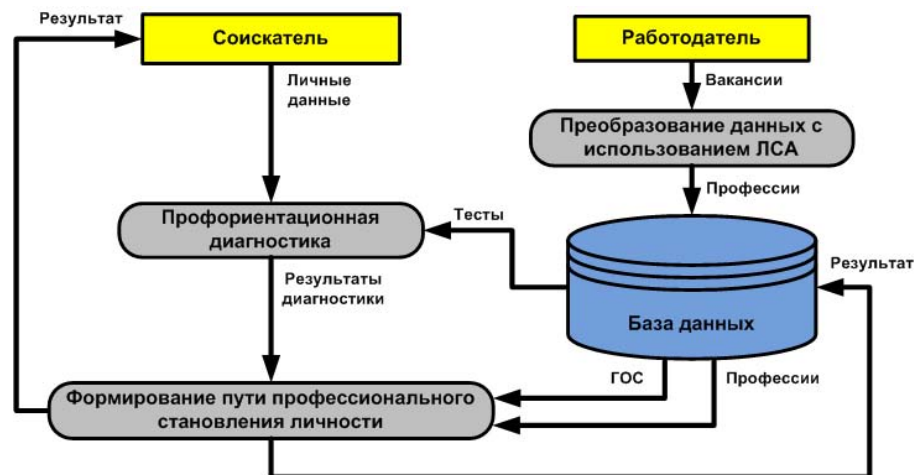


Рисунок 4 – Модель управления профориентационными процессами

Рассмотрим процесс формирования пути профессионального становления личности более подробно. В данном случае под этим понимается представление нескольких профессий, включая и разновидности профессий, наиболее подходящих соискателю по результатам проведенной профессиональной диагностики, а также варианты получения образования, необходимого для той или иной профессии. Путь профессионального становления личности формируется путем выполнения следующих этапов:

- запрос из базы данных тех профессий (и их разновидностей), которые соответствуют полученному результату профориентационной диагностики;
- выбор наиболее соответствующих полученным профессиям государственных образовательных стандартов;
- предоставление пользователю полученных результатов в виде иерархии;
- занесение полученных результатов в базу данных.

Наиболее соответствующие профессиям ГОСы определяются путем нахождения максимального значения коэффициента схожести (значения расположены в интервале от 0 до 100 баллов). Для большей гибкости системы выбираются также те стандарты, которые отличаются от максимального значения не более, чем на 5 баллов.

В качестве демонстрации работоспособности данной модели в рассматриваемой информационной системе для диагностики пользователей была использована наиболее популярная профориентационная методика Е.А. Климова [12]. Согласно полученным в ходе тестирования результатам, для конкретного пользователя наиболее предпочтительными являются профессии типа «Человек-человек», а именно врач, юрист, менеджер и учитель, имеющиеся в базе данных системы. Сформированный путь профессионального становления личности по вышеописанному алгоритму представлен на рисунке 5.

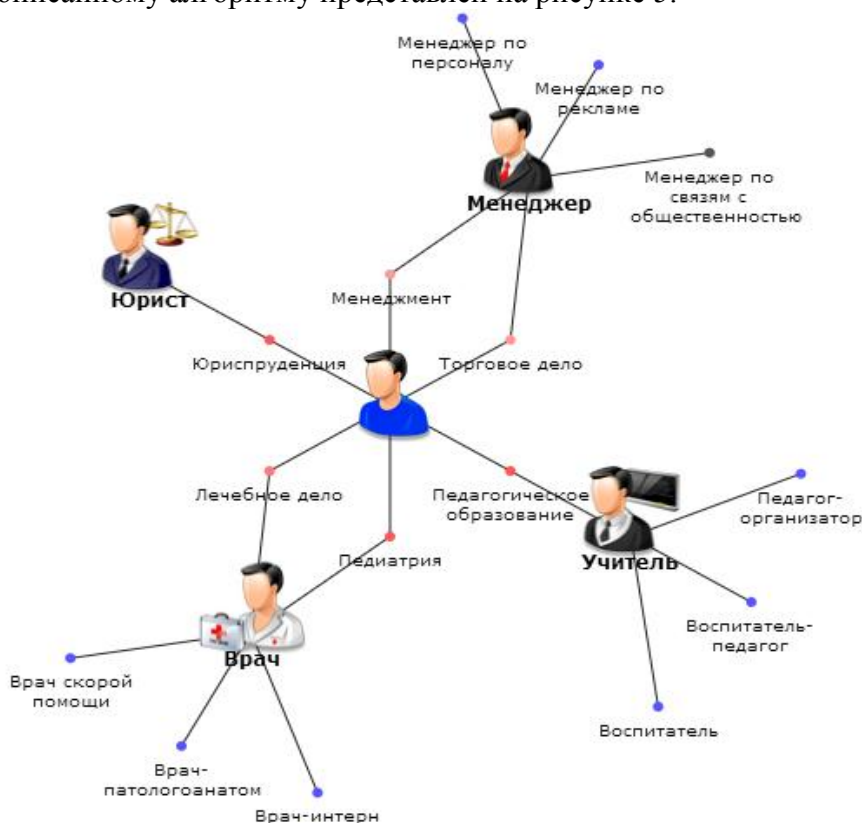


Рисунок 5 – Результат формирования пути профессионального становления личности на основе проведенной диагностики

ЗАКЛЮЧЕНИЕ

В данной статье была предложена модель управления профориентационными процессами, позволяющая устранить существующие недостатки профориентационной

работы за счет применения латентно-семантического анализа. Использование данного метода позволило определить те направления обучения, которые наилучшим образом подходят для освоения той или иной профессии, а также устранить противоречия в понимании образа профессии за счет объединения схожих видов деятельности в группы, тем самым облегчая профессиональный выбор.

На основе представленной модели и алгоритмов, а также используя результаты процесса проектирования системы [13], была реализована информационная система поддержки профориентационных процессов. Тестирование работоспособности данной системы показало, что установленные с помощью латентно-семантического анализа соответствия между образовательными стандартами и профессиями корректны, а сформированный путь профессионального становления личности является интуитивно понятным и простым для восприятия.

СПИСОК ЛИТЕРАТУРЫ

1. Живага А.Ю. Особенности профессиональных ориентаций студенческой молодежи (на материалах Сахалинской области): автореф. дис. ... канд. соц. наук. – М., 2008. – 31 с.
2. Ивашина А.В. Модели и методы построения систем интеллектуальной поддержки профориентации: автореф. дис. ... канд. техн. наук. – Красноярск, 2008. – 21 с.
3. Мутырова А.С. Педагогическая система профориентационной работы ВУЗа в условиях непрерывного образования: автореф. дис. ... пед. наук. – Астрахань, 2008. – 22 с.
4. Самойло И.В. Математические модели и алгоритмы профессиональной ориентации и управления знаниями: автореф. дис. ... канд. техн. наук. – М., 2010. – 25 с.
5. Сибгатова К.И. Профориентационная работа в интегрированной системе «школа-колледж-предприятие»: автореф. дис. ... канд. пед. наук. – Казань, 2011. – 24 с.
6. Диагностический компьютерный тестовый комплекс «Профориентатор» [Электронный ресурс]. – URL: <http://teletesting.ru/modules/tests>.
7. Комплекс Effecton Studio. Пакет «Профориентация» [Электронный ресурс]. – URL: <http://www.effecton.ru/730.html>.
8. Компьютерная система психодиагностики СМОЛ-Эксперт [Электронный ресурс]. – URL: http://www.nmc-radix.com/r_se.html.
9. ПДК «Эксперт+» [Электронный ресурс]. – URL: <http://www.expert-plus.ru>.
10. Программный комплекс КАТАРСИС [Электронный ресурс]. – URL: <http://katarsisprog.ru>.
11. Макарова М.Ю. Применение латентно-семантического анализа для автоматизации процесса профориентации // Системный анализ в науке и образовании, 2012. – № 3. – [Электронный ресурс]. – URL: <http://www.sanse.ru/archive/25> (дата обращения: 08.06.2013).
12. Климов Е.А. Как выбирать профессию: кн. для учащихся. – М.: Просвещение, 1984. – 160 с.
13. Макарова М.Ю. Проектирование информационной системы поддержки профориентационных процессов различных уровней // Наука и образование, 2012. – № 9. [Электронный ресурс]. – URL: <http://technomag.edu.ru/doc/452460.html> (дата обращения: 08.06.2013).

Макарова Мария Юрьевна

Муромский институт (филиал) ФГБОУ ВПО «Владимирский государственный университет имени Александра Григорьевича и Николая Григорьевича Столетовых», г. Муром
Аспирант кафедры «Физика и прикладная математика»
Тел.: 8 920 623 67 75
E-mail: marisabel.science@gmail.com

M.Yu. MAKAROVA (*Post-graduate Student of the Department «Physics and Applied Mathematics»*)
Murom Institute (Branch) State Institution of Higher Professional Education «Vladimir State University named after Alexander and Nicholas Stoletovs», Murom

THE CONSTRUCTION OF MODEL OF VOCATIONAL GUIDANCE
PROCESSES CONTROL

This article presents an example of using latent semantic analysis method for establishing correspondences between the competences described in the state educational standards and job descriptions for the professions, and for grouping similar types of activities. On basis this approach, a model for management of vocational guidance processes is built, which can eliminate the contradiction between the demands of employers and ambitions of job seekers.

Keywords: management of vocational guidance processes; latent semantic analysis method.

BIBLIOGRAPHY (TRANSLITERATED)

1. Zhivaga A.Yu. Osobnosti professional'ny'x orientacij studencheskoj molodyozhi (na materialax Saxalinskoj oblasti): avtoref. dis. ... kand. socz. nauk. – M., 2008. – 31 s.
2. Ivashina A.V. Modeli i metody' postroeniya sistem intellektual'noj podderzhki proforientacii: avtoref. dis. ... kand. texn. nauk. – Krasnoyarsk, 2008. – 21 s.
3. Mut'yrova A.S. Pedagogicheskaya sistema prforientacionnoj raboty' VUZa v usloviyax neprery'vnogo obrazovaniya: avtoref. dis. ... ped. nauk. – Astrazan', 2008. – 22 s.
4. Samojlo I.V. Matematicheskie modeli i algoritmy' professional'noj orientacii i upravleniya znaniyami: avtoref. dis. ... kand. texn. nauk. – M., 2010. – 25 s.
5. Sibgatova K.I. Proforientacinnaya rabota v integrirovannoj sisteme «shkola-kolledzh-predpriyatie»: avtoref. dis. ... kand. ped. nauk. – Kazan', 2011. – 24 s.
6. Diagnosticheskij komp'yuternyj testovyj kompleks «Proforientator» [E'lektronny'j resurs]. – URL: <http://teletesting.ru/modules/tests>.
7. Kompleks Effecton Studio. Paket «Proforientaciya» [E'lektronny'j resurs]. – URL: <http://www.effecton.ru/730.html>.
8. Kmp'yuternaya sistema psixodiagnostiki SMOL-E'kspert [E'lektronny'j resurs]. – URL: http://www.nmc-radix.com/r_se.html.
9. PDK «E'kspert+» [E'lektronny'j resurs]. – URL: <http://www.expert-plus.ru>.
10. Programmny'j kompleks KATARSIS [E'lektronny'j resurs]. – URL: <http://katarsisprog.ru>.
11. Makarova M.Yu. Primenenie latentno-semanticheskogo analiza dlya avtomatizacii processa proforientacii // Sistemny'j analiz v nauke i obrazovanii, 2012. – № 3. – [E'lektronny'j resurs]. – URL: <http://www.sanse.ru/archive/25> (data obrashheniya: 08.06.2013).
12. Klomov E.A. Kak vy'birat' professiyu: kn. dlya uchashhixsya. – M.: Prosveshhenie, 1984. – 160 s.
13. Makarova M.Yu. Proektirovanie informacionnoj sistemy' podderzhki proforientacionny'x processov razlichny'x urovnej // Nauka i obrazovanie, 2012. – № 9. [E'lektronny'j resurs]. – URL: <http://technomag.edu.ru/doc/452460.html> (data obrashheniya: 08.06.2013).

УДК 004.912

А.А. ТОЛКУНОВ

**АЛГОРИТМЫ ФОРМИРОВАНИЯ ВЫБОРОК
ТИПИЧНЫХ И ПОЛЕЗНЫХ КОММЕНТАРИЕВ
В ХОДЕ ОБРАБОТКИ РЕЗУЛЬТАТОВ
ОБЩЕСТВЕННОГО ОБСУЖДЕНИЯ ЗАКОНОПРОЕКТОВ**

Рассмотрены подходы к решению задачи повышения оперативности выделения высказываний при обработке массива комментариев интернет-пользователей, полученного в результате общественного обсуждения законопроекта. Предложены алгоритмы формирования выборок комментариев при заданном бюджете времени, учитывающие неравномерность распределения высказываний в комментариях. Приведены результаты экспериментов, показывающие эффективность предложенных алгоритмов в сравнении с существующим подходом к обработке массива комментариев.

Ключевые слова: общественное обсуждение законопроектов; массив комментариев; мнение; высказывание; модель оперативной аналитической обработки текстов; детализация; выборка типичных комментариев; выборка полезных комментариев; полнота выделения высказываний.

ВВЕДЕНИЕ

В настоящее время активно развивается направление совершенствования законопроектов методом краудсорсинга, а именно с использованием процедуры общественного осуждения на основе информационных порталов в сети Интернет. В их число входят порталы министерств, ведомств, а также специально созданные порталы в целях проведения общественного обсуждения, такие, как: zakonproekt2012.ru, regulation.gov.ru, на которых функционируют специальные интерфейсы для ознакомления с текстом законопроекта и создания комментариев интернет-пользователей [1, 2].

Анализ аналитического обзора, сформированного экспертами к законопроекту «О полиции», показал, что в ходе обработки полученных массивов комментариев группа экспертов классифицирует полученные комментарии по статьям, тональности, выделяет типичные мнения, которые подкрепляет полезными комментариями. Поскольку массив комментариев представляет собой массив текстов, приобретение знаний о содержании предложений осуществляется экспертами по итогам последовательного изучения массива комментариев.

Данные обстоятельства указывают на то, что решение задачи обработки комментариев осуществляется в условиях, существующих в области обработки текста проблем: понимания и моделирования знаний, при этом обработке подвергается специфический класс текстов, содержащих мнения, направленные на улучшение законопроектов. В этих условиях возникает потребность в моделировании аналитической обработки комментариев (рис. 1).

Задаче анализа мнений посвящено направление Opinion Mining, развивающееся в рамках методологии Text Mining. В рамках данного направления активно применяются методы классификации по эмоциональной и оценочной тональностям, а также распространены методы анализа мнений на основе извлечения ключевых слов и понятий, содержащихся в тексте [3]. Однако в доступной литературе не описаны модели аналитической обработки мнений, направленных на улучшение обсуждаемого объекта, позволяющие провести их обработку с обоснованной эффективностью. В то же время потребность в такой модели существует и по причине трудоемкости данного процесса, например, для изучения комментариев к законопроекту «О полиции» она составила:

$$T = N_c \cdot \bar{N}_w \cdot \bar{t}_{1w} = 20\,000 \cdot 73,5 \cdot 0,3 \approx 123(\text{часа}),$$

где N_c – количество комментариев, \bar{N}_w – среднее количество слов в комментарии, \bar{t}_{1w} – среднее время чтения одного слова. Таким образом, можно отметить существующие в области обработки данного вида мнений противоречия (рис. 2).



Рисунок 1 – Потребность в моделировании аналитической обработки комментариев

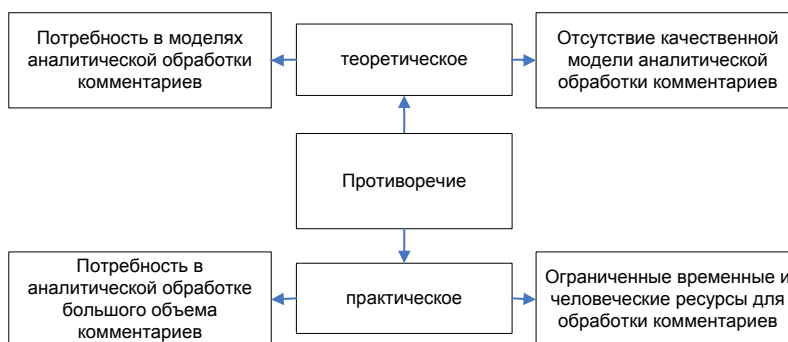


Рисунок 2 – Противоречие в исследуемой предметной области

Анализ источников, характеризующих понятие «мнение», позволяет ввести для него теоретико-множественное представление в виде кортежа:

$$M = \langle W, F, V, O, \{A_i\}, D, C, S, T, Q, E, U \rangle,$$

где W – текст высказывания, F – факт, относительно которого высказывают мнение, V – оценка факта, O – отношение к факту, $\{A_i\}$ – автор(ы) мнения, D – степень доверия мнению (обоснованность мнения), C – цель высказывания мнения, S – оценка стиля высказывания, T – время, Q – аспект факта, по которому высказывают мнение, E – интенсивность проявления мнения, U – оценка полезности мнения. Многомерность такого формального описания позволяет применять для обработки мнений модели оперативной аналитической обработки текстов, однако результаты анализа доступной литературы показывают, что предложенные методы погружения с использованием ключевых слов [4–5] также не в полной мере соответствуют решаемой задаче. Обработка мнений требует введения специфических операций погружения. Описанные условия позволяют провести постановку задачи.

ПОСТАНОВКА ЗАДАЧИ

Даны: $SENT = \bigcup_{m=1}^n sent_m$ – множество высказываний, $COM = \bigcup_{j=1}^k com_j$ – множество комментариев,

$L = \{l_j\}$ – множество весов комментариев, определяемых количеством

голосов респондентов, отданных за комментарий, $A = \bigcup_{s=1}^z a_s$ – упорядоченное множество атрибутов комментариев, где z – мощность кортежа, $z=7$; a_1 – глава законопроекта, к которой относится комментарий, $a_1 \in \overline{1, a_{1\max}}$; a_2 – статья законопроекта, к которой относится комментарий, $a_2 \in \overline{1, a_{2\max}}$; a_3 – идентификатор комментария, $a_3 \in \overline{1, a_{3\max}}$; a_4 – функциональный стиль комментария, $a_4 \in \{\text{научный, официально-деловой, газетно-публицистический, художественно-беллетристический, разговорный}\}$; a_5 – оценочная тональность комментария, $a_5 \in \{\text{мажорная, минорная}\}$; a_6 – эмоциональная тональность комментария, $a_6 \in \{\text{высокая, средняя, низкая}\}$; a_7 – количество комментариев. $\Omega = \bigcup_{v=1}^w \omega_v$ – множество классов комментариев, формируемых с использованием декартова произведения подмножества атрибутов из множества A . $\Psi = \bigcup_{j=1}^k \psi_j$, где $\psi_j : SENT \rightarrow COM$ – множество функций синтеза комментариев респондентами. Тогда процесс синтеза комментариев в ходе обсуждения законопроекта можно представить как последовательное добавление вершин $sent_{v,m_v}$, com_{v,j_v} и ребер $(sent_{v,m_v}, com_{v,j_v})$, где $v \in [1, w]$, $l_v \in [1, n_v]$, $j_v \in [1, k_v]$ в двудольный граф $G = (\{\{SENT_v\}, \{COM_v\}\}; \{(sent_{v,m_v}, com_{v,j_v})\})$ при добавлении в множество Ψ очередной функции ψ_{v,j_v} .

Требуется: для графа $G = (\{\{SENT_v\}, \{COM_v\}\}; \{(sent_{v,m_v}, com_{v,j_v})\})$ предложить модель обработки элементов множеств $\{COM_v\}$ при заданном директивном времени обработки Td на основе определения подмножеств $\{com_{v,j_v}\} \subseteq \{COM_v\}$, обеспечивающую выполнение требования:

$$\begin{cases} \frac{\Delta R_{\{com_{v,j_v}\}}}{\Delta t} \geq \frac{\Delta R_{\{COM_v\}}}{\Delta t}; \\ t \in [0; Td], \end{cases}$$

где $R_{\{com_{v,j_v}\}} = \frac{|SENT_{\{com_{v,j_v}\}}|}{|SENT_v|}$ – полнота присутствия высказываний в отобранных комментариях множества $\{com_{v,j_v}\}$ в сравнении с комментариями множества COM_v , $R_{\{COM_v\}} = \frac{|SENT_{\{COM_v\}}|}{|SENT_v|}$ – полнота присутствия высказываний в последовательно взятых за бюджет времени Td комментариях исходного множества COM_v , t – время.

РЕШЕНИЕ ЗАДАЧИ

Решение данной задачи с применением оперативной аналитической обработки текстов требует решения следующих подзадач:

1. Сформировать на основе множества значимых атрибутов $A = \bigcup_{s=1}^z a_s$ гиперкубовую модель $hc : a_1 \times \dots \times a_s \rightarrow a_r$, $s, r \in [1, z]$, $s \neq r$ и описать функции детализации

$f_{UCS} : COM \rightarrow UCS$ и $f_{TCS} : COM \rightarrow TCS$ для обеспечения операции качественного погружения в массив текстовых комментариев.

2. Пусть $p_{v,j_v} = d(com_{v,j_v})$ – количество высказываний в комментарии com_{v,j_v} , определяемое степенью соответствующей вершины графа G , $u_{v,j_v} = f(p_{v,j_v}, l_{v,j_v})$ – свёртка показателей p_{v,j_v} и l_{v,j_v} , характеризующая полезность комментария com_{v,j_v} .

Необходимо определить вид свёртки u_{v,j_v} и найти такое множество комментариев $UCS = \{com_{v,j_v}\}$, $UCS \subseteq COM$, для которого выполняется система требований:

$$\begin{cases} |UCS| \rightarrow \min; \\ \sum_{v=1}^w \sum_{j_v=1}^{h_v} u_{v,j_v} \rightarrow \max; \\ Td = const, \end{cases}$$

где $h_v = |UCS_v|$ – мощность множества UCS_v (рис. 3(а)).

3. Пусть $y_{v,m_v} = d(sent_{v,m_v})$ – количество комментариев, содержащих высказывание $sent_{v,m_v}$, определяемое степенью соответствующей вершины графа G , y_{v,m_v} характеризует типичность высказывания $sent_{v,m_v}$.

Необходимо найти такое множество комментариев $TCS = \{com_{v,j_v}\}$, $TCS \subseteq COM$, для которого выполняется система требований:

$$\begin{cases} |TCS| \rightarrow \min; \\ \sum_{v=1}^w \sum_{l_v=1}^{g_v} y_{v,m_v} \rightarrow \max; \\ Td = const, \end{cases}$$

где $g_v = |TCS_v|$ – мощность множества TCS_v (рис. 3(б)).

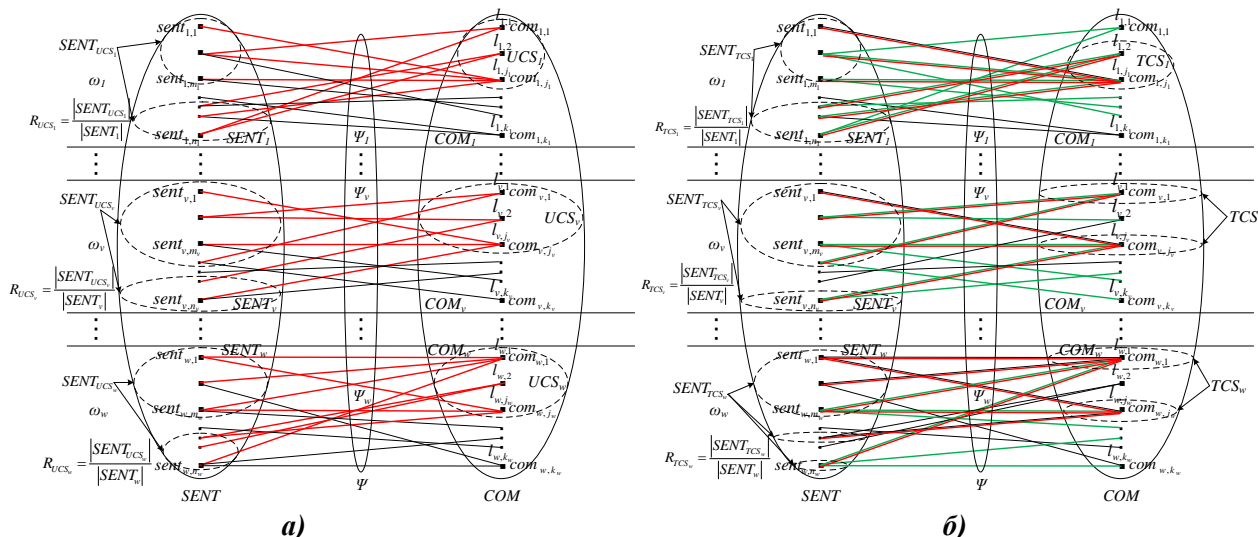


Рисунок 3 – Графы нахождения выборов а) полезных и б) типичных комментариев

В данной статье описаны алгоритмы решения 2 и 3 подзадач.

Алгоритм решения 2 подзадачи представлен на рисунке 4. Сущность данного алгоритма заключается в определении необходимых параметров и решении оптимизационной задачи с использованием метода множителей Лагранжа.

Применение метода множителей Лагранжа обусловлено возможностью построения аналитических выражений для убывающих функций предельной полезности комментариев в каждом классе, наличием временной стоимости обработки комментариев, связанной с темпом чтения и длиной комментария, и возможностью сформировать уравнение бюджетного ограничения (в данной задаче стоимость измеряется в единицах времени).

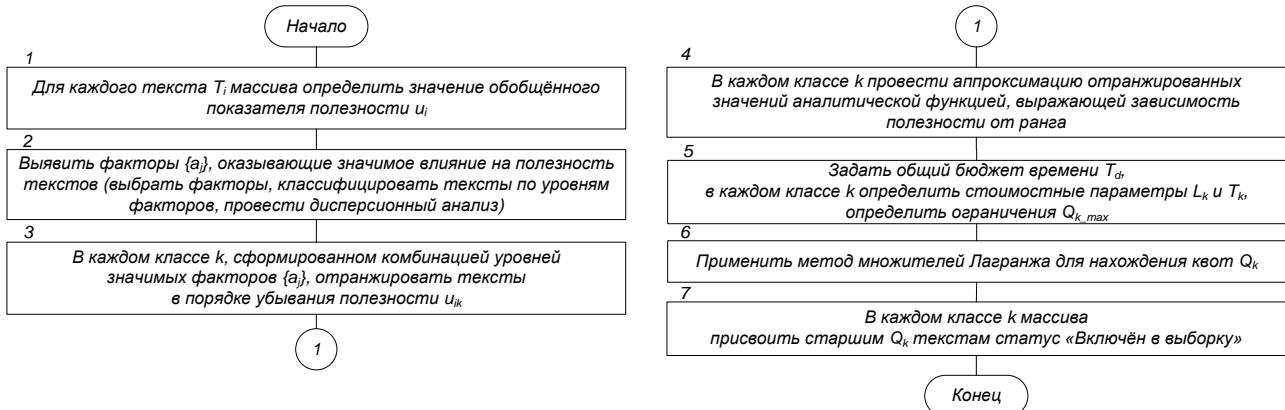


Рисунок 4 – Алгоритм формирования выборки полезных комментариев

На шаге 1 для определения обобщенного показателя полезности применяется аддитивная свертка числа голосов, отданных за комментарий, и числа конструктивных предложений; с учетом распределений данных параметров фрагмент расчета представлен в таблице 1.

Таблица 1 – Результаты расчета обобщенного показателя полезности комментариев к ЗП «О полиции»

Идентификатор комментария	Голосов	Конструктивных предложений	Голосов нормир.	Предложений нормир.	Свертка
...
11099	1	3	0,00108	0,157895	0,013468
11100	1	1	0,00108	0,052632	0,005152
11101	0	4	0	0,210526	0,016632
11102	0	0	0	0	0
11103	0	1	0	0,052632	0,004158
...

На шаге 2 классы, оказывающие значимое влияние на обобщенный показатель полезности, определяются с использованием дисперсионного анализа (табл. 2).

Таблица 2 – Результаты дисперсионного анализа для комментариев к ЗП «О полиции»

Эффект	SS	Степ. свободы	MS	F	P
Статья	0,064	56	0,001	2,75	0,000*
Стиль	0,15	4	0,037	89,6	0,000*
Тон(эмоц)	0,001	2	0	0,77	0,462
Тон(оцен)	0,003	1	0,003	6,56	0,010*
Статья	0,064	56	0,001	2,75	0,000*

Шаги 3 и 4 необходимы для получения аналитических выражений функций предельной полезности. На шаге 3 в выявленных на шаге 2 значимых классах производится ранжирование комментариев по убыванию обобщенного показателя полезности, полученного для каждого комментария на шаге 1.

На шаге 4 полученные кривые аппроксимируются логарифмической функцией (рис. 5). Результатом шага 4 является совокупность параметров всех функций предельной полезности.

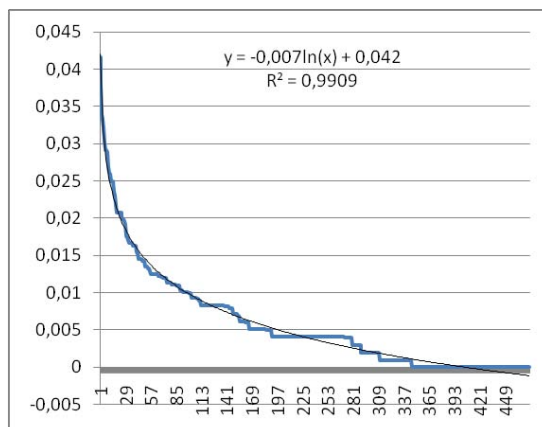


Рисунок 5 – Пример аппроксимации графика предельной полезности комментариев класса «2-Науч-Минор» к законопроекту «О полиции» логарифмической кривой с параметрами и достоверностью аппроксимации

На шаге 5 определяются стоимостные параметры и ограничения, характеризующие массив комментариев: максимальное число комментариев в значимых классах $Q_k \max$, средняя длина комментария в классе L_k , задается бюджет времени Td . Результаты выполнения шагов 4 и 5 представлены в таблице 3.

Таблица 3 – Фрагмент таблицы с параметрами полезности к ЗП «О полиции»

j	Статья	Стиль	Тон.(оцен.)	A	B (-B*lnX)	Достов. аппр.	Qmax	L
...
132	18	газет.-публ.	минорная	0,0303	0,009	0,9234	28	71,32
133	19	газет.-публ.	минорная	0,0176	0,005	0,9626	36	49,94
134	20	газет.-публ.	минорная	0,0221	0,006	0,9284	25	49,08
135	21	газет.-публ.	минорная	0,023	0,008	0,9344	20	60,7
...

Совокупность параметров, полученных на шаге 4 и 5 используется на шаге 6 в решении системы (1), здесь представлен фрагмент системы уравнений с неизвестными параметрами выборки полезных комментариев к законопроекту «О полиции»:

$$\begin{cases}
 0,5 \cdot 348800 = \sum_{v=1}^{472} L_v \cdot 0,3 \cdot Q_v \\
 \frac{0,076 - 0,013 \cdot \ln(Q_1)}{97,77 \cdot 0,3} = \frac{0,042 - 0,007 \cdot \ln(Q_2)}{73,45 \cdot 0,3} \\
 \frac{0,042 - 0,007 \cdot \ln(Q_2)}{73,45 \cdot 0,3} = \frac{0,051 - 0,009 \cdot \ln(Q_3)}{84,82 \cdot 0,3} \\
 \dots \\
 \frac{0,003 - 0,004 \cdot \ln(Q_{471})}{10 \cdot 0,3} = \frac{0,001 - 0,001 \cdot \ln(Q_{471})}{15 \cdot 0,3} \\
 1 \leq Q_1 \leq 496 \\
 \dots \\
 1 \leq Q_{472} \leq 1
 \end{cases} \quad (1)$$

На шаге 7 полученные на шаге 6 квоты применяются к полученным на шаге 3 ранжировкам, вошедшие в квоту комментарии выбираемого аналитиком класса составляют выборку полезных комментариев.

Алгоритм решения 3 подзадачи представлен на рисунке 6. Сущность данного алгоритма заключается в определении множества комментариев, обладающих наилучшим сходством с типичными представителями классов комментариев, в качестве которых выступают сводные рефераты классов.

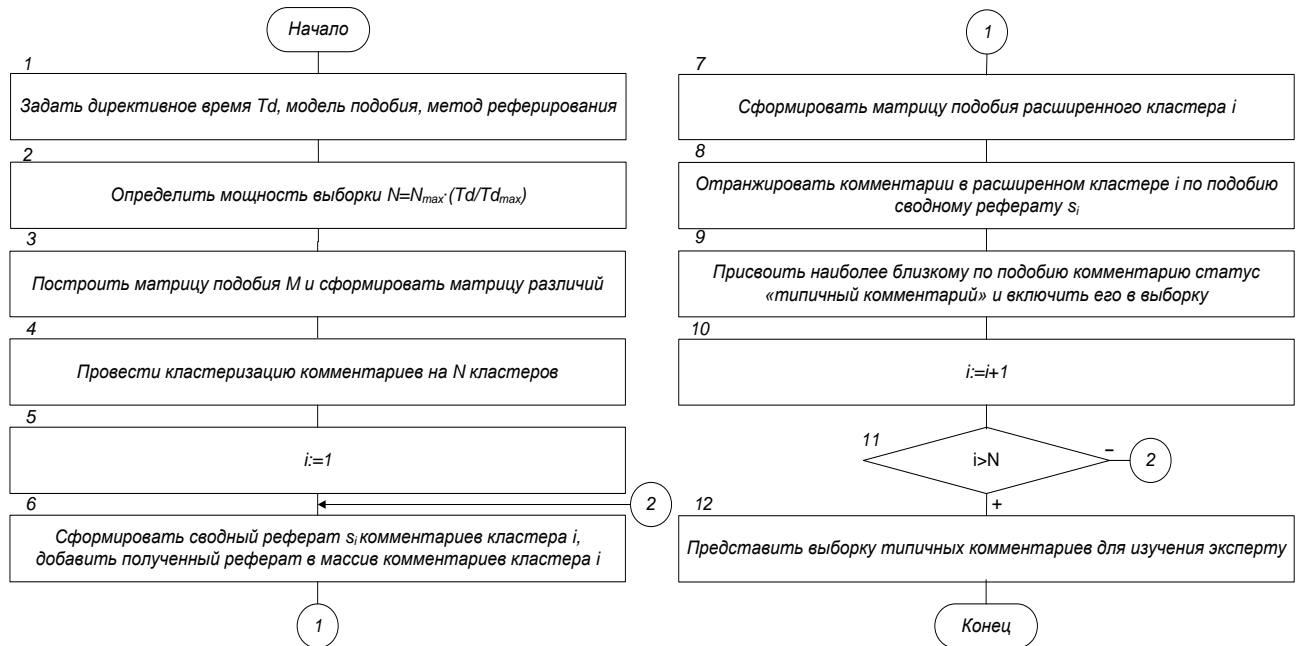


Рисунок 6 – Алгоритм формирования выборки типичных комментариев

В ходе экспериментов было установлено, что сводные рефераты обладают наилучшим средним подобием в классе комментариев (табл. 4). Данное свойство было положено в основу алгоритма определения типичных комментариев.

Таблица 4 – Свойство наилучшего среднего подобия сводных рефератов комментариев различных классов на примере ЗП «Об образовании» (модель подобия – марковская [6], рефераты получены методом сканирующих статистик)

Текст (класс «34-ГП-Н-Ми»)	Ср. коэф. подобия	Текст (класс «11-ОД-Н-Ми»)	Ср. коэф. подобия	Текст (класс «1-ОД-Н-Ми»)	Ср. коэф. подобия
edu 34-3.txt	0,626469	edu 11-4.txt	0,163084	edu 1-3.txt	0,324475
edu 34-4.txt	0,62749	edu 11-1.txt	0,163153	edu 1-1.txt	0,326142
edu 34-1.txt	0,629656	edu 11-2.txt	0,163153	edu 1-2.txt	0,3335
4166	0,986917	2343	0,226228	7732	0,600645
520	1,071068	2823	0,289161	5520	0,658391
458	1,078332	2346	0,306828	6276	0,667973
..

На шаге 1 алгоритма необходимо определить бюджет времени, выбрать модель подобия и метод реферирования.

На шаге 2, исходя из заданного бюджета времени, определяется мощность выборки типичных комментариев.

На шаге 3 для выбранного класса комментариев строится матрица подобия, на основе которой в зависимости от используемой модели подобия формируется матрица различий (табл. 5).

Таблица 5 – Фрагмент матрицы подобия (модель подобия – марковская, класс «7-42-Н-В-Ми» ЗП «О полиции»)

Идент. комм.	19074	359	1658	1779	2715	2839	3875	...
19074	0	2,7175	3,0272	3,0839	3,1179	2,5739	4,2059	...
359	2,6959	0	0,9163	1,0651	0,3536	0,156	0,8233	...
1658	2,9231	0,9522	0	1,1987	1,3682	0,1889	2,7109	...
1779	2,8662	1,0529	1,206	0	0,2912	0,926	2,8638	...
2715	3,0401	0,3536	1,384	0,2912	0	1,1362	3,1515	...
2839	2,6374	0,156	0,1889	0,901	1,1147	0	0,4452	...
3875	4,1056	0,8233	2,5657	2,6757	2,8837	0,4452	0	...
...

На шаге 4 к полученной матрице различий применяем метод иерархического объединения для кластеризации комментариев на требуемое количество классов в соответствии с заданным бюджетом времени (рис. 7).

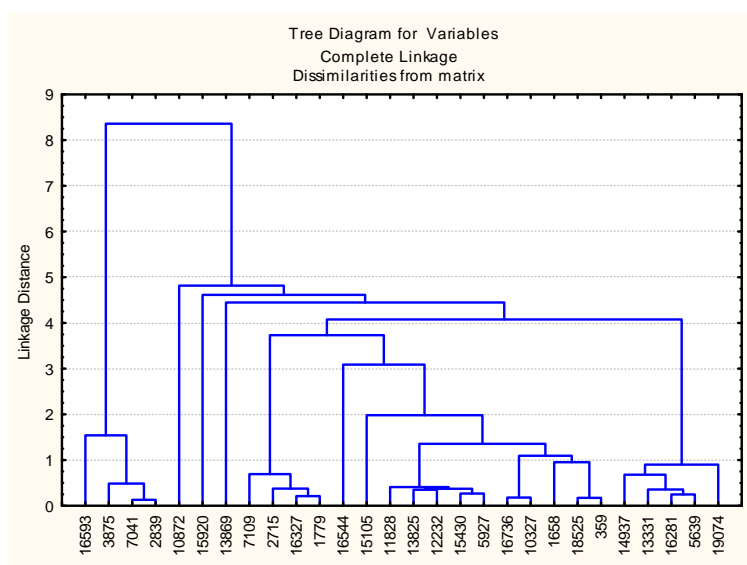


Рисунок 7 – Результат кластеризации комментариев (метод кластеризации – иерархическое объединение, класс «7-42-Н-В-МИ» ЗП «О полиции»)

На шаге 6 в полученных кластерах строим сводные рефераты комментариев и включаем их в массив соответствующего кластера.

На шаге 7 формируем матрицы подобия в расширенных полученных сводными рефератами кластерах.

На шаге 8 определяем комментарии, обладающие наилучшим сходством со сводным рефератом своего кластера (табл. 6).

Таблица 6 – Выбор типичного комментария в 15 кластере (модель подобия – марковская, класс «7-42-Н-В-Ми» ЗП «О полиции»)

Текст	5639.txt	13331.txt	14937.txt	16281.txt	19074.txt	all15ref3.txt
all15ref3.txt	0,2838	0,0591	2,7037	0,0834	3,1937	0
13331.txt	0,2624	0	0,6804	0,3556	0,8036	0,0591
16281.txt	0,2474	0,3556	0,674	0	0,8227	0,0834

5639.txt	0	0,2624	0,6736	0,2474	0,8242	0,2838
14937.txt	0,6736	0,6804	0	0,674	0,8996	2,5486

Шаги 6-10 повторяем для каждого кластера, полученного на шаге 4.

На шаге 12 совокупность полученных комментариев составляет выборку типичных комментариев выбираемого аналитиком класса, которые подлежат изучению в порядке включения в выборку при увеличении числа кластеров (табл. 7).

Таблица 7 – Формирование выборки типичных комментариев при увеличении числа кластеров (модель подобию – марковская, класс «7-42-Н-В-Ми» ЗП «О полиции»)

К-во комм.	Номер кластера																											Поряд. включ.
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	
1	2839	359	10327	1779	5639	5927	12232	5639	5927	1779	5927	2839	5639	1779	19074	359	359	359	2839	359	359	359	19074	19074	19074	19074	19074	15
2	7041	18525	16736	16327	16281	15430	13825	16281	15430	16327	15430	7041	16281	16327	5639	18525	18525	18525	7041	18525	18525	18525	5639	5639	5639	5639	5639	24
3																												23
4																												6
5																												17
6																												27
7																												14
8																												12
9																												26
10																												13
11																												21
12																												7
13																												19
14																												22
15																												11
16																												9
17																												8
18																												2
19																												25
20																												20
21																												16
22																												5
23																												4
24																												3
25																												1
26																												28
27																												18

ОЦЕНКА ЭФФЕКТИВНОСТИ АЛГОРИТМОВ

Для оценивания эффективности алгоритмов были исследованы статьи законопроектов «О полиции» и «Об образовании». В ходе исследования было установлено, что статьи законопроектов содержат различное количество пунктов и подпунктов, что обуславливает большее либо меньшее разнообразие высказываний. При этом распределение статей по количеству объектов комментирования подчинено экспоненциальному закону.

Для экспериментов были отобраны классы комментариев к статьям, содержащим наибольшее, наименьшее и среднее количество объектов комментирования, т.е. по три класса для двух законопроектов. В данных классах был произведен экспертный анализ комментариев, выделены высказывания (табл. 8) и составлены распределения высказываний по комментариям класса (табл. 9).

Таблица 8 – Фрагмент таблицы с высказываниями класса «34-Науч.-Низ.-Минор.» к ЗП «О полиции»

m_v	Содержание высказывания $sent_{v,m_v}$, $v = "6-34-Н-Н-Ми"$
1	Исключить привлечение к оперативно-розыскной деятельности сотрудников внутренней службы, юстиции и стажёров.
2	Конкретизировать требования (статус, права и обязанности, уровень образования, степень подготовки, экипировка и пр.) к привлекаемым к охране общественного порядка и пр. задачам иных сотрудников органов внутренних дел (ОВД) и стажёров
3	Включить в перечень оснований привлечения сотрудников ОВД объявление режима ЧС, природные и техногенные катастрофы, ликвидацию последствий стихийных действий, проведение контртеррористических операций, проведение массовых мероприятий.

4	Уточнить категории сотрудников ОВД, на которые распространяется действие статьи
5	Предоставить гражданину право удостовериться, что привлечённый сотрудник уполномочен совершать в отношении него действия
...	...

Таблица 9 – Пример распределения высказываний в классе «34-Науч.-Низ.-Минор.» к ЗП «О полиции»

Ид. комм.	Ид. высказывания							
	1	2	3	4	5	6	7	8
2393	+	+						
2505		+	+					
5601				+				
6205		+			+			
6239		+				+		
14322		+						
15793		+						
15907							+	+
17329				+				

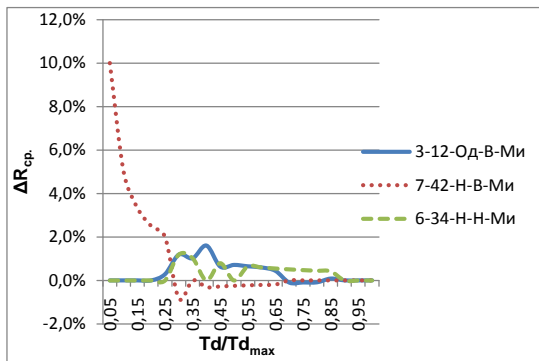
После этого в соответствии с алгоритмами были определены множества комментариев, предоставляемые для изучения в соответствии с выделенным бюджетом времени с шагом $0,05 \cdot T_{d_{max}}$. Полученные множества пошагово сравнивались с множествами не отранжированных по полезности и типичности комментариев при одинаковом бюджете времени по четырем видам полноты (табл. 10).

Таблица 10 – Виды полноты отбора высказываний

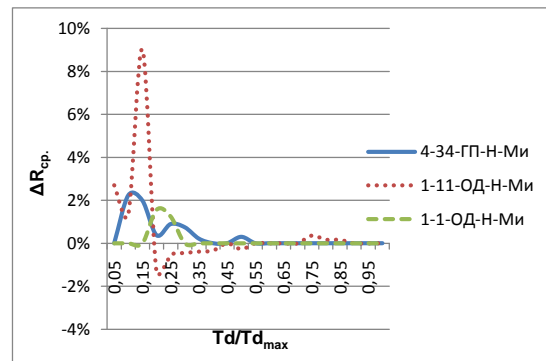
По учету веса высказываний	По идентификации повторов высказываний	
	Без идентификации повторов высказываний	С идентификацией повторов высказываний
Без учета веса высказываний	R1	R3
С учетом веса высказываний	R2	R4

В ходе экспериментов было установлено, что средний прирост полноты зависит от способа её расчета, в зависимости от учета веса высказываний и идентификации повторов высказываний. Оцениваемый алгоритм формирования выборки полезных комментариев показал наилучшие результаты при оценке по среднему приросту полноты вида R1 (рис. 8), а алгоритм формирования выборки типичных комментариев – при оценке по среднему приросту полноты вида R4 (рис. 9).

Эксперименты показали, что для различных законопроектов и статей с различным количеством объектов комментирования предложенные алгоритмы могут обеспечить опережающий темп изучения высказываний аналитиком лишь в первую половину от максимального бюджета времени, что связано с тем, что значительная часть комментариев содержит по одному высказыванию.



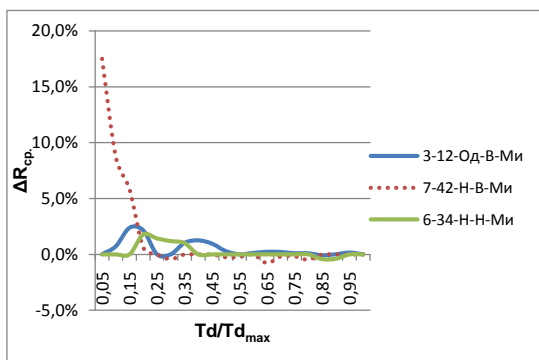
а)



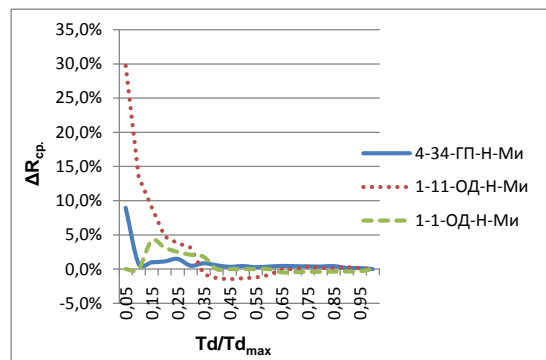
б)

Рисунок 8 – Абсолютный выигрыш в среднем приросте полноты вида R1 при $dt=0,5 \cdot Td_{max}$ для выборки полезных комментариев к трем классам с различным числом объектов комментирования

а) ЗП «О полиции» б) ЗП «Об образовании»



а)



б)

Рисунок 9 – Абсолютный выигрыш в среднем приросте полноты вида R4 при $dt=0,5 \cdot Td_{max}$ для выборки типичных комментариев к трем классам с различным числом объектов комментирования

а) ЗП «О полиции» б) ЗП «Об образовании»

ВЫВОДЫ

Моделирование процессов аналитической обработки мнений, представленных в текстовом виде, требует учета таких особенностей мнений, как многомерность их теоретико-множественного представления, неравномерность распределения высказываний, специфика цели выражения мнения.

При обработке комментариев, полученных в ходе обсуждения законопроектов, модели обработки мнений должны предоставлять возможности аналитику по оперативному ознакомлению с типичными и полезными комментариями, содержащимися в массиве.

Предложенные алгоритмы формирования выборок типичных и полезных комментариев обеспечивают операции детализации при применении к массиву комментариев модели оперативной аналитической обработки текстов и позволяют повысить эффективность изучения высказываний в выбираемых аналитиком классах.

СПИСОК ЛИТЕРАТУРЫ

1. Общественное обсуждение законопроектов [Электронный ресурс]. – URL: <http://www.zakonoproekt2012.ru/> (дата обращения: 12.01.2012).

2. Единый портал раскрытия информации о подготовке федеральными органами исполнительной власти проектов нормативных правовых актов и результатах их общественного обсуждения [Электронный ресурс]. – URL: <http://regulation.gov.ru/> (дата обращения: 17.07.2013).
3. Pang B., Lee L. Opinion Mining and Sentiment Analysis // Foundations and Trends in Information Retrieval. – V. 2. – № 1-2, January, 2008. – P.1-135.
4. Method for Performing Effective Drill-down Operations in Text Corpus Visualization and Exploration Using Language Model Approaches for Key Phrase Weighting: заявка 2008/0243482 A1 США: МПК G06F 17/27 / Michal Skubacz, Cai-Nicolas Ziegler. – № 11/797632; заявл. 04.05.2007; опублик. 02.10.2008. – 13 с.
5. Text Cube: Computing IR Measures for Multidimensional Text Database Analysis. Cindy Xide Lin, Bolin Ding, Jiawei Han и др. [Электронный ресурс]. – URL: www.cs.uiuc.edu/~hanj/pdf/icdm08_xlin.pdf (дата обращения: 22.12.2010).
6. Бородащенко А.Ю., Яковлев В.А. Алгоритм фильтрации текстовой информации на основе Марковской модели // Информационные технологии, 2011. – № 5. – С. 2-5.

Толкунов Александр Александрович
Академия ФСО России, г. Орел
Адъюнкт
E-mail: altifact@mail.ru

A.A. TOLKUNOV (*Adjunct*)
Academy of Federal Agency of Protection of the Russian Federation, Orel

ALGORITHMS FOR GENERATING TYPICAL AND UTILITY COMMENTS SAMPLES DURING THE PROCESSING OF THE RESULTS OF PUBLIC DISCUSSION OF BILL

The article describes the approaches to solving the problem of increasing the operability of revealing of statements when processing an array of Internet users comments received as a result of public discussion of bill. These approaches include the algorithms of sampling comments at a given time budget, taking into account the uneven distribution of statements in the comments. The article also includes the results of experiments showing the effectiveness of the proposed algorithms in comparison with the existing approach to the treatment of an array of comments.

Keywords: *public discussion of bills; an array of comments; opinion; statement, a model for online analytical processing of text; drill-down; typical comments sample; utility comments sample; statements revealing recall.*

BIBLIOGRAPHY (TRANSLITERATED)

1. Obshhestvennoe obsuzhdenie zakonproektov [E'lektronny'j resurs]. – URL: <http://www.zakonoproekt2012.ru/> (data obrashheniya: 12.01.2012).
2. Ediny'j portal raskry'tiya informacii o podgotovke federal'ny'mi organami ispolnitel'noj vlasti proektov normativny'x pravovy'x aktov i rezul'tatax ix obshhestvennogo obsuzhdeniya [E'lektronny'j resurs]. – URL: <http://regulation.gov.ru/> (data obrashheniya: 17.07.2013).
3. Pang B., Lee L. Opinion Mining and Sentiment Analysis // Foundations and Trends in Information Retrieval. – V. 2. – № 1-2, January, 2008. – P.1-135.
4. Method for Performing Effective Drill-down Operations in Text Corpus Visualization and Exploration Using Language Model Approaches for Key Phrase Weighting: zayavka 2008/0243482 A1 SShA: MPK G06F 17/27 / Michal Skubacz, Cai-Nicolas Ziegler. – № 11/797632; zayavl. 04.05.2007; opubl. 02.10.2008. – 13 s.
5. Text Cube: Computing IR Measures for Multidimensional Text Database Analysis. Cindy Xide Lin, Bolin Ding, Jiawei Han и др. [E'lektronny'j resurs]. – URL: www.cs.uiuc.edu/~hanj/pdf/icdm08_xlin.pdf (data obrashheniya: 22.12.2010).
6. Brodashhenko A.Yu., Yakovlev V.A. Algoritm fil'tracii tekstovoj informacii na osnove Markovskoj modeli // Informacionny'e tehnologii, 2011. – № 5. – S. 2-5.

УДК 004.73

А.В. ЕРЕМЕНКО, А.Н. ОСИПОВ, Д.С. МИШИН,
Н.И. ПЕНЬКОВ, А.Г. ПОЛЯКОВА, С.Н. ИВАНОВА

ПРОБЛЕМЫ ОБМЕНА ДАННЫМИ В ИНТЕГРИРОВАННОЙ СРЕДЕ АСУ ПРОМЫШЛЕННЫХ ПРЕДПРИЯТИЙ

Использование мировых стандартов TCP/IP, JAVA, OPC и DNA в изделиях является инструментом для построения новой стратегии глобальной инфраструктуры предприятия и свидетельствует также о новой определяющей роли информационных технологий в области промышленной автоматизации.

***Ключевые слова:** стратегия Transparent Factory; инфраструктура предприятия; промышленная автоматизация.*

Развитие сети Internet привело к серьезному изменению области автоматизации производства. Несмотря на распространение Internet во всех областях жизни и многочисленные заявления компаний об использовании web-технологий в своих изделиях промышленной автоматизации, наибольший интерес представляет глобальная стратегия Transparent Factory / Open for Business (Прозрачное производство / Открытость для бизнеса), объединяющая ряд продуктов, подчиняющихся единой философии открытости, прозрачности и универсальности. Эта концепция означает новый этап в развитии фирм, характеризующийся отходом от своих сетевых технологий и принятием распространенных мировых стандартов, таких, как TCP/IP, JAVA, OPC и DNA [1].

Большинство предприятий для автоматизации производства используют в качестве базового протокола своих сетей и полевых шин протокол Ethernet TCP/IP, а в качестве основного средства доступа к информации – web-технологии и интерфейс OPC.

Концепция TF (Transparent Factory) способна решить многие проблемы внутри межсетевых обмена на уровнях АСУТП и АСУП. Рассмотрим основные трудности, наиболее часто встречающиеся в области промышленной автоматизации [2].

ПРОБЛЕМА 1. МНОГООБРАЗИЕ СЕТЕЙ И ПОЛЕВЫХ ШИН

Если протокол Ethernet TCP/IP практически утвердился как базовый для сети на уровне предприятия (АСУП), то на уровне АСУТП существует более десятка сетей и полевых шин, конкурирующих за право считаться основной. Протоколы Profibus, Interbus-S, Modbus, WordFip, DeviceNet, Fieldbus Foundation и другие являются широко распространенными и хотя обмен между ними теоретически возможен, практически это требует от пользователя больших дополнительных расходов по их стыковке и дальнейшему обслуживанию. Каждый из этих протоколов активно лоббируется своим производителем, что не способствует процессу какой-либо унификации [3, 7].

Попытки международных организаций, в частности, МЭК, принять единый стандарт полевой шины (в частности, IEC-61158) пока не увенчались успехом. Вместе с тем, развитие сети Ethernet и ее нынешние характеристики практически позволяют использовать эту сеть в качестве локальной промышленной сети. Повышение пропускной способности сети, связанное с новыми возможностями сетевых коммутаторов и маршрутизаторов, повышение скорости обмена информацией, возможность сегментации сетей – все это привело к тому, что протокол Ethernet действительно может быть открытым и эффективным решением не только для систем автоматизации офисной деятельности, но и для организации промышленных сетей [4, 8].

С решением проблемы обеспечения необходимого детерминизма Ethernet данная технология может обеспечить превосходство над другими промышленными ЛВС (LAN) и промышленными сетями (fieldbus) по ряду параметров.

ПРОБЛЕМА 2. МЕЖСЕТЕВОЙ ОБМЕН В МАСШТАБАХ ПРЕДПРИЯТИЯ

Речь идет об обмене данными между сетями уровней 1, 2 и 3, то есть о том, как сделать прозрачным и простым доступ к данным устройств низовой автоматики (уровень 1, полевая шина), контроллерам и системам SCADA (уровень 2) и, соответственно, офисным приложениям (уровень 3). Так как требования к сетевым характеристикам этих уровней разные (различные конструктивные параметры сетей, способы адресации, проверки ошибок, различные конфигурация и диагностика), всегда существовала проблема организации их взаимодействия. Типичным решением данной проблемы является организация сбора данных на серверах различного уровня с последующей передачей их в другую сеть через шлюзы. Однако такой подход имеет ряд серьезных недостатков. При конфигурировании шлюзов существует вероятность потери информации между различными промышленными шинами, а установка и последующее обслуживание устройств и программного обеспечения связаны с дополнительными, зачастую огромными расходами [5, 9].

Необходимо отметить, что данные хранятся в различных базах данных серверов (с разными форматами, способами доступа, инструментами и т.д.). Растущее количество клиентов на уровне 3 (АСУП) и их требования к СУБД и хранилищам данных вызывают, в свою очередь, повышенные требования к пропускной способности сетей каждого уровня и в целом межсетевого обмена. Таким образом, в современной инфраструктуре именно стыки между сетями разных уровней вызывают перегрузки всей информационной системы. Эти узкие места, своеобразные пробки должны быть удалены при создании единой и прозрачной информационной инфраструктуры всего предприятия [6].

В настоящее время доступ к информации в пределах одного предприятия (сети предприятий одной компании) производится, как правило, через корпоративную сеть Intranet. Если раньше эта сеть использовалась, в основном, для доступа к данным АСУП, то сегодня речь идет об использовании Intranet для доступа к производственным данным и процессам. В качестве сетевой среды для промышленных сетей Intranet рынок диктует применение Ethernet, в качестве информационной среды все чаще используются web-технологии [10-12].

ПРОБЛЕМА 3. НЕОБХОДИМОСТЬ ПОЛУЧЕНИЯ ДАННЫХ С ПРОИЗВОДСТВА В РЕАЛЬНОМ ВРЕМЕНИ

В большинстве приложений для сбора производственных данных на уровне АСУП используется пакетный способ обработки, когда данные передаются в конце смены или в другое незагруженное время дня. Другие архивные данные для обработки или анализа могут передаваться на уровень АСУП из системы SCADA. Между тем, просмотр или обработка архивных данных уже не устраивают управленцев на предприятии. Для того, чтобы гибко и эффективно управлять производством, им нужны текущие значения выхода готовой продукции, параметров производственных процессов, состояния оборудования и т.п., то есть данные реального времени. Такие данные могут быть извлечены из SCADA-сервера отдельного производственного участка [7].

Проблема состоит в том, что в таком случае сервер SCADA является лишь промежуточным звеном между АСУП и контроллерами, которые содержат данные реального времени с производства. Как любое дополнительное звено в системе, сам сервер SCADA может быть источником ошибок или сбоев при передаче данных на вышестоящий уровень. Если добавить к этому стоимость такой системы, необходимость шлюзов для перехода в другую сеть обслуживания, то возникает идея устранить это промежуточное звено там, где это возможно. Приемлемым решением в среде Internet/Intranet является получение данных со встроенных web-серверов. Любое устройство, имеющее встроенный web-сервер, может не

только поставлять данные по запросу, но и само инициировать передачу сообщений (e-mail). Именно такое решение было предложено компанией Schneider Electric: web-сервер, встроенный в промышленный программируемый контроллер, может обеспечить данными реального времени любого клиента в пределах сети Intranet [10].

В этом случае частично или полностью отпадает необходимость в SCADA-системе и непосредственные данные могут быть получены на различных уровнях АСУП без дополнительных издержек. При этом не имеет значения, на какой компьютер, с какой операционной системой передаются данные. Web-технология предлагает действительную универсальность и прозрачность. Безусловно, система SCADA/HMI (Human Machine Interface – человеко-машинный интерфейс) на уровне АСУТП выполняет ряд других функций и там, где это нужно, без нее не обойтись.

Концепция TF компании Schneider Electric (рис. 1) базируется на трех основных положениях [1, 2]:

1. Следующим поколением сетей, используемых для автоматизации промышленности, интегрированным с сетями офисной деятельности, становится Ethernet.

2. Прозрачный и легкий доступ к получению данных модульных структур их сбора и обработки от датчиков/исполнительных механизмов и до уровней планирования и управления предприятием возможен при наличии стандартной, интегрированной сети, обеспечивающей множество сервисных уровней. Такой сетью становится корпоративная сеть предприятия Intranet, построенная по принципу клиент/сервер и обеспечивающая создание единого информационного пространства.

3. Будущие системы управления будут использовать стандартные и открытые, объектно-ориентированные средства управления и доступа к информации. В качестве таких средств предлагаются встроенные web-серверы и интерфейсы автоматизации технологических процессов.

Ключевыми положениями концепции Transparent Factory являются прозрачность и открытость. Для пользователя выгода заключается в том, что они могут выбирать лучшие устройства различных производителей и сохранять уверенность в том, что эти устройства способны взаимодействовать между собой. Это обязательство гарантирует, что потребители не будут привязаны к какому-либо производителю, а будут применять новейшие продукты по мере их возникновения. Открытость означает готовность к тому, что в случае, если продукты и услуги одной фирмы перестанут удовлетворять требованиям потребителя, их могут заменить продукты и услуги других поставщиков.

Прозрачность компонентов Transparent Factory (TF) обеспечивает возможность работы в соответствии с принципом plug&play в максимально возможном на сегодняшний день объеме. При использовании встроенных web-серверов неважно, где и как сервер подключен к сети. Так же, как мы набираем номер телефона, здесь достаточно ввести адрес web-сервера и получить полную информацию.

С использованием концепции Transparent Factory впервые в АСУ ТП становится возможным построение сложной системы с непосредственным прямым доступом на все уровни к каждому устройству (интеллектуальному датчику, частотно-управляемому приводу, модулю ввода/вывода и т.д.). В этом и заключается основное отличие от традиционного иерархического дерева, в котором непосредственный контакт с узлами нижнего уровня имел только один узел верхнего уровня – управляющий контроллер, мост или супервизорное программное обеспечение на РС (рис. 2).

Повышение надежности и управляемости с применением элементов концепции TF достигается [1, 2, 4, 5]:

– в случае отказа узла сети или сегмента линии связи в традиционной схеме становятся недоступными все подключенные к нему устройства. С использованием TF-компонентов они остаются достигаемыми как минимум для удаленной диагностики происшедшего;

– процедура полной инсталляции супервизора на новый компьютер занимает не один час, даже при наличии свежей резервной копии проекта, а запуск браузера с любого компьютера корпоративной сети – несколько минут;

– разработчик ПО для контроллеров будет испытывать серьезные трудности в отладке собственной программы, если он оторван от своего рабочего места и инструмента. Использование универсальных механизмов доступа к компонентам системы не заменит пакета программирования, но доступно всегда и везде. Снижение стоимости на этапе инсталляции системы из TF-компонентов возможно за счет разгрузки домена технологических серверов, применения аппаратных платформ меньшей стоимости, уменьшения объема или полного отказа от закупок специализированного ПО и коммерческих ОС;

– применение массовых компонентов Ethernet вместо дорогостоящих каналообразующих элементов промышленных сетей, если нет агрессивных внешних воздействий;

– широкое применение существующей инфраструктуры каналов связи, серверов и программного обеспечения, в том числе сторонних фирм.

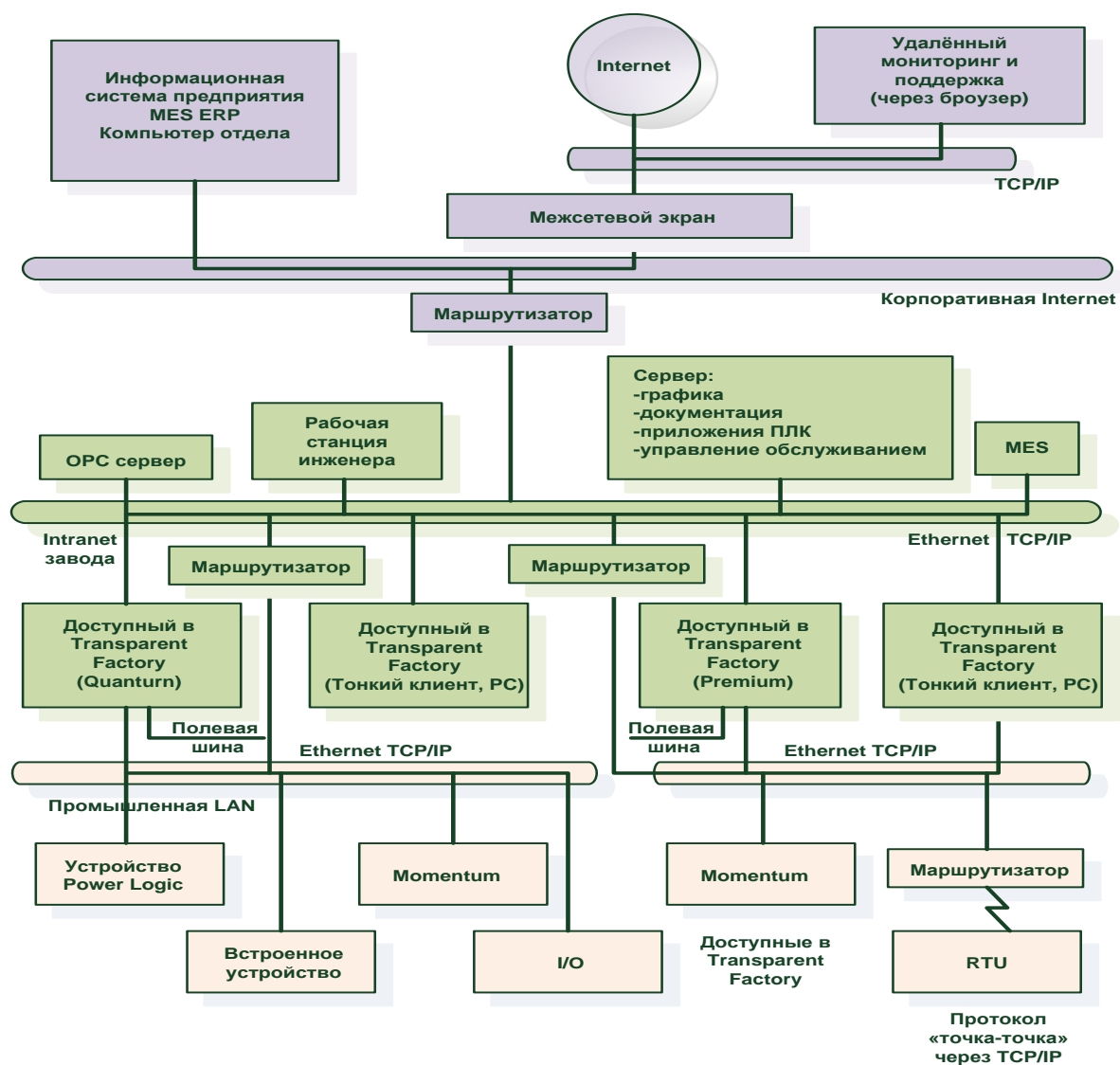


Рисунок 1– Internet-технологии в автоматизации производственных процессов

ПЕРСПЕКТИВЫ ИСПОЛЬЗОВАНИЯ

Многие отечественные пользователи заинтересованы во внедрении данной технологии. Укажем некоторые категории пользователей, проявляющих интерес к подобным TF-решениям. Это [1, 8, 9]:

- отделы АСУТП и АСУП крупных предприятий в металлургии, транспортировке газа, нефтепереработке, химии и нефтехимии, пищевой промышленности. Внедрение концепции TF открывает доступ к производственным данным для самых разных специалистов таких предприятий;

- системные интеграторы, реализующие свои системы на значительном удалении от своих офисов. Учитывая все накладные расходы, вопросы дальнейшего обслуживания и диагностики систем могут быть эффективно решены с помощью TF. В эту же категорию можно добавить крупные машиностроительные заводы, экспортирующие свое оборудование в другие страны;

- предприятия, нуждающиеся в мониторинге объектов, разбросанных на значительные расстояния. Концепция TF может эффективно решить проблемы создания АСУД облэнерго, крупных городских радиоэлектронных средств и других объектов коммунального хозяйства;

- интересной является идея использования web-серверов для всех случаев, где необходим мониторинг производственных данных для многих пользователей. Эта технология может заменить установку дорогостоящих пакетов SCADA или им подобных.

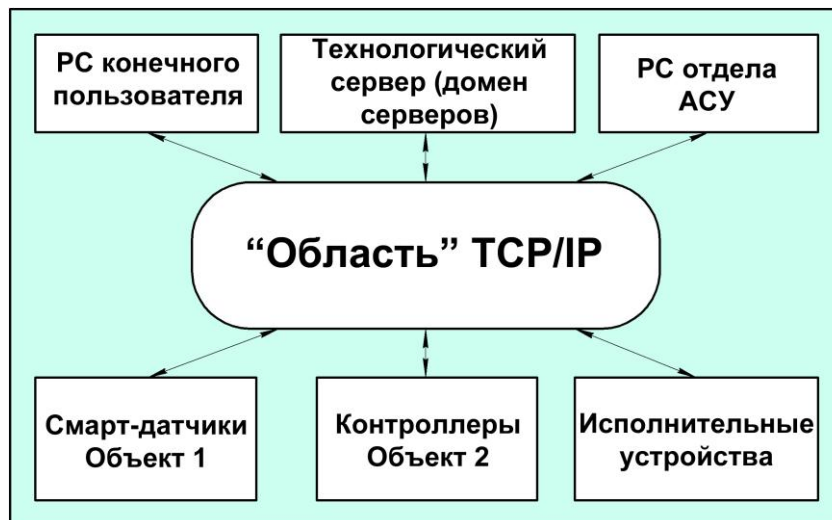


Рисунок 2– Архитектура сбора и обработки данных в концепции Transparent Factory

ЗАКЛЮЧЕНИЕ

Использование технологии Ethernet на уровне единой коммуникационной инфраструктуры предприятия позволяет собирать и обрабатывать информацию с различных уровней для применения в различных приложениях, включать в контур управления различные устройства контроля, обеспечивать централизованное управление устройствами автоматизации технологических процессов и производств, унифицировать оборудование, снизить стоимость эксплуатационного обслуживания и обучения персонала.

Разработка совместимых технологий сбора и обработки данных в настоящее время осуществляется на основе Единого стандарта обмена данными и межсерверных коммуникаций для Industrial Ethernet и унаследованных промышленных протоколов OPC Data Exchange Standard for Ethernet, что позволяет получить единые спецификации на сетевые компоненты и кабельную инфраструктуру.

СПИСОК ЛИТЕРАТУРЫ

1. Тютюнник М.Н. Концепция Transparent Factory: web-технология в автоматизации производства [Электронный ресурс]. – URL: <http://www.mka.ru/?p=40134>.
2. Тютюнник М.Н. Transparent Factory. Первые шаги [Электронный ресурс]. – URL: <http://asutp.ru/?p=600099>
3. Еременко В.Т. Оптимизация ресурсов и управление процессами информационного обмена в сетях АСУТП на основе полевых шин / В.Т. Еременко, С.И. Афонин, С.А. Максаков, А.И. Куленич // Вестник компьютерных и информационных технологий, 2011. – № 9. – С. 46-49.
4. Еременко В.Т. Моделирование информационных потоков в сетях передачи данных интегрированных АСУ / В.Т. Еременко, С.И. Афонин, Т.М. Парамохина, Л.В. Кузьмина, Д.А. Плащенков // Информационные системы и технологии, 2011. – № 6. – С. 35-42.
5. Еременко В.Т., Тютякин А.В. Методологические аспекты выбора профилей сбора и обработки данных в системах неразрушающего контроля и диагностики технических объектов // Контроль. Диагностика, 2013. – № 1. – С. 24-31.
6. Еременко В.Т. Метод проектирования сетей передачи данных, совместимых с неблокируемой маршрутизацией / В.Т. Еременко, А.И. Офицеров, С.А. Черепков // Вестник компьютерных и информационных технологий, 2012. – № 4. – С. 38-46.
7. Проблемы автоматизации управления предприятием // ВПТЕ-Россия, 2009. – № 9 [Электронный ресурс]. – URL: <http://www.bytemag.ru/articles/detail.php?ID=6822>.
8. Орлов С.А. Ethernet в системах промышленной автоматизации [Электронный ресурс]. – URL: <http://www.osp.ru/lan/2002/06/136286/>.
9. Точилин Н.В. Применение Internet- и Intranet-технологий для обмена электронными данными // Научно-экономический сборник «Газовая промышленность». – М.: ООО «ИРЦ Газпром», 2000. – № 6. – С. 22-25.
10. Турчин С. Общие проблемы автоматизации управления предприятием [Электронный ресурс]. – URL: http://www.iteam.ru/publications/it/section_53/article_1371/.
11. Парк Дж., Маккей С. Сбор данных в системах контроля и управления. Практическое руководство. – М.: ООО «Группа ИТД», 2006. – 504 с.
12. Парк Дж. Передача данных в системах контроля и управления. Практическое руководство / Дж. Парк, С. Маккей, Э. Райт. – М.: ООО «Группа ИТД», 2007. – 480 с.

Еременко Алексей Владимирович

ФГБОУ ВПО «Госуниверситет – УНПК», г. Орел

Аспирант кафедры «Электроника, вычислительная техника и информационная безопасность»

Тел.: 8 920 812 89 90

E-mail: alex_er@bk.ru

Осипов Алексей Николаевич

Академия ФСО России, г. Орел

Преподаватель

E-mail: osipov2008@inbox.ru

Мишин Дмитрий Станиславович

ФГКОУ ВПО «Орловский юридический институт МВД России им. В.В. Лукьянова», г. Орел

Кандидат юридических наук, старший преподаватель кафедры «Информационные технологии в деятельности внутренних дел»

Тел.: 8 903 880 23 45

E-mail: mishinds@mail.ru

Пеньков Николай Геннадьевич

ФГБОУ ВПО «Госуниверситет – УНПК», г. Орел

Аспирант кафедры «Электроника, вычислительная техника и информационная безопасность»

Тел.: 8 906 664 24 42

E-mail: corsa1r_89@mail.ru

Полякова Алёна Геннадьевна

ФГБОУ ВПО «Госуниверситет – УНПК», г. Орел

Студент кафедры «Электроника, вычислительная техника и информационная безопасность»
Тел.: 8 953 611 60 91
E-mail: studentkalenka@mail.ru

Иванова Светлана Николаевна

ФГБОУ ВПО «Госуниверситет – УНПК», г. Орел
Магистр кафедры «Электроника, вычислительная техника и информационная безопасность»
Тел.: 8 953 413 83 10
E-mail: isn140181@mail.ru

A.V. ERYOMENKO (*Post-graduate Student of the Department «Electronics, Computer Facilities and Information Security»
State University – ESPC, Orel*)

A.N. OSIPOV (*Teacher*)
Academy of Federal Agency of Protection of the Russian Federation, Orel

D.S. MISHIN (*Candidate of Juridical Sciences, Senior Teacher of the Department «Information Technologies in Working of Internal Affairs»
Law Institute of the Russian Interior Ministry named V.V. Luk'yanov, Orel*)

N.I. PEN'KOV (*Post-graduate Student of the Department «Electronics, Computer Facilities and Information Security»*)

A.G. POLYAKOVA (*Student of the Department «Electronics, Computer Facilities and Information Security»*)

S.N. IVANOVA (*Undergraduate of the Department «Electronics, Computer Facilities and Information Security»
State University – ESPC, Orel*)

PROBLEMS OF DATA EXCHANGE IN THE INTEGRATED ENVIRONMENT OF THE AUTOMATED CONTROL SYSTEMS OF THE INDUSTRIAL ENTERPRISES

Application of world standards TCP/IP, JAVA, OPC and DNA in products is the tool for construction of new strategy of a global infrastructure enterprise strategy; also it is evidence of a new, defining role of information technologies in the field of industrial automation.

Keywords: *Transparent Factory strategy; enterprise infrastructure; industrial automation.*

BIBLIOGRAPHY (TRANSLITERATED)

1. Tyutyunnik M.N. *Koncepciya Transparent Factory: web-texnologiya v avtomatizacii proizvodstva* [E'lektronny'j resurs]. – URL: <http://www.mka.ru/?p=40134>.
2. Tyutyunnik M.N. *Transparent Factory. Pervy'e shagi* [E'lektronny'j resurs]. – URL: <http://asutp.ru/?p=600099>
3. Eryomenko V.T. *Optimizaciya resursov i upravlenie processami informacionnogo obmena v setyax ASUTP na osnove polevy'x shin* / V.T. Eryomenko, S.I. Afonin, S.A. Maksakov, A.I. Kulenich // *Vestnik komp'yuterny'x i informacionny'x texnologij*, 2011. – № 9. – S. 46-49.
4. Eryomenko V.T. *Modelirovanie informacionny'x potokov v setyax peredachi danny'x integrirovanny'x ASU* / V.T. Eryomenko, S.I. Afonin, T.M. Paramoxina, L.V. Kuz'mina, D.A. Plashhenkov // *Informacionny'e sistemy' i texnologii*, 2011. – № 6. – S. 35-42.
5. Eryomenko V.T., Tyutyakin A.V. *Metodolgicheskie aspekty' vy'bora profilej sbora i obrabotki danny'x v sistemax nerazrushayushhego kontrolya i diagnostiki texnicheskix ob'ektov* // *Kontrol'. Diagnostika*, 2013. – № 1. – S. 24-31.
6. Eryomenko V.T. *Metod proektirovaniya setej peredachi danny'x, sovmestimy'x s neblokiruemoj marshrutizaciej* / V.T. Eryomenko, A.I. Oficerov, S.A. Cherepkov // *Vestnik komp'yuterny'x i informacionny'x texnologij*, 2012. – № 4. – S. 38-46.

7. Problemy' avtomatizacii upravleniya predpriyatiem // BITE-Rossiya, 2009. – № 9 [E'lektronny'j resurs]. – URL: <http://www.bytemag.ru/articles/detail.php?ID=6822>.
8. Orlov S.A. Ethernet v sistemax promy'shlennoj avtomatizacii [E'lektronny'j resurs]. – URL: <http://www.osp.ru/lan/2002/06/136286/>.
9. Tochilin N.V. Primenenie Internet- i Intranet-texnologij dlya obmena e'lektronny'mi danny'mi // Nauchno-e'konomicheskij sbornik «Gazovaya prmy'shlennost'». – M.: OOO «IRC Gazprom», 2000. – № 6. – S. 22-25.
10. Turchin S. Obshhie problemy' avtomatizacii upravleniya predpriyatiem [E'lektronny'j resurs]. – URL: http://www.iteam.ru/publications/it/section_53/article_1371/.
11. Park Dzh., Makkej S. Sbor danny'x v sistemax kontrolya i upravleniya. Prakticheskoe rukovodstvo. – M.: OOO «Gruppa ITD», 2006. – 504 s.
12. Park Dzh. Peredacha danny'x v sistemax kontrolya i upravleniya. Prakticheskoe rukovodstvo / Dzh. Park, S. Makkej, E'. Rajt. – M.: OOO «Gruppa ITD», 2007. – 480 s.

МАТЕМАТИЧЕСКОЕ И ПРОГРАММНОЕ ОБЕСПЕЧЕНИЕ
ВЫЧИСЛИТЕЛЬНОЙ ТЕХНИКИ И АВТОМАТИЗИРОВАННЫХ СИСТЕМ

УДК 004.05

Е.И. НОВИКОВ

МЕТОДИКА ОЦЕНКИ ОПЕРАТИВНОСТИ
АВТОМАТИЗИРОВАННЫХ ИНФОРМАЦИОННЫХ СИСТЕМ

Предложена методика экспериментальной оценки эффективности автоматизированных систем по свойству оперативности выполнения автоматизированных функций. Особенностью методики является учет случайных факторов, влияющих на корректную оценку оперативности и использование усеченного нормального распределения времени выполнения автоматизированных функций. Представлена процедура планирования эксперимента, позволяющая обосновать достаточное количество реплик и достоверность результатов обработки и анализа экспериментальных данных.

Ключевые слова: эффективность; свойство; показатель; эксперимент; реплика; дисперсионный анализ; статистическая гипотеза; случайная величина; закон распределения; функция распределения.

ВВЕДЕНИЕ

Среди всех стадий и этапов проектирования автоматизированных информационных систем (АИС) особое место занимает этап оценки ее пригодности к использованию по назначению или превосходства над аналогами. Это обусловлено тем, что проектировщик должен объективно доказать, что характеристики предлагаемой им системы действительно соответствуют заявленным в техническом задании значениям.

Целенаправленность процесса функционирования АИС является одной из важных ее особенностей, позволяющих отнести ее к классу сложных систем. Формализация целей функционирования в виде некоторой системы критериев связана с определением всех ее значимых свойств и введением их количественной меры – системы показателей качества. Качество может быть выражено как в собственных свойствах системы, так и посредством достижения максимальной эффективности реализуемого ею целенаправленного процесса.

Оперативность является одной из наиболее объективных характеристик функционирования автоматизированных информационных систем, так как характеризует ее способность своевременно выполнять заданные функции. Исходя из того, что время выполнения функций АИС является случайной величиной, показателем оперативности является вероятность того, что время выполнения функции ($t_{вф}$) не превысит допустимого времени ($t_{вф}^{доп}$): $P(t_{вф} \leq t_{вф}^{доп})$. При этом выполнение функций следует рассматривать в широком смысле (проведение расчетов, выдача данных по запросам пользователей, формирование выходных документов и т.п.).

Предлагаемая методика оценивания эффективности АИС по свойству оперативности базируется на методах теории вероятностей и математической статистики и планирования эксперимента и включает в себя следующие этапы:

- планирование и проведение эксперимента;
- обработка и анализ экспериментальных данных;
- оценка оперативности системы.

ПЛАНИРОВАНИЕ И ПРОВЕДЕНИЕ ЭКСПЕРИМЕНТА

Организация статистического планирования эксперимента позволяет собрать необходимые данные, применить для их анализа статистические методы и сделать

правильные и объективные выводы. На этапе планирования эксперимента необходимо выполнить нижеперечисленные действия.

1. Сформулировать цели проведения эксперимента. При оценке оперативности АИС целями являются:

- определение эмпирического распределения времени выполнения функций;
- выявление влияния факторов на время выполнения функций.

2. Определить факторы, подлежащие учету в ходе эксперимента. Под фактором эксперимента понимается совокупность условий функционирования системы, влияющая на значение переменной отклика, которой является время выполнения функций АИС. В зависимости от числа учитываемых факторов различают однофакторные и многофакторные эксперименты. Факторы, исследуемые в ходе эксперимента, могут носить как количественный, так и качественный характер. В качестве факторов, влияющих на время выполнения функций АИС, могут выступать типы технического и программного обеспечений АИС, квалификация пользователей АИС и др. Для каждого фактора необходимо определить его уровни, характеризующие различные условия функционирования АИС.

3. Оценить достаточное число реплик эксперимента. Под репликацией понимается повторение основного эксперимента. Повторные опыты обладают двумя важными свойствами. Во-первых, они позволяют получить оценку ошибки эксперимента (случайной погрешности). Эта оценка является основной мерой при определении того, являются ли различия в действительности статистически различными. Во-вторых, если выборочное среднее используется для оценивания эффекта фактора в эксперименте, то повторные наблюдения позволяют последовательно получить более точную оценку этого эффекта.

Оценка достаточности числа реплик эксперимента базируется на оценке мощности дисперсионного анализа. Мощностью называется вероятность того, что количество реплик будет достаточным для выявления влияния фактора. Формально мощность можно представить в виде следующего выражения [1]:

$$1 - \beta = P(\text{отклонить } H_0 / H_0 \text{ ложная}),$$

где β – вероятность ошибки второго рода.

Для оценки мощности дисперсионного анализа используется семейство оперативных характеристик [1]. Важным аспектом применения оперативных характеристик для определения мощности дисперсионного анализа является выбор с их помощью такого числа реплик, чтобы планируемый эксперимент оказался чувствительным к важным потенциальным различиям уровней факторов. Эти характеристики представляют собой графики зависимости вероятности ошибки второго рода β от параметра Φ . Далее в статье этапы методики рассматриваются для случая однофакторного дисперсионного анализа, что не влияет на ее общность. Тогда выражение для определения параметра Φ имеет вид [1]:

$$\Phi^2 = \frac{n}{m \cdot \sigma^2} \cdot \sum_{i=1}^m \tau_i^2, \quad (1)$$

где n – искомое количество реплик; m – количество уровней фактора; σ^2 – дисперсия ошибки (обусловленная различиями внутри уровней); τ_i – отклонения среднего на i -ом уровне фактора от общего среднего.

Безусловно, целесообразно провести оценку достаточности количества реплик до проведения эксперимента. В связи с этим возникает проблема, обусловленная невозможностью рассчитать значения σ^2 и τ_i . Для решения этой проблемы необходимо

задаться величиной отношения $\frac{\sum_{i=1}^m \tau_i^2}{\sigma^2}$, описывающего различие между общей дисперсией и

отклонениями от общего математического среднего отклика. Чем больше заданное отношение, тем меньше будет заметно различие между уровнями фактора, т.е. используемый аппарат становится менее чувствительным к небольшим различиям между уровнями фактора. Поэтому целесообразно выбирать не большие значения этого отношения. Также можно сначала провести эксперимент с произвольным числом реплик, а затем оценить его мощность с использованием выражения (1).

Аналогичным образом осуществляется оценка достаточного числа реплик при проведении многофакторного дисперсионного анализа.

4. Провести эксперимент. Проведение эксперимента заключается в измерении времени выполнения функций АИС. Причем в случае наличия факторов, влияющих на время выполнения функций АИС, необходимо проводить измерения для различных уровней фактора. Обязательным условием корректного проведения эксперимента является обеспечение рандомизации, т.е. распределение экспериментального материала и порядок, в котором должны проводиться отдельные измерения, устанавливаются случайным образом. Это обусловлено тем, что для применения статистических методов требуется, чтобы наблюдения были независимыми случайными переменными.

Значения наблюдений времени выполнения функций АИС $t_{вф}$ заносятся в таблицу.

Вид таблицы зависит от количества исследуемых факторов. В случае однофакторного дисперсионного анализа результаты измерений представляются таблицей 1.

Таблица 1 – Результаты измерения времени выполнения функций АИС

Фактор	Номер реплики					
	1	2	...	j	...	n
уровень 1	t_{11}	t_{12}	...	t_{1j}	...	t_{1n}
уровень 2	t_{21}	t_{22}	...	t_{2j}	...	t_{2n}
.....
уровень i	t_{i1}	t_{i2}	...	t_{ij}	...	t_{in}
.....
уровень m	t_{m1}	t_{m2}	...	t_{mj}	...	t_{mn}

При этом t_{ij} – время выполнения функции АИС, полученное в j -м наблюдении при i -м уровне фактора. Число реплик определяется с помощью процедуры, описанной выше. Количество уровней фактора задается исследователем, причем для анализа целесообразно использовать модель постоянных эффектов, т.к. нижняя и верхняя границы уровней фактора остаются неизменными.

ОБРАБОТКА И АНАЛИЗ ЭКСПЕРИМЕНТАЛЬНЫХ ДАННЫХ

Полученные в ходе эксперимента значения времени выполнения функций АИС необходимо очистить от случайных выбросов и проанализировать на предмет наличия влияния факторов и оценки вида распределения.

1. Обработать экспериментальные данные. После проведения эксперимента необходимо проверить экспериментальные значения на аномальность, т.е. выявить выбросы, обусловленные случайными причинами. Для решения этой задачи используется критерий Диксона, включающий следующие шаги:

– формирование ранжированного ряда экспериментальных данных по каждому уровню фактора и по всем репликам;

– расчет критерия Диксона по выражениям, представленным в таблице 2 [2], в зависимости от числа наблюдений и оцениваемого значения экспериментальных данных (максимального или минимального);

– сравнение полученного коэффициента с табличным значением.

Если рассчитанный коэффициент меньше табличного, то проверяемое значение не является аномальным. В обратном случае проверяемое значение исключается из дальнейшего анализа и проверяется ближайшее к исключенному значение переменной отклика (времени выполнения функций АИС).

Таблица 2 – Выражения критерия Диксона

Число реплик n	Коэффициент Диксона	Для наименьшего экспериментального значения	Для наибольшего экспериментального значения
3..7	r_{10}	$\frac{t_2 - t_1}{t_n - t_1}$	$\frac{t_n - t_{n-1}}{t_n - t_1}$
8..10	r_{11}	$\frac{t_2 - t_1}{t_{n-1} - t_1}$	$\frac{t_n - t_{n-1}}{t_n - t_2}$
11..13	r_{21}	$\frac{t_3 - t_1}{t_n - t_1}$	$\frac{t_n - t_{n-2}}{t_n - t_2}$
14..30	r_{22}	$\frac{t_3 - t_1}{t_{n-2} - t_1}$	$\frac{t_n - t_{n-2}}{t_n - t_3}$

2. Провести дисперсионный анализ экспериментальных данных. Дисперсионный анализ проводится для выявления влияния выделенных факторов на время выполнения функций АИС. Основная идея дисперсионного анализа состоит в сравнении «факторной дисперсии», порождаемой воздействием фактора, и «остаточной дисперсии», обусловленной случайными причинами. Если различие между этими дисперсиями значимо, то фактор оказывает существенное влияние на время выполнения функций. В этом случае средние наблюдаемых значений на каждом уровне (групповые средние) различаются также значимо.

Для сравнительной оценки дисперсий необходимо проверить гипотезу о влиянии фактора на время выполнения функции АИС [3]:

$$H_0 : \tau_1 = \tau_2 = \dots = \tau_m = 0; H_1 : \tau_i \neq 0 \text{ хотя бы для одного } i. \quad (2)$$

Дисперсионный анализ применяется, чтобы установить однородность нескольких выборочных совокупностей, образуемых при различных условиях функционирования АИС (уровни фактора). Неоднородность совокупностей позволяет сделать вывод о влиянии фактора на время выполнения функций. Однородные же совокупности объединяются в одну и тем самым используется более полная информация и получаются более надежные выводы.

Если гипотеза, описываемая выражением (2), справедлива, то каждое наблюдение времени выполнения функций складывается из математического ожидания общего среднего и реализации случайной ошибки. Общая изменчивость данных характеризуется общей скорректированной суммой квадратов ($SS_{\text{общ}}$), состоящей из суммы квадратов отклонений средних по каждому уровню от общего среднего ($SS_{\text{ур}}$) и суммы квадратов отклонений всех наблюдений по каждому уровню от своего среднего ($SS_{\text{ош}}$) [3]:

$$SS_{\text{общ}} = SS_{\text{ур}} + SS_{\text{ош}}.$$

Для расчета соответствующих сумм квадратов используются следующие выражения [3]:

$$SS_{\text{общ}} = \sum_{i=1}^m \sum_{j=1}^n t_{ij}^2 - \frac{1}{N} \cdot \left(\sum_{i=1}^m \sum_{j=1}^n t_{ij} \right)^2,$$

$$SS_{\text{ур}} = \frac{1}{n} \cdot \sum_{i=1}^m \left(\sum_{j=1}^n t_{ij} \right)^2 - \frac{1}{N} \cdot \left(\sum_{i=1}^m \sum_{j=1}^n t_{ij} \right)^2,$$

$$SS_{\text{ош}} = SS_{\text{общ}} - SS_{\text{ур}},$$

где N – общее число наблюдений.

При этом $SS_{\text{общ}}$ обладает $N-1$ степенями свободы, $SS_{\text{ош}}$ обладает $m \cdot (n-1)$ степенями свободы, $SS_{\text{ур}}$ обладает $m-1$ степенями свободы. Тогда для проверки нулевой гипотезы используется статистический критерий, имеющий распределение Фишера [3]:

$$F = \frac{SS_{\text{ур}} / (m-1)}{SS_{\text{ош}} / (N-m)}. \quad (3)$$

После расчета наблюдаемого значения статистического критерия необходимо определить его критическое значение при условии правосторонней критической области. Критическое значение определяется в зависимости от выбранного уровня значимости α и значений числа степеней свободы числителя $\nu_1 = m-1$ и знаменателя $\nu_2 = N-m$. Для определения критического значения статистического критерия используются специальные таблицы с табулированными значениями распределения Фишера. Если наблюдаемое значение критической статистики больше критического значения, то нулевая гипотеза отклоняется и принимается гипотеза о существовании различия между уровнями фактора и о влиянии фактора на время выполнения функций АИС.

3. Оценить закон распределения времени выполнения функций. При достаточно большом числе наблюдений согласно центральной предельной теореме распределение времени выполнения функций АИС должно подчиняться нормальному закону. Для проверки этого предположения целесообразно использовать специальные критерии согласия (Пирсона, Смирнова, Колмогорова и т.п.), основанные на сопоставлении эмпирических и теоретических частот. Однако для их применения необходимо, чтобы объем выборочной совокупности (число реплик) составлял не менее 100 единиц, что, как правило, на практике не выполняется, особенно если существуют влияющие факторы. Поэтому для доказательства нормального распределения времени выполнения функций АИС предлагается применить правило трех сигм: если случайная величина распределена по нормальному закону, то абсолютная величина её отклонения от математического ожидания не превосходит утроенного среднего квадратического отклонения и вероятность выполнения этого условия равна 0,9973 [4]:

$$P\left(|t_{ij} - M(t)| < 3 \cdot \sigma\right) = 0,9973, \quad (4)$$

где $M(t)$ – математическое ожидание времени выполнения функций АИС; σ – среднее квадратическое отклонение времени выполнения функций АИС.

Поскольку значения $M(t)$ и σ неизвестны, их необходимо заменить оценками, полученными по выборке значений времени выполнения функций АИС – выборочным средним \bar{t}_B и исправленной дисперсией s соответственно.

В случае, если фактор оказывает влияние на время выполнения функций АИС, то значения времен по каждому уровню необходимо рассматривать как отдельные выборочные совокупности и правило (4) нужно применять для каждой совокупности в отдельности.

ОЦЕНКА ОПЕРАТИВНОСТИ СИСТЕМЫ

Задача оценки оперативности выполнения функций АИС является задачей определения эмпирической функции распределения вероятностей или функции плотности вероятностей времени выполнения функций. На основании центральной предельной теоремы и с учетом результатов применения правила трех сигм можно предполагать нормальный закон распределения вероятностей. Однако это предположение является не совсем корректным, поскольку корректность использования классического нормального распределения времени достигается при выполнении условия $M(t) \geq 3 \cdot \sigma$. При малом значении $M(t)$ и большом σ может возникнуть ситуация, когда кривая распределения времени выполнения функций АИС «покрывает» своей левой ветвью область отрицательных времен. Таким образом, нормальное распределение, являясь общим случаем распределения времени выполнения функций в диапазоне $(-\infty; +\infty)$, лишь в частности (при определенных условиях) может быть использовано для моделей оперативности.

С другой стороны, полученные в ходе проведения эксперимента значения времени выполнения функций АИС, очевидно, будут лежать в некотором диапазоне $(t_{\text{вф}}^{\text{мин}}; t_{\text{вф}}^{\text{макс}})$ и значения, меньшие или большие граничных значений, не будут использоваться при формировании функции распределения вероятностей времени выполнения функций АИС. Другими словами, при проведении эксперимента выбор значений времени производится не из всей генеральной совокупности, а только из ее части. Тогда оценка оперативности сводится к решению задачи определения вероятности попадания случайной величины – времени выполнения функций АИС в заданный интервал, нижней границей которого является значение минимального времени выполнения функций, а верхней границей – допустимое время выполнения функций: $P(t_{\text{вф}}^{\text{мин}} \leq t_{\text{вф}} < t_{\text{вф}}^{\text{доп}})$.

Для определения нижней границы $t_{\text{вф}}^{\text{мин}}$ предлагается использовать результаты экспериментальной оценки времени выполнения функций. При этом в случае влияния факторов на переменную отклика необходимо определять граничные значения для совокупности времен, соответствующих каждому уровню. Верхняя граница $t_{\text{вф}}^{\text{доп}}$ задается заказчиком АИС либо формируется на основе анализа аналогов.

Исходя из описанных положений, для оценки оперативности автоматизированных систем предлагается использовать усеченное нормальное распределение, которое позволяет учесть обозначенные выше особенности.

Функция плотности усеченного нормального распределения $f^y(t)$ связана с соответствующей функцией классического нормального распределения выражением вида [5]:

$$f^y(t) = c \cdot f(t),$$

где c – нормирующий множитель; $f(t) = \frac{1}{\sigma \cdot \sqrt{2\pi}} \cdot \exp\left(-\frac{(t - M(t))^2}{2\sigma^2}\right)$ – функция плотности

классического нормального распределения.

Нормирующий множитель определяется из условия, что площадь под кривой функции $f^y(t)$ равна 1, т.е.:

$$\int_{t_{\text{вф}}^{\text{мин}}}^{t_{\text{вф}}^{\text{доп}}} f^y(t) dt = \int_{t_{\text{вф}}^{\text{мин}}}^{t_{\text{вф}}^{\text{доп}}} c \cdot f(t) dt = c \cdot \int_{t_{\text{вф}}^{\text{мин}}}^{t_{\text{вф}}^{\text{доп}}} f(t) dt = 1.$$

Откуда

$$c = \frac{1}{\int_{t_{\text{вф}}^{\text{мин}}}^{t_{\text{вф}}^{\text{доп}}} f(t) dt}.$$

Исходя из теоремы о равенстве вероятности того, что непрерывная случайная величина примет значение, принадлежащее некоторому интервалу, определенному интегралу от плотности распределения, взятому в пределах от границ этого интервала и свойств функции распределения, можно записать:

$$P(t_{\text{вф}}^{\text{мин}} \leq t_{\text{вф}} < t_{\text{вф}}^{\text{доп}}) = \int_{t_{\text{вф}}^{\text{мин}}}^{t_{\text{вф}}^{\text{доп}}} f(t) dt = F(t_{\text{вф}}^{\text{доп}}) - F(t_{\text{вф}}^{\text{мин}}),$$

где $F(t_{\text{вф}}^{\text{доп}})$ и $F(t_{\text{вф}}^{\text{мин}})$ – значения функции распределения при значениях времени $t_{\text{вф}}^{\text{макс}}$ и $t_{\text{вф}}^{\text{доп}}$ соответственно.

Переходя от случайной величины $T = (t_1, t_2, \dots, t_n)$ к нормированной величине $X = (x_1, x_2, \dots, x_n)$, получаем [4]:

$$x^{\text{доп}} = \frac{(t_{\text{вф}}^{\text{доп}} - \bar{t}_{\text{в}})}{s}, \quad x^{\text{мин}} = \frac{(t_{\text{вф}}^{\text{мин}} - \bar{t}_{\text{в}})}{s}.$$

Используя функцию Лапласа, можно записать:

$$\int_{t_{\text{вф}}^{\text{мин}}}^{t_{\text{вф}}^{\text{доп}}} f(t) dt = F(t_{\text{вф}}^{\text{доп}}) - F(t_{\text{вф}}^{\text{мин}}) = 0,5 + \Phi(x^{\text{доп}}) - 0,5 - \Phi(x^{\text{мин}}) = \Phi(x^{\text{доп}}) - \Phi(x^{\text{мин}}).$$

Отсюда нормирующий множитель равен:

$$c = \frac{1}{\Phi(x^{\text{доп}}) - \Phi(x^{\text{мин}})}. \quad (5)$$

Из определения оперативности и на основании выражения, связывающего функцию распределения и функцию плотности распределения вероятностей, следует:

$$P(t_{\text{вф}} \leq t_{\text{вф}}^{\text{доп}}) = F(t_{\text{вф}}^{\text{доп}}) = c \cdot \int_{-\infty}^{t_{\text{вф}}^{\text{доп}}} f(t) dt = 0,5 + \Phi(x).$$

Тогда оперативность АИС при использовании усеченного нормального распределения определяется следующим выражением:

$$P^{\text{ус}}(t_{\text{вф}}^{\text{мин}} \leq t_{\text{вф}} < t_{\text{вф}}^{\text{доп}}) = \int_{-\infty}^{t_{\text{вф}}^{\text{доп}}} c \cdot f(t) dt = c \cdot \int_{-\infty}^{t_{\text{вф}}^{\text{доп}}} f(t) dt = c \cdot (0,5 + \Phi(x)), \quad (6)$$

где множитель c вычисляется по выражению (5).

Теоретически можно использовать не двухстороннее усечение, а усеченное слева нормальное распределение с нижней границей $t_{\text{вф}}^{\text{мин}} = 0$. Однако это приведет к завышению искомой вероятности, поскольку практически не будет значений времени выполнения функций АИС, лежащих в диапазоне от 0 до минимального времени, полученного в результате эксперимента $t_{\text{вф}}^{\text{мин}}$.

ВЫВОД

Описанная методика за счет применения методов планирования и анализа экспериментальных данных, а также использования усеченного нормального распределения для описания времени выполнения функций АИС, позволяет повысить достоверности оценки оперативности АИС.

СПИСОК ЛИТЕРАТУРЫ

1. Монтомгери Д.К. Планирование эксперимента и анализ данных: пер. с англ. Л.П. Денисова. – М.: Судостроение, 1980. – 384 с.
2. Блохин В.Г. Современный эксперимент: подготовка, проведение, анализ результатов / В.Г. Блохин, О.П. Глудкин, А.И. Гуров, М.А. Ханин. – М.: Радио и связь, 1997. – 232 с.
3. Кремер Н.Ш. Теория вероятностей и математическая статистика: учеб. для ВУЗов. – 2-е изд., перераб. и доп. – М.: ЮНИТИ, 2006. – 550 с.
4. Гмурман В.Е. Теория вероятностей и математическая статистика: учебное пособие для ВУЗов. – М.: Высш. шк., 2004. – 479 с.
5. Крамер Г. Математические методы статистики: пер. с англ. А.С. Монина и А.А. Петрова. – М.: Мир, 1975. – 648 с.

Новиков Евгений Иванович

Академия ФСО России, г. Орел

Кандидат технических наук, доцент кафедры № 32

Тел.: 8 953 812 44 22

E-mail: nei05@rambler.ru

E.I. NOVIKOV (*Candidate of Engineering Science, Associate Professor of the Department № 32*)
Academy of Federal Agency of Protection of the Russian Federation, Orel

EVALUATION METHODS OF EFFICIENT PROPERTY OF AUTOMATED INFORMATION SYSTEM

The article describes the technique of experimental evaluation of the effectiveness of automated systems for the property of the operability of automated functions. The basic feature of the technique is the consideration of random factors that affect the correct assessment of the efficiency and the use of a truncated normal distribution of the automated functions run-time. The article also reveals the procedure for planning experiments to prove a sufficient number of replicas and the validity of the results of the processing and analysis of experimental data.

Keywords: *effectiveness; property; factor; experiment; replica; analysis of variance; statistical hypothesis; random variable; distribution law; distribution function.*

BIBLIOGRAPHY (TRANSLITERATED)

1. Montgomeri D.K. Planirovanie e'ksperimenta i analiz danny'x: per. s angl. L.P. Denisova. – М.: Sudostroenie, 1980. – 384 s.
2. Bloxin V.G. Sovremenny'j e'ksperiment: podgotovka, provedenie, analiz rezul'tatov / V.G. Bloxin, O.P. Gludkin, A.I. Gurov, M.A. Xanin. – М.: Radi i svyaz', 1997. – 232 s.
3. Kremer N.Sh. Teoriya veroyatnostej i matematicheskaya statistika: ucheb. dlya VUZv. – 2-e izd., pererab. i dop. – М.: YuNITI, 2006. – 550 s.
4. Gmurman V.E. Teoriya veroyatnostej i matematicheskaya statistika: uchebnoe posbie dlya VUZov. – М.: Vy'ssh. shk., 2004. – 479 s.
5. Kramer G. Matematicheskie metody' statistiki: per. s angl. A.S. Monina i A.A. Petrva. – М.: Mir, 1975. – 648 s.

УДК004.056.5, 004.89

А.Ю. ОЛАДЬКО, В.С. АТКИНА

**МНОГОАГЕНТНАЯ СИСТЕМА ЗАЩИТЫ ИНФОРМАЦИИ
В ОПЕРАЦИОННЫХ СИСТЕМАХ,
ПОСТРОЕННАЯ С ИСПОЛЬЗОВАНИЕМ ТЕХНОЛОГИЙ
ИММУНОКОМПЬЮТИНГА**

Проанализированы основные причины нарушения информационной безопасности в операционных системах. Рассмотрены методы и средства защиты информации, выделены их достоинства и недостатки. Обоснована необходимость создания адаптивных систем защиты в операционных системах. Приведена формальная модель многоагентной системы защиты информации в операционной системе, построенная с использованием технологии иммунокомпьютинга. Представлены результаты экспериментальных исследований разработанной системы, сделаны выводы о возможности применения ее на практике.

Ключевые слова: операционная система; система защиты; искусственные иммунные системы; многоагентные системы.

ВВЕДЕНИЕ

Обеспечение безопасности информации в современном бизнесе является одной из ключевых задач. А поскольку деятельность практически любой организации тесно связана с обработкой электронной информации средствами персональных компьютеров, автоматизированных рабочих мест пользователей, объединённых в корпоративные сети и информационные системы различного уровня и масштаба, то успешность выполнения ее ключевых бизнес-процессов будет напрямую зависеть от обеспечения таких свойств информации, как доступность, целостность и конфиденциальность. А так как при работе пользователя и обработки бизнес-процессов всегда используются операционные системы, то защита информации на их уровнях – одна из приоритетных задач.

Анализ данных статистики (рис. 1) позволяет сделать вывод о том, что несмотря на развитие средств защиты, таких, как брандмауэры, системы аутентификации и разграничения доступа, количество атак злоумышленников на операционные системы (ОС) как на одни из ключевых элементов корпоративных информационных систем организаций возрастает с каждым годом.

Увеличение числа уязвимостей и атак, как известных, так и неизвестных, на ресурсы ОС приводит к необходимости разработки и применения средств и систем защиты, способных адаптироваться и проводить свою реконфигурацию в зависимости от состояния операционной среды и событий, происходящих в ней. Как правило, способность к адаптации и обучению таких систем защиты обеспечивается за счет использования технологий искусственного интеллекта, распределенного анализа данных, модульности и многоагентного подхода.

Следовательно, актуальным направлением является исследование и решение задач, связанных с разработкой интеллектуальных и адаптивных систем защиты информации в ОС.

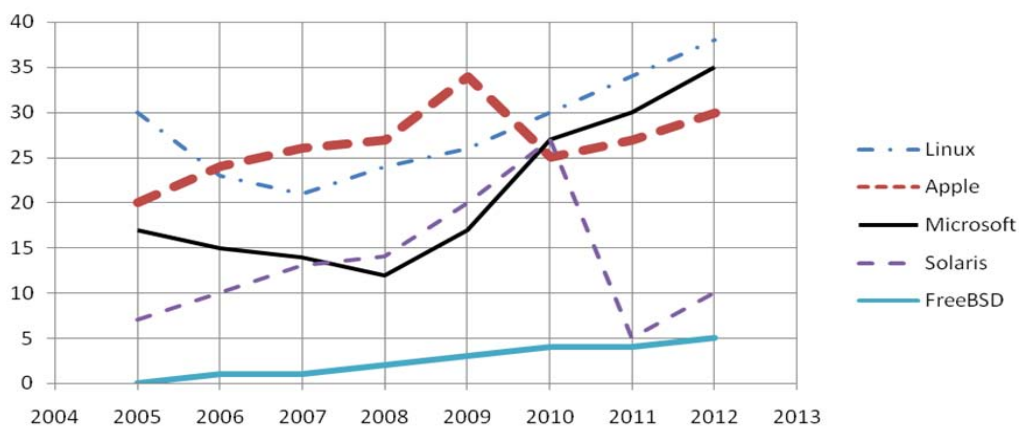


Рисунок 1 – Статистика незакрытых уязвимостей, найденных в операционных системах, представленная IBM-X Force в период 2005-2012 гг.

УГРОЗЫ БЕЗОПАСНОСТИ В ОПЕРАЦИОННЫХ СИСТЕМАХ

Анализ литературных источников [1, 2], процесса функционирования и архитектуры ОС различного вида показал, что наиболее часто воздействию злоумышленника подвергаются такие элементы, входящие в состав ОС, как:

- ядро ОС;
- системные утилиты;
- прикладные программы;
- пользовательские данные;
- сетевые интерфейсы.

Анализ атак показал, что процесс реализации атаки осуществляется в 4 этапа: рекогносцировка, вторжение, стадия атакующего воздействия, стадия дальнейшего развития атаки. При этом наиболее характерными и по статистике самыми часто встречающимися категориями атак для ОС являются:

- атаки, приводящие к отказу (DenialofService) в обслуживании (системный сбой);
- вредоносное программное обеспечение;
- атаки, позволяющие несанкционированно запустить исполняемый код.

Проведенный анализ встроенных (штатных) средств защиты, а также дополнительных сервисов и подсистем защиты информации в ОС позволил выделить следующие наиболее распространенные методы и механизмы защиты:

- средства аутентификации и идентификации;
- криптографические сервисы;
- ролевое управление доступом;
- изолированная программная среда;
- межсетевые экраны;
- системы обнаружения атак и вторжений (COA и COB);
- системы мониторинга и аудита;
- защита от выполнения кода в стеке;
- антивирусное программное обеспечение.

Однако, как было показано в [3], данные средства защищают преимущественно от таких видов атак, как: несанкционированные операции чтение/запись; атак, способных обойти установленные разграничения прав доступа; DOS-атак; атак, связанных с несанкционированным запуском программного кода. В то же время большинство из перечисленных средств не позволяют защититься от атак, использующих недокументированные возможности, от атак, непосредственно воздействующих на ядро ОС, и от некоторых видов вредоносного программного обеспечения. Также штатные средства

большинства ОС не позволяют защититься от уязвимостей в прикладном программном обеспечении, что приводит к удаленному или локальному выполнению кода (например, ошибка переполнения `heap`'а и ошибка формата строки).

Таким образом, можно сделать вывод, что использование одних только встроенных средств и сервисов защиты, реализованных в ОС, недостаточно для защиты от всего спектра возможных угроз и атак. Основным недостатком дополнительно устанавливаемых подсистем защиты информации в ОС является то, что после первоначальной конфигурации они не обладают достаточной гибкостью к постоянно изменяемым условиям и событиям, протекающим в операционной среде. А системы активного аудита СОА и СОВ, хотя и обладают возможностью обнаруживать атаки и детектировать аномальные события и состояния в ОС в режиме реального времени, в большинстве случаев лишь сигнализируют о возможном вторжении и активности злоумышленника, не вырабатывая при этом своевременной защитной реакции в ответ на обнаруженный вид деструктивного воздействия. Следовательно, возникает необходимость в создании и применении таких механизмов защиты, которые могли бы обнаруживать и противодействовать не только наиболее типовым для ОС атакам, но и их новым видам, что может быть достигнуто за счет применения современных адаптивных технологий.

ЗАДАЧИ И ФОРМАЛЬНАЯ МОДЕЛЬ СИСТЕМЫ ЗАЩИТЫ ИНФОРМАЦИИ В ОПЕРАЦИОННОЙ СИСТЕМЕ

При создании системы защиты информации в ОС, способной обнаруживать и идентифицировать аномальные состояния и атаки злоумышленника, а также адаптироваться в зависимости от вида обнаруженного воздействия, авторами был проведен анализ адаптивных методов, таких, как: нейронные сети, иммунокомпьютинг, статистический анализ, кластерный анализ, поведенческая биометрия. По результатам проведенных исследований [4-6] был сделан вывод, что данные методы обладают примерно одинаковой вычислительной сложностью, однако одной из наиболее развивающихся и перспективных в областях распределённого анализа, диагностики аномальных состояний системы, управления и выработки решений является технология иммунокомпьютинга.

Таким образом, для повышения эффективности защиты информации в ОС была разработана система защиты, в основе которой лежит многоагентная технология для распределенного анализа и сбора данных о процессах и событиях, происходящих в ОС, и принцип технологии искусственной иммунной системы для обеспечения адаптивности и управления процессом принятия решений по защите. Данная система защиты по своим функциональным возможностям обеспечивает решение следующих задач:

- диагностика известных атак и аномальных состояний в ОС;
- регистрация, выявление и оценка серьезности событий, имеющих признаки инцидента;
- идентификация инцидента на основе оперативного анализа доказательств, принятие решения в условиях неполной определенности имеющейся информации и при необходимости генерация сигнала тревоги;
- выработка решений по защите;
- обработка и устранение последствий инцидента путем введения в действие соответствующих ресурсов безопасности.

Формально модель адаптивной системы защиты в ОС описывается следующим кортежем:

$$SysAP = \{ \{S\}, \{Ag\}, \{SP\}, \{REP\} \}, \quad (1)$$

где $\{S\}$ – множество возможных состояний системы; $\{Ag\}$ – множество мобильных интеллектуальных агентов, осуществляющих сбор данных об объектах системы, процессах, зарегистрированных пользователях и перехват системных вызовов, классификацию состояний системы, идентификацию аномалий, выработку защитной реакции; $\{SP\}$ –

множество защитных механизмов; $\{REP\}$ – отчет о результатах работы и принятых решениях системы защиты в ОС, содержит информацию о выявленных несоответствиях между «эталонным» и текущим состоянием ОС, действиях пользователей в системе и зафиксированных потенциально возможных нарушениях, а также защитных реакциях.

Каждое состояние $S_i \in \{S\}$ ОС описывается следующим кортежем:

$$S_i = \{\{O\}, \{U\}, \{PR\}\} \quad (2)$$

где $\{O\}$ – множество объектов ОС, причем $\{O\} = \{F\} \cup \{Pu\} \cup \{Ps\}$, где $\{F\}$ – множество файлов, $\{Pu\}$ – множество прикладных программ, установленных на данной ОС, $\{Ps\}$ – системное программное обеспечение и модули ядра ОС; $\{U\}$ – множество пользователей, зарегистрированных в ОС; $\{PR\}$ – множество запущенных процессов в ОС.

В свою очередь, каждый процесс $PR_j \in \{PR\}$ будет описываться кортежем $PR_j = \{UID, PID, \{API\}, \{NCP\}\}$, где UID – идентификатор пользователя, запустившего процесс, согласно данному идентификатору определяются права процесса при выполнении API функций ОС; PID – идентификатор процесса в системе; $\{API\}$ – множество API функций операционной системы, которые были вызваны приложением, включая параметры этих вызовов; $\{NCP\}$ – множество сетевых соединений, инициированных конкретным приложением.

При описании состояний ОС можно выделить:

- текущее состояние системы $S_{TOS} \in \{S\}$, которое формируется на основании данных собранных агентами, осуществляющими мониторинг процессов и событий в ОС в процессе «реального» времени;

- множество профилей, описывающих «эталонные» состояния ОС $\{S_{TOS}\} \in \{S\}$, составленных на основании статистических данных, данных, полученных в процессе обучения системы, когда все состояния системы считаются условно «безопасными», или заданных при конфигурации системы администратором безопасности;

- все остальные состояния ОС, определяемые как $S_{DOS} = \{S\} \setminus S_{SOS}$, считаются «опасными», т.е. теми, в которых возможно потенциальное нарушение информационной безопасности. При этом верно, что

$$S_{TOS} \in S_{DOS} \wedge S_{TOS} \notin \{S_{SOS}\} \vee S_{TOS} \notin S_{DOS} \wedge S_{TOS} \in \{S_{SOS}\}.$$

Поскольку в основе разработанной системы защиты лежит иммунный подход, то все агенты, входящие в состав многоагентной системы, в зависимости от их функционального назначения делятся на следующие группы: агенты-реакторы – $\{AgR\}$, агенты-рецепторы – $\{AgM\}$, агент-координатор – $\{AgC\}$, агенты-идентификаторы – $\{AgI\}$. При этом верно, что $Ag = \{AgR\} \cup \{AgM\} \cup \{AgC\} \cup \{AgI\}$.

Агенты иммунной системы функционируют в режиме реального времени, осуществляют распределенный сбор данных о процессах и событиях, происходящих в ОС, классифицируют их и в процессе своего взаимодействия путем обмена сообщениями вырабатывают решения и управляющие воздействия. Структура разработанной иммунной системы и описание алгоритмов работы агентов представлены в [1].

При разработке системы защиты ОС на базе иммуннокомпьютинга в соответствие клеткам организма были поставлены процессы.

По аналогии с функционированием иммунной системы были реализованы 3 основных направления действия системы защиты:

1. Идентификация – выявление и классификация аномальной активности в ОС.
2. Иммунизация делится на две подфункции:
 - неспецифический (врожденный) иммунитет, обуславливающий однотипные реакции на любые чужеродные вторжения. За данную подфункцию отвечают текущие настройки средств и систем защиты $\{SP\}$ в ОС.
 - специфический иммунитет – система данных и алгоритмов, а также средств

защиты {SP}, которые формируются и настраиваются в процессе работы иммунной системы; запускается и реализуется при обнаружении вторжения в качестве средства противодействия.

3. Апоптоз – процесс остановки вторжения в ОС.

Процесс идентификации злоумышленной активности заключается в анализе текущего состояния системы S_{TOS} на предмет его принадлежности к множеству нормальных состояний. Для этого множество агентов-идентификаторов $\{AgI\}$ осуществляют классификацию состояний ОС по вектору признаков в соответствии со следующим правилом:

$$F: S_{TOS} \rightarrow Y. \quad (3)$$

Функция F отображает пространство векторов-признаков S_{TOS} в пространство векторов меток Y . В этом случае $Y=[0,1]$, где 1 соответствует «безопасному» состоянию системы, а 0 – «опасному», т.е., тому, где было зафиксировано возможное нарушение.

Профиль текущего состояния системы S_{TOS} сравнивается агентами-идентификаторами $\{AgI\}$ с элементами $S_{SOS_i} \in \{S_{SOS}\}$ из множества профилей, описывающих «эталонные» состояния ОС, в качестве меры сходства применяется коэффициент Жаккара, описанный формулой:

$$K_J(S_{TOS}, S_{SOS_j}) = \frac{n(S_{TOS} \cap S_{SOS_j})}{n(S_{TOS} \cup S_{SOS_j})}. \quad (4)$$

Считается, что S_{TOS} входит во множество «эталонных» состояний системы $\{S_{SOS}\}$, т.е. система признает данное состояние безопасным и нарушений не обнаруживается, если коэффициент Жаккара K_J больше порогового значения K_{Jmin} : $K_{Jmin} \leq K_J(S_{TOS}, S_{SOS_j}) \leq 1$, в противном случае считается, что в ОС произошло нарушение информационной безопасности и состояние системы в соответствии с формулой (3) классифицируется агентами-идентификаторами как опасное:

$$AgI(K_J(S_{TOS}, S_{SOS_j}), K_{Jmin}) = \begin{cases} 1, & \text{если } K_{Jmin} \leq K_J(S_{TOS}, S_{SOS_j}) \leq 1 \\ 0, & \text{если } K_J(S_{TOS}, S_{SOS_j}) \leq K_{Jmin} \end{cases}. \quad (5)$$

После того, как состояние ОС было распознано как «опасное», агентом-координатором AgC запускается процедура иммунного ответа.

В отношении процессов, идентифицированных как злоумышленные, под влиянием агентов-реакторов AgR производится программируемое разрушение. И, как следствие, апоптоз – остановка обнаруженного вторжения.

Далее по полученным результатам система формирует отчет {RPT} с выявленными несоответствиями текущих параметров системы, с «эталонными» значениями параметров в соответствующих им позициях в профилях «эталонных» состояний ОС, а также данные о типе зафиксированного (выявленного) нарушения, пользователе, совершившем его, объекте и времени совершения, а также механизмах защиты, устранивших несоответствия и нарушения.

РЕЗУЛЬТАТЫ ЭКСПЕРИМЕНТАЛЬНЫХ ИССЛЕДОВАНИЙ МНОГОАГЕНТНОЙ СИСТЕМЫ ЗАЩИТЫ В ОС СЕМЕЙСТВА LINUX

Описанная система защиты информации была реализована для ОС семейства Linux. Выбор данной ОС был обусловлен следующими факторами:

– ОС Linux на сегодняшний день имеет широкое распространение среди серверных платформ, также она охватывает определенную часть рынка пользовательских операционных систем; анализ данных статистики показывает, что в настоящее время системы Linux лидируют на рынках смартфонов (Android занимает 64,1% рынка), интернет-серверов (60%), самых мощных суперкомпьютеров (91%), а также, согласно LinuxFoundation, в дата-центрах и на предприятиях имеют значительную долю рынка нетбуков (32%), а на рынке домашних компьютеров Linux занимает 3 место с 1-2,5%.

– Президентом РФ Д.И. Медведевым и Правительством Российской Федерации 1 апреля 2011г. на заседании комиссии по высоким технологиям и инновациям в рамках федеральной программы «Информационное общество (2011-2020 годы)» было утверждено создание национальной программной платформы (НПП). В основу реализации, основным этапом которой является разработка и внедрение проекта «Национальная операционная система», ляжет базовая платформа и дистрибутивы Linux. В связи с введением данной программы наметилась тенденция перехода на открытое программное обеспечение, в рамках которой, в частности, идет установка и апробация ОС Linux в школах и государственных учреждениях.

Поэтому вопросы информационной безопасности данной ОС имеют большую актуальность.

Архитектура системы защиты и ее основные функциональные возможности в ОС Linux представлены на рисунке 2 и состоят из интерфейса, двух баз данных и четырех взаимодействующих функциональных модулей, реализующих процедуры мониторинга состояний ОС, диагностики и обнаружения вторжений в ОС Linux, а также функций подбора наиболее оптимальных средств и методов защиты.

Модуль мониторинга состояний системы отслеживает текущее состояние операционной системы и выполняемых процессов.

Модуль управления средствами защиты на основе данных, полученных от координирующего модуля, и данных о возможном типе вторжения осуществляет управление встроенными средствами защиты, в противодействие обнаруженной атаке.

База данных средств защиты содержит информацию обо всех имеющихся в данной ОС средствах защиты, их настройке и конфигурации, а также функциональном назначении.

База данных антигенов представляет собой множество шаблонов нормальных состояний системы, а также множество шаблонов выявленных атак для инициализации процедуры вторичного иммунного ответа.

Модуль выявления чужеродных антигенов на основе данных, полученных в результате работы модуля мониторинга состояний системы и множества шаблонов нормальных состояний системы, обнаруживает аномалии в поведении системы, место их возникновения, а также осуществляет диагностику атак.

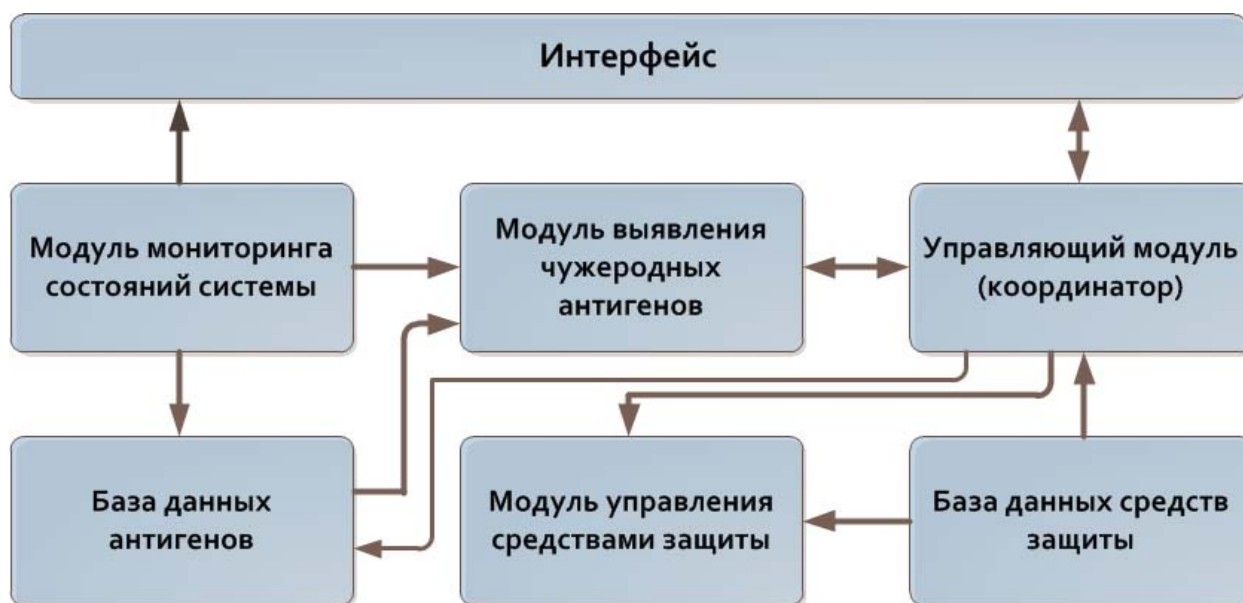
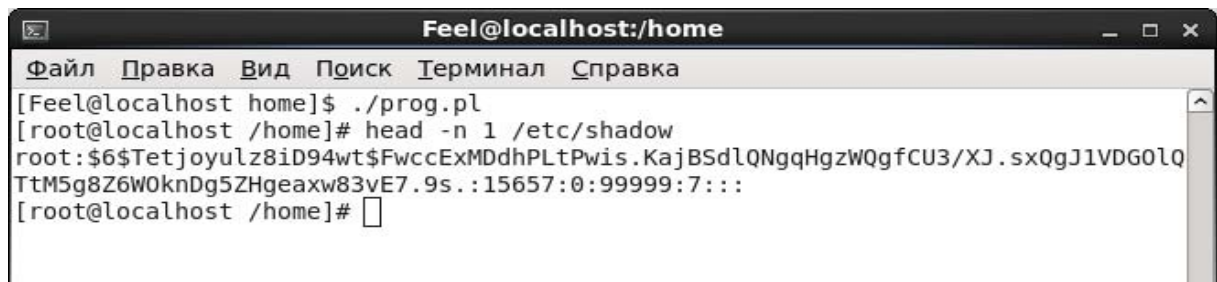


Рисунок 2 – Архитектура многоагентной системы защиты в ОС Linux с использованием технологии иммунокомпьютинга

Для проверки работоспособности и адекватности разработанной модели системы защиты информации было проведено 1000 экспериментальных исследований на ОС Linux с различными вариантами конфигураций и моделированием нормальной активности, а также различных видов атак и нарушений нормального состояния ОС.

Ниже приведен пример одного из экспериментов, ориентированного на оценку результатов проведения злоумышленником атаки, направленной на повышение привилегий через модель уязвимого приложения. Для проведения эксперимента было написано приложение для операционной системы Linux, которое моделировало уязвимую программу, позволявшую несанкционированно запустить исполняемый код с целью повышения привилегий в атакуемой системе. Данное приложение позволяет запустить командный интерпретатор `/bin/bash` от имени суперпользователя `root`.

При проведении данного эксперимента моделировались действия злоумышленника, который через уязвимое к несанкционированному выполнению кода приложение повышает свои привилегии в системе до суперпользователя. Затем злоумышленник получает хэш пароля суперпользователя из файла `/etc/shadow`. Результат проведения атаки представлен на рисунке 3.



```
Feel@localhost:/home
Файл Правка Вид Поиск Терминал Справка
[Feel@localhost home]$ ./prog.pl
[root@localhost /home]# head -n 1 /etc/shadow
root:$6$Tetjoyulz8iD94wt$FwccExMDdhPLtPwis.KajBSdlQNgqHgzWQgfCU3/XJ.sxQgJ1VDG0lQ
TtM5g8Z6W0knDg5ZHgeaxw83vE7.9s.:15657:0:99999:7:::
[root@localhost /home]#
```

Рисунок 3 – Результат проведения злоумышленником атаки, направленной на повышение привилегий через модель уязвимого приложения (экранная копия)

Как из него следует, злоумышленник смог через уязвимое приложения получить права суперпользователя и получил хэш пароля пользователя `root`.

Однако попытка скачать вредоносное программное обеспечение была оборвана системой защиты и сеанс работы под `root` был завершен. Сразу после этого злоумышленник вновь попытался повысить свои привилегии через уязвимую программу, но после запуска командного интерпретатора `/bin/bash` его работа была завершена системой защиты (рис. 4).

Как видно из рисунка, программой была обнаружена атака, однако команды просмотра содержимого файла паролей не были определены как несанкционированные действия, что связано с тем, что это стандартные команды, которыми пользовался пользователь `root` в своей деятельности.

Также система защиты применила защитное средство, которое представляет из себя межсетевой экран, предоставляемый ОС. Это средство защиты было применено в связи с обнаружением чужеродной сетевой активности. При повторном проведении атаки злоумышленником она была сразу предотвращена, что является действием вторичного иммунного ответа. В данном случае иммунная система защиты уже имела информацию о проведенной атаке через приложение `prog`.

При проведении экспериментальных исследований, направленных на проверку адекватности результатов работы адаптивной системы защиты, было проведено 1000 экспериментов. В качестве гипотезы H_0 было выдвинуто предположение, что события, действия пользователей и процессы, происходящие в ОС, соответствуют нормальному состоянию системы. В качестве альтернативной гипотезы H_1 было выдвинуто предположение об аномальности состояния и наличии нарушений со стороны злоумышленника.

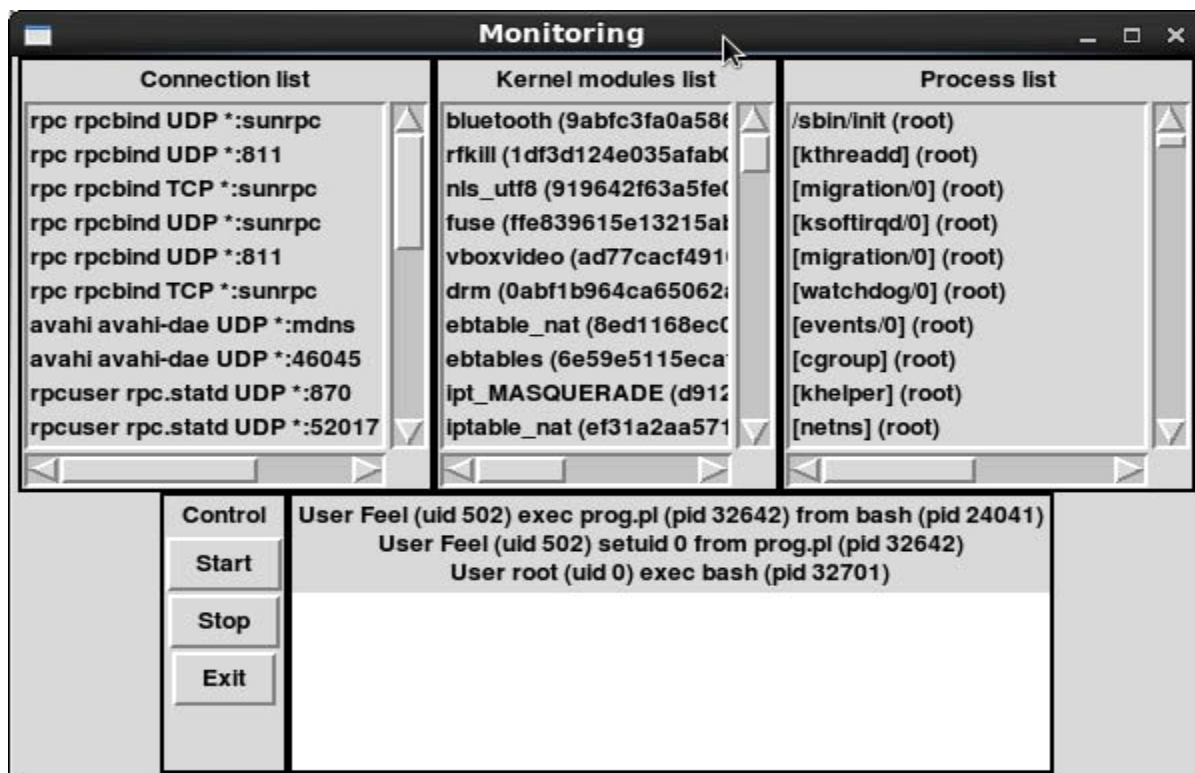


Рисунок 4 – Результат работы системы защиты при проведении злоумышленником атаки, направленной на повышение привилегий через модель уязвимого приложения (экранный снимок)

Для проверки гипотез на ошибки первого и второго эксперименты были разделены на 2 группы:

- 500 экспериментов, моделирующих активность злоумышленника в ОС Linux;
- 500 экспериментов, моделирующих нормальную активность пользователей и процессов в ОС Linux.

Результаты представлены в таблице 1.

Таблица 1 – Результаты проверки модели адаптивной системы защиты информации на ошибки первого и второго рода

		Верная гипотеза	
		H0	H1
Результат применения критерия	H0	450	30 (ошибка второго рода)
	H1	50 (ошибка первого рода)	470

Таким образом, на ошибки первого и второго рода пришлось 80 случаев, что составляет около 8% и является допустимым. Следовательно, разработанную систему защиты в ОС Linux можно применять на практике как в качестве самостоятельного средства, так и в сочетании с другими механизмами защиты для повышения эффективности защиты информации в ОС и противодействия как известным, так и неизвестным видам атак.

СПИСОК ЛИТЕРАТУРЫ

1. Оладько А.Ю. Модель адаптивной многоагентной системы защиты // Известия ЮФУ. Технические науки. Тематический выпуск «Информационная безопасность», 2011. – № 12(125). – С. 210-217.
2. Вахалия Ю. Unix изнутри. – СПб.: Питер, 2003. – 844 с.

3. Оладько А.Ю. Подсистема мониторинга и аудита информационной безопасности в операционной системе Linux // Известия ЮФУ. Технические науки. Тематический выпуск «Информационная безопасность», 2012. – № 12(137). – С. 22-28.
4. Аткина В.С. Мониторинг состояний катастрофоустойчивой информационной системы с помощью гибридной иммунной сети // Известия ЮФУ. Технические науки. «Информационная безопасность», 2012. – № 12(137). – С. 90-96.
5. Bidyuk P.I., Litvinenko V.I., Gasanov A.S. Immune Network Based Method for Identification of Turbine Engines Urging // Кафедра математического и системного анализа [Электронный ресурс]. – URL: <http://www.mmsa.kpi.ua> (дата обращения: 01.09.2013).
6. Гладыш С.В. Иммунотомпьютинг в управлении инцидентами информационной безопасности // Искусственный интеллект, 2008. – № 1. – С. 123-130.

Оладько Алексей Юрьевич

Волгоградский государственный университет, г. Волгоград
Старший преподаватель
E-mail: bor-x@yandex.ru

Аткина Владлена Сергеевна

Волгоградский государственный университет, г. Волгоград
Старший преподаватель
E-mail: atkina.vladlena@yandex.ru

A.Yu. OLAD'KO (*Senior Teacher*)

V.S. ATKINA (*Senior Teacher*)
Volgograd State University, Volgograd

**MULTI-AGENT SYSTEMS OF INFORMATION PROTECTION IN THE OPERATING SYSTEM
BUILT USING THE IMMUNOCOMPUTING TECHNOLOGY**

The main causes of information security violation in the operating systems are analyzed. The methods and resources of information security protection are considered, some advantages and disadvantages are highlighted. In this article the necessity of creating adaptive security systems in operating systems is defined. The formal model of multi-agent system of information security protection in the operating system built using immunocomputing technology is given. Moreover, research results of the system's experimental studies are presented; the conclusions about the possibility of practical application are outlined.

Keywords: *operating system; protection system; artificial immune systems; multi-agent systems.*

BIBLIOGRAPHY (TRANSLITERATED)

1. Olad'ko A.Yu. Model adaptivnoj mnogoagentnoj sistemy' zashhity' // Izvestiya YuFU. Texnicheskie nauki. Tematicheskij vy'pusk «Informacionnaya bezopasnost'», 2011. – № 12(125). – S. 210-217.
2. Vaxaliya Yu. Unix iznutri. – SPb.: Piter, 2003. – 844 s.
3. Olad'ko A.Yu. Podсистема monitoringa i audita informacionnoj bezopasnosti v operacionnoj sisteme Linux // Izvestiya YuFU. Texnicheskie nauki. Tematicheskij vy'pusk «Informacionnaya bezopasnost'», 2012. – № 12(137). – S. 22-28.
4. Atkina V.S. Monitoring sostoyanij katastrofoustojchivoj informacionnoj sistemy' s pomoshh'yu gibridnoj imunnoj seti // Izvestiya YuFU. Texnicheskie nauki. Tematicheskij vy'pusk «Informacionnaya bezopasnost'», 2012. – № 12(137). – S. 90-96.
5. Bidyuk P.I., Litvinenko V.I., Gasanov A.S. Immune Network Based Method for Identification of Turbine Engines Urging // Kafedra matematicheskogo i sistemnogo analiza [E'lektronny'j resurs]. – URL: <http://www.mmsa.kpi.ua> (data obrashheniya: 01.09.2013).
6. Glad'y'sh S.V. Imunnokomp'yuting v upravlenii incidentami informacionnoj bezopasnosti // Iskusstvenny'j intellekt, 2008. – № 1. – S. 123-130.

УДК 004.724.4

О.М. БОРИСОВ, П.Н. РЯЗАНЦЕВ,
В.В. КУЗЬМИН, Н.И. ФОМИН

АЛГОРИТМЫ ОБЕСПЕЧЕНИЯ ЖИВУЧЕСТИ ВИРТУАЛЬНОЙ ЧАСТНОЙ СЕТИ АВТОМАТИЗИРОВАННОЙ СИСТЕМЫ УПРАВЛЕНИЯ ПРОМЫШЛЕННОГО ПРЕДПРИЯТИЯ (ВЧС АСУ ПП) НА ОСНОВЕ ОПТИМАЛЬНОГО РАСПРЕДЕЛЕНИЯ СЕТЕВЫХ РЕСУРСОВ

В статье описана математическая модель потока данных АСУ ПП, базирующаяся на теории неориентированных графов и отличающаяся оптимальным распределением сетевых ресурсов для симметричного потока данных, реализации стратегии защиты звена и пути.

Ключевые слова: виртуальные частные сети; автоматизированная система управления; математическая модель; теория графов; стратегия защиты звена; стратегия защиты пути.

ВВЕДЕНИЕ

Виртуальная частная сеть (ВЧС) – обобщенное название технологий, позволяющих обеспечить одно или несколько сетевых соединений поверх другой сети. Несмотря на то, что коммуникации могут осуществляться по сетям с меньшим или неизвестным уровнем доверия (в том числе по публичным сетям), уровень доверия к построенной логической сети не зависит от уровня доверия к базовым сетям благодаря использованию средств криптографии (шифрования, аутентификации, инфраструктуры открытых ключей, средств для защиты от повторов и изменений, передаваемых по логической сети сообщений). При должном уровне реализации и использовании специального программного обеспечения сеть ВЧС может обеспечить высокий уровень шифрования передаваемой информации. Примером защищённых ВЧС являются IPsec и OpenVPN [8].

IPsec является набором стандартов Интернет и своего рода «надстройкой» над IP-протоколом. Его ядро составляют три протокола:

– Authentication Header (AH); обеспечивает целостность передаваемых данных, аутентификацию источника информации и функцию по предотвращению повторной передачи пакетов.

– Encapsulating Security Payload (ESP); обеспечивает шифрование передаваемой информации, ограничение потока конфиденциального потока данных.

– Internet Security Association and Key Management Protocol (ISAKMP); используется для первичной настройки соединения, взаимной аутентификации конечными узлами друг друга и обмена секретными ключами. Протокол предусматривает использование различных механизмов обмена ключами, включая задание фиксированных ключей, использование таких протоколов, как Internet Key Exchange, Kerberized Internet Negotiation of Keys (RFC 4430) или записей DNS типа IPSECKEY (RFC 4025).

OpenVPN – свободная реализация технологии ВЧС с открытым исходным кодом для создания зашифрованных каналов типа «точка-точка» или «сервер-клиенты» между компьютерами. Для обеспечения безопасности управляющего канала и потока данных OpenVPN использует библиотеку OpenSSL. Благодаря этому задействуется весь набор алгоритмов шифрования, доступных в данной библиотеке. Также может использоваться пакетная авторизация HMAC для обеспечения большей безопасности и аппаратное ускорение для улучшения производительности шифрования. Эта библиотека использует OpenSSL, а точнее протоколы SSLv3/TLSv1.

На данный момент разработано несколько моделей оптимального распределения ресурсов ВЧС с целью обеспечения заданного качества обслуживания. Предпочтительной является потоковая модель, в которой взаимодействие каждой конечной точки ВЧС с сетью

называется потоком и характеризуется полосой пропускания, требуемой для совокупного входящего и исходящего потоков данных между конечной точкой и сетью [1].

Под задачей обеспечения живучести ВЧС понимают определение множества возможных обходных маршрутов для каждого множества отказавших ресурсов и выбор оптимального варианта по критерию минимальной стоимости [2, 3]. Обычно для уменьшения вычислительной сложности задачу упрощают, рассматривая отказ только одного типа ресурса – звена данных, связывающего два смежных узла сети [1]. Так, в работе [3] рассматривается алгоритм проектирования защищенной ВЧС иерархической структуры для модели с кратковременными отказами одного звена. В модели предполагается, что в один момент времени отказать может только одно звено и что его работоспособность будет восстановлена до того, как откажет другое звено.

Существуют две основные стратегии обеспечения живучести ВЧС АСУ ПП: защита звена (line protection) и защита пути (path protection) [1, 2].

Защита звена предполагает построение обходного маршрута, связывающего конечные точки отказавшего звена. Поток данных будет направляться по обходному маршруту при достижении одной из конечных точек отказавшего звена.

Защита пути подразумевает построение нового маршрута между конечными точками ВЧС, по которому поток данных будет направляться в обход неисправного участка сети.

В ВЧС АСУ ПП, в отличие от обычной пользовательской сети, стратегии защиты звена и пути используют перечисленные выше протоколы для обеспечения конфиденциальности передаваемых данных.

ПОСТАНОВКА ЗАДАЧИ И МЕТОДЫ РЕШЕНИЯ

Критерием задачи оптимизации ВЧС АСУ ПП является максимальная загрузка сети.

Многие проблемы, возникающие в различных областях знания, могут быть сформулированы как задачи теории графов [7].

В теории неориентированных графов потоковая модель представляется в виде неориентированного графа $G = (V, E)$, где V – совокупность вершин графа (например, маршрутизаторы/коммутаторы сети), E – совокупность ребер графа. Пусть C – стоимость резервирования полосы пропускания в сети для ВЧС. Пусть Q – совокупность конечных точек ВЧС, которая является подмножеством совокупности вершин графа V , $Q \subseteq V$. Каждая конечная точка ВЧС $v \in Q$ характеризуется двумя полосами пропускания: B_v^{in} (максимальное значение потока данных, которое может принять конечная точка v от других конечных точек ВЧС) и B_v^{out} (максимальное значение потока данных, которое может отправить конечная точка v другим конечным точкам ВЧС).

Рассмотрим иерархическую топологию ВЧС. Если считать, что каждая вершина $v \in Q$ является конечной точкой дерева T , то существует отдельное ребро, инцидентное вершине v . Пусть (u, v) – любое ребро дерева T , тогда граф в виде дерева легко может быть разделен на два подграфа $(T_u^{(u,v)}$ и $T_v^{(u,v)})$ с помощью удаления ребра (u, v) . Будем считать граф $T_u^{(u,v)}$ поддеревом дерева T , содержащим вершину u , когда ребро (u, v) удалено из дерева T . Резервируемая полоса пропускания на таком ребре должна быть равна максимальному суммарному потоку данных между двумя множествами конечных точек ВЧС, которые связаны этим ребром. Необходимо отметить, что в дереве T весь поток данных из одной конечной точки ВЧС в другую передается по единственному пути.

Рассмотрим ребро (u, v) , которое связывает два множества конечных точек $Q_u^{(u,v)}$ и $Q_v^{(u,v)}$, где $Q_u^{(u,v)}$ – совокупность конечных точек поддерева $T_u^{(u,v)}$. Поток данных из вершины u в вершину v не может превышать минимальное из двух значений: суммарный исходящий поток данных конечных точек множества $Q_u^{(u,v)}$ и суммарный входящий поток данных множества конечных точек $Q_v^{(u,v)}$. Следовательно, резервируемая полоса пропускания

$C_T(u, v)$ на ребре (u, v) дерева T в направлении от вершины u к вершине v задается выражением

$$C_T(i, j) = \min \left\{ \sum_{i \in Q_u^{(u,v)}} B_i^{out}, \sum_{i \in Q_v^{(u,v)}} B_i^{in} \right\}. \quad (1)$$

Из (1) следует, что суммарная резервируемая полоса пропускания C_T на всех ребрах дерева T рассчитывается по выражению

$$C_T = \sum_{(u,v) \in T} C_T(u, v). \quad (2)$$

При этом ребра (u, v) и (v, u) – два различных ребра в дереве T .

В работе [4] предложено использовать дерево Штейнера (наикратчайшее дерево, стягивающее множество из заданных точек [5]) для нахождения оптимальной топологии ВЧС. Необходимо отметить, что дерево Штейнера не всегда оптимально с точки зрения резервируемой полосы пропускания. Подробный анализ проведен в работе [1].

Рассмотрим частный случай, когда поток данных каждой конечной точки v ВЧС симметричен, то есть $B_v^{in} = B_v^{out}$, поэтому можно опустить верхние индексы *in*, *out* и обозначить поток данных конечной точки v , создаваемый в обоих направлениях как B_v . Тогда для ВЧС с симметричным потоком данных конечных точек выражение (1) примет вид

$$C_T(u, v) = \min \left\{ B(Q_u^{(u,v)}), B(Q_v^{(u,v)}) \right\}. \quad (3)$$

Очевидно, что $C_T(u, v) = C_T(v, u)$.

Пусть для дерева T и некоторой вершины r в нем определена следующая величина:

$$R(T, r) = 2 \sum_{i \in Q} B_i l_T(r, i), \quad (4)$$

где $l_T(r, i)$ – количество ребер между вершинами r и i в дереве T (длина единственного пути). Тогда можно показать, что резервируемая полоса пропускания C_T любого дерева T , конечные точки которого являются вершинами из совокупности Q , удовлетворяет следующим свойствам [1]:

1. Существует такая вершина $i \in T$, что $C_T = R(T, i)$.
2. Для всех вершин $v \in T$ справедливо неравенство $C_T \leq Q(T, v)$.

Для нахождения оптимального дерева с минимальным значением резервируемой полосы пропускания необходимо построить деревья для каждой конечной точки графа и определить дерево T_r с минимальным значением полосы пропускания $R(T_r, r)$.

Далее необходимо провести проверку возможности резервирования полосы пропускания $C_T(e)$ на ребре e . Очевидно, что после удаления ребра e образуется два поддерева T_1 и T_2 . Тогда между любой парой вершин $u \in T_1$ и $v \in T_2$ существует только один уникальный путь P_{uv} , при этом $e \in P_{uv}$.

Следовательно, максимально возможный поток данных ребра e для любой допустимой матрицы потока данных Y определяется выражением

$$\begin{aligned} & \sum_{u \in Q \cap Q_u^{(u,v)}, v \in Q \cap Q_v^{(u,v)}} y(u, v) + y(v, u) \leq \\ & \min \left\{ \sum_{u \in Q \cap Q_u^{(u,v)}} B_u^{in}, \sum_{u \in Q \cap Q_v^{(u,v)}} B_u^{out} \right\} + \\ & \min \left\{ \sum_{v \in Q \cap Q_u^{(u,v)}} B_v^{out}, \sum_{v \in Q \cap Q_v^{(u,v)}} B_v^{in} \right\}, \end{aligned} \quad (5)$$

где $y(u, v)$ – поток данных между вершинами u и v в дереве T .

Из формулы (5) можно сделать вывод, что для любого ребра $e \in E(T)$ требуемая резервируемая полоса пропускания $C_T(e)$ на ребре e для любой допустимой матрицы потока данных Y должна удовлетворять условию

$$\begin{aligned} C_T(e) \geq & \min \left\{ \sum_{u \in Q \cap Q_u^{(u,v)}} B_u^{in}, \sum_{u \in Q \cap Q_v^{(u,v)}} B_u^{out} \right\} + in \left\{ \sum_{v \in Q \cap Q_u^{(u,v)}} B_v^{out}, \sum_{v \in Q \cap Q_v^{(u,v)}} B_v^{in} \right\}. \quad (6) \\ C_T(e) \geq & \min \left\{ \sum_{u \in Q \cap Q_u^{(u,v)}} B_u^{in}, \sum_{u \in Q \cap Q_v^{(u,v)}} B_u^{out} \right\} + \\ & in \left\{ \sum_{v \in Q \cap Q_u^{(u,v)}} B_v^{out}, \sum_{v \in Q \cap Q_v^{(u,v)}} B_v^{in} \right\}. \end{aligned}$$

Следовательно, для получения минимально возможного резервирования полосы пропускания $C_T(e)$ в дереве T для всех ребер $e \in E(T)$ должно выполняться равенство (6).

Задачей обеспечения живучести СПД АСУ ПП при одиночном отказе какого-либо звена является задача выбора набора резервных ребер $f \in G - T$ и решение вопросов маршрутирования потока данных по запасному пути, при этом в запасном пути различаются ребра, принадлежащие и не принадлежащие дереву T .

Пусть P_{uv} – путь, который проходит от вершины $u \in T$ до вершины $v \in T$, при этом путь проходит через вершины графа $G(V, E)$, не принадлежащие дереву T , то есть путь $P_{uv} \in G - T$. Следовательно, граф $G' = G \cup P_{uv}$ содержит цикл c_{uv} , при этом, если из графа G' удалить любое ребро e из пути P_{uv} , инцидентное вершине u или v , то получим дерево $T' = G' - e$, по которому может передаваться поток данных конечных точек ВЧС. Следовательно, ребра $f \in P_{uv}$ выступили в роле защиты ребер $e \in T \cap c_{uv}$ дерева T .

Очевидно, что любое ребро $f \in P_{uv}$ может входить в другой путь $P_{xy}: x \in T, y \in T$, соответственно, быть защитным для ребер $e \in T \cap f_{uv}$.

Рассмотрим способ расчета величины защитной полосы пропускания. Пусть $E(f)$ – множество ребер $e \in T$, для которых ребро f может быть защитным, $\tilde{E}(f)$ – множество ребер $e \in T$, для которых ребро f является защитным. Тогда выберем направление в цикле c_{uv} , сориентируем ребра цикла c_{uv} в соответствии с выбранным направлением и получим контур \bar{c}_{uv} , в котором любое ребро (x, y) направлено от вершины x к вершине y . Будем считать, что ребро $f = (x, y)$ и ребра $(u, v) \in \tilde{E}(f)$ направлены одинаково в соответствующих контурах (т.е. от вершины x к вершине y и от вершины u к вершине v). Тогда величина защитной полосы пропускания, которая должна быть выделена на этом ребре $f = (x, y)$, определяется:

$$b^{\text{заш}}(x, y) = \max_{(u, v) \in \tilde{E}(f)} \{b(v, u)\}. \quad (7)$$

Отметим, что в случае симметричного потока данных $b^{\text{заш}}(x, y) = b^{\text{заш}}(y, x)$. Для наглядности рассмотрим рисунок 1.

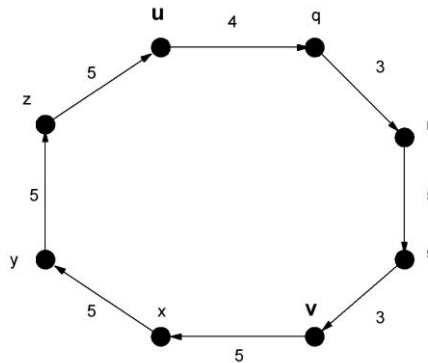


Рисунок 1 – Пример контура \bar{c}_{uv} для выделения защитной полосы пропускания

Контур \bar{c}_{uv} , изображенный на рисунке 1, включает в себя 8 ребер: (u, q) , (q, r) , (r, s) , (s, v) , (v, x) , (x, y) , (y, z) и (z, u) . Для ребра (x, y) множество защищаемых ребер $\tilde{E}(f) = \{(u, q), (q, r), (r, s), (s, v)\}$. Соответственно, защитная полоса на ребре (x, y) равна $b^{\text{заш}}(x, y) = \max\{b(v, u)\} = \max\{4, 3, 5, 3\} = 5$. Аналогичным образом определяется полоса пропускания на остальных защитных ребрах.

В стратегии защиты звена защитный путь должен связать конечные вершины отказавшего ребра. Пусть произошел отказ на ребре $(u, v) \in T$ и существует путь P_{xy} в $G - T$, такой, что все защитные ребра $f \in P_{xy}$ являются защитными для звена (u, v) . Тогда все ребра $f \in P_{xy}$ вместе с ребром $(u, v) \in T$ входят в цикл c_{xy} , следовательно, необходимо

дополнительно выделить $b_{uv}^{доп}(i, j) = b(v, u)$ и $b_{uv}^{доп}(j, i) = b(u, v)$ на всех ребрах e пути P_{xy} , удовлетворяющих условию $e \in T \cap P_{xy} / (u, v)$. Будем считать, что ребро (u, v) и ребра $e \in T \cap P_{xy} - (u, v)$ направлены одинаково в контуре \bar{c}_{xy} [1].

Защитная полоса пропускания для всех ребер $(u, v) \in T$ находится с помощью описанного ниже выражения:

$$b_T^{заш}(x, y) = \max_{\forall (u, v) \in T - (x, y)} \{b_{uv}^{доп}(x, y)\}, \quad (8)$$

Стратегия защиты звена для ВЧС с симметричным потоком данных конечных точек проиллюстрирована на рисунке 2.

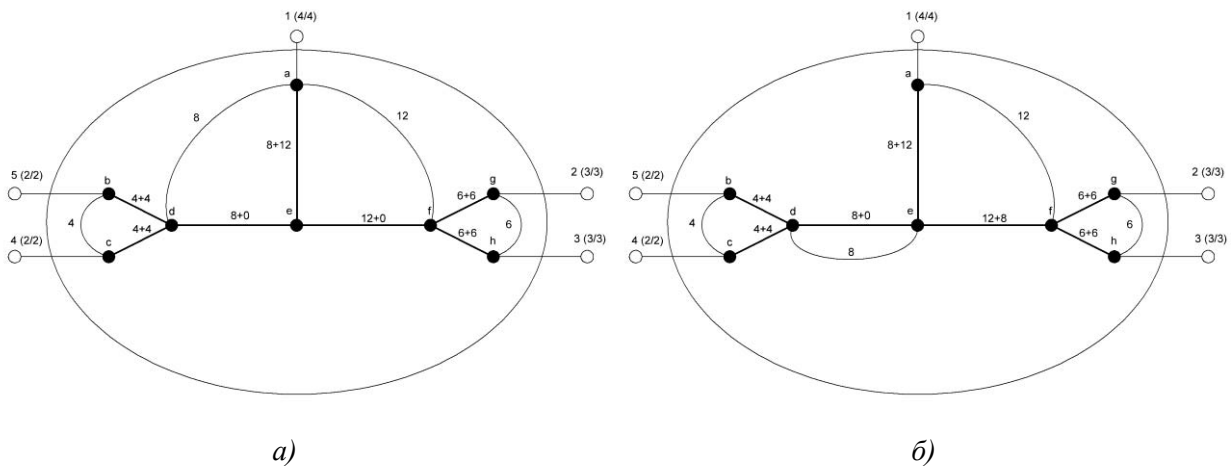


Рисунок 2 – Пример реализации стратегий защиты звена и защиты пути:
 а) – стратегия защиты звена; б) – стратегия защиты пути

На рисунке 2 (а) показано дерево, к которому добавлено 4 звена – (a, f) , (b, c) , (d, e) , (g, h) . Выпишем ребра, защищаемые этими звеньями: $\bar{E}(a, f) = \{(a, e), (e, f)\}$, $\bar{E}(b, c) = \{(b, d), (c, d)\}$, $\bar{E}(d, e) = \{(d, e)\}$, $\bar{E}(g, h) = \{(g, f), (h, f)\}$. Тогда по формуле (7) получаются следующие значения защитных полос пропускания на этих звеньях: $b^{заш}(a, f) = \max\{b(a, e), b(e, f)\} = 12$, $b^{заш}(b, c) = \max\{b(b, d), b(c, d)\} = 4$, $b^{заш}(d, e) = 8$, $b^{заш}(g, h) = \max\{b(g, f), b(h, f)\} = 6$.

Тогда дополнительная полоса пропускания для всех ребер $e \in T$, для всех защитных путей: $b_{ae}^{доп}(e, f) = 12$, $b_{ae}^{доп}(f, e) = 8$, $b_{bc}^{доп}(b, d) = 4$, $b_{bc}^{доп}(c, d) = 4$, $b_{de}^{доп}(d, e) = 0$, $b_{gh}^{доп}(g, f) = 6$, $b_{gh}^{доп}(h, f) = 6$.

После этого необходимо воспользоваться выражением (8) и найти максимальные значения дополнительной полосы пропускания для каждого ребра $e \in T$.

Просуммировав результаты, полученные из выражений (7) и (8), мы получим величину суммарной защитной полосы пропускания для обеспечения живучести всего дерева T методом защиты звена. В нашем примере она равна 70.

В стратегии защиты пути защитный путь должен связать два поддерева $T(u, v)$ и $T(v, u)$ дерева T в случае отказа ребра (u, v) . Пусть существует путь P_{xy} в $G - T$, такой, что все защитные ребра $f \in P_{xy}$ являются защитными для звена (u, v) , то есть пакеты в ВЧС будут передаваться по образованному дереву $T_{uv} = T - (u, v) + P_{xy}$. В этом случае на защитных ребрах $f \in T - (u, v)$ необходимо выделить дополнительную полосу пропускания, рассчитываемую по выражению:

$$b_{uv}^{доп}(x, y) = \max\{0, \min\{\sum_{i \in Q \cap T_{uv}(x, y)} b_{out}(i), \sum_{j \in Q \cap T_{uv}(y, x)} b_{in}(j)\} - b(x, y)\} \quad (9)$$

Защитная полоса пропускания для всех ребер $(u, v) \in T$ находится с помощью описанного ниже выражения:

$$b_T^{\text{зам}}(x, y) = \max_{\forall (u, v) \in T - (x, y)} \{b_{uv}^{\text{доп}}(x, y)\}, \quad (10)$$

На рисунке 2 (б) показано дерево, к которому добавлено 4 звена – (a, f) , (b, c) , (d, e) , (g, h) – (a, f) , (a, d) , (b, c) , (g, h) . Рассмотрим расчет дополнительной полосы пропускания.

Рассчитаем дополнительную полосу пропускания, используя выражение (9), в случае отказа ребра (f, e) дерева T . Так как поток данных симметричный, то для ускорения расчетов мы будем считать суммарный поток данных (входящий и исходящий).

$$b_{f,e}^{\text{доп}}(a, e) = \max\{0, \min\{20, 8\} - 8\} = 12,$$

$$b_{f,e}^{\text{доп}}(f, g) = \max\{0, \min\{6, 22\} - 6\} = 0,$$

$$b_{f,e}^{\text{доп}}(f, h) = \max\{0, \min\{6, 22\} - 6\} = 0$$

и т. д.

Аналогичным образом рассчитываются значения дополнительной полосы пропускания для отказа каждого ребра дерева T . После применения к полученным данным выражения (10) мы получим значения защитной полосы пропускания на каждом ребре. Просуммировав эти значения, значение дополнительной резервируемой полосы пропускания для дерева T равно 62.

Для определения оптимальных защитных ребер $e \in G - T$ в литературе описан аппроксимационный алгоритм [1], который имеет очевидный недостаток – функция стоимости не учитывает полосу пропускания, которая должна быть дополнительно выделена на ребрах дерева ВЧС.

В работе [2] предлагается решение задачи методом смешанного целочисленного линейного программирования, которая отдельно формулируется для резервирования звена и для резервирования пути, который имеет плохую сходимость для сетей больших размеров, но в небольших сетях он более эффективен, чем аппроксимационный алгоритм.

Результаты применения описанных выше алгоритмов приведены в работе [2] для сетей с различной топологией при различном числе конечных точек ВЧС. Из них следует, что стратегия защиты пути наиболее эффективна, аппроксимационный алгоритм наименее эффективен.

ВЫВОДЫ

С использованием теории неориентированных графов описана потоковая модель ВЧС, приведена методика расчета минимально возможной резервируемой полосы пропускания, описана методика расчета необходимой дополнительной полосы пропускания для стратегии защиты звена и стратегии защиты пути, обеспечивающих защиту ВЧС в случае отказа какого-либо звена.

Анализируя зависимость занимаемой полосы пропускания от топологии сети, количества конечных точек ВЧС и используемой стратегии резервирования, можно сделать вывод, что в общем случае наибольшей эффективностью обладают стратегия защиты звена и стратегия защиты пути. Исходя из этого, целесообразна разработка гибридной стратегии защиты ВЧС, использующей резервирование звена и резервирование пути в зависимости от различных факторов, которые необходимо определить в дальнейшей работе.

СПИСОК ЛИТЕРАТУРЫ

1. Росляков А.В. Виртуальные частные сети. Основы построения и применения. – М.: Экотрендз, 2006. – 304 с.
2. Balasubramanian A., Sasaki G. Bandwidth Requirements for the Protected VPNs in the Hose Model // International Symposium on Information Theory. – Kanagawa, 2003. – P. 89-90.
3. Italiano G.F., Rastogi R., Yener B. Restoration Algorithms for Virtual Private Networks in the Hose Model // IEEE INFOCOM, 2002.

4. Duffield N.G. A Flexible Model for Resource Management in Virtual Private Networks / N. Duffield, P. Goyal, A. Greenberg, P. Mishra, K. Ramakrishnan, J.E. van der Merwe // ACM SIGCOMM Computer Communication Review, 1999. – V. 29. – № 4. – P. 95-108.
5. Басакер Р., Саати Т. Конечные графы и сети. – М.: Наука, 1974. – 367 с.
6. Нуштаев А.В., Росляков А.В. Алгоритмы построения отказоустойчивых виртуальных частных сетей // Доклады 60-й научной сессии, посвященной Дню Радио. – М., 2005. – С. 54-57.
7. Ор О. Теория графов. – М.: Наука, 1968. – 336 с.
8. Виртуальные частные сети [Электронный ресурс]. – URL: <http://ru.wikipedia.org/wiki/VPN> (дата обращения: 29.10.2013).
9. Еременко В.Т. Анализ моделей управления трафиком в сетях АСУП на основе технологии MPLS / В.Т. Еременко, С.В. Еременко, Д.В. Анисимов, С.А. Черепков, А.А. Лякишев // Информационные системы и технологии, 2013. – № 1. – С. 106-112.

Борисов Олег Михайлович

ФГБОУ ВПО «Госунiversитет – УНПК», г. Орел
Аспирант кафедры «Электроника, вычислительная техника и информационная безопасность»
Тел.: 8 (4862) 45-57-57
E-mail: atomoleg@mail.ru

Рязанцев Павел Николаевич

ФГБОУ ВПО «Госунiversитет – УНПК», г. Орел
Аспирант кафедры «Электроника, вычислительная техника и информационная безопасность»
Тел.: 8 953 615 04 19
E-mail: evtib_gu_unpk@mail.ru

Кузьмин Виктор Викторович

ФГБОУ ВПО «Госунiversитет – УНПК», г. Орел
Магистр кафедры «Электроника, вычислительная техника и информационная безопасность»
Тел.: 9 960 641 45 36
E-mail: evtib_gu_unpk@mail.ru

Фомин Николай Иванович

ФГБОУ ВПО «Госунiversитет – УНПК», г. Орел
Аспирант кафедры «Электроника, вычислительная техника и информационная безопасность»
Тел.: 8 953 610 63 47
E-mail: evtib_gu_unpk@mail.ru

O.M. BORISOV (*Post-graduate Student of the Department «Electronics, Computer Sciences and Information Security»*)

P.N. RYAZANCEV (*Post-graduate Student of the Department «Electronics, Computer Sciences and Information Security»*)

V.V. KUZ'MIN (*Master of the Department «Electronics, Computer Sciences and Information Security»*)

N.I. FOMIN (*Post-graduate Student of the Department «Electronics, Computer Sciences and Information Security»*)
State University – ESPC, Orel

SURVIVABILITY ALGORITHMS OF VIRTUAL PRIVATE NETWORK OF AUTOMATED CONTROL SYSTEMS INDUSTRIAL ENTERPRISE (VPN ACS IE) BASED ON THE OPTIMAL ALLOCATION OF NETWORK RESOURCES

The article describes a mathematical model of the traffic ACS of industrial enterprise based on the theory of undirected graphs and characterized the optimal allocation of network resources of symmetric traffic, providing path protection and line protection.

Keywords: *virtual private networks; mathematical model; theory of graphs; path protection; line protection.*

BIBLIOGRAPHY (TRANSLITERATED)

1. Roslyakov A.V. Virtual'ny'e chastny'e seti. Osnovy' postroeniya i primeneniya. – M.: E'ko-Trendz, 2006. – 304 s.
2. Balasubramanian A., Sasaki G. Bandwidth Requirements for the Protected VPNs in the Hose Model // International Symposium on Information Theory. – Kanagawa, 2003. – P. 89-90.
3. Italiano G.F., Rastogi R., Yener B. Restoration Algorithms for Virtual Private Networks in the Hose Model // IEEE INFOCOM, 2002.
4. Duffield N.G. A Flexible Model for Resource Management in Virtual Private Networks / N. Duffield, P. Goyal, A. Greenberg, P. Mishra, K. Ramakrishnan, J.E. van der Merwe // ACM SIGCOMM Computer Communication Review, 1999. – V. 29. – № 4. – P. 95-108.
5. Basaker R., Saati T. Konechny'e grafy' i seti. – M.: Nauka, 1974. – 367 s.
6. Nushtaev A.V., Roslyakov A.V. Algoritmy' postreniya otkazoustojchivy'x virtual'ny'x chastny'x setej // Doklady' 60-j nauchnoj sessii, posvyashhyonnoj Dnyu Radio. – M., 2005. – S. 54-57.
7. Or O. Teoriya grafov. – M.: Nauka, 1968. – 336 s.
8. Virtual'ny'e chastny'e seti [E'lektronny'j resurs]. – URL: <http://ru.wikipedia.org/wiki/VPN> (data obrashheniya: 29.10.2013).
9. Eryomenko V.T. Analiz modelej upravleniya trafikom v setyax ASUP na osnove texnologii MPLS / V.T. Eryomenko, S.V. Eryomenko, D.V. Anisimov, S.A. Cherepkov, A.A. Lyakishev // Informacionny'e sistemy' i texnologii, 2013. – № 1. – S. 106-112.

ПРЕДСТАВЛЕНИЕ САМООРГАНИЗУЮЩЕЙСЯ ИТ-ИНФРАСТРУКТУРЫ В ВИДЕ МОДЕЛИ ЖИЗНЕСПОСОБНОЙ СИСТЕМЫ

В работе описана модель самоорганизующейся ИТ-инфраструктуры предприятия в виде модели жизнеспособной системы. Такая модель интегрирует в одно представление как непосредственно саму ИТ-инфраструктуру, предоставляющую сервисы автоматизации, так и системы долгосрочного и оперативного управления ее в процессе жизненного цикла с учетом рекурсивно-иерархического их принципа организации. Представление самоорганизующейся ИТ-инфраструктуры в виде модели жизнеспособной системы позволяет идентифицировать компоненты системы управления и их функции.

Ключевые слова: ИТ-инфраструктура; административное управление; архитектура; эффективность; модель жизнеспособной системы; системная динамика; жизненный цикл.

ВВЕДЕНИЕ

ИТ-инфраструктура предприятия в процессе своего жизненного цикла управляется соответствующей системой управления, называемой административной. Долгосрочное административное управление заключается в создании и утилизации компонентов ИТ-инфраструктуры в зависимости от потребностей предприятия и изменения условий внешней среды. Системная сложность управления развитием ИТ-инфраструктуры из-за высокой неопределенности и сложности объекта управления предъявляет соответствующие требования к организации системы административного управления (САУ), обоснование и разработка архитектуры которой является актуальной задачей.

С точки зрения кибернетики возможно совместное рассмотрение ИТ-инфраструктуры и ее системы управления. При таком объединении рассматривается самоорганизующаяся ИТ-инфраструктура, способная структурно и функционально развиваться с течением времени. Для описания архитектур рекурсивно-организованных систем, таких как ИТ-инфраструктуры предприятий, с точки зрения кибернетического подхода применяется модель жизнеспособной системы С. Бира. Необходимость обоснования архитектуры системы административного управления определяет актуальность представления самоорганизующейся ИТ-инфраструктуры в соответствии с моделью жизнеспособной системы, в рамках которой выделяются формализованные компоненты управления различного уровня.

АНАЛИЗ ИЗВЕСТНЫХ ПОДХОДОВ К ПРЕДСТАВЛЕНИЮ САМООРГАНИЗУЮЩИХСЯ ОРГАНИЗАЦИОННО-ТЕХНИЧЕСКИХ СИСТЕМ В ВИДЕ МОДЕЛИ ЖИЗНЕСПОСОБНОЙ СИСТЕМЫ

Для описания сложных организационно-технических систем в кибернетике предложена модель жизнеспособной системы (МЖС) [1]. В соответствии с указанной моделью любую самоорганизующуюся систему (жизнеспособную систему) можно представить в виде рекурсивной структуры, содержащей пять (шесть) основных подсистем:

- «система 1» представляет собой жизнеспособную систему и реализует одну из функций основного вида деятельности системы (функциональная подсистема самоорганизующейся системы);
- «система 2» предоставляет собой каналы управления и передачи информации между системами «1» и «3», благодаря чему реализуется регуляция всех «систем 1»;
- «система 3» реализует оперативное управление (регуляцию) «систем 1» (путем оптимизации управления ресурсами, их распределением и расходом), также выделяется «система 3*», реализующая независимый аудит «систем 1»;

- «система 4» реализует адаптивное управление функциональными подсистемами за счет учета требований и изменений внешней среды;
- «система 5» отвечает за политики управления и тем самым реализует долгосрочное управление сложной системой.

Модель жизнеспособной системы нашла широкое применение в моделировании организационных и организационно-технических систем [2]. В литературе [2-4] описаны различные варианты архитектур систем управления предприятием на основе указанной модели. Так, в [3] рассмотрено применение методологии жизнеспособных систем для управления архитектурой предприятия, построенного по иерархичному принципу. В рамках такого подхода в работе описаны функции всех пяти систем МЖС, реализуемые в рамках управления архитектурой на различных уровнях управления. В [4] рассмотрено применение МЖС для управления продукто-ориентированными системами, создание которых произошло вследствие уменьшения длительности жизненного цикла сложных продуктов. Следует отметить, что ИТ-инфраструктуру предприятия также можно рассматривать как интеллектуальную сложную продукто-ориентированную систему. В рамках такого подхода интеллектуальный продукт рассматривается как жизнеспособная система. Для указанных продуктов приведены спецификации компонентов жизнеспособной системы, на основе которых предложена архитектура системы производственного планирования. Для управления виртуальными предприятиями и, в частности, виртуальными сообществами в [5] предложено использовать модель жизнеспособной системы. При этом рассмотрены вопросы контролирования и координации сообщества со стороны центра управления, функции которого определены путем идентификации всех компонентов МЖС.

Совершенствование методологии жизнеспособных систем связано с учетом их динамики, а также с разработкой методов анализа и формальных языков моделирования. Развитие теории организационной кибернетики за счет объединения двух ее аспектов – системной динамики и моделей жизнеспособных систем – рассмотрено в [6]. Такое объединение позволяет перейти от статических моделей самоорганизующихся систем к их динамическим моделям, предполагающим учет развития объекта в процессе всего жизненного цикла. Для разработки интегрированных моделей организационных систем в соответствии МЖС в [7] рассмотрены шаблоны компонентов модели жизнеспособной системы и связей между ними. Выделено 9 основных типов взаимодействий, реализованных в виде шаблонов: выделенное (*separate*) управление, операционное (*operational*) управление, регуляция (*regulatory center*), аудит (*sporadic audit*), адаптивное (*adaptive*) управление, супервизорный (*supervisory*) контроль, тревоги (*alerts*), рекурсивное объединение (*recursive composition*), гомеостатическая петля (*homeostatic loop*). Представлены модели и формальные описания указанных шаблонов в UML-виде. В работе представлено сравнение указанных шаблонов, применяемых в рамках методологии жизнеспособных систем, с аналогичными, в том числе связанными и с другими функциями проектирования: шаблоном проектирования адаптивной стратегии (*adaptive strategy*), шаблонами проектирования архитектуры – программной тектоники (*software tectonics*), гибкой базы (*flexible foundations*) и метаморфизма (*metamorphosis*), шаблонами в рамках динамической объектной модели (*dynamic object model*) и адаптивной объектной модели (*adaptive object model*). Разработаны методы анализа и проектирования жизнеспособных систем, в частности, в [8] рассмотрено применение методологии VEA (*viable enterprise analysis*) как средства анализа развития предприятия, предполагающего его представление в виде модели жизнеспособной системы. Целью такого анализа является выявление информационных факторов (требований), важных для информационной системы стратегического управления предприятием. В рамках методологии для анализа предприятия предложена модель анализа жизнеспособности (*viable analysis model – ViAM*).

ИТ-инфраструктура с системой административного управления является многоуровневой иерархической организационно-технической системой [9]. Для создания и

управления виртуальными системами управления такого типа применяется рекурсивно-иерархический подход [10], предполагающий создание виртуальных систем административного управления для управления всеми выделенными функциональными ИТ-компонентами на всем их жизненном цикле. Такое представление ИТ-инфраструктуры и ее системы административного управления определяет возможность ее рассмотрения как жизнеспособной системы.

В литературе [11-14] применение методологии жизнеспособных систем применительно к организации ИТ-инфраструктур предприятий нашло широкое применение. Модели жизнеспособных систем используются как для описания как всей ИТ-инфраструктуры в целом [11], так и ее выделенных аспектов: программного обеспечения [12, 13], информационных систем [15], компонентов управления проектами [14] и политиками [16] в ИТ-сфере. Так, в [11] представлен подход к проектированию и созданию ИТ-инфраструктуры крупного образовательного учреждения на основе модели жизнеспособной системы. В рамках такого подхода, учитывая иерархический принцип организации систем управления, представлена модель центра управления ИТ-сферой ВУЗа. Центр управления ИТ-инфраструктурой, в свою очередь, сам представлен жизнеспособной системой в соответствии с методологией ITIL. Для программных систем рассматривалось применение МЖС для решения задач реинжиниринга архитектур программных систем [12]. В [13] представлена архитектура жизнеспособной системы и методы ее использования для проектирования программных систем. Отмечены основные особенности жизнеспособных программных систем, такие, как: автономность, адаптивность, рекурсивность, иерархичность, инвариантность и самоподобность (self-reference). Для разработки жизнеспособного программного обеспечения предложена компонентная методология, включающая абстрактную архитектуру компонентов. Для реализации методологии описаны основные интерфейсы абстрактной архитектуры, структура которых спроектирована в соответствии с МЖС.

В [15] рассмотрено применение МЖС для обеспечения гибкости информационных систем. Отмечена важность обеспечения гибкости информационных системы к изменению внешних условий, подтверждаемая требованиями набора практических рекомендаций CobIT, которые ставят целью обеспечение гибкости. Разрыв между теорией и практикой в области гибкости информационных систем заключается в узкости понятия практической гибкости, невозможности измерения гибкости в рамках соответствующей обеспечивающей функции, сложности обеспечения поддержки таких опций, которые могут быть немедленно применены. Применение МЖС для проектирования гибких информационных систем показала значимость применения сформулированных рекомендаций на практике. В [14] рассмотрено применение МЖС для управления проектами области создания ИТ-компонентов. В [15] обосновывается применение МЖС для анализа и разработки политики управления ИТ-инфраструктурой организации на долгосрочный период.

Анализ применения методологии жизнеспособных систем в представленных работах показывает статичность объекта управления – моделей функциональных подсистем ИТ-инфраструктуры – «систем 1» МЖС, что не учитывает их изменение в соответствии с особенностями жизненного цикла. Необходимость управления ИТ-инфраструктурой в целом и всеми ее компонентами на всем жизненном цикле [17] определяет актуальность адаптации модели жизнеспособной системы в соответствии с 4D структурой объекта управления (эволюционирующим характером ИТ-инфраструктуры) системы административного управления. Это позволит идентифицировать компоненты САУ в соответствии с моделью жизнеспособной системы.

ФОРМАЛИЗАЦИЯ САМООРГАНИЗУЮЩЕЙСЯ ИТ-ИНФРАСТРУКТУРЫ В ВИДЕ МОДЕЛИ ЖИЗНЕСПОСОБНОЙ СИСТЕМЫ

С целью обеспечения гибкости управления ИТ-инфраструктурой на всем жизненном цикле рассмотрим ее представление в виде модели жизнеспособной системы. В рамках разработки такой модели применим сервис-ориентированный подход [18] к описанию самоорганизующейся ИТ-инфраструктуры, в рамках которого предполагается, что она состоит из набора высокоуровневых прикладных ИТ-сервисов. ИТ-сервис предоставляет одну услугу (автоматизации, информатизации или телекоммуникационную) потребителям из состава расширенного предприятия. Каждый ИТ-сервис в процессе своего функционирования задействует ресурсы и возможности, представляемые сервисами более низкого уровня.

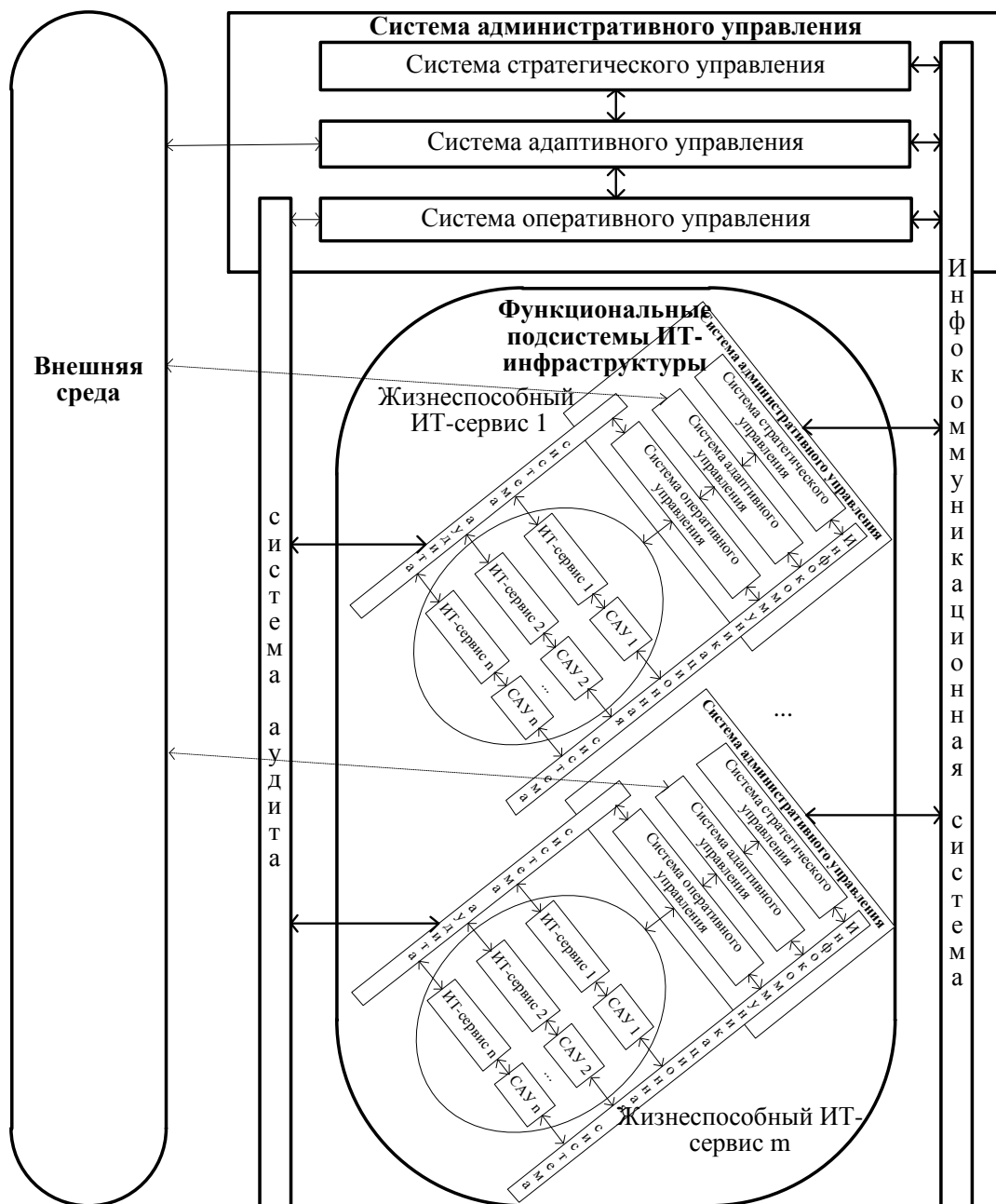


Рисунок 1 – Формализация САУ ИТ-инфраструктурой в виде модели жизнеспособной системы

В рамках представления самоорганизующейся ИТ-инфраструктуры в нотации жизнеспособных систем (рис. 1) в статическом виде она включает в свой состав систему административного управления (объединяющую системы «3», «4» и «5»), объект управления (управляемую ИТ-инфраструктуру в виде совокупности самоорганизующихся ИТ-сервисов – множество «систем 1»), инфокоммуникационную систему мониторинга и управления, а также независимый канал аудита. В процессе функционирования жизнеспособная ИТ-инфраструктура реализует взаимодействие с внешней средой, к изменению которой осуществляется адаптация всей системы.

Рассмотрим соответствие компонентов самоорганизующейся ИТ-инфраструктуры и, в первую очередь, ее системы административного управления виртуального типа с абстрактной моделью жизнеспособной системы. В соответствии с применяемым методом исследования для формализации модели должны быть идентифицированы все компоненты жизнеспособной системы административного управления объектами ИТ-инфраструктуры. Результаты такого представления для верхнего уровня иерархии приведены в таблице 1.

Таблица 1 – Отображение компонентов САУ ИТ-инфраструктурой виртуального типа на модель жизнеспособной системы

Компонент жизнеспособной системы	Основная функция	Состав жизнеспособной ИТ-инфраструктуры		Реализация функции в САУ ИТ-инфраструктуры
Система 5	Стратегическое управление	Система административного управления	Руководство организации, ИТ-директор	Управление политикой ИТ-инфраструктуры организации
Система 4	Адаптивное управление		ИТ-директор, руководители функциональных подразделений	Изменение ИТ-инфраструктуры под изменение внешних условий и внутренних потребностей
Система 3	Оперативное управление		ИТ-менеджеры, менеджеры функциональных подразделений	Поддержание ИТ-инфраструктуры в требуемом состоянии
Система 3*	Аудит		Подразделение аудита	Определение эффективности информатизации организации
Система 2	Регуляция		Информационно-коммуникационная среда взаимодействия	Мониторинг и передача воздействий
Система 1	Функциональная	ИТ-инфраструктура	Множество прикладных ИТ-сервисов	Предоставление и поддержка ИТ-сервиса

Предложенный вариант представления самоорганизующейся ИТ-инфраструктуры в виде модели жизнеспособной системы и результаты аналогичных исследований [11, 15] позволяют сделать вывод о возможности такой формализации ИТ-инфраструктуры вместе с системой административного управления. Представление ИТ-инфраструктуры в таком виде позволяет повысить адекватность ее описания как самоорганизующейся системы с учетом ее иерархичности. Основным выводом от описания САУ вместе с ИТ-инфраструктурой является обоснование необходимости интеграции управления на трех уровнях (стратегическом, адаптивном и оперативном) в рамках интегрированной системы

управления. Интеграция такого управления с наделением компонентов соответствующими функциональными возможностями позволяет повысить эффективность управления развитием ИТ-инфраструктуры и степень ее соответствия потребностям и изменению внешней среды.

В соответствии с сервис-ориентированным представлением ИТ-инфраструктуры она состоит из множества прикладных ИТ-сервисов. В рамках такого подхода жизнеспособный ИТ-сервис представляет собой систему из предоставляемой пользователям инфокоммуникационной услуги, обеспечивающую предоставление такой услуги инфраструктуры и компонента управления, ответственного за развитие ИТ-сервиса на всем его жизненном цикле. Методология жизнеспособных систем определяет самоподобие систем, находящихся на разных уровнях иерархии. Однако, несмотря на такое самоподобие, состав компонентов жизнеспособной системы для уровня прикладного ИТ-сервиса и всей ИТ-инфраструктуры будет различным. Это определяет необходимость идентификации компонентов прикладного жизнеспособного ИТ-сервиса. Типовой состав компонентов ИТ-сервиса в виде МЖС представлен в таблице 2.

Таблица 2 – Типовой состав жизнеспособного ИТ-сервиса

Компонент жизнеспособной системы	Основная функция	Состав жизнеспособного ИТ-сервиса	Реализация функций ИТ-сервисе
Система 5	Стратегическое управление	ИТ-директор, руководители функциональных подразделений	Управление жизненным циклом ИТ-сервиса
Система 4	Адаптивное управление	ИТ-менеджеры, менеджеры функциональных подразделений	Изменение ИТ-сервиса (модификации/реинжиниринг) под изменение внешних условий и внутренних потребностей
Система 3	Оперативное управление	ИТ-администраторы, специалисты функциональных подразделений	Поддержание ИТ-сервиса в требуемом состоянии
Система 3*	Аудит	Аудиторы	Определение эффективности сервиса на всем жизненном цикле
Система 2	Регуляция	Информационно-коммуникационная среда взаимодействия	Мониторинг и передача воздействий по управлению
Система 1	Функциональная	Множество системных ИТ-сервисов	Предоставление и поддержка системных ИТ-сервисов

Представленный в таблице 2 типовой состав ИТ-сервиса показывает понижение уровня лиц (структур) управления, отвечающих за реализацию, предоставление и развитие сервиса по сравнению с уровнем всей ИТ-инфраструктуры.

Рисунок 1, таблицы 1 и 2 определяют архитектуру и идентифицируют состав модели самоорганизующейся ИТ-инфраструктуры в виде жизнеспособной системы. Применение такой статической модели позволяет идентифицировать контуры управления ИТ-инфраструктурой и соотнести их с функциями управления, специфицируемыми моделью жизнеспособной системы. Это позволяет при разработке ИТ-инфраструктуры и ее системы управления осуществлять назначение требуемых для адекватного управления полномочий, возможностей и ресурсов.

МОДЕЛЬ СИСТЕМНОЙ ДИНАМИКИ ЖИЗНЕСПОСОБНОЙ ИТ-ИНФРАСТРУКТУРЫ

Необходимость управления ИТ-инфраструктурой и ее компонентами на всем жизненном цикле определяет необходимость учета системной динамики при построении модели жизнеспособной системы. В организационной кибернетике для учета развития системы применяют совместно с моделью жизнеспособной системы также и модель системной динамики [6]. Использование такого подхода предполагает модификацию модели ИТ-инфраструктуры в виде жизнеспособной системы (рис. 1) с учетом особенностей ее развития в процессе жизненного цикла.

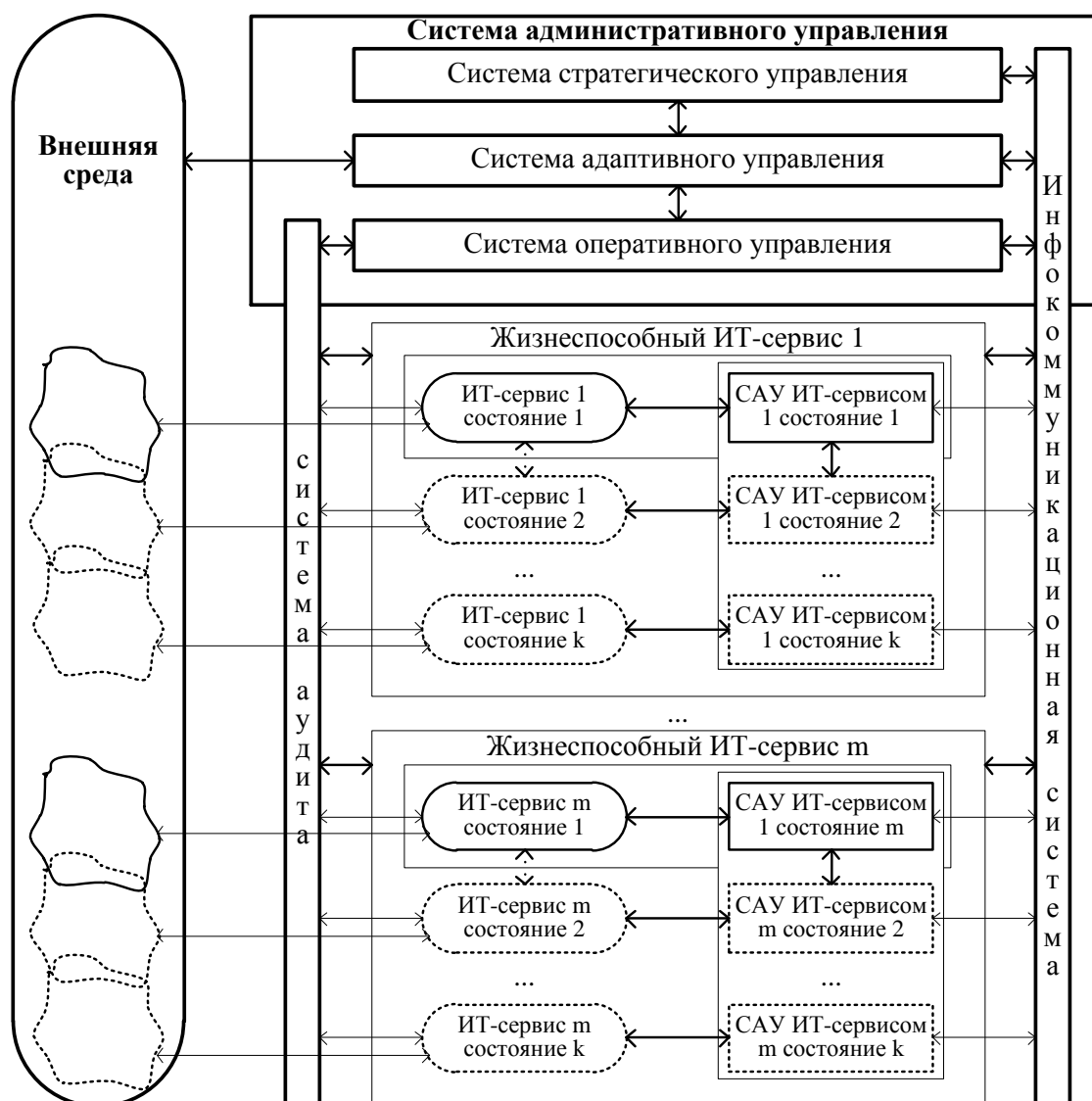


Рисунок 2 – Жизнеспособная ИТ-инфраструктура с учетом системной динамики

В соответствии с используемым подходом функциональные компоненты жизнеспособной системы (ИТ-компоненты) под воздействием управляющих компонентов (системы «3», «4» и «5») изменяют свое состояние с целью соответствия потребностям надсистемы и внешней среде. Разработанная модифицированная модель жизнеспособной ИТ-инфраструктуры (рис. 2) отражает системную динамику ИТ-компонентов. В соответствии с этой моделью «система 4» реализует адаптацию множества жизнеспособных ИТ-сервисов («систем 1»). При этом реализуется соответствующее изменение систем «3», «3*» и «2» с целью обеспечения поддержки всех функциональных подсистем. «Система 2»

должна обеспечивать взаимодействие в процессе развития на требуемом качественном уровне, обеспечивая мониторинг и регуляцию «систем 1», инвариантно к структуре и составу функциональных компонентов.

Модель системной динамики ИТ-инфраструктуры в виде жизнеспособной системы отражает развитие ИТ-инфраструктуры в целом (самоорганизацию системы). Перевод ИТ-сервисов из одного состояния в другое реализуется системой адаптивного управления САУ. Она на основе анализа внешней среды с учетом политики определенной системы долгосрочного управления («система 5») осуществляет перевод системы из одного состояния в другое. На более низких уровнях организации учет развития в процессе жизненного цикла реализуется аналогично.

АДАПТИВНОЕ УПРАВЛЕНИЕ ЖИЗНЕСПОСОБНОЙ ИТ-ИНФРАСТРУКТУРОЙ

Необходимость управления жизнеспособной ИТ-инфраструктурой на всем жизненном цикле [17] обосновывает актуальность разработки методов административного управления. Модель жизнеспособной системы определяет рекурсивный характер построения сложных самоорганизующихся систем. Применение такого подхода к управлению ИТ-инфраструктурой [10] требует описания методов управления. На уровне всего жизненного цикла важным аспектом управления является адаптивное (реализуемое «системой 4»). Адаптивное управление ИТ-инфраструктурой заключается в обеспечении ее соответствия потребностям предприятия и внешней среды. В соответствии с моделью жизнеспособной системой самоорганизующаяся ИТ-инфраструктура в статическом виде описывается кортежем из совокупности систем $S_1 \dots S_5$ и связей r между ними:

$$IT = \langle S_5, S_4, S_3, S_3^*, S_2, \{S_{1i}\}, r \rangle.$$

В соответствии с используемой моделью управление жизненным циклом реализуется подсистемой адаптивного управления («системой 4») S_4 . Поскольку управляемый компонент ИТ-инфраструктуры $\{S_{1i}\}$ в сервис-ориентированном представлении включает множество n ИТ-сервисов: $\{s_1, \dots, s_i, \dots, s_n\}$, то управление ими возможно путем реализации дискретных управляющих воздействий. При этом каждый i -й ИТ-сервис может находиться во множестве j -й состояний: s_i^j . Адаптивное управление в ИТ-инфраструктуре предполагает реализацию трех типов управляющих воздействий (создание, модификация и утилизация i -го ИТ-сервиса). Типы управляющих воздействий определяют характер выполняемых системой оперативного управления («системой 3») операций по модификации ИТ-инфраструктуры. Каждое управляющее воздействие переводит управляемый компонент ИТ-инфраструктуры $\{S_{1i}\}$ из состояния S^t в состояние S^{t+1} . Формальное описание управляющих воздействий имеет следующий вид:

– управляющее воздействие (u_c) по созданию s_i^j ИТ-сервиса трансформирует ИТ-инфраструктуру путем добавления нового сервиса ко множеству имеющихся: $S^t \xrightarrow{u_c} S^{t+1}$, где $S^{t+1} = S^t \cup s_i^j$;

– управляющее воздействие (u_u) по утилизации (выводе из эксплуатации) предоставляемого s_i^j ИТ-сервиса приводит к снижению функциональности ИТ-инфраструктуры путем утилизации ненужного сервиса: $S^t \xrightarrow{u_u} S^{t+1}$, где $S^{t+1} = S^t \setminus s_i^j$;

– управляющее воздействие (u_m) по модификации s_i^j ИТ-сервиса реализует его перевод из одного j -го состояния в другое ($j+1$)-е (более соответствующее внешней среде

и потребностям предприятия): $S^t \xrightarrow{u_m} S^{t+1}$, где $S^{t+1} = S^t \setminus s_i^j \cup s_i^{j+1}$. С формальной точки зрения операция модификации сервиса может быть представлена совокупностью операций по созданию нового ИТ-сервиса и удалению старого. С технологической точки зрения реализуется изменение или качества предоставляемого сервиса, или предоставляющей такой сервис инфраструктуры.

Реализация управляющих воздействий осуществляется по мере необходимости на основе анализа внешней среды в дискретном виде. На практике управление имеет некоторое запаздывание, связанное с необходимостью осуществления преобразования ИТ-инфраструктуры в новое состояние (рис. 3). Адаптивное управление в представленном виде представляет последовательность воздействий по трансформации исходного множества предоставляемых ИТ-сервисов. Ресурсы на адаптацию жизнеспособной ИТ-инфраструктуры затрачиваются из имеющихся в наличии. При необходимости дополнительные ресурсы запрашиваются у вышестоящей системы.

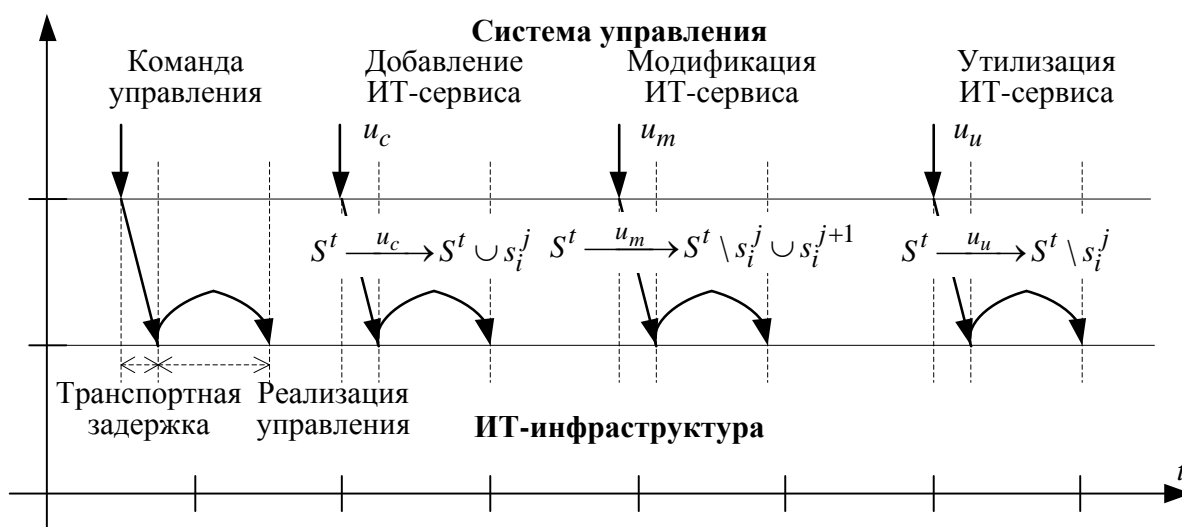


Рисунок 3 – Адаптивное управление жизнеспособной ИТ-инфраструктурой

В соответствии с представленной схемой реализации управления осуществляется адаптивное управление развитием ИТ-инфраструктурой посредством трех управляющих воздействий. При этом схема управления инварианта к уровню управления ИТ-компонентами.

ЗАКЛЮЧЕНИЕ

В работе представлена модель организации самоорганизующейся ИТ-инфраструктуры, включающей систему административного управления в качестве центра управления и множество ИТ-сервисов в качестве функциональных компонентов, представленная в виде жизнеспособной системы. Для управления жизнеспособной ИТ-инфраструктурой на всем жизненном цикле разработана модифицированная модель жизнеспособной системы, учитывающая системную динамику функциональных компонентов ИТ-инфраструктуры. На основе такой модели представлена схема адаптированного управления ИТ-компонентами, предполагающая три типа управляющих воздействий: создание, утилизация и модификация ИТ-компонентов. Представление ИТ-инфраструктуры совместно с системой управления в виде жизнеспособной системы позволяет применять методы и модели организационной кибернетики для управления развитием ИТ-инфраструктуры в процессе жизненного цикла. Направлением дальнейших исследований является разработка методов административного управления в соответствии с моделью жизнеспособной системой на различных уровнях управления.

СПИСОК ЛИТЕРАТУРЫ

1. Beer S. The viable system model: its provenance, development, methodology and pathology // Journal of the Operational Research Society, 1984. – Vol. 35. – № 1. – P. 7-25.
2. Leonard A. The Viable System Model and Its Application to Complex Organizations // Systemic Practice and Action Research, August 2009. – Vol. 22. – Issue 4. – P. 223-233.
3. Buckl S., Matthes F., Schweda C.M. A Viable System Perspective on Enterprise Architecture Management // 2009 IEEE International Conference on Systems, Man, and Cybernetics, San Antonio, USA, 2009. – P. 1483-1488.
4. Herrera C., Thomas A., Belmokhtar S., Pannequin R. Viable System Model for Product-Driven Systems // International Conference on Industrial Engineering and Systems Management. IESM 2011, Metz - France, May 25 – May 27. – 10 p.
5. Rosenkranz C., Feddersen C. Managing Virtual Communities. A Case Study of a Viable System // AMCIS 2007 Proceedings. 2007. Paper 34. DOI=<http://aisel.aisnet.org/amcis2007/34>.
6. Schwaninger M., Rios J.P. System dynamics and cybernetics: a synergetic pair // System Dynamics Review. – Vol. 24. – № 2. – P. 145-174.
7. Herring C. The Pattern of the Viable System and its Language // KoalaPLoP, 2001. – Melbourne, Australia. – 38 p.
8. Ahmad R., Yusoff M. A Viable System Approach to Tackle Complex Enterprise Situation for SISP // Malaysian Journal of Computer Science, 2006. – № 19(1). – P. 87-103.
9. Христенко Д.В., Гришаков В.Г., Логинов И.В. Многоуровневая система административного управления ИТ-инфраструктурой предприятия // Информатика и системы управления, 2012. – № 3. – С. 68-78.
10. Гришаков В.Г. Виртуализация системы административного управления ИТ-инфраструктурой на основе рекурсивно-иерархического подхода // Информационные системы и технологии, 2013. – № 1. – С. 80-88.
11. Baquero P.J., Aguilar R.M., Ayala A. Design of the ICT infrastructure of an educational system // Workshop on Complexity in Design and Engineering. University of Glasgow, 2005. – 10 p.
12. Herring C., Kaplan S. Viable systems: the control paradigm for software architecture revisited // Software Engineering Conference, 2000. Proceedings. Australian. – P. 97-105.
13. Herring C., Kaplan S. The Viable System Architecture // Proceedings of the 34th Hawaii International Conference on System Sciences. University of Queensland, 2001. 10 p.
14. Murad R.S.A.; Cavana R.Y. Applying the viable system model to ICT project management // Int. J. of Applied Systemic Studies, 2012. – Vol. 4. – № 3. – P. 186-205.
15. Hobbs G., Scheepers R. Identifying capabilities for the it function to create agility in information systems // PACIS 2009 Proceedings. Paper 20. – 13 p. DOI=<http://aisel.aisnet.org/pacis2009/20>.
16. Alinaghian R., Rahman A.A., Ibrahim R. Information and Communication Technology (ICT) Policy: Significances, Challenges, Issues and Future Research Framework // Australian Journal of Basic and Applied Sciences, 2011. – № 5(12). – P. 963-969.
17. Гришаков В.Г., Логинов И.В., Христенко Д.В. Управление модернизацией АСУ предприятием на основе информационной поддержки ее жизненного цикла // Информационно-управляющие системы, 2012. – Т. 3. – № 58. – С. 84-90.
18. Логинов И.В., Христенко Д.В., Гришаков В.Г. Сервис-ориентированный подход к модернизации информационно-вычислительной сети ВУЗа // Программные продукты и системы, 2011. – № 1(93). – С. 103-106.

Гришаков Вадим Геннадьевич
Академия ФСО России, г. Орел
Кандидат технических наук
Тел.: 8 903 637 21 91
E-mail: vg@academ.msk.rsnet.ru

V.G. GRIShAKOV (Candidate of Engineering Science)
Academy of Federal Agency of Protection of the Russian Federation, Orel

MODELING SELF-ORGANIZATION IT-INFRASTRUCTURE AS A VIABLE SYSTEM MODEL

In the article the model of self-organized IT-infrastructure of corporation as a viable system model is offered. This model integrated into one describing IT-infrastructure, that is delivered automated services, and its administrative and operative management systems on all them lifecycle with recursive-hierarchical organization principles. Modeling self-organization IT-infrastructure as a viable system model allows to identification components of management systems and its functions.

Keywords: *IT-infrastructure; administrative management; architecture; effectiveness; viable system mode; system dynamics; lifecycle.*

BIBLIOGRAPHY (TRANSLITERATED)

1. Beer S. The viable system model: its provenance, development, methodology and pathology // Journal of the Operational Research Society, 1984. – Vol. 35. – № 1. – P. 7-25.
2. Leonard A. The Viable System Model and Its Application to Complex Organizations // Systemic Practice and Action Research, August 2009. – Vol. 22. – Issue 4. – P. 223-233.
3. Buckl S., Matthes F., Schweda C.M. A Viable System Perspective on Enterprise Architecture Management // 2009 IEEE International Conference on Systems, Man, and Cybernetics, San Antonio, USA, 2009. – P. 1483-1488.
4. Herrera C., Thomas A., Belmokhtar S., Pannequin R. Viable System Model for Product-Driven Systems // International Conference on Industrial Engineering and Systems Management. IESM 2011, Metz - France, May 25 – May 27. – 10 p.
5. Rosenkranz C., Feddersen C. Managing Virtual Communities. A Case Study of a Viable System // AMCIS 2007 Proceedings. 2007. Paper 34. DOI=<http://aisel.aisnet.org/amcis2007/34>.
6. Schwaninger M., Rios J.P. System dynamics and cybernetics: a synergetic pair // System Dynamics Review. – Vol. 24. – № 2. – P. 145-174.
7. Herring C. The Pattern of the Viable System and its Language // KoalaPLoP, 2001. – Melbourne, Australia. – 38 p.
8. Ahmad R., Yusoff M. A Viable System Approach to Tackle Complex Enterprise Situation for SISP // Malaysian Journal of Computer Science, 2006. – № 19(1). – P. 87-103.
9. Kristenko D.V., Grishakov V.G., Loginov I.V. Mnogourovnevaya sistema administrativnogo upravleniya IT-infrastrukturoj predpriyatiya // Informatika i sistemy'upravleniya, 2012. – № 3. – S. 68-78.
10. Grishakov V.G. Virtualizaciya sistemy' administrativnogo upravleniya IT-infrastrukturoj na osnove rekursivno-ierarxicheskogo podxoda // Infrmacionny'e sistemy' i texnologii, 2013. – № 1. – S. 80-88.
11. Baquero P.J., Aguilar R.M., Ayala A. Design of the ICT infrastructure of an educational system // Workshop on Complexity in Design and Engineering. University of Glasgow, 2005. – 10 p.
12. Herring C., Kaplan S. Viable systems: the control paradigm for software architecture revisited // Software Engineering Conference, 2000. Proceedings. Australian. – P. 97-105.
13. Herring C., Kaplan S. The Viable System Architecture // Proceedings of the 34th Hawaii International Conference on System Sciences. University of Queensland, 2001. 10 p.
14. Murad R.S.A.; Cavana R.Y. Applying the viable system model to ICT project management // Int. J. of Applied Systemic Studies, 2012. – Vol. 4. – № 3. – P. 186-205.
15. Hobbs G., Scheepers R. Identifying capabilities for the it function to create agility in information systems // PACIS 2009 Proceedings. Paper 20. – 13 p. DOI=<http://aisel.aisnet.org/pacis2009/20>.
16. Alinaghian R., Rahman A.A., Ibrahim R. Information and Communication Technology (ICT) Policy: Significances, Challenges, Issues and Future Research Framework // Australian Journal of Basic and Applied Sciences, 2011. – № 5(12). – P. 963-969.
17. Grishakov V.G., Loginov I.V., Kristenko D.V. Upravlenie modernizaciej ASU predpriyatiem na osnove informacionnoj podderzhki eyo zhiznennogo cikla // Informacionno-upravlyayushhie sistemy', 2012. – T. 3. – № 58. – S. 84-90.
18. Loginov I.V., Kristenko D.V., Grishakov V.G. Servis-orientirovannyj podxod k modernizacii infrmacionno-vy'chislitel'noj seti VUZa // Programmny'e pordukty' i sistemy', 2011. – № 1(93). – S. 103-106.

УДК 004.732

А.А. ЛЯКИШЕВ

СПОСОБЫ И ПРИЕМЫ ОПТИМИЗАЦИИ ПРОЦЕССОВ ИНФОРМАЦИОННОГО ОБМЕНА В МОДУЛЬНЫХ СТРУКТУРАХ СИСТЕМ СБОРА И ОБРАБОТКИ ДАННЫХ АСУ ГАЗОТРАНСПОРТНОГО ПРЕДПРИЯТИЯ

В статье рассматривается методика оптимизации процессов информационного обмена в модульных структурах систем сбора и обработки данных в АСУ ГТП, базирующаяся на использовании замкнутой экспоненциальной модели сети массового обслуживания и отличающаяся возможностью оперирования нечёткими множествами.

Ключевые слова: оптимизация; сети массового обслуживания; нечёткие множества.

ВВЕДЕНИЕ

Функционирование современных сетей ГТП определяется стохастическим характером нагрузки, а также недетерминированной обработкой их в узлах коммутации и каналах связи. Это определяет использование моделей теории массового обслуживания для описания и моделирования нагрузок в сетях передачи данных (СПД) ГТП. Кроме этого, существующие в настоящий момент телекоммуникационные сети, предназначенные для объединения большого количества разнородных пользователей, в основном имеют достаточно сложную структуру и предназначены для решения нескольких классов задач, каждый из которых характеризуется своими особенностями [1-5, 7-11].

Оптимальное проектирование модульных структур систем сбора и обработки данных газотранспортных предприятий (ГТП) невозможно без процедуры принятия решений на различных этапах процесса проектирования. Проектирование СПД ГТП характеризуется большим количеством этапов, на каждом из которых необходимо осуществить принятие решения как на основе каких-либо математических или эвристических методов, так и на основе опыта проектировщика.

МЕТОДИКА ОПТИМИЗАЦИИ ПРОЦЕССОВ ИНФОРМАЦИОННОГО ОБМЕНА В МОДУЛЬНЫХ СТРУКТУРАХ СИСТЕМ СБОРА И ОБРАБОТКИ ДАННЫХ

Предлагаемая методика состоит из следующих основных этапов, представленных на рисунке 1.

Первый этап представлен в виде алгоритма оптимизация пропускной способности в модульных структурах систем сбора и обработки данных в АСУ ГТП.

Задача оптимизации модульной структуры систем сбора и обработки данных состоит в максимизации ее производительности с учетом того, что она может быть представлена в виде замкнутой однородной СМО при стоимости, не превосходящей заданной, или минимизации стоимости сети при производительности не ниже заданной [6].

Максимум производительности сети не может достигаться внутри области ограничения на стоимость, следовательно, должно выполняться равенство $S(\mu) = S^*$, где S^* – ограничение по стоимости СМО. Аналогично минимальная стоимость сети достигается при выполнении ограничения на производительность в виде равенства $\lambda(\mu) = \lambda^*$. Таким образом, задача оптимизации замкнутой экспоненциальной СМО может быть сформулирована в одной из следующих постановок:

1. Найти:

$$\max \lambda = e_1 G_M(N-1) / G_M(N) \quad (1)$$

$$\text{при ограничении } S = \sum_{i=1}^M c_i \mu_i^{a_i} = S^*, \mu > 0 \quad (2)$$

2. Найти:

$$\min S = \sum_{i=1}^M c_i \mu_i^{a_i} \quad (3)$$

при ограничении $\lambda = e_1 G_M(N-1) / G_M(N) = \lambda$. (4)

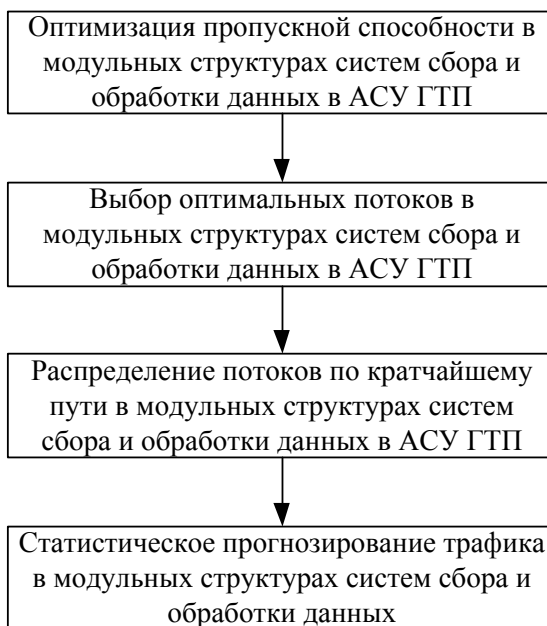


Рисунок 1 – Этапы методики оптимизации процессов информационного обмена в модульных структурах систем сбора и обработки данных АСУ газотранспортного предприятия

Функции $S(\mu)$ и $\lambda(\mu)$ – выпуклые, поэтому любой локальный максимум задачи в постановке 1 является также и глобальным максимумом; любой локальный минимум в постановке 2 является также и глобальным минимумом.

Оптимальное решение задач (1), (3) будем искать методом неопределённых множителей Лагранжа. Составим функцию Лагранжа $Q = \lambda + \gamma(S - S^*)$, где γ – множитель Лагранжа [12].

Получаем окончательную систему нелинейных уравнений относительно переменных μ_i :

$$\mu_i^{a_i} = \frac{S}{\left\{ c_1 \left[1 + \frac{1}{L_1(N-1) \sum \frac{a_1}{a_i} (L_i(N) - L_i(N-1)) \right] \right\}}$$

$$\mu_i^{a_i} = \frac{a_2 c_2}{a_i c_i} \mu_1^{a_i} \frac{L_i(N) - L_i(N-1)}{L_2(N) - L_2(N-1)}, i = (\overline{2, M}) \quad (5)$$

Вектор μ^* , являющийся решением системы уравнений (5), доставляет максимум целевой функции $\lambda(\mu)$ при выполнении ограничения $S(\mu) = S^*$

Решая аналогичным образом задачу для постановки 2, получаем:

$$\mu_1 = \lambda^* / U_1(N)$$

$$\mu_1^{a_i} = \frac{a_2 c_2}{a_i c_i} \mu_1^{a_i} \frac{L_i(N) - L_i(N-1)}{L_2(N) - L_2(N-1)}, i = (\overline{2, M}) \quad (6)$$

Таким образом, задача оптимизации замкнутой однородной сети МО сведена к решению системы нелинейных уравнений (5) или (6).

Второй этап представлен в виде алгоритма выбора оптимального потока в модульных структурах систем сбора и обработки данных в АСУ ГТП.

Задача выбора оптимальных маршрутов относится к классу многопродуктовых задач с выпуклой целевой функцией и выпуклым множеством ограничений. Следовательно, существует единственный локальный минимум данной задачи, являющийся глобальным

минимумом, для нахождения которого разработано достаточно большое число вычислительных методов [6].

Наиболее известным методом решения данной задачи является метод отклонения потока. Данный алгоритм является частным случаем метода Франка-Вольфе для решения более общих задач нелинейного программирования с выпуклым множеством ограничений. Доказано, что метод отклонения потока в пределе уменьшает значение целевой функции до минимума, хотя по мере приближения к оптимуму скорость сходимости существенно замедляется [13].

Для решения поставленной задачи воспользуемся модифицированным алгоритмом отклонения потока.

Шаг 1. Определить «веса» линий связи w_{kl} и проинициализировать потоки в линиях связи f_{kl} :

$$w_{kl} := \begin{cases} \left[\frac{\partial T}{\partial f_{kl}} \right]_{f_{kl}=0} = \frac{1}{d_{kl}}; k, l = 1, 2, \dots, N; d_{kl} > 0 \\ \infty, & k, l = 1, 2, \dots, N; d_{kl} = 0; \end{cases}$$

$$f_{kl} := 0; k, l = 1, 2, \dots, N.$$

Шаг 2. Используя «веса» линий связи w_{kl} , определить кратчайшие пути π_{ij} между всеми парами узлов «источник-адресат». Для нахождения кратчайших путей в этом случае наиболее подходящим является алгоритм Флойда.

Шаг 3. Распределить потоки по кратчайшим путям:

$$\forall i, j = 1, 2, \dots, N; \forall (k, l) \in \pi_{ij}; f_{kl} := f_{kl} + \frac{y_{ij}}{\mu}.$$

Шаг 4. Вычислить:

$$T_{old} = \frac{1}{\gamma} \sum_{k=1}^N \sum_{l=1}^N \frac{f_{kl}}{d_{kl} - f_{kl}}.$$

Шаг 5. Положить: $\gamma^{(1)} := \gamma$.

Шаг 6. Положить:

$$\gamma^{(2)} := \min \left\{ \gamma, \frac{\gamma^{(1)}}{\rho_{max}} \right\}, \text{ где } \rho_{max} = \max \left\{ \frac{f_{kl}}{d_{kl}} \right\}; \forall (k, l): d_{kl} > 0.$$

Шаг 7. Пересчитать потоки в линиях связи:

$$f_{kl} := f_{kl} * \frac{\gamma^{(2)}}{\gamma^{(1)}}; k, l = 1, 2, \dots, N.$$

Шаг 8. Определить «веса» линий связи w_{kl} , инициализировать потоки по кратчайшим путям φ_{kl} :

$$w_{kl} := \begin{cases} \frac{\partial T}{\partial f_{kl}} = \frac{d_{kl}}{(d_{kl} - f_{kl})^2}; k, l = 1, 2, \dots, N; d_{kl} > 0 \text{ и } f_{kl} < d_{kl} \\ \infty, & k, l = 1, 2, \dots, N; d_{kl} = 0 \text{ или } f_{kl} \geq d_{kl}; \end{cases}$$

$$\varphi_{kl} := 0; k, l = 1, 2, \dots, N.$$

Шаг 9. Используя «веса» линий связи w_{kl} , определить кратчайшие пути π_{ij} между всеми парами узлов «источник-адресат».

Шаг 10. Распределить потоки по кратчайшим путям:

$$\forall i, j = 1, 2, \dots, N; \forall (k, l) \in \pi_{ij}; \varphi_{kl} := \varphi_{kl} + \gamma_{ij} * \frac{\gamma^{(2)}}{\mu * \gamma}.$$

Шаг 11. Найти величину $\beta \in [0; 1]$, минимизирующую функцию:

$$T(\beta) = \frac{1}{\gamma^{(2)}} \sum_{l=1}^N \sum_{k=1}^N \frac{\beta * \varphi_{kl} + (1 - \beta) * f_{kl}}{d_{kl} - \beta * \varphi_{kl} - (1 - \beta) * f_{kl}}.$$

Поиск величины β можно осуществить любым из известных методов одномерного поиска, например, методом Фибоначчи.

Шаг 12. Выполнить отклонение потока на величину β :

$$f_{kl} = \beta * \varphi_{kl} + (1 - \beta) * f_{kl}; k, l = 1, 2, \dots, N.$$

Шаг 13. Вычислить:

$$T_{\text{new}} = \frac{1}{\gamma^{(2)}} \sum_{l=1}^N \sum_{k=1}^N \frac{f_{kl}}{d_{kl} - f_{kl}}$$

Шаг 14. Если $|T_{\text{old}} - T_{\text{new}}| \leq \varepsilon$, то STOP:

Если $\gamma^{(2)} < \gamma$, то допустимых решений нет;

Если $\gamma^{(2)} = \gamma$, то получено оптимальное решение с заданной точностью ε .

Иначе:

1) положить: $T_{\text{old}} := T_{\text{new}}, \gamma^{(1)} := \gamma^{(2)}$;

2) если $\gamma^{(1)} < \gamma$, то перейти к шагу 6; иначе перейти к шагу 8.

Приведенный алгоритм решения задачи выбора оптимальных потоков и определения оптимальных маршрутов в случае альтернативной маршрутизации объединяет в себе шаги построения начального допустимого потока (шаги 1-14) и задачи минимизации средней задержки сети (шаги 8-14).

Третий этап представлен в виде алгоритма оптимизации процесса маршрутизации в модульных структурах систем сбора и обработки данных в АСУ ГТП.

В случае K -путевой маршрутизации используем следующие дополнительные ограничения.

Введем переменную:

$$v_{kl}^{(j)} = \begin{cases} 1, & \text{если } \sum_{j=1}^N x_{kl}^{(i,j)} > 0 \\ 0, & \text{если } \sum_{i=1}^N x_{kl}^{(i,j)} = 0; \\ & j, k, l = 1, 2, \dots, N. \end{cases} \quad (9)$$

Иными словами, переменная $v_{kl}^{(j)} = 1$, если линия связи (k, l) используется для передачи потока в узел-адресат j хотя бы от одного узла-источника и равно 0 в противном случае.

Тогда ограничение на число исходящих линий (K), используемых для передачи данных из каждого узла к узлу-адресату j , можно записать в следующем виде:

$$\sum_{i=1}^N v_{kl}^{(j)} \leq K; j, k = 1, 2, \dots, N. \quad (10)$$

Четвертый этап представлен в виде алгоритма статистического прогнозирования трафика в модульных структурах систем сбора и обработки данных в АСУ ГТП.

Процессы развития нагрузки сети, подлежащие статистическому прогнозированию, обычно носят случайный характер. В общем случае процедура прогнозирования трафика выглядит следующим образом:

Шаг 1. Определить рамки временного периода, для которого делается прогноз развития трафика.

Шаг 2. Определить факторы, влияющие на трафик в течение исследуемого временного периода, и провести прогнозирование на основе анализа зависимости трафика от величины тарифов на основе регрессионных моделей.

Шаг 3. В зависимости от выбранной процедуры и на основании имеющейся статистической информации произвести прогнозирование и оценить его достоверность.

Рассмотрим подробнее содержание части процедур прогнозирования трафика.

Пусть N – число всех потенциальных пользователей сети передачи данных в рассматриваемой области. Далее, пусть $m(t)$ – число пользователей, подключенных к сети к моменту t . Определим вероятность $P_n(t)$ того, что $m(t)$ равно n :

$$P_n(t) = P(m(t) = n), \quad n = 1, \dots, N. \quad (11)$$

Сделаем предположение, что количество лиц, желающих подключиться, но не имеющих доступа к сети, пропорционально количеству абонентов, уже подключенных к сети. Точнее, если к моменту t существует ровно n абонентов, то вероятность для каждого будущего абонента подключиться к сети в течение промежутка времени h равна:

$$P = \lambda \frac{n}{N} h + o(h). \quad (12)$$

Процедура прогнозирования трафика описывается выражением с использованием средней нагрузки сети в момент времени t для m разновидностей услуг:

$$Y = \sum_{i=2}^m \frac{k_i N_i}{2} (1 + th(\omega_i t - p_i)), \quad (13)$$

где k_i – нагрузка, создаваемая одним элементом i -й разновидности услуг.

Для выявления основной тенденции изучаемого процесса чаще всего используется метод наименьших квадратов. Согласно методу наименьших квадратов, неизвестные параметры функции выбираются таким образом, чтобы сумма квадратов отклонений эмпирических данных от значений, вычисленная с помощью выражения, была равна:

$$S = \sum_{i=1}^t (v_{\text{э.мин}} - f(t))^2 \rightarrow \min. \quad (14)$$

Согласно необходимому условию экстремума функции нескольких переменных, для нахождения минимального значения параметров частные производные $f(t)$ приравниваются нулю. После этого полученная система уравнений решается. Для функции линейного вида система уравнений будет иметь вид:

$$\begin{cases} a_0 n + a_1 \sum_{t=1}^n t = \sum_{t=1}^n y_t \\ a_0 \sum_{t=1}^n t + a_1 \sum_{t=1}^n t^2 = \sum_{t=1}^n y_t t \end{cases} \quad (15)$$

а в случае параболы $f(t) = a_0 + a_1 t + a_2 t^2$ следующий вид:

$$\begin{cases} a_0 n + a_1 \sum_{t=1}^n t + a_2 \sum_{t=1}^n t^2 = \sum_{t=1}^n y_t \\ a_0 \sum_{t=1}^n t + a_1 \sum_{t=1}^n t^2 + a_2 \sum_{t=1}^n t^3 = \sum_{t=1}^n y_t t \\ a_0 \sum_{t=1}^n t^2 + a_1 \sum_{t=1}^n t^3 + a_2 \sum_{t=1}^n t^4 = \sum_{t=1}^n y_t t^2 \end{cases} \quad (16)$$

ЗАКЛЮЧЕНИЕ

Решение задачи оптимизации модульных структур систем сбора и обработки данных сводится к решению системы нелинейных уравнений, которая не содержит частных производных. Такой подход позволяет расширить область применения методов оптимизации модульных структур и снизить вычислительную сложность алгоритмов.

При применении метода наименьших квадратов для оценки параметров оптимизации на основе экспоненциальной функции или функции Гомперца возникают сложности с решением получаемой системы уравнений. В работе используются логарифмические преобразования этих функций.

Полученная методика оптимизации процессов информационного обмена в модульных структурах систем сбора и обработки данных в АСУ ГТП позволяет произвести оценку их наиболее важных показателей, выбрать на основании предложенного критерия оптимизации маршрут и произвести прогнозирование дальнейшего развития нагрузки на исследуемую модульную структуру.

СПИСОК ЛИТЕРАТУРЫ

1. Еременко В.Т. Оптимизация ресурсов и управление процессами информационного обмена в сетях АСУТП на основе полевых шин / С.И. Афонин, В.Т. Еременко, С.А. Максаков, А.И. Куленич // Вестник компьютерных и информационных технологий, 2011. – № 9. – С. 46-49.
2. Еременко В.Т. Математическая модель оценки производительности беспроводной вычислительной сети АСУ предприятия / В.Т. Еременко, С.И. Афонин, Д.А. Краснов и др. // Информационные системы и технологии, 2011. – № 5. – С. 11-20.
3. Моделирование информационных потоков в сетях передачи данных интегрированных АСУ / С.И. Афонин, В.Т. Еременко, Т.М. Парамохина, Л.В. Кузьмина, Д.А. Плащенков // Информационные системы и технологии, 2011. – № 6. – С. 35-42.
4. Еременко В.Т. Методы решения задач распределения информационных потоков в сетях передачи данных предприятия на основе резервирования ресурсов / С.И. Афонин, В.Т. Еременко, Л.В. Кузьмина и др. // Информационные системы и технологии, 2012. – № 1. – С. 78-84.
5. Еременко В.Т., Афонин С.И. Создание теоретических основ автоматизации и построения технологической составляющей АСУ территориально распределенных предприятий, В.Т. Еременко // Информационные системы и технологии, 2012. – № 2. – С. 99-105.
6. Еременко В.Т. Рекурсивный алгоритм оценки минимальной величины канального ресурса в сети передачи данных / В.Т. Еременко, Л.В. Кузьмина, Д.А. Плащенков, Д.А. Краснов // Информационные системы и технологии, 2012. – № 4. – С. 97-102.
7. Еременко В.Т. Метод проектирования сетей передачи данных, совместимых с неблокируемой маршрутизацией / В.Т. Еременко, А.И. Офицеров, С.А. Черепков // Вестник компьютерных и информационных технологий, 2012. – № 4. – С. 38-46.
8. Еременко В.Т. Решение задач управления сетевыми ресурсами в условиях динамического изменения конфигурации беспроводной сети АСУП / В.Т. Еременко, Д.В. Анисимов, Д.А. Плащенков, Д.А. Краснов, С.А. Черепков, А.Е. Георгиевский // Информационные системы и технологии, 2012. – № 6. – С. 114-119.
9. Еременко В.Т. Анализ моделей управления трафиком в сетях АСУП на основе технологии MPLS / В.Т. Еременко, С.В. Еременко, Д.В. Анисимов, С.А. Черепков, А.А. Лякишев // Информационные системы и технологии, 2013. – № 1. – С. 106-112.
10. Еременко В.Т. Моделирование пропускной способности сегмента беспроводной сети АСУП на базе стандарта 802.11 / В.Т. Еременко, Д.В. Анисимов, С.А. Черепков, А.А. Лякишев, П.А. Чупахин // Информационные системы и технологии, 2013. – № 2. – С. 82-86.
11. Еременко В.Т. Математическое моделирование беспроводного сегмента вычислительной сети АСУ ПП / В.Т. Еременко, Д.В. Анисимов, Т.М. Парамохина, А.А. Лякишев // Информационные системы и технологии, 2013. – № 3. – С. 67-72.
12. Вишневский В.М., Федотов Е.В. Анализ методов маршрутизации при проектировании сетей пакетной коммутации // 3-rd I.S. «Teletraffic Theory and Computing Modeling». – София, 1992.
13. Рындин А.А. Автоматизация проектирования сетей передачи данных распределенных информационно-телекоммуникационных систем / А.А. Рындин, С.В. Сапегин, А.В. Хаустович // Воронеж: ВГТУ, 2001.

Лякишев Александр Александрович

ФГБОУ ВПО Госуниверситет – УНПК, г. Орел

Аспирант кафедры «Электроника, вычислительная техника и информационная безопасность»

Тел.: 8 919 777 06 64

E-mail: stiplchez@list.ru

A.A. LYAKISHEV (*Post-graduate Student of the Department «Electronics, Computer Facilities and Information Security»
State University – ESPC, Orel*)

METHODS AND TECHNIQUES OF OPTIMIZATION OF THE PROCESS INFORMATION EXCHANGE IN THE MODULAR STRUCTURE OF DATA COLLECTION AND PROCESSING SYSTEMS IN THE ACS OF A GAS TRANSPORTATION COMPANY

The article describes the technique of optimization of information exchange in the modular structure of data collection and processing systems in the ACS of a gas transportation company based on the using of the exponential model of a closed queuing network and differed by the ability to operate on fuzzy sets.

Keywords: optimization; queuing networks; fuzzy sets.

BIBLIOGRAPHY (TRANSLITERATED)

1. Eryomenko V.T. Optimizaciya resursov i upravlenie processami informacionnogo obmena v setyax ASUTP na osnove polevy'x shin / S.I. Afonin, V.T. Eryomenko, S.A. Maksakov, A.I. Kulenich // Vestnik komp'yuterny'x i informacionny'x texnologij, 2011. – № 9. – S. 46-49.
2. Eryomenko V.T. Matematicheskaya model' ocenki proizvoditel'nosti besprovodnoj vy'chislitel'noj seti ASU predpriyatiya / V.T. Eryomenko, S.I. Afonin, D.A. Krasnov i dr. // Informacionny'e sistemy' i texnologii, 2011. – № 5. – S. 11-20.
3. Modelirovanie informacionny'x potokov v setyax peredachi danny'x integrirovanny'x ASU / S.I. Afonin, V.T. Eryomenko, T.M. Paramoxina, L.V. Kuz'mina, D.A. Plashhenkov // Informacionny'e sistemy' i texnologii, 2011. – № 6. – S. 35-42.
4. Eryomenko V.T. Metody' resheniya zadach raspredeleniya informacionny'x potokov v setyax peredachi danny'x predpriyatiya na osnove rezervirovaniya resursov / S.I. Afonin, V.T. Eryomenko, L.V. Kuz'mina i dr. // Informacionny'e sistemy' i texnologii, 2012. – № 1. – S. 78-84.
5. Eryomenko V.T., Afonin S.I. Sozdanie teoreticheskix osnov avtomatizacii i postroeniya texnologicheskoy sostavlyayushhej ASU territorial'no raspredelyonny'x predpriyatij // Informacionny'e sistemy' i texnologii, 2012. – № 2. – S. 99-105.
6. Eryomenko V.T. Rekursivny'j algoritm ocenki minimal'noj velichiny' kanal'nogo resursa v seti peredachi danny'x / V.T. Eryomenko, L.V. Kuz'mina, D.A. Plashhenkov, D.A. Krasnov // Informacionny'e sistemy' i texnologii, 2012. – № 4. – S. 97-102.
7. Eryomenko V.T. Metod proektirovaniya setej peredachi danny'x, sovmestimy'x s neblokiruemoj marshrutizaciej / V.T. Eryomenko, A.I. Oficerov, S.A. Cherepkov // Vestnik komp'yuterny'x i informacionny'x texnologij, 2012. – № 4. – S. 38-46.
8. Eryomenko V.T. Reshenie zadach upravleniya setevy'mi resursami v usloviyax dinamicheskogo izmeneniya konfiguracii besprovodnoj seti ASUP / V.T. Eryomenko, D.V. Anisimov, D.A. Plashhenkov, D.A. Krasnov, S.A. Cherepkov, A.E. Georgievskij // Informacionny'e sistemy' i texnologii, 2012. – № 6. – S. 114-119.
9. Eryomenko V.T. Analiz modelej upravleniya trafikom v setyax ASUP na osnove texnologii MPLS / V.T. Eryomenko, S.V. Eryomenko, D.V. Anisimov, S.A. Cherepkov, A.A. Lyakishev // Informacionny'e sistemy' i texnologii, 2013. – № 1. – S. 106-112.
10. Eryomenko V.T. Modelirovanie propusknoj sposobnosti segmenta besprovodnoj seti ASUP na baze standarta 802.11 / V.T. Eryomenko, D.V. Anisimov, S.A. Cherepkov, A.A. Lyakishev, P.A. Chupaxin // Informacionny'e sistemy' i texnologii, 2013. – № 2. – S. 82-86.
11. Eryomenko V.T. Matematicheskoe modelirovanie besprovodnogo segmenta vy'chislitel'noj seti ASUP PP / V.T. Eryomenko, D.V. Anisimov, T.M. Paramoxina, A.A. Lyakishev // Informacionny'e sistemy' i texnologii, 2013. – № 3. – S. 67-72.
12. Vishnevskij V.M., Fedotov E.V. Analiz metodov marshrutizacii pri proektirovanii setej paketnoj kommutacii // 3-rd I.S. «Teletraffic Theory and Computing Modeling». – Sofiya, 1992.
13. Ry'ndin A.A. Avtomatizaciya proektirovaniya setej peredachi danny'x raspredelyonny'x informacionno-telekommunikacionny'x sistem / A.A. Ry'ndin, S.V. Sapegin, A.V. Xaustovich // Voronezh: VGTU, 2001.

УДК 621.391

В.Е. ФИСЕНКО, А.П. ФИСУН

ФОРМИРОВАНИЕ ТРЕБОВАНИЙ К ПОКАЗАТЕЛЯМ НАДЕЖНОСТИ ЭЛЕМЕНТОВ ТЕЛЕКОММУНИКАЦИОННЫХ СИСТЕМ

В статье предложен новый подход к формированию требований к показателям надежности процесса функционирования телекоммуникационных систем на основе метода приоритетных путей, обеспечивающий значительное снижение затрат для достижения требуемого качества процесса функционирования телекоммуникационных систем.

Ключевые слова: телекоммуникационная система; информационное направление; направление связи; нормирование надежности; частично коммутируемые сети; коэффициент готовности.

ВВЕДЕНИЕ

При нормировании надежности процесса функционирования элементов телекоммуникационных систем (ТКС), то есть установлении в нормативно-технической документации количественных и качественных требований к надежности ТКС и составляющих ее элементов, как правило, не оговаривается, на какие ТКС эти требования распространяются: некоммутируемые, коммутируемые или частично коммутируемые. При этом также не оговаривается и не учитывается состав информационных направлений (ИН), входящих в ТКС, и составляющих ИН направлений связи (НС). Поэтому на практике оперативно-техническим составом данные требования распространяются, как правило, и на пути установления соединения, составляющие ИН и НС. Не говоря уже о том, что обеспечить столь высокие показатели надежности для путей, составляющих НС, практически невозможно или это требует значительных затрат. Необходимо подчеркнуть и то, что данный подход по своей сути является некорректным, так как не учитывает вероятностный характер используемых показателей надежности. С учетом вышесказанного предлагается для совершенствования программного обеспечения процесса функционирования ТКС использовать метод приоритетных путей, который позволяет устранить вышеуказанные недостатки, получать оперативно-техническим составом и применять на практике научно обоснованные показатели надежности не только ко всему ИН, но и путям, его образующим.

ПОСТАНОВКА ЗАДАЧИ

В ходе исследований [1, 2] было выявлено, что в рамках решаемой задачи наибольшую значимость для получения объективной оценки надежности направлений связи имеют количественные характеристики наиболее часто применяемых на практике показателей надежности имеющих вероятностный характер. К таким показателям в первую очередь относятся коэффициент готовности и вероятность безотказной работы.

Достаточно часто коэффициент готовности направления связи K_r задается не менее 0,98, при этом не оговаривается и не учитывается состав направления связи. Поэтому на практике оперативно-техническим составом данное требование распространяется и на пути установления соединения, составляющие данное НС.

Однако постоянное развитие методологии оценки надежности телекоммуникационных систем [3] и, в частности, метод приоритетных путей оценки показателей надежности ТКС [1] позволяют по-новому подойти к решению проблемы задания требований к показателям надежности элементов некоммутируемых и частично коммутируемых ТКС.

Адаптивный метод оценки надежности информационных направлений с уменьшенной вычислительной сложностью (метод приоритетных путей) включает в себя два этапа [1].

Первый этап. Выделение из массива путей, составляющих данное информационное

направление, параллельно–последовательной системы, состоящей из приоритетных путей данного информационного направления.

Второй этап. Определение надежности информационного направления как надежности параллельно-последовательной системы, состоящей из приоритетных путей данного информационного направления.

Рассмотрим этот подход более подробно для частично коммутируемых ТКС. Как было рассмотрено в [2], информационные направления частично коммутируемых ТКС являются системами с параллельно-последовательным соединением элементов (из параллельно соединенных подсистем). Для параллельно-последовательной системы при предположении о независимости отказов надежность $P(e_w)$ вычисляется по формуле

$$P(e_w) = 1 - \prod_{z=1}^{Z_w} \left(1 - \prod_{u=1}^{U_z} p(l_u) \right),$$

где $p(l_u)$ – показатели надежности отдельных элементов, составляющих параллельно–последовательную систему; U_z – количество элементов в Z -й последовательной подсистеме параллельно-последовательной системы; Z_w – количество параллельно соединенных последовательных подсистем в W -й параллельно-последовательной системе.

Если информационное направление описывается параллельной системой, при предположении о независимости отказов, то надежность $P(e)$ вычисляется по формуле

$$P(e) = 1 - \prod_{z=1}^Z (1 - p(n_z)),$$

где $p(n_z)$ – показатели надежности отдельных подсистем, составляющих параллельную систему; Z – количество подсистем в параллельной системе.

Таким образом, можно решить обратную задачу, а именно, зная требуемый показатель надежности информационного направления частично коммутируемой ТКС, можно определить задаваемые требования к показателям надежности составляющих его путей. Это можно сделать по формуле (1), преобразовав ее к следующему виду:

$$P_{E_w \Pi w}^{np} = \sqrt[Z_{E_w}^{np}]{1 - P_{E_w}^{зад}}, \quad (1)$$

где $P_{E_w \Pi w}^{np}$ – показатели надежности отдельных подсистем (в нашем случае приоритетных путей в НС), составляющих параллельную систему; $Z_{E_w}^{np}$ – количество подсистем (в нашем случае количество приоритетных путей в НС E_w) в параллельной системе; $P_{E_w}^{np}$ – требуемый (заданный) показатель надежности НС.

МЕТОДИКА РАСЧЕТА ЗАДАВАЕМЫХ ТРЕБОВАНИЙ К ПОКАЗАТЕЛЯМ НАДЕЖНОСТИ НС ЧАСТИЧНО КОММУТИРУЕМЫХ ТКС

Разработанная на основе метода приоритетных путей методика расчета задаваемых требований к показателям надежности НС частично коммутируемых ТКС состоит в выполнении следующих действий.

На первом шаге с учетом требования к показателю надежности НС и количества независимых путей в НС по формуле (3) определяются задаваемые требования к показателям надежности приоритетных путей НС частично коммутируемой ТКС заданной структуры.

На втором шаге с учетом структуры приоритетных путей в НС определяются надежность ребер, их образующих, по формуле

$$P_{E_w \Pi z L_u}^{np} = 1 - \sqrt[U_{E_w \Pi z}]{1 - \sqrt[Z_{E_w}^{np}]{1 - P_{E_w}^{зад}}},$$

где $P_{E_w \Pi z L_u}^{np}$ – показатели надежности L_u ребер Π_z -го приоритетного пути в E_w -м

информационном направлении; $U_{Ew\Pi z}^{np}$ – количество ребер L_u в Π_z -м приоритетном пути информационного направления E_w .

Предложенная методика задания требований к показателям надежности элементов частично коммутируемых ТКС на основе адаптивного метода приоритетных путей обладает следующими важными достоинствами по сравнению с существующими:

- позволяет на основе заданных требований к показателям надежности НС частично коммутируемых ТКС, количества независимых путей в них, заданной (требуемой) величины погрешности получаемого результата рассчитывать требуемую надежность линий доступа (линий привязки) НС;

- позволяет задавать требуемую надежность ребер линий доступа (линий привязки) для каждого узла НС в зависимости от структуры его линий доступа (линий привязки) к частично коммутируемой ТКС;

- позволяет получать и применять на практике оперативно-техническим составом научно обоснованные показатели надежности не только ко всему НС, но и путям, его образующим;

- требует незначительного вычислительного ресурса.

Количественные характеристики показателей надежности с учетом структуры НС, рассчитанные с использованием вышеприведенной методики на основе метода приоритетных путей [2], приведены в таблице 1.

Таблица 1 – Количественные характеристики показателей надежности с учетом структуры НС

Показатели	Условия (количество приоритетных путей в НС)			
	1	2	3	4
Коэффициент готовности НС, $K_r^{НС}$	0.98	0.98	0.98	0.98
Коэффициент готовности путей НС, $K_r^{ПНС}$	0.98	0.86	0.73	0.62

Из результатов расчетов, приведенных в таблице 1, можно сделать очень важный вывод о том, что применение на практике оперативно-техническим составом требований, предъявляемых к коэффициенту готовности НС, к коэффициенту готовности путей, составляющих данное НС, является корректным только в случае, когда структура направления связи включает только один приоритетный путь. Однако на практике такое встречается редко [1]. Следовательно, применение на практике оперативно-техническим составом требований, предъявляемых к коэффициенту готовности НС, к коэффициенту готовности путей, составляющих данное НС, является не только некорректным, но и ведущим к необоснованному завышению требований к путям, составляющим данное НС. Что, в свою очередь, ведет к увеличению затрат на обеспечение требуемого качества процесса функционирования ТКС.

ЗАКЛЮЧЕНИЕ

В данной статье предлагается новый подход к совершенствованию программного обеспечения процесса функционирования информационно-телекоммуникационных систем за счет формирования требований к показателям надежности элементов ТКС на основе метода приоритетных путей. Этот подход позволяет получать и применять на практике оперативно-техническим составом научно обоснованные показатели надежности не только ко всему информационному направлению, но и путям, его образующим. Это, в свою очередь, ведет к значительному снижению затрат на обеспечение требуемого качества процесса функционирования ТКС.

СПИСОК ЛИТЕРАТУРЫ

1. Фисун А.П., Фисенко В.Е., Митяев В.В. Оценка надежности информационно-телекоммуникационных систем в реальном масштабе времени на основе приоритетных методов: монография. – Орел: Изд-во ОГУ, 2007. – 191 с.
2. Фисун А.П., Фисенко В.Е., Минаев В.А., Митяев В.В. и др. Развитие методологических основ информатики и информационной безопасности систем: монография. – Орел: ОГУ, 2004. – 253 с. – Деп. в ВИНТИ 07.07.04. № 1165-В2004.
3. Фисенко В.Е. Новая концепция оценки надежности больших информационно-телекоммуникационных систем // Известия ТулГУ. Сер. «Вычислительная техника. Информационные технологии. Системы управления». – Вып. 1. – Тула: Изд-во ТулГУ, 2006. – С. 161-165.

Фисенко Виктор Евгеньевич

ФГБОУ ВПО Государственный университет – УНПК, г. Орел

Кандидат технических наук, доцент, доцент кафедры «Информационные системы»

Тел.: (4862) 36-47-36, 8 910 303 52 72

E-mail: fisenko_v@mail.ru

Фисун Александр Павлович

ФГБОУ ВПО Государственный университет – УНПК, г. Орел

Доктор технических наук, профессор, профессор кафедры «Электроника, вычислительная техника и информационная безопасность»

Тел.: (4862) 42-15-56, 8 910 307 00 81

E-mail: fisun01@pisem.net

V.E. FISENKO (*Candidate of Engineering Science, Associate Professor, Associate Professor of the Department «Information Systems»*)

A.P. FISUN (*Doctor of Engineering Science, Professor, Professor of the Department «Electronics, Computer Engineering and Information Security» State University – ESPC, Orel*)

FORMATION OF REQUIREMENTS TO RELIABILITY OF ELEMENTS OF TELECOMMUNICATION SYSTEMS

In article the new approach to formation of requirements to indicators of reliability of process of functioning of telecommunication systems on the basis of a method of the priority ways, providing considerable decrease in expenses for achievement of demanded quality of process of functioning of telecommunication systems is offered.

Keywords: *telecommunication system; information direction; communication direction; reliability rationing; partially switched networks; readiness factor.*

BIBLIOGRAPHY (TRANSLITERATED)

1. Fisun A.P., Fisenko V.E., Mityaev V.V. Ocenka nadyozhnosti informacionno-telekommunikacionny'x sistem v real'nom masshtabe vremeni na osnove prioritetny'x metodov: monografiya. – Oryol: Izd-vo OGU, 2007. – 191 s.
2. Fisun A.P., Fisenko V.E., Minaev V.A., Mityaev V.V. i dr. Razvitie metodologicheskix osnov informatiki i informacionnoj bezopasnosti sistem: monografiya. – Oryol: OGU, 2004. – 253 s. – Dep. v VINITI 07.07.04. № 1165-V2004.
3. Fisenko V.E. Novaya koncepciya ocenki nadyozhnosti bol'shix informacionno-telekommunikacionny'x sistem // Izvestiya TulGU. Ser. «Vy'chislitel'naya texnika. Informacionny'e tehnologii. Sistemy' upravleniya». – Vy'p. 1. – Tula: Izd-vo TulGU, 2006. – S. 161-165.

УДК 004.057.4

Р.Н. ШУЛЬГИН

**МАТЕМАТИЧЕСКОЕ МОДЕЛИРОВАНИЕ
ПРОЦЕССОВ СБОРА И ОБРАБОТКИ ДАННЫХ
О НАРУШЕНИЯХ ФУНКЦИОНИРОВАНИЯ
В ВЫЧИСЛИТЕЛЬНЫХ СЕТЯХ ГАЗОТРАНСПОРТНОГО ПРЕДПРИЯТИЯ**

В статье рассматривается математическая модель сбора и обработки данных о нарушениях функционирования в вычислительных сетях ГТП, базирующаяся на их потоковом представлении трафика данных контроля и управления с помощью аппарата ненаправленных графов и отличающаяся процедурами разделения путей и выявления нарушений для синхронных и асинхронных сообщений, позволяющая гарантировать характеристики доставки.

Ключевые слова: нарушения; сетевой трафик; процесс информационного обмена; газотранспортное предприятие.

ВВЕДЕНИЕ

Современное газотранспортное предприятие – объект со сложной инфраструктурой как в части функционального назначения структурных подразделений, так и в части их географической распределенности. Эти факторы оказывают существенное влияние на экономическую эффективность производственной деятельности, повышение которой – приоритетная задача любого крупного предприятия. Газотранспортные предприятия отличает четкое разделение функций административного аппарата и функций структурных подразделений (филиалов). Административный аппарат осуществляет управление финансово-хозяйственной деятельностью предприятия и координацию работы его филиалов, распределенных на значительной территории. На филиалы возложены функции исполнения производственных планов и экономических показателей. Каждый филиал имеет, в свою очередь, и административную, и цеховую структуру. Зачастую количество уровней управления, разделяющих администрацию предприятия и производственные подразделения, различается для определенных географических зон. Речь идет о региональных управлениях, объединяющих подразделения одного функционального назначения в пределах региона. Причина возникновения множества уровней управления состоит все в той же географической распределенности, затрудняющей коммуникации, и, как следствие, оперативное управление подразделениями [1].

Из-за прогресса в технологиях сбора и обработки данных число конечных точек в вычислительных сетях АСУ ГТП растет очень быстро и модели связи между конечными точками таких больших сетей становятся трудными для анализа и прогнозирования. На стадии заказа услуги вычислительных сетей АСУ ГТП чаще всего для пользователей затруднено или просто невозможно определить нагрузку между каждой парой конечных точек. Соответственно, при использовании отношений «точка-точка» существенно увеличивается сложность определения требований к качеству обслуживания. В этом случае предлагается более гибкая модель услуги вычислительных сетей АСУ ГТП, которая названа потоком [2].

Поток обеспечивает гарантию характеристик передачи трафика от данной конечной точки к набору других конечных точек и к данной конечной точке от набора всех других конечных точек ВС АСУ ГТП. С точки зрения пользователя, поток эквивалентен звену доступа в сеть. Использование модели потока позволяет пользователю передавать трафик в сеть без необходимости определения всех нагрузок типа «точка-точка».

В настоящее время в рамках Отраслевой системы оперативно-диспетчерского управления (ОСОДУ) Единой системы газоснабжения (ЕСГ) России заканчивается внедрение унифицированной интегрированной АСУ («ГОФО-2») газотранспортным предприятиям с отработкой технических решений на основе современных информационных

технологий и программно-технических средств. Интеграция осуществляется по следующим направлениям: по этапам технологических цепей управления, программно-технических средств и организационно-экономических подсистем. Она обеспечит комплексное использование на газотранспортном предприятии локальных автоматизированных и информационных систем, баз данных (БД), разнотипных средств автоматизации компрессорных станций (КС) и систем передачи данных, локальных вычислительных сетей (ЛВС), персональных электронно-вычислительных машин (ПЭВМ) для повышения эффективности системы управления в целом. ИАСУ – важный фактор ускорения оперативности, экономичности и надежности управления; это основной технологический процесс дальнего транспорта газа в целях бесперебойного снабжения потребителей [2].

Диагностика состояний и защита сетевых ресурсов в компонентах вычислительной сети газотранспортного предприятия требует оперативного выявления состояний компонентов, приводящих к потере ее работоспособности, являющихся следствием отказов оборудования, сбоев программного обеспечения и приложений случайного характера или результатом попыток получения несанкционированного доступа к сетевым ресурсам, проникновения вирусов. Обнаружение таких состояний позволяет своевременно устранить их причину, а также предотвратить возможные последствия.

Проблемам информационного обмена в интегрированной информационной среде промышленных предприятий посвящен целый ряд работ [3-12]. Однако за рамками их рассмотрения остаются вопросы критериев оценки нарушений функционирования и своевременности их обнаружения.

ПОСТАНОВКА ЗАДАЧИ

Сущность модели сбора и обработки данных о нарушениях функционирования в вычислительных сетях ГТП заключается в занятии такой полосы пропускания в сети, при которой будет обеспечена реализация любой допустимой матрицы трафика между конечными точками ВС АСУ ГТП. Матрица трафика является допустимой тогда и только тогда, когда она отображает нагрузку каждой пары конечных точек (m, v) в множестве Q при положительном запросе полосы пропускания d_{uv} так, что для каждой конечной точки v выполняются соотношения

$$\sum_{u \in Q} d_{ux} \leq B_v^{in} \text{ и } \sum_{u \in Q} d_{vu} \leq B_v^{out}.$$

Очевидно, что отсутствует трафик между одноименными конечными точками ВС АСУ ГТП, т.е. $d_{vv}=0$ для всех $v \in Q$.

Полосу пропускания, зарезервированную для ВС АСУ ГТП на ребре $e \in E$, обозначим через y_e . Тогда полная стоимость резервирования полосы пропускания в сети для ВС АСУ ГТП $C = \sum_{e \in E} c_e y_e$, а целью оптимального построения ВС АСУ ГТП является нахождение резервирования соответствующей полосы пропускания с минимальной стоимостью.

МАТЕМАТИЧЕСКОЕ МОДЕЛИРОВАНИЕ ПОТОКА ДАННЫХ О НАРУШЕНИЯХ ФУНКЦИОНИРОВАНИЯ В ВЫЧИСЛИТЕЛЬНЫХ СЕТЯХ

Рассмотрим ситуацию, в которой трафик конечных точек ВС АСУ ГТП асимметричный (произвольный). Это означает, что для конечной точки v значения трафика на входе B_v^{in} и выходе B_v^{out} различны. При асимметричном трафике усложняется расчет даже древовидной топологии ВС АСУ ГТП, так как резервируемая полоса пропускания на ребре (i, j) может быть неодинаковой в разных направлениях, т.е. для дерева T , соединяющего конечные точки ВС АСУ ГТП, $C_T(i, j)$ может быть не равно $C_T(j, i)$. Это объясняется тем, что

$$C_T(i, j) = \min \left\{ \sum_{v \in P_i^{(i, j)}} B_v^{out}, \sum_{v \in P_j^{(i, j)}} B_v^{in} \right\},$$

$$C_T(j, i) = \min \left\{ \sum_{v \in p_i^{(i,j)}} B_v^{in}, \sum_{v \in p_j^{(i,j)}} B_v^{out} \right\}.$$

Пусть дерево для реализации ВС АСУ ГТП, изображенное на рисунке 1, соединяет конечные точки ВС АСУ ГТП $P = \{0, 1, \dots, 4\}$. Трафик конечных точек ВС АСУ ГТП определен следующим образом: для конечных точек 0 и 1 $B_{in} = 3, B_{out} = 6$, а для конечных точек 2, 3 и 4 $B_{in} = 3, B_{out} = 4$. Распределение резервируемых полос пропускания в обоих направлениях для ребер дерева в виде стрелок с соответствующими числами показано на рисунке 1. Например, для ребра (5,6) $C_m(5,6) = 9$ (так как $\sum_{v \in p_5^{(5,6)}} B_v^{out} = 12$ больше, чем $\sum_{v \in p_6^{(5,6)}} B_v^{out} = 9$), $C_T(6,5) = 6$ (так как $\sum_{v \in p_5^{(6,5)}} B_v^{out} = 12$ больше, чем $\sum_{v \in p_6^{(6,5)}} B_v^{out} = 6$). Аналогично можно показать, что для ребра (6,7) $C_T(6,7) = 6$ (так как $\sum_{v \in p_6^{(6,7)}} B_v^{out} = 16$ больше, чем $\sum_{v \in p_7^{(6,7)}} B_v^{out} = 6$), а $C_T(7,6) = 8$ (так как $\sum_{v \in p_7^{(7,6)}} B_v^{out} = 8$ меньше, чем $\sum_{v \in p_6^{(7,6)}} B_v^{out} = 9$).

Для случая асимметричного трафика задача определения древовидной топологии ВС АСУ ГТП с минимальной резервируемой полосой пропускания по сложности не сложнее расчета дерева Штейнера. Поэтому решение данной задачи возможно только аппроксимационными методами.

Задача нахождения оптимальной древовидной топологии ВС АСУ ГТП в случае асимметричного трафика конечных точек может быть сформулирована как задача целочисленного программирования. Определим свойства древовидной топологии графа, соединяющего конечные точки с асимметричной полосой пропускания.

Для ребра (i, j) дерева T будем считать, что оно смещено по направлению к вершине i , если выполняются два условия:

1. $(\sum_{v \in p_i^{(i,j)}} B_v^{in} < \sum_{v \in p_j^{(i,j)}} B_v^{out})$ или $(\sum_{v \in p_i^{(i,j)}} B_v^{in} = \sum_{v \in p_j^{(i,j)}} B_v^{out})$ и $P_i^{(i,j)}$

содержит специальную вершину v).

2. $(\sum_{v \in p_i^{(i,j)}} B_v^{out} < \sum_{v \in p_j^{(i,j)}} B_v^{in})$ или $(\sum_{v \in p_i^{(i,j)}} B_v^{out} = \sum_{v \in p_j^{(i,j)}} B_v^{in})$ и $P_i^{(i,j)}$ содержит специальную вершину v).

Ребро (i, j) считается смещенным, если оно смещено по направлению к вершине i или к вершине j . Ребро считается равновесным, если оно не имеет смещения. Будем считать вершину дерева T корневой, если равновесное ребро подключается к ней.

Возьмем пример сети, изображенной на рисунке 1. Ребро (5,6) равновесное (так как $\sum_{v \in p_5^{(5,6)}} B_v^{in} = 6$ меньше, чем $\sum_{v \in p_6^{(5,6)}} B_v^{out} = 12$, а $\sum_{v \in p_6^{(5,6)}} B_v^{in} = 9$ меньше, чем $\sum_{v \in p_5^{(5,6)}} B_v^{out} = 12$). Таким образом, вершины 5 и 6 являются корневыми. Однако ребро (6,7) смещено по направлению к вершине 7, так как $\sum_{v \in p_6^{(6,7)}} B_v^{out} = 6$ меньше, чем $\sum_{v \in p_7^{(6,7)}} B_v^{out} = 16$, а $\sum_{v \in p_7^{(5,7)}} B_v^{out} = 8$ меньше, чем $\sum_{v \in p_6^{(7,6)}} B_v^{out} = 9$.

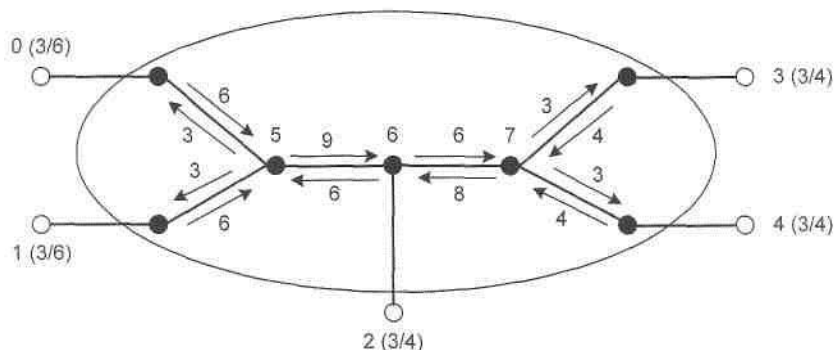


Рисунок 1 – Пример дерева с асимметричным трафиком конечных точек ВС АСУ ГТП

Определим основные свойства равновесных ребер в ориентированном графе. Как было показано выше, в случае симметричного трафика резервируемая полоса пропускания на любом ребре (i, j) определяется выражением

$$C_T(i, j) = \min\{\sum_{v \in P} B_v^{in}, \sum_{v \in P} B_v^{out}\}. \quad (2)$$

Очевидно, что сумма резервируемых полос пропускания на равновесном ребре (i, j) дерева T в обоих направлениях равна $M = C_T(i, j) + C_T(j, i)$. Так, в вышеприведенном примере (рис. 1) суммарная резервируемая полоса пропускания на равновесном ребре $(5, 6)$ $M = \sum_{v \in P} B_v^{in} = 15$.

Очевидно, что использование в дереве T только равновесных ребер позволяет образовать связанную компоненту. Отсюда следует, что корневые вершины образуют дерево, полностью состоящее из равновесных ребер. Таким образом, если удалить равновесные ребра из дерева T , то в каждой из получающихся связанных компонент останется только одна корневая вершина. Обозначим через P_u – компоненту, включающую корневую вершину u . Однако, если дерево T не содержит равновесных ребер, то будет только одна компонента. При смещении всех ребер в компоненте можно показать, что существует уникальная вершина v , такая, что каждое ребро, инцидентное с вершиной v , смещается от нее. При этом считается, что данная вершина v является корневой для компоненты.

Пусть компонента P_u соответствует корневой вершине u . В этом случае каждое ребро (i, j) в совокупности P_u смещается по направлению к вершине j (предполагается, что вершина j располагается дальше от вершины u , чем вершина i). При этом суммарную полосу пропускания компоненты P_u можно определить по формуле

$$C_u = \sum_{v \in (C_u \cap P)} d_T(u, v) (B_v^{in} + B_v^{out}), \quad (3)$$

где $d_T(u, v)$ – длина пути (количество ребер в пути) от вершины u к вершине v в дереве T .

Таким образом, дерево T состоит из совокупности корневых вершин, связанных равновесными ребрами, и связанных компонент P_u для каждой корневой вершины, включающих только смещенные ребра. Обозначим через $S(T)$ совокупность корневых вершин в дереве T , а через $H(T)$ – совокупность равновесных ребер в дереве T . С учетом (3) полоса пропускания дерева T определяется выражением

$$C_S = M|H(T)| + \sum_{u \in S(T)} \sum_{v \in (C_u \cap P)} d_T(u, v) (B_v^{in} + B_v^{out}). \quad (4)$$

Отсюда следует, что каждое дерево ВС АСУ ГТП в общем случае может быть полностью охарактеризовано своей совокупностью корневых вершин. Обозначим такую совокупность вершин через S . Тогда суммарная резервируемая полоса пропускания для совокупности корневых вершин S будет выражаться следующим образом:

$$C_S = Mb + \sum_{v \in P} \min_{u \in S} \{d_T(u, v)\} (B_v^{in} + B_v^{out}), \quad (5)$$

где b – число ребер в дереве Штейнера, соединяющего вершины в совокупности S .

Чтобы построить дерево $T(S)$ для совокупности вершин S таких, что $C_{T(S)} \leq C_S$, во-первых, соединим вершины совокупности S деревом Штейнера и добавим все ребра дерева Штейнера к дереву $T(S)$. Затем объединим все вершины в совокупности S в одну супервершину и построим дерево с поиском в ширину, подключенное к супервершине и соединяющее все конечные точки ВС АСУ ГТП в совокупности P . Добавим ребра дерева с поиском в ширину к дереву $T(S)$. Тогда задача определения оптимальной ВС АСУ ГТП с древовидной топологией эквивалентна определению совокупности вершин S , для которых резервируемая полоса пропускания C_S минимальна. Справедливо и обратное, если T_{opt} – оптимальное дерево, то для него полоса пропускания минимальна, т.е. всегда $C_S \leq C_{T_{opt}}$. Следовательно, определив совокупность узлов S , всегда можно построить оптимальное

дерево $T(S)$. Таким образом, при асимметричном трафике для определения оптимального дерева ВС АСУ ГТП необходимо найти совокупность вершин S , для которых резервируемая полоса пропускания в соединяющих их ребрах минимальна.

Задача расчета совокупности вершин S с минимальной резервируемой полосой пропускания может быть сформулирована как задача целочисленного программирования, если известна одна из вершин в совокупности S . Пусть $x_{i,j}$, y_i и z_e – двоичные переменные, которые принимают значения 0 или 1, причем

- $y_i = 1$, если вершина i входит в совокупность S ;
- $x_{i,j} = 1$, если конечная точка ВС АСУ ГТП инцидентна вершине i ;
- $z_e = 1$, если ребро e принадлежит дереву Штейнера, соединяющего вершины в совокупности S .

Обозначим через $\delta(V)$ совокупность ребер, соединяющих группы вершин V и $(V - V)$ в графе G . Трафик, проходящий через конечную точку j , $B_j = B_j^{in} + B_j^{out}$. Предположим, что заранее известно, что вершина $v \in S$. Тогда формулировка задачи целочисленного программирования (ЦП) для нахождения оптимальной совокупности вершин S , содержащих вершину v , будет иметь вид:

минимизировать величину

$$\sum_{i \in v, j \in P} d_G(i, j) B_j x_{i,j} + M \sum_{e \in E} z_e; \quad (6)$$

с учетом следующих ограничений:

$$\begin{aligned} \forall j \in P: \sum_{i \in E} x_{i,j} &\geq 1 \\ \forall i \in V, \forall j \in P: y_i - x_{i,j} &\geq 0 \\ \forall v \subset V, v \in V, j \in P: \sum_{e \in \delta(v)} z_e - \sum_{i \in v} x_{i,j} &\geq 0 \\ x_{i,j}, z_e &\in \{0, 1\}. \end{aligned}$$

Первые два ограничения определяют, что каждая конечная точка j ВС АСУ ГТП должна быть назначена, как минимум, одной вершине из совокупности S . Третье ограничение гарантирует, что вершины в совокупности S соединяются деревом Штейнера. Это достигается требованием того, что если конечная точка j ВС АСУ ГТП сопоставляется вершине i (данная вершина i принадлежит совокупности S), то вершина i соединяется с вершиной v ребрами дерева Штейнера. Соответственно функция, которую минимизировали, и есть резервируемая полоса пропускания совокупности вершин S . Таким образом, зная, что совокупность S должна содержать вершину из совокупности V , можно определить оптимальное дерево, выполняя следующие шаги.

1) для каждой вершины $v \in V$ выполняется алгоритм целочисленного программирования для расчета оптимальной совокупности вершин S_v , содержащей вершину v (S_v состоит из вершин i , для которых $y_i=1$ в решении задачи целочисленного программирования);

2) определяется дерево $T(S_v)$, у которого резервируемая полоса пропускания минимальна.

Отметим, что хотя задача определения совокупности вершин с минимальной резервируемой полосой пропускания имеет несколько сходств с хорошо известной проблемой распределения ресурсов [5, 7, 8, 10, 11], однако есть одно существенное отличие. Если мы будем рассматривать отдельные вершины в совокупности V как ресурсы, то в рассматриваемом случае полоса пропускания для каждого индивидуального ресурса (вершины) равна 0. Однако выбранные ресурсы как единое целое имеют суммарную величину полосы пропускания, так как они объединяются деревом Штейнера с соответствующими величинами пропускных способностей ребер M . Таким образом, полоса пропускания каждого отдельного ресурса (вершины) определяется полосой пропускания в дереве

Штейнера, объединяющего эти выбранные ресурсы (вершины).

В алгоритме древовидной маршрутизации также возможно отклонение запросов на реализацию ВС АСУ ГТП, даже если имеющихся свободных ресурсов в сети достаточно для их обработки. Это происходит из-за того, что данный алгоритм формирует из звеньев оптимальное дерево без учета имеющейся в наличии свободной полосы пропускания в сети. Если полосы пропускания на выбираемых звеньях для оптимального дерева ВС АСУ ГТП недостаточно, то очевидно, что алгоритм древовидной маршрутизации будет повышать вероятность отклонения запросов.

ЗАКЛЮЧЕНИЕ

При использовании модели потока пользователь ВС АСУ ГТП определяет набор конечных точек, которые должны быть соединены с общей гарантией качества «точка-точка».

Математическое моделирование потока данных о нарушениях функционирования в вычислительных сетях ГТП существенно упрощает технические требования к этим сетям. Поток гарантирует определенную полосу пропускания и позволяет передавать и принимать трафик без необходимости точного его определения от конкретной конечной точки к каждой другой конечной точке ВС АСУ ГТП.

СПИСОК ЛИТЕРАТУРЫ

1. Якименко А.А. Особенности совершенствования системы управления газотранспортным предприятием [Электронный ресурс]. – URL: <http://do.gendocs.ru/docs/index-380553.html>.
2. Автоматизация диспетчерского управления газотранспортным предприятием [Электронный ресурс]. – URL: <http://www.informtransgas.ru/acstp-gofo-2.html>.
3. Еременко В.Т. Синтез сетей передачи данных автоматизированных систем управления на основе критерия неблокируемой маршрутизации / В.Т. Еременко, С.И. Афонин, А.И. Офицеров, О.О. Басов // НАУЧНЫЕ ВЕДОМОСТИ Белгородского государственного университета. Серия «История. Политология. Экономика. Информатика», 2011. – № 7(102). – Выпуск 18/1. – С. 168-176.
4. Еременко В.Т. Оптимизация ресурсов и управление процессами информационного обмена в сетях АСУТП на основе полевых шин / В.Т. Еременко, С.И. Афонин, С.А. Максаков, А.И. Куленич // Вестник компьютерных и информационных технологий, 2011. – № 9. – С. 46-49.
5. Еременко В.Т. Моделирование информационных потоков в сетях передачи данных интегрированных АСУ / В.Т. Еременко, С.И. Афонин, Т.М. Парамохина, Л.В. Кузьмина, Д.А. Плащенков // Информационные системы и технологии, 2011. – № 6. – С. 35-42.
6. Еременко В.Т. Моделирование взаимодействия протокольных реализаций TCP RENO и TCP VEGAS в сети с ограниченной производительностью // Информационные системы и технологии, 2010. – № 1. – С. 109-114.
7. Еременко В.Т. Методы решения задач распределения информационных потоков в сетях передачи данных предприятия на основе резервирования ресурсов / В.Т. Еременко, С.И. Афонин, Л.В. Кузьмина и др. // Информационные системы и технологии, 2012. – № 1. – С. 78-84.
8. Еременко В.Т., Афонин С.И. Создание теоретических основ автоматизации и построения технологической составляющей АСУ территориально распределенных предприятий // Информационные системы и технологии, 2012. – № 2. – С. 99-105.
9. Еременко В.Т. Рекурсивный алгоритм оценки минимальной величины канального ресурса в сети передачи данных / В.Т. Еременко, Л.В. Кузьмина, Д.А. Плащенков, Д.А. Краснов // Информационные системы и технологии, 2012. – № 4. – С. 97-102.
10. Еременко В.Т. Метод проектирования сетей передачи данных, совместимых с неблокируемой маршрутизацией / В.Т. Еременко, А.И. Офицеров, С.А. Черепков // Вестник компьютерных и информационных технологий, 2012. – № 4. – С. 38-46.

11. Еременко В.Т. Решение задач управления сетевыми ресурсами в условиях динамического изменения конфигурации беспроводной сети АСУП / В.Т. Еременко, Д.В. Анисимов, Д.А. Плащенков, Д.А. Краснов, С.А. Черепков, А.Е. Георгиевский // Информационные системы и технологии, 2012. – № 6. – С. 114-119.
12. Еременко В.Т. Анализ моделей управления трафиком в сетях АСУП на основе технологии MPLS / В.Т. Еременко, С.В. Еременко, Д.В. Анисимов, С.А. Черепков, А.А. Лякишев // Информационные системы и технологии, 2013. – № 1. – С. 106-112.

Шульгин Роман Николаевич

Академия ФСО России, г. Орел

Преподаватель

Тел.: 8 953 612 03 11

E-mail: romanspartak1977@mail.ru

R.N. ShUL'GIN (*Teacher*)

Academy of Federal Agency of Protection of the Russian Federation, Orel

**MATHEMATICAL MODELING OF COLLECTING AND DATA PROCESSING
ABOUT OPERATIONAL OCCURRENCES IN COMPUTER NETWORKS
OF GAS TRANSMISSION COMPANY**

In article the mathematical model of collecting and data processing about functioning violations in the GTP computer networks, based on their stream representation of a traffic of data of control and management by means of the device of not directed counts and differing by procedures of division of ways and identifications of violations for the synchronous and asynchronous messages, allowing to guarantee delivery characteristics is considered.

Keywords: *violations; network traffic; process of information exchange; gas transmission company.*

BIBLIOGRAPHY (TRANSLITERATED)

1. Yakimenko A.A. Osobennosti sovershenstvovaniya sistemy' upravleniya gazotransportny'm predpriyatiem [E'lektronny'j resurs]. – URL: <http://do.gendocs.ru/docs/index-380553.html>.
2. Avtomatizaciya dispetcherskogo upravleniya gazotransportny'm prespriyatiem [E'lektronny'j resurs]. – URL: <http://www.informtransgas.ru/acstp-gof-2.html>.
3. Eryomenko V.T. Sintez setej peredachi danny'x avtomatizirovanny'x sistem upravleniya na osnove kriteriya neblokiruemoj marshrutizacii / V.T. Eryomenko, S.I. Afonin, A.I. Oficerov, O.O. Basov // NAUCHNY'E VEDOMOSTI Belgorodskogo gosudarstvennogo universiteta. Seriya «Istoriya. Politlogiya. E'konomika. Informatika», 2011. – № 7(102). – Vy'pusk 18/1. – S. 168-176.
4. Eryomenko V.T. Optimizaciya resursov i upravlenie processami informacionnogo obmena v setyax ASUTP na osnove polevy'x shin / V.T. Eryomenko, S.I. Afonin, S.A. Maksakov, A.I. Kulenich // Vestnik komp'yuterny'x i informacionny'x tehnologij, 2011. – № 9. – S. 46-49.
5. Eryomenko V.T. Modelirovanie informacionny'x potokov v setyax peredachi danny'x integrirovanny'x ASU / V.T. Eryomenko, S.I. Afonin, T.M. Paramoxina, L.V. Kuz'mina, D.A. Plashhenkov // Informacionny'e sistemy' i tehnologii, 2011. – № 6. – S. 35-42.
6. Eryomenko V.T. Modelirovanie vzaimodejstviya protokol'ny'x realizacij TCP RENO i TCP VEGAS v seti s ogranichennoj proizvoditel'nost'yu // Informacionny'e sistemy' i tehnologii, 2010. – № 1. – S. 109-114.
7. Eryomenko V.T. Metody' resheniya zadach raspredeleniya informacionny'x potokov v setyax peredachi danny'x predpriyatiya na osnove rezervirovaniya resursov / V.T. Eryomenko, S.I. Afonin, L.V. Kuz'mina i dr. // Informacionny'e sistemy' i tehnologii, 2012. – № 1. – S. 78-84.
8. Eryomenko V.T., Afonin S.I. Sozdanie teoreticheskix osnov avtomatizacii i postroeniya texnologicheskoy sostavlyayushhej ASU territorial'no raspredelyonny'x predpriyatij // Informacionny'e sistemy' i tehnologii, 2012. – № 2. – S. 99-105.
9. Eryomenko V.T. Rekursivny'j algoritm ocenki minimal'noj velichiny' kanal'nogo resursa v seti peredachi danny'x / V.T. Eryomenko, L.V. Kuz'mina, D.A. Plashhenkov, D.A. Krasnov // Informacionny'e sistemy' i tehnologii, 2012. – № 4. – S. 97-102.
10. Eryomenko V.T. Metod proektirovaniya setej peredachi danny'x, sovместimy'x s neblokiruemoj marshrutizaciej / V.T. Eryomenko, A.I. Oficerov, S.A. Cherepkov // Vestnik komp'yuterny'x i informacionny'x tehnologij, 2012. – № 4. – S. 38-46.

11. Eryomenko V.T. Reshenie zadach upravleniya setevy'mi resursami v usloviyax dinamicheskogo izmeneniya konfiguracii besprovodnoj seti ASUP / V.T. Eryomenko, D.V. Anisimov, D.A. Plashhenkov, D.A. Krasnov, S.A. Cherepkov, A.E. Gergievsckij // *Informacionny'e sistemy' i tehnologii*, 2012. – № 6. – S. 114-119.
12. Eryomenko V.T. Analiz modelej upravleniya trafikom v setyax ASUP na osnove tehnologii MPLS / V.T. Eryomenko, S.V. Eryomenko, D.V. Anisimov, S.A. Cherepkov, A.A. Lyakishev // *Informacionny'e sistemy' i tehnologii*, 2013. – № 1. – S. 106-112.

ТРЕБОВАНИЯ к оформлению статьи для опубликования в журнале «Информационные системы и технологии»

ОБЩИЕ ТРЕБОВАНИЯ

Объем материала, предлагаемого к публикации, измеряется страницами текста на листах формата А4 и содержит от 4 до 9 страниц; все страницы рукописи должны иметь сплошную нумерацию.

В одном сборнике может быть опубликована только одна статья одного автора, включая соавторство.

Плата с аспирантов за публикацию рукописей не взимается.

Аннотации всех публикуемых материалов, ключевые слова, информация об авторах, списки литературы будут находиться в свободном доступе на сайте соответствующего журнала и на сайте Российской научной электронной библиотеки – РУНЭБ (Российский индекс научного цитирования).

ТРЕБОВАНИЯ К СОДЕРЖАНИЮ НАУЧНОЙ СТАТЬИ

Научная статья, предоставляемая в журналы, должна иметь следующие **обязательные** элементы:

- постановка проблемы или задачи в общем виде;
- анализ достижений и публикаций, в которых предлагается решение данной проблемы или задачи, на которые опирается автор, выделение научной новизны;
- исследовательская часть;
- обоснование полученных результатов;
- выводы по данному исследованию и перспективы дальнейшего развития данного направления;
- библиография.

ТРЕБОВАНИЯ К ОФОРМЛЕНИЮ НАУЧНОЙ СТАТЬИ

• Статья должна быть набрана шрифтом Times New Roman, размер 12 pt с одинарным интервалом, текст выравнивается по ширине; абзацный отступ – 1,25 см, правое поле – 2 см, левое поле – 2 см, поля внизу и сверху – 2 см.

• **Обязательные элементы:**

- УДК
- заглавие (на русском и английском языках)
- аннотация (на русском и английском языках)
- ключевые слова (на русском и английском языках)
- список литературы, на которую автор ссылается в тексте статьи.

ТАБЛИЦЫ, РИСУНКИ, ФОРМУЛЫ

• Все таблицы, рисунки и основные формулы, приведенные в тексте статьи, должны быть пронумерованы.

• **Формулы** следует набирать в редакторе формул Microsoft Equation 3.0 с размерами: обычный шрифт – 12 pt, крупный индекс – 10 pt, мелкий индекс – 8 pt. **Формулы, внедренные как изображение, не допускаются!** Русские и греческие буквы, а также обозначения тригонометрических функций набираются прямым шрифтом, латинские буквы – *курсивом*.

• **Рисунки** и другие иллюстрации (чертежи, графики, схемы, диаграммы, фотоснимки) следует располагать непосредственно после текста, в котором они упоминаются впервые. Рисунки, число которых должно быть логически оправданным, представляются в виде отдельных файлов в формате *.eps (Encapsulated PostScript) или TIF размером не менее 300 dpi.

СВЕДЕНИЯ ОБ АВТОРАХ

• В конце статьи приводятся набранные 10 pt сведения об авторах в такой последовательности: фамилия, имя, отчество (полужирный шрифт); учреждение или организация, ученая степень, ученое звание, должность, адрес, телефон, электронная почта (обычный шрифт).

Сведения об авторах предоставляются отдельным файлом и обязательно дублируются на английском языке.