

**XV ЕЖЕГОДНАЯ МЕЖДУНАРОДНАЯ
НАУЧНО-ТЕХНИЧЕСКАЯ КОНФЕРЕНЦИЯ**

IT-ТЕХНОЛОГИИ: РАЗВИТИЕ И ПРИЛОЖЕНИЯ

СБОРНИК ДОКЛАДОВ



12-14 декабря 2018 года. Владикавказ.

МИНИСТЕРСТВО НАУКИ И ВЫСШЕГО ОБРАЗОВАНИЯ РФ
СЕВЕРО-КАВКАЗСКИЙ ГОРНО-МЕТАЛЛУРГИЧЕСКИЙ
ИНСТИТУТ
(ГОСУДАРСТВЕННЫЙ ТЕХНОЛОГИЧЕСКИЙ
УНИВЕРСИТЕТ)

IT-технологии: развитие и приложения

XV ЕЖЕГОДНАЯ МЕЖДУНАРОДНАЯ
НАУЧНО-ТЕХНИЧЕСКАЯ КОНФЕРЕНЦИЯ

Сборник докладов

12–14 декабря 2018 г.

Владикавказ

УДК 004
ББК 73
И92

И92 **ИТ-Технологии: развитие и приложения:** XV Ежегодная Международная научно-техническая конференция: Сборник докладов. – Владикавказ: Северо-Кавказский горно-металлургический институт (государственный технологический университет), 2018. – 340 с.

При финансовой поддержке Российского Фонда фундаментальных исследований в соответствии с грантовой заявкой № 18-07-20096

ISBN 978-5-9500071-1-8

В сборнике представлены материалы, обсуждаемые в рамках работы международной научно-практической конференции «ИТ-технологии: развитие и приложения».

Организационный комитет оставляет за авторами право на самостоятельное изложение своей точки зрения на научное обоснование и практическую реализацию рассматриваемой проблемы.

Точка зрения авторов может не совпадать с позицией оргкомитета.

Авторы несут полную ответственность за содержание представленных материалов.

В издании сохраняется авторское изложение текста: орфография, пунктуация, оформление рисунков, таблиц и формул.

Текст статей не должен противоречить законодательству РФ, а также принятым нормам этики и морали.

**УДК 004
ББК 73**

ISBN 978-5-9500071-1-8

© СКГМИ (ГТУ), 2018
© Дизайн издательство «Веста», 2018

ОРГАНИЗАЦИОННЫЙ КОМИТЕТ КОНФЕРЕНЦИИ

Председатель

Дмитрак Ю.В., д. т. н., проф. (Владикавказ, Россия)

Сопредседатели

Темираев Р.Б., д. с/х. н., проф. (Владикавказ, Россия)

Гроппен В.О., д. т. н., проф. (Владикавказ, Россия)

Члены оргкомитета

Хатагов А.Ч., к.т.н., проф. (Владикавказ, Россия)

Будаева А.А., к.т.н., доц. (Владикавказ, Россия)

Даурова А.А., к.т.н., доц. (Владикавказ, Россия)

Калининченко А. В., к. т. н. доц. (Владикавказ, Россия)

Секретариат

Давыдова Е. (Владикавказ, Россия)

362021 Россия, Владикавказ, ул. Николаева, 44/1-310.

Тел.: 8(8627) 407107, 407518. Факс: 8(8672) 407203

E-mail: groppen@skgtu.ru

СОДЕРЖАНИЕ

Будаева А.А. ПРОГНОЗИРОВАНИЕ ПЕРСОНАЛЬНОЙ УСПЕВАЕМОСТИ СТУДЕНТОВ В ВУЗЕ	9
Воронов М.В. СИСТЕМА ИНТЕЛЛЕКТУАЛИЗИРОВАННОЙ ПОДДЕРЖКИ ПРОЦЕССОВ САМОПОДГОТОВКИ	16
Даурова А. А. ИНФОРМАЦИОННЫЕ ТЕХНОЛОГИИ В ИНДИВИДУАЛИЗАЦИИ ОБУЧЕНИЯ	26
Степанов А.В. ФОРМИРОВАНИЕ ОБЪЕКТНО-ОРИЕНТИРОВАННОГО МЫШЛЕНИЯ ДЛЯ ПОВЫШЕНИЯ РЕЗУЛЬТАТИВНОСТИ ОБУЧЕНИЯ БАКАЛАВРОВ	31
Касьянов В.Н., Касьянова Е.В. WEB-СИСТЕМЫ ПОДДЕРЖКИ ПРИМЕНЕНИЯ ТЕОРЕТИКО-ГРАФОВЫХ МЕТОДОВ ДЛЯ КОНСТРУИРОВАНИЯ ЭФФЕКТИВНЫХ И НАДЕЖНЫХ ПРОГРАММ	38
Назаров С.В. ТЕХНОЛОГИИ ПРОЕКТИРОВАНИЯ ПРОГРАММНЫХ СИСТЕМ	44
Соколова Е.А. ЭКСПЕРИМЕНТАЛЬНОЕ ИССЛЕДОВАНИЕ ЭФФЕКТИВНОСТИ РАБОТЫ АЛГОРИТМОВ КIF	65
Томаев М.Х. ДИСКРЕТНЫЕ ПОДХОДЫ К ОПТИМИЗАЦИИ ПОЛЬЗОВАТЕЛЬСКИХ ПРОГРАММНЫХ КОДОВ	69
Мирошников А. С. МЕТОД ОПТИМИЗАЦИИ ПАРАЛЛЕЛЬНЫХ АЛГОРИТМОВ ДЛЯ SMP-СИСТЕМ	75
Небаев И.А. АНАЛИЗ ПОДДЕРЖКИ ПАРАЛЛЕЛЬНЫХ ВЫЧИСЛЕНИЙ И МНОГОЯДЕРНЫХ СИСТЕМ В СРЕДЕ ПАКЕТА ПРОГРАММНОГО ОБЕСПЕЧЕНИЯ МАТЛАВ ДЛЯ ЗАДАЧ ПОЛИНОМИАЛЬНЫХ ВЫЧИСЛЕНИЙ В РАСШИРЕННЫХ ПОЛЯХ ГАЛУА	82
Пекунов В.В. СОКРАЩЕННАЯ АЛГЕБРАИЧЕСКАЯ НОТАЦИЯ ЗАПИСИ ПАРАЛЛЕЛЬНЫХ ВЫЧИСЛИТЕЛЬНЫХ ТОПОЛОГИЙ В PLANNING C	88
Посохов А.С., Головина И.П. ПОДХОД К ОРГАНИЗАЦИИ МУЛЬТИПРОГРАММНОГО РЕЖИМА ФУНКЦИОНИРОВАНИЯ ВЫЧИСЛИТЕЛЬНЫХ СИСТЕМ	94
Анисимов А.В. ОБЕСПЕЧЕНИЕ БЕЗОПАСНОСТИ DNS	97

Козунова С.С., Бабенко А.А. ОБЕСПЕЧЕНИЕ ИНФОРМАЦИОННОЙ БЕЗОПАСНОСТИ В СЕГМЕНТЕ КОРПОРАТИВНОЙ ИНФОРМАЦИОННОЙ СИСТЕМЫ	103
Ибрагимов Б., Гаджиев Т. ЭФФЕКТИВНОСТЬ КРИПТОГРАФИЧЕСКИХ МЕТОДОВ АКТИВНОЙ ЗАЩИТЫ ИНФОРМАЦИИ В СИСТЕМЕ ТЕЛЕКОММУНИКАЦИЙ.....	107
Алгазин С. Д. МНОГОСЛОЙНЫЙ, НЕЯВНЫЙ АЛГОРИТМ ДЛЯ НЕОДНОРОДНОГО ДВУМЕРНОГО УРАВНЕНИЯ ТЕПЛОПРОВОДНОСТИ.	115
Альшакова Е.Л., Альшакова Е.А. ПРИМЕНЕНИЕ ИНФОРМАЦИОННЫХ ТЕХНОЛОГИЙ ДЛЯ ИССЛЕДОВАНИЯ КОНСТРУКТИВНЫХ И ОБЪЕМНО-ПЛАНИРОВОЧНЫХ РЕШЕНИЙ ЖИЛЫХ ЗДАНИЙ	120
Крутиков Д.И., Бабенко А.А. ВЕРОЯТНОСТНАЯ МОДЕЛЬ ОЦЕНКИ НАДЕЖНОСТИ ИНФОРМАЦИОННЫХ СИСТЕМ	129
Алексеева А.В., Клячкин В.Н. ПРОГРАММНОЕ ОБЕСПЕЧЕНИЕ ДЛЯ КОНТРОЛЯ СТАБИЛЬНОСТИ ВИБРАЦИЙ ГИДРОАГРЕГАТА	134
Lagunovich N.L. THE NEW PROCESS FLOW OF MAKING AND DEVICE-PROCESS SIMULATION OF BIPOLAR STATIC INDUCTION	141
Леонтьев А.Л.; Никитина А.В. РАЗРАБОТКА МОДЕЛИ ГИДРОДИНАМИКИ ГЛУБОКОВОДНОГО ВОДОЕМА НА ОСНОВЕ МОДЕЛИ ПЕРЕМЕННОЙ ПЛОТНОСТИ*	147
Никитина А.В., Семенякина А.А. МАТЕМАТИЧЕСКОЕ МОДЕЛИРОВАНИЕ ИЗМЕНЕНИЯ ГИДРОХИМИЧЕСКИХ ПОКАЗАТЕЛЕЙ КАЧЕСТВА ВОД ПРИБРЕЖНЫХ СИСТЕМ НА СУПЕР-ЭВМ*.....	161
Орлова Н.С. МОДЕЛИРОВАНИЕ ВИБРОКИПЕНИЯ ПОЛИДИСПЕРСНОЙ СРЕДЫ.....	175
V. Popovs, V. Skudnovs, A. Vasiljevs MATHEMATICAL MODELS AND ALGORITHMS OF RADIO WAVE PROPAGATION IN CELLULAR NETWORKS OF MOBILE COMMUNICATION	181
Бурлов В. В., Косолапов В. В., Косолапова Е. В., Ремонтова Л. В. МОДЕЛИРОВАНИЕ ПОВЕРХНОСТЕЙ НА ОСНОВЕ КРИВЫХ ВТОРОГО ПОРЯДКА	199
Гроппен В.О. ИСПОЛЬЗОВАНИЕ ИНФОРМАЦИОННЫХ ТЕХНОЛОГИЙ ДЛЯ АНАЛИЗА ДИАГРАММЫ ХАББЛА	214
Леонтьев А.Л. MATHCAD КАК СРЕДСТВО ВИЗУАЛИЗАЦИИ УРАВНЕНИЙ МАТЕМАТИЧЕСКОЙ ФИЗИКИ	220

Пылаев Н. А. МЕТОДИКА ОЦЕНКИ ВЕРОЯТНОСТНО-ВРЕМЕННЫХ ХАРАКТЕРИСТИК ПЕРЕДАЧИ СООБЩЕНИЙ В СОСТАВНЫХ РАДИОЛИНИЯХ С ИСПОЛЬЗОВАНИЕМ РЕТРАНСЛЯТОРОВ СВЯЗИ НА БЕСПИЛОТНЫХ ЛЕТАТЕЛЬНЫХ АППАРАТАХ.....	233
Черняев А.П., Черняева С.А. ПОСТАНОВКА И ЧИСЛЕННЫЙ АНАЛИЗ НЕКОТОРЫХ ЗАДАЧ ДЛЯ КНОИДАЛЬНОЙ ВОЛНЫ И СОЛИТОНА, КАК РЕШЕНИЙ УРАВНЕНИЯ КОРТЕВЕГА ДЕ ФРИЗА	239
V. Popovs, V. Skudnovs, A. Vasiljevs, A. Ševcenko PROBLEMS OF TELEMEDICINE IN LATVIA	248
Жуков Д.А., Клячкин В.Н. ОТБОР ЗНАЧИМЫХ ПОКАЗАТЕЛЕЙ ПРИ ДИАГНОСТИКЕ ТЕХНИЧЕСКОГО ОБЪЕКТА С ПРИМЕНЕНИЕМ МАШИННОГО ОБУЧЕНИЯ	261
Проскурин А. Е., Токарева И. В. ИСПОЛЬЗОВАНИЕ МОДЕЛИ «НЕОКОГНИТРОНА» В МЕТОДАХ РАСПОЗНАВАНИЯ ОБЪЕКТОВ.....	266
Датиев А.А. ЭКСПЕРИМЕНТАЛЬНЫЙ АНАЛИЗ ЭФФЕКТИВНОСТИ КОМПОЗИТНОЙ ВЕРСИИ МЕТОДА НЕЯВНОГО ПЕРЕБОРА ПРИМЕНИТЕЛЬНО К ЗАДАЧЕ О РАНЦЕ	271
Жуков Р.В., Жуков И.В. КОНЦЕПЦИЯ АВТОМАТИЗАЦИИ ПРОЦЕССА ПОДГОТОВКИ ДОКУМЕНТАЦИИ НА ИНФОРМАЦИОННЫЕ СИСТЕМЫ	277
Ибрагимов Б., Керимов В. ОБ ОДНОМ ПОДХОДЕ К ОЦЕНКЕ РЕСУРСОВ В МУЛЬТИСЕРВИСНЫХ СЕТЯХ СВЯЗИ НА БАЗЕ ЕДИНОГО ИНФОКОММУНИКАЦИОННОГО ПРОСТРАНСТВА.....	283
Кристаллинский В.Р., Колюсь А.Н. ОБ ИСПОЛЬЗОВАНИИ АССОЦИАТИВНЫХ ПРАВИЛ ДЛЯ НЕЧЕТКОГО МОДЕЛИРОВАНИЯ	294
Левин В.И. ИНТЕРВАЛЬНАЯ МАТЕМАТИКА И ВЫЧИСЛЕНИЯ В УСЛОВИЯХ НЕОПРЕДЕЛЕННОСТИ.....	298
Левин В.И. ИНТЕРВАЛЬНАЯ МАТЕМАТИКА И ПРИБЛИЖЕННЫЕ ВЫЧИСЛЕНИЯ.....	305
Носова Т.Н. СОЗДАНИЕ NET-ПРИЛОЖЕНИЯ ДЛЯ УВЕЛИЧЕНИЯ ЭФФЕКТИВНОСТИ ПОИСКОВЫХ ЗАПРОСОВ В БАЗАХ ДАННЫХ.....	312
Пименов В.И., Пименов И.В. МЕТОДИКА АНАЛИЗА БОЛЬШИХ ДАННЫХ В СИСТЕМАХ ПОДДЕРЖКИ ПРИНЯТИЯ РЕШЕНИЙ.....	321
Полещук О. М., Тумор С. В. ИСПОЛЬЗОВАНИЕ Z-ЧИСЕЛ (ЧИСЕЛ ЗАДЕ) ДЛЯ ПОДДЕРЖКИ ПРИНЯТИЯ РЕШЕНИЙ	333

СЕКЦИЯ 1. ИНФОРМАЦИОННЫЕ ТЕХНОЛОГИИ В ОБРАЗОВАНИИ

УДК 681.343.001

Будаева А.А.

к.т.н., доц., Северо-Кавказский горно-металлургический институт (государственный технологический университет), г. Владикавказ, Россия.

ПРОГНОЗИРОВАНИЕ ПЕРСОНАЛЬНОЙ УСПЕВАЕМОСТИ СТУДЕНТОВ В ВУЗЕ

Аннотация: проведен обзор существующих подходов к прогнозированию успеваемости. Предложена процедура прогнозирования персональной успеваемости студентов на базе методов таксономии и эталонов. Прогноз дальнейшего обучения, базируется на прецедентах. Приводятся соответствующие математические постановки.

Ключевые слова: персональная успеваемость, прогнозирование, студенты-аналоги, расстояние, оптимальная таксономия, эталон.

Budaeva A.A.

Ph.D., Assoc., North Caucasus Mining and Metallurgical Institute (State Technological University), Vladikavkaz, Russia.

FORECASTING THE PERSONAL PERFORMANCE OF STUDENTS IN THE UNIVERSITY

Annotation: A review of existing approaches to the prediction of performance. A procedure is proposed for predicting the student's personal performance on the basis of taxonomy and reference methods. The forecast of further education is based on precedents. Corresponding mathematical statements are given.

Keywords: personal performance, forecasting, students-analogues, distance, optimal taxonomy, standard

1. Введение

Процессы, протекающие в системе образования, требуют постоянной объективной оценки, корректировки и управления. Однако без прогнозирования управление невозможно. Поэтому возникает необходимость прогнозирования показателей качества как на завершающих этапах обучения, так и в ходе учебного процесса.

Прогнозирование успеваемости студентов по той или иной дисциплине позволяет сформировать индивидуальную траекторию работы обучаемого в семестре и тем самым повысить уровень их профессиональной подготовки.

Таким образом, актуальна цель работы – разработка и экспериментальная проверка методики, базирующейся на современных математи-

ческих подходах, позволяющей прогнозировать динамику развития основных показателей учебной деятельности студентов.

Приведем краткий обзор существующих подходов к прогнозированию успеваемости.

В работе [1] описывается подход на базе регрессионных моделей, позволяющий выявить связь между уровнем знаний и умений студентов по общепрофессиональным и специальным дисциплинам в зависимости от обеспечивающих курсов с помощью многофакторной линейной регрессионной модели.

В работе [2] показана возможность применения аппарата Марковских цепей для прогнозирования успеваемости студентов. На основании данных об успеваемости на начальных курсах строится прогноз об успеваемости на старших курсах. Расчетные данные сравниваются с экспериментальными посредством составления уравнения парной линейной регрессии $y=a+b*x$.

Статья [3] посвящена разработке алгоритма выявления факторов, влияющих на успеваемость студента. В основе этого подхода лежит математический инструментальный дискриминантного анализа, который реализует классификацию с «обучением». Для проведения классификации студентов по уровню успеваемости используются следующие дискриминантные переменные: средний балл аттестата, пол, тип населенного пункта, количество детей в семье, возраст отца, возраст матери, количество больничных листов в течение года, проводимое в интернете время; наличие родителей, количество друзей, количество прочитанных в месяц книг, средний балл в колледже. Корректировка некоторых из перечисленных факторов (напр., количество прочитанных в месяц книг и др.) позволит повысить качество подготовки учащихся. Дискриминантная модель оптимально разделяет множество объектов на подмножества и проводит классификацию новых объектов в тех случаях, когда неизвестно заранее, к какому из существующих классов они принадлежат.

В статье [4] рассматриваются аспекты применения радиальных базисных нейронных сетей для решения задачи прогнозирования успеваемости студентов. Прогнозируется успеваемость студентов по отдельным дисциплинам текущего семестра. На вход сети поступает следующая информация: средняя оценка студента по предшествующим смежным дисциплинам, освоение которых в полном объеме является необходимым условием; наличие свободного посещения; оценка за тестирование остаточных знаний, проводимое перед началом изучения данной дисциплины; наличие задолженностей по другим дисциплинам. На выходе сети получена прогнозируемая по дисциплине оценка.

Подход, описанный в [5], позволяет получить оценку перспективности абитуриентов на основе использования трехслойной нейронной сети. Цель подхода в выявлении как наиболее одаренных студентов, так

и тех, кому могут потребоваться дополнительные занятия. На входе сети используются следующие характеристики: место жительства; номер школы; выпускные оценки по физике и математике при окончании школы; профессия отца и матери; сведения о семейном доходе. На выходе сети получают оценку перспективности абитуриента.

В [6] предложена методика прогнозирования успеваемости студентов на основе методов кластерного анализа. В основе подхода лежит модификация метода k -средних Мак-Кина. Количество кластеров задается явно и равно 4, каждому кластеру ставится в соответствие типологическая группа, характеризующая успеваемость студентов: «отлично», «хорошо», «удовлетворительно» и «плохо». Прогнозирование выполняется по каждой отдельной дисциплине, в качестве оценочных параметров выступают факторы, значения которых можно оценить в начальный момент изучения дисциплины: уровень начальных знаний студентов (определяется в начале первого занятия по дисциплине); уровень компетенций, сформированных у студентов по первой теме дисциплины (определяется на первой лабораторной работе); количество пропусков занятий студентами на момент составления прогноза. Для прогнозирования задаются эталонные объекты, соответствующие центрам четырех заданных групп. Прогнозирование будущей типологической группы студента осуществляется по расстоянию до ближайшего эталона.

В литературе также описывается подход к прогнозированию успеваемости студентов в зависимости от мотивационных предпочтений учения на основе нечеткой логики [7]. В результате с учетом степени мотивации учения прогнозируется возможность перехода студентов между типологическими группами в зависимости от уровня успеваемости по выбранному предмету.

Анализ литературных источников [1-7] позволяет выделить следующие основные группы методов прогнозирования успеваемости студентов:

- методы, основанные на регрессионных моделях,
- методы, основанные на дискриминантных моделях,
- нейросетевые технологии,
- методы кластерного анализа и таксономии.

Как видно из приведенного выше обзора, подходы к прогнозированию успеваемости весьма разнообразны и, как правило, отличаются:

- конечной целью прогнозирования,
- набором оценочных факторов,
- периодом прогноза.

Одни – просто прогнозируют будущую типологическую группу студентов (по уровню успеваемости, по степени мотивации учения), другие – конечную оценку по отдельной дисциплине, третьи определяют факторы, влияющие на успеваемость, цель четвертых – оценка перспективности абитуриентов и т.д.

Ниже приводится описание подхода на базе методов таксономии и эталонов [8]. Методы таксономии позволяют рассматривать множество исходных данных практически произвольной природы, а также производить разбиение объектов не по одному параметру, а по целому набору признаков, что дает возможность анализа результатов учебной деятельности студентов за весь период обучения.

Отличительные особенности предлагаемого подхода состоят в следующем:

- прогнозируется успеваемость за весь оставшийся период обучения студента;
- в качестве оценочных факторов используются текущие результаты обучения студентов по всем ранее изученным дисциплинам;
- прогноз основывается на прецедентах и состоит в поиске группы студентов-аналога, обучающихся на старших курсах или закончивших обучение.

2. Процедура прогнозирования успеваемости студентов на базе методов таксономии и эталонов

В основе процедуры прогнозирования успеваемости лежит поиск студентов-аналогов выбранного студента, отстоящих от него в пространстве признаков на минимальное расстояние.

Прогноз дальнейшего обучения, базируется на прецедентах (т.е. на сравнении результатов обучения студента за пройденный этап с результатами других студентов, обучающихся на старших курсах, либо закончивших обучение).

Очевидно, что качество прогнозирования, зависит от объема выборки и от степени близости студентов и их аналогов. Поэтому, с целью повышения достоверности прогноза, были приняты следующие **допущения**:

1. Сравняются студенты одного направления
2. Прогнозирование проводится персонально по каждому студенту.
3. Выбранный студент сравнивается только со студентами, закончившими обучение или обучающимися на старших курсах
4. Поиск аналогов осуществляется в евклидовом пространстве.
5. В качестве аналогов рассматриваются студенты, попавшие в ближайший таксон.

Процедура прогнозирования успеваемости состоит из следующих **этапов**:

- 1) Таксономия студентов старших курсов и студентов, закончивших обучение;
- 2) Определение таксона, содержащего студентов-аналогов выбранного студента;
- 3) Прогнозирование успеваемости
- 4) Оценка достоверности прогноза

Суть *первого этапа* состоит в построении таксономии студентов,

удовлетворяющей следующему условию: суммарное расстояние между объектами, попавшими в один таксон, должно быть минимальным. Формально поставленная задача может быть записана в виде (1).

$$\left\{ \begin{array}{l} F(\vec{Z}) = \sum_{\forall i, t_i \in T} \sum_{p \in t_i} \sum_{\substack{q \in t_i, \\ q \neq p}} r(p, q) \cdot z(p, t_i) \cdot z(q, t_i) \rightarrow \min \\ \forall p : \sum_{i=1}^n z(p, t_i) = 1 \\ \forall p, \forall i : z(p, t_i) = 1, 0 \end{array} \right. , \quad (1)$$

где $F(\vec{Z})$ – критерий оптимальности таксономии; $z(p, t_i)$ – булева переменная, равная единице, если p -й объект принадлежит i -ому таксону, и равная нулю если не принадлежит; $r(p, q)$ – евклидово расстояние между p -м и q -м объектами.

Для разбиения на таксоны воспользуемся алгоритмом КРАВ. Алгоритмы семейства КРАВ оперируют отношениями расстояний между несколькими соседними точками, в качестве которых выступают объекты таксономии. Работа алгоритма начинается с построения графа, с минимальной суммарной длиной ребер без петель, иначе, кратчайшего незамкнутого пути, иначе минимального остовного дерева. Помимо характеристики расстояний каждому ребру графа приписывается характеристика напряженности по формуле (2).

$$c(i, j) = d(i, j) \cdot \frac{m_i}{m} \cdot \frac{m_j}{m}, \quad (2)$$

где c – напряженность, d – длина ребра, $\frac{m_i}{m}, \frac{m_j}{m}$ – доли точек, находящиеся по разные стороны от данного ребра.

Самыми напряженными являются длинные ребра, через которые проходят связи между точками двух подмножеств, в каждом из которых содержится половина всех точек исходного множества. Разрыв самого напряженного ребра обеспечивает следующее:

- граница между таксонами проходит по самым напряженным ребрам;
- средняя напряженность внутренних ребер в таксонах минимальна;
- таксоны имеют почти одинаковое число точек.

Разбиение будет проводится до тех пор, пока будет происходить улучшение критерия оптимальности $F(\vec{Z})$, заданного в (1).

Целью *второго этапа* является выбор таксона, содержащего студентов-аналогов для заданного студента h , на основании (3). В качестве таксона аналогов выбирается таксон, суммарное расстояние до студентов в

котором от заданного студента минимально.

$$\forall h \in S : \exists t_f \in T : \sum_{p \in I_f} r(h, p) = \min_i \sum_{q \in I_i} r(h, q), \quad (3)$$

где S – множество всех студентов; h – студент, для которого выполняется поиск студентов-аналогов; T – множество таксонов студентов; t_f – таксон аналогов выбранного студента; $r(h, p)$ – евклидово расстояние между выбранным студентом и его p -м аналогом.

На *третьем этапе* непосредственно выполняется прогнозирование успеваемости выбранного студента на основании анализа данных успеваемости найденных аналогов. В качестве прогнозной оценки по дисциплине используется среднее значение оценок студентов-аналогов в соответствии с (4).

$$\forall h \in S, \forall d \in D : P(h, d) = \frac{1}{m_f} \sum_{q \in I_f} X(q, d), \quad (4)$$

где S – множество всех студентов; h – студент, для которого выполняется прогнозирование успеваемости; t_f – таксон аналогов выбранного студента; $P(h, d)$ – прогноз успеваемости по d -ой дисциплине для h -го студента; $X(q, d)$ – оценка по d -ой дисциплине q -го аналога, m_f – количество аналогов, попавших в таксон t_f .

На *четвертом этапе* вычисляется оценка достоверности прогноза. Для этого выполняется прогнозирование уже известной успеваемости выбранного студента и оценивается максимальное отклонение полученного прогноза по сравнению с реальными данными в соответствии с (5).

$$\left\{ \begin{array}{l} \forall h \in S, \forall d \in D_1 : P'(h, d) = \frac{1}{m_f} \sum_{q \in I_f} X(q, d) \\ \forall h \in S : \delta_h = \max_d (|P'(h, d) - X(h, d)|) \end{array} \right. , \quad (5)$$

где S – множество всех студентов; h – студент, для которого выполняется прогнозирование успеваемости; t_f – таксон аналогов выбранного студента; D_1 – множество дисциплин, по которым у выбранного студента известна успеваемость; $P'(h, d)$ – значение прогнозируемой оценки по d -ой дисциплине для h -го студента; $X(q, d)$ – оценка по d -ой дисциплине q -го аналога, $X(h, d)$ – реальная оценка по d -ой дисциплине у h -го студента, m_f – количество аналогов, попавших в таксон t_f , δ_h – максимальная ошибка прогноза для h -го студента.

Экспериментальные исследования предложенной процедуры прогнозирования успеваемости проводились на кафедре автоматизированной обработки информации СКГМИ (ГТУ) на студентах-бакалаврах направления «Информатика и вычислительная техника» в 2016 году. Всего в эксперименте оценивались результаты обучения 307 студентов, в том

числе: 124 обучающихся и 183 выпускника. Для оценки достоверности прогноза выполнялось прогнозирование уже известных результатов обучения студентов. Максимальная ошибка прогноза составила 0,55 балла.

Стоит отметить, что существует ряд факторов, которые могут приводить к снижению качества прогноза [9, 10]. Так, например, поиск аналогов должен осуществляться только среди студентов, обучающихся на одном направлении. Другими словами, набор дисциплин у студентов должен максимально совпадать с набором дисциплин у их потенциальных аналогов. Однако постоянные изменения образовательных программ в вузах в связи с переходом на новые образовательные стандарты, приводит к тому, что количество общих дисциплин у студентов разных курсов уменьшается, что приводит к ухудшению качества прогнозирования, выполняемое по описанной процедуре. Качество прогноза также зависит от степени схожести студента с его аналогами.

3. Заключение

В работе представлена процедура прогнозирования персональной успеваемости студентов. Теоретической базой для процедуры служит аппарат таксономии и эталонов. Прогнозирование выполняется на основе данных об успеваемости у найденных студентов-аналогов. Аналоги выбираются из числа студентов старших курсов и студентов, закончивших обучение, что дает возможность строить прогноз на весь период обучения. Эксперименты показали удовлетворительное совпадение результатов расчета с экспериментальными данными.

Таким образом, предложенная процедура представляется перспективной, и может быть использована для прогнозирования успеваемости студентов.

Список литературы

1. Моисеев В.Б., Зубков А.Ф., Деркаченко В.Н. Прогнозирование успеваемости студентов по общепрофессиональным и специальным дисциплинам на основе регрессионных моделей. // Научно-технические ведомости СПбГПУ № 6, 2010, С. 169-173
2. Гоголева Н.Г., Тарасова О.Ю. Прогнозирование успеваемости студентов вуза на основе цепей Маркова (на примере СПб ГЭТУ «ЛЭТИ»). // Научные труды SWorld, 2016. Т. 5. № 44. С. 104-109.
3. Панова Н.Ф., Денисова Н.В. Классификация студентов по уровню успеваемости с помощью аппарата дискриминантного анализа. // Вестник Оренбургского государственного университета. 2014. № 8 (169). С. 33-36.
4. Прошкина Е.Н., Бальшова И.Ю. Анализ и прогнозирование успеваемости студентов на основе радиальной базисной нейронной сети. // Технические науки: традиции и инновации: материалы III Междунар. науч. конф. (г. Самара, март 2018 г.). – Казань: Молодой ученый, 2018. – iv, С. 24-28
5. Ясинский И.Ф., Семенова М.Б. Опыт прогнозирования успеваемости студентов при помощи нейросетевой технологии. // Вестник ИГЭУ, вып. 5, 2007, С.29-31

6. Шевченко В.А. Прогнозирование успеваемости студентов на основе методов кластерного анализа. // Вестник ХНАДУ, вып. 68, 2015, С. 15-18
7. Апатова Н.В., Гапонов А.И., Майорова А.Н. Прогнозирование успеваемости студентов на основе нечеткой логики. // Современные наукоемкие технологии, № 4, 2017, С. 7-11
8. Гроппен В.О., Будаева А.А. Эталоны как уникальный инструмент постановки и решения задач теории принятия решений. // В сборнике: IT-технологии: Теория и Практика. Материалы семинара. 2017. С. 3-20.
9. Будаева А.А. Об одном подходе к прогнозированию успеваемости студентов// В сборнике: IT-технологии: Теория и Практика. Материалы семинара. 2018. С. 28-33.
10. Будаева А.А. Прогнозирование успеваемости студентов вуза методами таксономии. // Вестник Института цивилизации. 2007. № 7. С. 124-127

УДК 378.1

Воронов М.В.

доктор технических наук, профессор,
Московский государственный психолого-педагогический университет,
Москва, Россия. mivoronov@yandex.ru.<http://mgppu.ru/>

СИСТЕМА ИНТЕЛЛЕКТУАЛИЗИРОВАННОЙ ПОДДЕРЖКИ ПРОЦЕССОВ САМОПОДГОТОВКИ

Аннотация: в работе приведены принципы создания системы интеллектуальной поддержки индивидуальной подготовки обучаемого (СИПИП). Приводятся содержательная и формальная постановки задачи, предметные области и условия наиболее эффективного функционирования СИПИП.

Ключевые слова: индивидуальная подготовка, структурирование знаний, система интеллектуальной поддержки самоподготовки обучаемого.

Voronov M.V.

Doctor of Technical Sciences, Professor,
Moscow State Psychological and Pedagogical University, Moscow, Russia
mivoronov@yandex.ru <http://mgppu.ru>

INTELLECTUALIZED SUPPORT OF SELF-TRAINING PROCESSES

Annotation: The paper presents the principles of creating a system of intellectual support for individual student training (CIPP). An informative and formal statement of the problem, subject areas and conditions for the most efficient functioning of the CIPP are given.

Keywords: individual training, knowledge structuring, intellectual support system for the student's self-preparation.

Введение

Развитие информационных технологий проникает во все сферы современной жизни. Весьма активно проникают они и в сферу образования, благотворно влияя на повышение его эффективности. Этот процесс начался с

«перевода в цифру» имеющихся учебных материалов, формирования различного рода информационных баз, внедрение в учебный процесс современных технологий доставки информации. В настоящее время все больше усилий направляется на создание интеллектуализированных средств поддержки процессов обучения в целом и самообучения в частности.

Благодаря развитию вычислительной техники и компьютерных технологий накопленные огромные массивы научных, учебных и учебно-методических сведений, стали потенциально доступны всем и каждому в любом месте и в любое время суток. Казалось бы, «бери и учись»! Однако ситуация свидетельствует о неутешительной тенденции: несмотря на существующий потенциал и высокие технологии, уровень профессиональной подготовки выпускников стремительно снижается.

Налицо усугубляющаяся проблемная ситуация: растущий объем знаний необходимый для успешной деятельности входит в противоречие с ограниченностью возможностей их освоения. Действительно, с одной стороны, объем выработанных знаний стремительно растет и им в достаточной мере должно овладевать новое поколение. Для этого им предоставляются все новые и новые технологии доставки знаний. С другой стороны психо-физиологические возможности человека по освоению знаний не могут «угнаться» за лавинообразным нарастанием этих знаний [1]

Несомненно, это противоречие объективно и оно всегда будет иметь место. Однако его острота в значительной мере ситуативна. Конечно, можно отобрать наиболее способных обучающихся, создать им благоприятные условия и обеспечить высокий уровень их мотивации к обучению. Однако в массовом порядке пойти по этому пути не позволяют имеющиеся у общества ресурсы. Как следствие, образовательное сообщество находится в постоянном поиске выхода из складывающейся ситуации.

На пути создания эффективных средств поддержки процессов обучения стоят и проблемы, обусловленные, в конце концов, незнанием механизмов мыслительной деятельности человека в целом и его способности к обучению в частности. Отсюда, в частности, и следует несовершенство многих существующих методик обучения и воспитания.

Одно из интенсивно развиваемых в настоящее время направлений в сфере образования связано с повышением личной активизации обучающихся, обусловленной использованием высоко адаптируемых средств поддержки самообучения (самоподготовки). Основу таких средств составляют модели представления знаний и алгоритмы освоения этих знаний, позволяющие строить соответствующие программные системы поддержки индивидуального обучения в самых различных жизненных ситуациях.

Постановка задачи

Учебная деятельность требует от обучаемого умения самостоятельно мыслить, определять цели своего обучения и формулировать учебные задачи, выбирать рациональные приемы и способы их решения, кон-

тролировать и оценивать свою работу. К сожалению, эти обстоятельства становятся барьером для многих обучаемых, и они оказываются не готовыми к выполнению своей роли. Причин тому много. Главные из них обусловлены тем, что обучение – творческая деятельность индивидуума, поскольку требуется прилагать усилия, чтобы осознанно воспринимать и осознать учебный материал, осмысливать объективные связи между понятиями, формировать новые для себя умения и навыки.

В большинстве случаев практики причиной возникновения такой ситуации задачи является отсутствие в данный момент у обучающихся некоторых необходимых для решения образовательных задач элементов знаний и, конечно же, недостаточный уровень мотивированности учиться должным образом. В этой непростой ситуации целесообразно предоставить им возможность активизировать имеющиеся у них знания и направить усилия по логически и содержательно понятному пути поиска нужной новой информации. Конечно, перед этим и в первую очередь необходимо идентифицировать сложившуюся ситуацию, выяснив, в чем причина затруднений?

Исследования показали, что, несмотря на эту крайне сложную в целом психолого-педагогическую ситуацию, в ее основе лежат два обстоятельства: нет понимания понятий, которые используются при изложении материала (постоянно возникают вопросы типа «что это?») и/или как собственно реализовать указанное действие (как это действие осуществить?).

Несомненно, в каждой конкретной ситуации обучаемый самостоятельно редко может четко определить, что конкретно он недопонимает. Но не вызывает сомнения то, что чем лучше структурирована и систематизирована совокупность знаний, подлежащих усвоению в процессе обучения, чем в большей степени обучающимся ясны цели изучения той или иной учебной дисциплины и значимость овладения системой содержащихся в ней знаний и умений, тем легче и прочнее эти знания и умения будут ими усваиваться [2].

Следовательно, если обучающимся в определенной последовательности при поддержке эффективного интерфейса будут заданы четкие вопросы, предложены и прокомментированы ответы на них и, если потребуется, указаны пути дальнейших действий, то они, как правило, смогут успешно продвигаться по своей образовательной траектории.

Для практической реализации выдвинутой гипотезы, и это крайне важно, на пути поиска нужных ему сведений обучаемый здесь и сейчас должен быть ситуационно принужден к проявлению своей интеллектуальной активности по схеме: ищи, объясни, сделай, если не можешь, подскажем как. Заметим, что примерно по такой схеме действует преподаватель в ходе педагогического диалога со своим обучаемым.

Современное развитие математики, информатики и информацион-

ных технологий позволяют разрабатывать соответствующие средства – систему интеллектуальной поддержки индивидуальной подготовки обучаемого (СИПИП). Эта система должна обеспечить обучаемому поддержку на всех этапах изучения данного курса: первичного ознакомления и проработки деталей, формирование умений навыков применять полученные знания, контроля и самоконтроля своей компетентности в данной сфере деятельности.

Психолого-педагогические основы построения СИПИП

Преподавателями часто произносятся слова: основная задача занятия заключается не столько в сообщении информации, сколько в обучении оперировать имеющимися и вновь приобретаемыми знаниями. Эффективно же оперировать знаниями можно только в том случае если имеет место понимание соответствующей информации, вне понимания полноценный учебный процесс невозможен.

Понимание, как операция мышления, связана с усвоением нового содержания, включением его в систему устоявшихся идей и представлений. Для понимания характерно ощущение ясной внутренней связанности, организованности рассматриваемых явлений, причем понимание достигается только на базе знаний и умений, уже добытых в предшествующем опыте[3]. В педагогической науке делается акцент на том, что понимание есть психический процесс включения новой информации о чем-либо в пережитый опыт, в усвоенные ранее знания и постижение на этой основе смысла и значения события. По-существу же это процесс установления связи неизвестного, нового с уже известным, составление о чем-либо правильного понятия [4]. Эта творческая составляющая процесса обучения при наличии эффективных технологий поиска и доставки знаний и становится объектом все более пристального внимания ученых и преподавателей.

В чем причины непонимания учебного материала? Обычно называют самые различные причины. Это, в частности:

- внешнее несоответствие представлений между вводимой и встречавшейся ранее информацией (использование иных обозначений, терминов и их интерпретации);
- неспособность обучаемого перейти на иной (обычно более высокий) уровень обобщения и абстракции;
- неготовность при рассмотрении данной материалы применять знания, полученные в ходе изучения иных учебных дисциплин;
- некомфортная психологическая обстановка во время занятий и др.

Понимание получаемой информации реализуется в ходе взаимодействия когнитивных и мнемических процессов, направленных на установление смысловых связей между вводимой информацией и имеющимися у обучаемого базисными знаниями. Через такую связь раскрывается сущность предметного содержания получаемых сведений, их смысл [5].

Именно поэтому в основе непонимания учебного материала обучаемым лежит наличие логических лакун, являющиеся следствием незнания (непонимания) им ряда фрагментов знаний (в первую очередь базисных), полученных ранее. Это обстоятельство обуславливает формирование у него фрагментарной картины мира и, как следствие, трудности при самостоятельном формировании необходимых логических умопостроений.

Практически всегда обучение происходит в ходе диалога обучаемого с авторами изучаемого материала или транслирующего его преподавателя, который может реализовываться различными способами. Несомненно, наиболее продуктивен персонифицированный прямой диалог. В прямом диалоге преподавателя с обучаемым реализуются процессы взаимодействия двух сознаний, в ходе которых происходит самоопределение и саморазвитие каждого участника процесса обучения. Преподаватель в ходе такого общения, оценивания и переосмысливая свою деятельность, может применить иные технологии (методики, подходы, формы) обучения, адаптируя их к сложившейся конкретной ситуации. К сожалению, в рамках традиционных форм обучения индивидуальное общение обучающего и обучаемого происходит фрагментарно, ибо оно по большому счету крайне дорого.

Примечание. Собственно передачей знаний процесс обучения не ограничивается. В процессе общения преподавателя с обучаемым происходит творческий диалог личностей, в ходе которого существенную роль играют артистизм, любопытство, воображение и другие свойственные человеку способности. Кроме того режим диалога невозможен без включения рефлексии обучаемого. В итоге в ходе диалога он вынужден оценивать свою собственную деятельность, и тем самым происходит формирование его тезауруса, в его сознании возникают новые связи между понятиями – реализуется процесс его самообучения.

Важно отметить следующее обстоятельство. В ходе самостоятельной работы, работая с текстами или с информацией, представленной в иной форме, также происходит диалог обучаемого с авторами этой информации, только последние здесь занимают пассивную позицию: они уже высказали свои суждения при ее разработке. Как следствие, нарушена обратная связь – связь обучаемого с обучаемым. Каким-то способом эту связь следует восстановить. Именно в этом заключается методологическая суть предлагаемой СИПИП.

Максимально структурированные знания обычно представляет собой совокупность идентификаторов объектов и свойств этих объектов с объединяющими их связями. В учебном курсе вводятся некоторые базисные понятия, например, путем описания их свойств, на основе которых и строится логика подачи все новых и новых знаний. Эту логику в аудиториях, в электронной среде обучения и в учебных материалах благодаря своим интеллектуальным возможностям и кропотливому труду воспроизводит преподаватель.

Обучающиеся должны отследить логику изложения, понять представленный материал и построить в своем сознании достаточно адекватный ее образ. Необходимость решения такого рода задач требует от обучающихся напряженной и сложной мыслительной деятельности. К сожалению, очень многие из обучаемых в достаточной мере это не демонстрируют. Как следствие на пути освоения учебных программ, у большинства обучающихся формируется барьер, наличие которого проявляется в снижении мотивации к обучению. При этом включается так называемая положительная обратная связь: обучающийся (по каким-то причинам) не освоил данный раздел, затем, как следствие, столкнулся с непониманием нового, что не только усложняет возможность понимания последующего материала, но и увеличивает его психологический дискомфорт. Постепенно возникает ситуация, когда он уже не может успешно продолжать учиться. Если в такой ситуации со стороны всех участников образовательного процесса не принимать дополнительных усилий, цели обучения (несмотря на формально пройденный учебный курс) не будут достигнуты.

Могут ли средства представления знаний и информационные технологии восполнить этот пробел? Нам представляется, что да, если они приобретут такие свойства интеллектуальности, которые позволят им на основе результатов анализа ситуации активно реагировать на действия обучаемого: помогать ему в понимании материала, направлять его мысли и побуждать его к цели достигающим действиям.

Сталкиваясь с непониманием обучаемым чего-то, перед преподавателем встает задача объяснения неизвестных ему (или не понимаемых им) элементов знаний. В ходе такого объяснения выявляется, как правило, непонимание обучаемым тех компонентов знаний, которые согласно программам должны быть освоены на предыдущих этапах получения образования. Анализ последних может привести к выявлению незнаний обучаемым еще более элементарных фрагментов знаний. При наличии желания и времени можно установить нижний уровень знаний обучаемого и на основании этих сведений построить процесс обучения. Эти соображения также были положены в теорию разработки СИПИП.

До каких пор идет этот процесс диалога преподавателя с обучаемым? Обычно до тех пор, пока данное понятие не будет состоять из понятых обучаемым элементов знаний и, что крайне важно, взаимосвязей между ними.

Явное представление логического перехода к компонентам формирующим данный фрагмент знаний (его наглядная декомпозиция) не только формирует у обучаемого логику его построения. Оно может способствовать восстановлению в памяти обучаемого ранее изучаемых им знаний, поскольку содержание и смысл информации осознается лучше, если логические связи в ней выражены достаточно явно, особенно между пред-

метом мысли (субъектом) и тем, что о нем говорится (предикатом). Этим активизируется мыслительная деятельность обучаемого, стимулируется постановка им вопросов к изучаемой информации и поиск ответов на них. Конечно, целостное понимание учебного материала зависит не только от объективной его сложности и наличия в словаре обучаемого соответствующей лексики. Решающее влияние на понимание оказывают объем и глубина его знаний в той области действительности, которой посвящен данный материал. Вот почему в учебном процессе при переходе к новому, сложному материалу очень большое внимание должно уделяться восстановлению в памяти учащихся содержания базисных знаний, необходимых для понимания нового.

Для эффективного ведения процесса обучения совершенно необходима перманентная оценка уровня понимания учебного материала, чтобы иметь представление о том, какой уровень подготовки достигнут обучаемым. Для оценки уровня понимания учебного материала применяется такая характеристика качества знаний, как осознанность. Используя этот термин, имеют в виду уровень понимания связей между понятиями, суждениями и умозаключениями, умение обосновывать этапы решения задач и как следствие, представления о возможности применения полученных знаний. К сожалению, понимание как педагогическая категория не определена операционально. В этой связи преподаватель вынужден определять уровень освоения материала интуитивным путем, подбирая, например, систему вопросов и анализируя ответы на них. Система интеллектуализированной поддержки должна обладать способностью выявлять «узкие места» знаний обучаемого и рекомендовать рациональные шаги к их устранению. Иначе говоря, интеллектуализация средств поддержки процессов самоподготовки в первую очередь и заключается в том, чтобы в автоматизированном режиме проводить экспертизу создавшейся педагогической ситуации и операционализировать последующие действия обучаемого.

Математическая модель СИПИП

Будем исходить из того, что рассматриваемый учебный курс прошел полномасштабную разработку, и весь необходимый учебный, учебно-методический научный материал данного учебного курса в вербальной форме находится в нашем распоряжении. Основная идея построения СИПИП заключается в интерпретации учебного процесса как процесса технологического [6].

Действительно, любой процесс обучения реализуется согласно определенной технологии. Последняя понимается нами, как совокупность знаний о том, каким образом преобразовать конкретно данное (в данном случае исходное состояние субъекта как обучающегося) в требуемое (результатирующее состояние этого обучающегося). Понятно, что каждый процесс изложения учебного материала по конкретному курсу может

быть представлен в виде упорядоченной совокупности частных подпроцессов – процессов изучения частей курса. Последние в свою очередь могут быть декомпозированы на процессы изучения еще более мелких (частных) фрагментов курса и т.д. (см. рис.1). Процесс декомпозиции должен продолжаться до таких малых порций знаний, которые в учебно-методическом материале рассматриваются как элементарные (т.е. не подлежащие последующей декомпозиции).

Независимо от уровня обобщения процесс получения некоего результата TP_i может быть записан в виде совокупности четырех множеств

$$TP_i = \langle TD_i, X_i, Y_i, U_i \rangle \quad (1)$$

где X_i – множество входных компонентов, которые при условиях U_i обеспечивают в ходе технологического действия TD_i получение результата Y_i .

Структуризация исходного учебного материала заключается в том, чтобы весь учебный курс описать в виде упорядоченного множества такого рода формализмов. При этом принципиально важно, чтобы в результате декомпозиции все конечные элементы учебного процесса были представлены в виде элементарных для данного рассмотрения и отображаемых моделями вида (1) концептов [7]

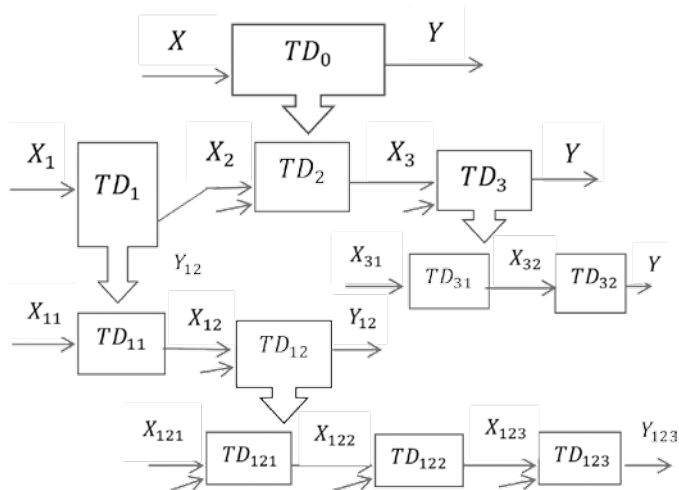


Рис. 1. Декомпозиционная схема логики изложения учебного материала

Компоненты в (1) разделим на две группы. К первой отнесем предикативные концепты (концепты действий), которые описывают отдельные технологические действия и выражаются, как правило, различными глагольными формами. В рамках конкретной учебной программы совокупность всех до-

ступных пользователю концептов действий составит базу концептов действий (БКД) данной предметной области. Важно подчеркнуть, что суть технологического подхода к структуризации учебного материала заключается в том, что именно концепты действия выступают в роли признака декомпозиции. Иначе говоря, декомпозиция технологии есть не что иное, как разбиение обобщенного действия на упорядоченную последовательность действий его составляющих.

В структуре про позиции текста (предложения) за предикатом признают ведущую роль уже многие современные исследователи, поскольку именно он однозначно указывает на определенный тип отношения между сущностями. Более того, семантическое ядро каждого высказывания определяет глагол. Недаром говорят, что в предложении нет подлежащего, дополнения, обстоятельства, если нет сказуемого.

Ко второй группе отнесем все остальные, главным образом предметные и признаковые концепты. Совокупность такого рода концептов составят базу предметных концептов (БПК). Каждый ее участвующий в реализации конкретного технологического действия компонент относится либо к базовому (исходному) понятию, либо выступает результатом другого, «предыдущего» технологического действия. В СИПИП всю информацию о предметных концептах предлагается формализовать в виде соответствующей концептуальной модели предметной области описанной в работе [8].

Проведя декомпозицию исходного учебного материала и дав формальное описание всех компонентов соответствующих элементарных концептов (как показывает опыт, предпочтительнее в виде фреймов), получим структурированную модель учебного материала, которую можно интерпретировать как граф. Организация построения на этом графе соответствующих путей и реализация движений по ним будет отвечать логике освоения данного учебного материала. При этом с учебно-методической точки зрения исходный учебный материал и предметно-педагогическое содержание построенной модели будут идентичны.

Опишем теперь некоторые составляющие алгоритма организации использования СИПИП, снабженной соответствующим множеством описанных моделей и баз знаний. Пусть обучаемый обращается к СИПИП за помощью в рассмотрении какой-то части курса. В этом случае ему будет предоставлена возможность самостоятельно воспроизвести элементарное действие, которое в соответствии с логикой учебного процесса является заключительным элементарным действием этой части курса. В том случае если он проходит соответствующий тест, то делается вывод о достаточном уровне владения данным фрагментом курса. В противном случае СИПИП «спускается» по схеме логики учебного курса (по графу знаний) на один шаг вниз и тест повторяется, но уже с предыдущим по логике изложения элементарным концептом действия. Такая процедура повторяется до тех пор, пока обучаемый не сможет выполнить тест на предоставленное элементарное действие или оно в данном разделе

отсутствует. Последнее означает, что обучаемый не обладает базовыми знаниями, необходимыми исходными для изучения данного раздела.

Прохождение теста на выполнение данного элементарного действия (или некоторого их множества) означает, что обучаемый определил нижнюю точку своих знаний по данному разделу курса и может начинать осваивать последующие (вышестоящие по логике изложения курса) элементарные действия, которые ранее оказывались ему незнакомыми. Поскольку каждое последующее из предлагаемых действие базируется исключительно на знаниях, которыми он, как показывают тесты прохождения предыдущих этапов, обладает, эта процедура (процедура подъема вверх по схеме логики учебного курса) должна им выполняться. Возможность реализации данного действия (или выхода из цикла повторения) по самой своей сути является безусловной, поскольку оно является элементарным и зиждется исключительно на знаниях, которыми обучаемый обладает. Если же данное действие не может быть выполнено, то система вернет его к выше описанному нисходящему движению – к повторению более простых (в смысле логики изложения учебного материала) действий. Эта ситуация возникает тогда и там, где и когда выясняется, что он не в достаточной мере освоил (или забыл) предыдущий материал.

Заключение

В учебных материалах их автор закладывает свою методику изложения, преследуя определенные учебные цели. Задачу выявления соответствующих сведений и на их основе активизации целевой деятельности обучающегося призваны решать интеллектуализированных систем поддержки самоподготовки. Для этого, исходя из описания процесса обучения как технологии, предлагается сформировать такую модель структуризации вербально представленной информации, чтобы имелась возможность разработать алгоритмы фиксации и анализа деятельности обучаемого, синтезировать правила вывода его рациональных последующих действий и создав комфортную для его самоподготовки среду, обеспечить повышение эффективности самостоятельной подготовки.

Такого рода системы, объединив интеллект разработчиков учебных и учебно-методических материалов с возможностями перспективных систем искусственного интеллекта, могут составить содержательную основу будущей образовательной среды информационного общества.

Литература

1. Воронов М.В. Письменский Г.И. Закономерности, тенденции и противоречия в системе современного образования. //Иновации в образовании, №4, 2012 С. 10-28.
2. Михеев В.И. Моделирование и методы теории измерений в педагогике. – М.: УРСС, 2010.
3. Коджаспирова Г. М., Коджаспиров А. Ю. К 57 Педагогический словарь: Для студ. высш. и сред.пед. учеб. заведений. – М.: Издательский центр «Академия», 2003. – 176 с.

4. Иванова С. С. Развитие понимания текста младшими школьниками через театрализованную деятельность // Молодой ученый. №1, 2011.–С. 207-208.
5. Калмыкова З.И. Понимание школьниками учебного материала. //Вопросы психологии. №1, 1986. С 87-94.
6. Воронов М.В. Разработка методов формализации знаний (технологический подход): Монография. – М.: Изд-во СГУ, 2016. 263с.
7. Андреев. Д. А., Воронов М. В.. Метод построения онтологии технологических действий. // Вестник СГТУ. № 3 (67). 2012. –С. 160–168.
8. Антонов И.В., Воронов М.В. Метод построения онтологии предметной области. //Вестник Санкт-Петербургского государственного университета технологии и дизайна. – Серия 1. Естественные и технические науки. № 2, 2010. – С. 28-32.

УДК 300-399-330.65

Даурова А. А.

к.т.н., доцент каф. ИВТ СКГМИ(ГТУ). Владикавказ, Россия, albina_daurova@mail.ru

ИНФОРМАЦИОННЫЕ ТЕХНОЛОГИИ В ИНДИВИДУАЛИЗАЦИИ ОБУЧЕНИЯ

Аннотация: В статье рассматривается применение математического моделирования для поиска оптимальной образовательной траектории, учитывающей индивидуальные особенности обучаемого. Приводятся математические модели задач поиска методов подачи учебного материала, позволяющих минимизировать время его усвоения и обеспечить заданный уровень его усвоения.

Ключевые слова: информационные технологии, эффективность обучения, индивидуализация обучения, математическое моделирование, оптимизация

Daurova A. A.

Ph.D., associate professor ICT SKGMI (GTU). Vladikavkaz, Russia. albina_daurova@mail.ru

INFORMATION TECHNOLOGIES FOR PERSONALIZED LEARNING

Abstract: The article discusses the use of mathematical modeling to find the optimal educational trajectory, taking into account the individual characteristics of the student. Mathematical models of the tasks of searching for methods for presenting educational material are presented, which make it possible to minimize the time of its mastering and to ensure a given level of its mastering.

Keywords: information technologies, learning efficiency, individualization of learning, mathematical modeling, optimization

Информатизация образования предполагает целый ряд направлений, отличающихся своими задачами и тем, какая роль в учебном процессе отводится компьютеру и средствам коммуникации. Это подготовка учащихся к жизни в современных информационных условиях; развитие дистанционных форм образования; внедрение ИТ-технологий в организацию и управление педагогическим процессом и обра-

зовательными системами; разработка и внедрение информационных технологий обучения, когда компьютер является средством обучения. В рамках второго направления особый интерес представляет исследование эффективности применения информационных технологий в решении задач индивидуализации обучения и издание соответствующих методических рекомендаций по их практическому использованию.

Психолого-педагогическая наука относит проблему учета индивидуальных особенностей учащихся к традиционным, поскольку на практике процесс обучения в основном ориентируется на средний уровень обученности и развития способностей к учению, из-за чего реализовать свои потенциальные возможности может не каждый ученик. Готовя учащихся к усвоению нового материала, преподаватель определяет какие знания и каким образом должны быть привлечены, разрабатывает систему заданий различной степени трудности и объема, выделяет основной и варьируемый материал, готовит индивидуальные и групповые задания и прорабатывает целый ряд других методов работы. Преподавателю также необходимо заранее предвидеть возможные затруднения учащегося и рекомендовать пути их преодоления. [1,2]

Для повышения эффективности процесса обучения следует использовать предварительное моделирование, в частности, методы математического моделирования, без которого решение задач в образовательной системе в настоящее время становится практически невозможным.

Исходными параметрами математической модели учебного процесса должны стать, в первую очередь, обучаемость и логическая структура учебного курса, а кроме того, квалификация преподавателя, материально-техническое обеспечение учебного процесса. Выходными же параметрами служат характеристики обратной связи. Вообще говоря, выбор вида модели зависит от текущей образовательной задачи.

Рассмотрим примеры построения математических моделей двух задач.

1. Задача первая

1.1. Содержательная постановка задачи

Дано: В ходе занятия n учащихся должны усвоить учебный материал, представляющий собой K понятий. В распоряжении преподавателя имеется m методов подачи материала, относящихся к каждому понятию. Предполагается, что преподаватель может прогнозировать время усвоения каждым учащимся каждого понятия при использовании каждого метода подачи материала.

Требуется: Выбрать методы подачи материала таким образом, чтобы суммарное время его усвоения было минимальным.

1.2. Математическая модель задачи

Введем следующие обозначения:

$z(i,j,k)$ – булева переменная, равная единице, если для формирова-

ния k -го понятия у j -го учащегося преподаватель пользуется i -м методом, и равная нулю в противном случае;

$t(i, j, k)$ – прогнозируемое время формирования k -го понятия у j -го ученика при использовании i -го метода;

m – количество используемых преподавателем методов обучения;

n – количество учащихся;

K – количество формируемых понятий.

Тогда формально задача представляется следующим образом:

$$\left\{ \begin{array}{l} \sum_k \max_i \max_j z(i, j, k) \cdot t(i, j, k) \rightarrow \min; \\ \forall k, \forall j: \sum_i z(i, j, k) = 1; \\ \forall k, \exists i: \sum_{j=1}^{j=n} z(i, j, k) = n; \\ \forall i, \forall j, \forall k: z(i, j, k) = 1, 0. \end{array} \right. \quad (1)$$

1.3. Форма представления исходных данных и решение

Для каждого k -го понятия формируется матрица M_k . Строки матрицы соответствуют различным методам обучения, которые могут быть использованы преподавателем при формировании k -го понятия, а столбцы – учащимся. Тогда в ячейке $M_k(i, j)$ содержится прогнозируемое время формирования k -го понятия i -м методом у j -го учащегося.

Пусть, например, имеется матрица M_k :

5	8	1	12
4	11	3	7
6	9	5	10
7	14	4	5
8	16	2	20

Поиск решения задачи прост. В каждой строке находим максимальный элемент и выбираем из них минимум. Очевидно, минимальное прогнозируемое время усвоения всеми студентами k -го понятия равно 10. Для достижения этого результата следует использовать третий метод обучения (третья строка).

2. Задача вторая

2.1. Содержательная постановка задачи

Дано: В ходе занятия n учащихся должны усвоить учебный материал, представляющий собой K понятий. В распоряжении преподавателя

имеется m методов подачи материала, относящихся к каждому понятию. Предполагается, что преподаватель может прогнозировать время на усвоения каждым учащимся каждого понятия при использовании каждого метода подачи материала применительно к оценкам «3», «4» и «5».

Требуется: Выбрать методы подачи материала, чтобы суммарное время его усвоения было минимально, а средний балл был не ниже заданной величины S .

2.2. Математическая модель задачи

Введем следующие обозначения:

$z(i,j,k,q)$ – булева переменная, равная единице, если для формирования k -го понятия у j -го учащегося преподаватель пользуется i -м методом с прогнозной оценкой q , и равная нулю в противном случае;

$t(i,j,k,q)$ – прогнозируемое время формирования k -го понятия у j -го учащегося при использовании i -го метода с прогнозной оценкой q ;

m – количество используемых преподавателем методов обучения;

n – количество учащихся;

K – количество формируемых понятий;

S – требуемое значение среднего балла.

Тогда формально задача представляется следующим образом:

$$\left\{ \begin{array}{l} \sum_k \max_i \max_j \sum_q z(i, j, k, q) \cdot t(i, j, k, q) \rightarrow \min; \\ \forall k, \forall j : \sum_i \sum_q z(i, j, k, q) = 1; \\ \forall k, \exists i : \sum_{j=1}^{j=n} \sum_{q=3}^{q=5} z(i, j, k, q) = n; \\ \sum_i \sum_j \sum_k \sum_q q \cdot z(i, j, k, q) = S \cdot n \cdot K; \\ \forall i, \forall j, \forall k, \forall q : z(i, j, k, q) = 1, 0. \end{array} \right. \quad (2)$$

2.3. Форма представления исходных данных и решение

Для каждого k -го понятия формируется матрица M_k . Строки матрицы соответствуют различным методам обучения, которые могут быть использованы преподавателем при формировании k -го понятия, а столбцы – учащимся. В каждой ячейке $M_k(i,j)$ содержатся три числа: первое число – прогнозируемое время формирования k -го понятия i -м методом у j -го учащегося на “тройку”, второе – на “четверку”, третье – на “пятерку”.

5,7,9	7,8,11	1,3,5	12,13,16
4,5,6	11,13,16	3,4,8	7,9,12

6,8,10	9,10,11	5,8,11	10,13,15
7,9,14	14,15,16	4,7,10	5,10,15
8,11,14	16,19,22	2,5,8	20,22, 24

Пусть дана приведенная выше матрица M_k , тогда нижняя граница прогнозируемого времени усвоения всеми студентами k -го понятия равна **10**, верхняя граница – **24**. В первом случае средний балл равен «3», во втором – «5».

2.4. Алгоритм поиска решения

Шаг 1. T = нижняя граница времени формирования k -го понятия.

Шаг 2. Выделяются значения $z(i,j,k,q)$, удовлетворяющие системе (2) и, одновременно, максимизирующие сумму баллов S_1 .

Шаг 3. Если $S_1=S$ или $S_1>S$, то перейти к шагу 6, в противном случае – к шагу 4.

Шаг 4. Если T меньше верхней границы, то перейти к шагу 5, в противном случае – к шагу 7.

Шаг 5. $T = T + 1$, перейти к шагу 2.

Шаг 6. Вывод вектора переменных Z , времени T . Перейти к шагу 8.

Шаг 7. Вывод «Нет решения».

Шаг 8. Конец алгоритма.

Пусть требуемый средний балл $S=4.75$. Тогда решение задачи будет следующим.

Первая итерация. $T = 10$; $S_1 = 4$.

Вторая итерация. $T=11$; $S_1 = 4,5$.

Третья итерация. $T=12$; $S_1 = 4,5$.

Четвертая итерация. $T=13$; $S_1 = 4,75$.

Таким образом, минимальное время, обеспечивающее среднюю успеваемость на уровне 4.75 баллов, составляет 13 единиц времени, и достигается первым методом подачи материала (первая строка).

Как можно видеть, моделирование задачи позволяет найти оптимальную образовательную траекторию и, тем самым, интенсифицировать все уровни образовательного процесса.

Литература:

1. Аткинсон Р., Бауэр Г., Кротерс Э. Введение в математическую теорию обучения. М.: Мир, 1969. – 488с.
2. Горбунова Л. И., Субботина Е. А. Использование информационных технологий в процессе обучения // Молодой ученый. – 2013. – №4. – С. 544-547. – URL <https://moluch.ru/archive/51/6685/>.

УДК 004.942:001.891

Степанов А.В.

доктор технических наук, профессор кафедры информатики и вычислительной техники Новокузнецкого института (филиала) Кемеровского государственного университета, Новокузнецк, Россия, stepal.rdtc.ru

ФОРМИРОВАНИЕ ОБЪЕКТНО-ОРИЕНТИРОВАННОГО МЫШЛЕНИЯ ДЛЯ ПОВЫШЕНИЯ РЕЗУЛЬТАТИВНОСТИ ОБУЧЕНИЯ БАКАЛАВРОВ

Аннотация: статья посвящена технологии формирование объектно-ориентированного мышления у бакалавров по направлению подготовки 09.03.01 «Информатика и вычислительная техника». Приведены основные этапы предлагаемой технологии и рекомендации по ее внедрению.

Ключевые слова: объектно-ориентированное мышление, обучение, бакалавриат, стратегия поиска решений.

Stepanov A.V.

Doctor of Technical Sciences, Professor, Department of Informatics and Computer Engineering, Novokuznetsk Institute (branch), Kemerovo State University, Novokuznetsk, Russia, stepal.rdtc.ru

FORMATION OF OBJECT-ORIENTED THINKING TO INCREASE THE EFFICIENCY OF TEACHING BACHELORS

Abstract: The article is devoted to the technology of the formation of object-oriented thinking among bachelors in the direction of training 09.03.01 "Computer Science and Engineering". The main stages of the proposed technology and recommendations for its implementation are given.

Key words: object-oriented thinking, learning, bachelor, solution search strategy.

Введение

Бакалавриат – первый уровень высшего образования, основное назначение которого – практико-ориентированное обучение. Бакалавр получает фундаментальную подготовку в широкой области знаний по выбранному направлению и не имеет узкой специализации. Бакалавр по направлению подготовки 09.03.01 «Информатика и вычислительная техника» должен уметь ставить и решать практические задачи из различных областей знаний с использованием современных вычислительных средств. Получение профессиональных навыков основывается на изучении фундаментальных законов, численных методов, языков программирования и других дисциплин. Тем не менее, это является недостаточным без формирования особого подхода к решению инженерных задач – объектно-ориентированного мышления.

О термине «объектно-ориентированное мышление»

Термин – это важнейший инструмент профессионального научного общения. Непоследовательность и неточность в применении терминов может оказаться причиной отсутствия взаимопонимания между сту-

дентами и преподавателями в процессе обучения и совместной исследовательской деятельности. Именно они определяют информационное содержание специального текста, являясь своеобразными ключами, организующими, структурирующими и кодирующими накопленную информацию. Одной из отличительных черт вузовской информатики является умение правильно использовать и толковать специальные термины, умение отвечать на вопрос: что это?

Многие из используемых технических терминов и словосочетаний являются многозначными, суть и значение которых зависят от совокупности различных факторов, необходимых для понимания и/или объяснения какого-либо явления действительности или контекста. В простейшем случае контекст это текстовое или словесное окружение.

Термин “объектно-ориентированное мышление” (ООМ) тесно связан с объектно-ориентированным подходом (ООП) – совокупностью компьютерных технологий, использующих объектно-ориентированные методы. Понятие “объектно-ориентированные методы”, в свою очередь, является очень обширным. В частности оно может означать объектно-ориентированное проектирование, программирование, анализ, объектную технологию и другое, то есть фактически целую философию разработки систем и представлений знаний на базе современных достижений в технологии проведения вычислительных экспериментов [1]. Однако ООМ – это нечто особое, не сформулированное ранее, хотя существует даже книга с таким названием, в которой автор Мэтт Вайсфельд пытается помочь читателю разобраться в концепциях проектирования, кроющихся за кодом. Эти концепции он и называет объектно-ориентированным мышлением. [2].

На наш взгляд, ООМ – это инструмент для решения практических задач в различных областях науки и техники, основанный на широком спектре представлений, позволяющих строить модели с использованием виртуальных объектов. Под виртуальным объектом понимается несуществующий объект, имеющий имя и его назначение в системе.

Место технологии формирования ООМ в учебном процессе

Базис, необходимый для формирования ООМ, бакалавры получают в первом семестре из дисциплины «Информатика». На первой лекции определяется специфика вузовской информатики. Студент должен научиться грамотно толковать и использовать специальные термины, четко отвечать на вопрос, – что это? Вводятся понятия многозначности, абстракций, предпочтений, контекста и многих других, являющихся базой объектно-ориентированного мышления. Иллюстрируется многозначность термина информация и уместность его применения там, где речь идет о данных. Обсуждая величие и сложность реального мира, вводится понятие объекта, класса, события, совокупности свойств и др.

В дисциплине «Программирование», изучаемой во втором учебном семестре, рассматриваются примеры использования объектно-ориентированных концепций, но основной «мозговой штурм» проводится по окончании семестра на двухнедельной учебной практике, когда изучение учебных дисциплин второго семестра закончено.

Руководитель практики «по получению первичных профессиональных умений и навыков, в том числе первичных умений и навыков научно-исследовательской деятельности» предлагает решить нетривиальную задачу одного из научных направлений, которая не решается известными методами. Студенты при этом пользуются всеми современными средствами коммуникации, знакомятся со статьями руководителя и других ученых, участвуют в дискуссиях, разрабатывают собственные методы решения поставленной задачи.

Пример задачи, для решения которой необходимо ООМ

В качестве примера необходимости использования объектно-ориентированного мышления можно привести процесс подготовки компьютерной программы для расчета состава кинематических цепей в задаче компьютерного синтеза структур механизмов. Он основан на получении целочисленных решений системы, состоящей из трех уравнений. Упрощенный вид системы для плоских кинематических цепей с одноподвижными вращательными кинематическими парами или парами пятого класса p_5 имеет следующий вид

$$\begin{cases} p_5 = \tau + (\tau - 1)n_{\tau-1} + \dots + i \cdot n_i + \dots + n_1, \\ n = 1 + n_{\tau-1} + \dots + n_i + \dots + n_1, \\ W = 3n - 2p_5 \end{cases}$$

Здесь: τ – наиболее сложное звено кинематической цепи, образующее τ кинематических пар с другими звеньями; n_i – число звеньев, добавляющих в цепь i кинематических пар; W – подвижность кинематической цепи [3].

Решения этой системы для небольшого количества звеньев получают вручную с помощью различного рода математических преобразований. С увеличением значения n ручные расчеты становятся невозможными, а при расширении классов применяемых кинематических пар система еще и несколько усложняется. В ней появляется переменная m , принимающая целые значения от нуля до четырех.

$$\begin{cases} p = \tau + (\tau - 1)n_{\tau-1} + \dots + i \cdot n_i + \dots + n_1, \\ n = 1 + n_{\tau-1} + \dots + n_i + \dots + n_1, \\ W = (6 - m) \cdot n - \sum_{k=m+1}^5 (k - m) \cdot p_k \end{cases}$$

Здесь t – число общих связей, наложенных на систему.

Предварительный анализ системы уравнений позволяет сделать вывод о том, что для нее не существует классических численных методов решения. Это система с физической неделимостью переменных и объектов расчета. Методы решения таких уравнений и систем лежат в области целочисленного программирования или дискретной оптимизации. Поисковые методы здесь тоже неприемлемы, поскольку невозможно разработать функцию отклика.

Для разработки технологии решения задачи и стратегии поиска целочисленных решений познакомимся с терминами предметной области. Что такое кинематическая цепь? Кинематическая цепь (англ. kinematic chain) – это связанная совокупность относительно подвижных звеньев, образующих между собой парные соединения или кинематические пары. Пример кинематической цепи представлен на рисунке 1. На этом рисунке фигуры n_2 , n_3 и n_4 представляют собой обозначения звеньев, соединенных между собой с помощью специальным образом подготовленных поверхностей. Два соединенных друг с другом звена образуют кинематическую пару (обозначенную небольшим кружком). Индексы в обозначениях звеньев соответствуют числу их геометрических элементов. На схеме под геометрическими элементами понимают число сторон и углов фигур, а в реальном механизме – те поверхности звеньев, которые образуют кинематические пары. Кинематическая цепь должна обладать заданной степенью подвижности W , выражающейся целым положительным числом.

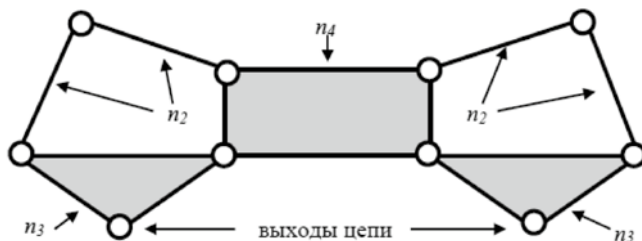


Рис. 1 – Схематическое изображение кинематической цепи

Разработка стратегии решения задачи

Рассматривая нижнее уравнение системы, можно сделать вывод о том, что заданная подвижность цепи W может быть обеспечена только с помощью определенных (вполне конкретных) наборов звеньев и пар. Количество звеньев набора известно (оно равно n), а количество образуемых кинематических пар – нет. Это дает нам основание считать, что процесс решения системы надо начинать с того уравнения, в котором есть заданное значение n . Заменяя бесконечные ряды, имеющиеся в системе на конечные суммы, представим ее в следующем виде.

$$n = 1 + \sum_{i=1}^{\tau-1} n_i \quad (1)$$

$$W = (6 - m) \cdot n - \sum_{k=m+1}^5 (k - m) \cdot p_k \quad (2)$$

$$p = \tau + \sum_{i=1}^{\tau-1} i \cdot n_i \quad (3)$$

Если в правой части уравнения (1) каким-то образом сформировать набор различных по сложности звеньев, общее число которых равно n , и подставить их количества в уравнение (2), то можно получить общее число кинематических пар, необходимых для создания цепи. Всего разновидностей кинематических пар (классов) – пять. Подставляя различные наборы пар, получаемые путем простого перебора, вкупе с набором звеньев в уравнение (3) можно получить фактическое значение подвижности цепи для каждого из вариантов. В случае, когда значение правой части уравнения (3) равно заданному значению W , – это одно из возможных целочисленных решений системы. Его нужно зарегистрировать и продолжить перебор. Если ни для одного из наборов пар не получено решения, то нужно повторить эту процедуру с другим набором звеньев. Ее продолжают до тех пор, пока не будут проведены вычислительные эксперименты со всеми возможными наборами звеньев, общее число которых равно n . На этом можно считать, что стратегия решения задачи определена, и можно приступить к созданию виртуального (не существующего) специализированного вычислительного устройства для получения целочисленных решений системы.

Разработка виртуального специализированного устройства

Прежде всего, определим номенклатуру необходимых узлов, входящих в состав устройства и их назначение. Для создания различных вариантов наборов звеньев, представляющих правую часть уравнения (1) необходим генератор наборов звеньев. Для получения общего числа кинематических пар по уравнению (2) необходим калькулятор. Для создания различных вариантов состава кинематических пар, общее число которых равно p , необходим генератор наборов пар. Для сравнения полученного значения подвижности цепи с заданным значением W , необходимо устройство сравнения или компаратор. Для фиксации варианта целочисленного решения необходимо регистрирующее устройство – регистратор. Для вывода всех полученных решений нужен визуализатор. Возможно, потребуются и какие-то логические элементы для организации цикличности процесса поиска решений.

Переходим к построению структурной схемы специализированного вычислительного устройства (СВУ). Для этого соединяем выбранные компоненты друг с другом в соответствии с разработанной стратегией решения задачи (рис. 2).

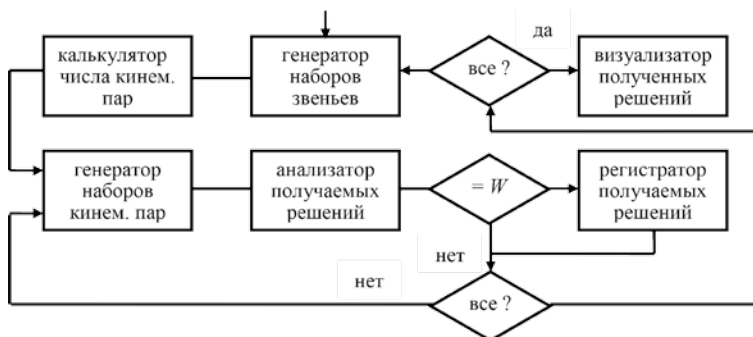


Рис. 2 – Структурная схема специализированного вычислительного устройства

Программная реализация СВУ

Работу разработанного специализированного вычислительного устройства удобно эмулировать с помощью компьютерной программы. После инициализации переменных в работу включается генератор наборов звеньев. В каждом такте своей работы он должен выдавать целое число, состоящее из τ разрядов, в каждом из которых записаны числа звеньев, привносящих в цепь от одной до τ кинематических пар. Этот узел может быть представлен в виде виртуального счетчика, содержащего τ разрядов и работающего в системе счисления с основанием $n + 1$, поскольку в каждом из разрядов может быть записано любое число от нуля до n . Добавляя единицу в младший разряд счетчика, можно обеспечить полный перебор всех возможных вариантов наборов звеньев, отбирая из них такие, для которых сумма разрядов счетчика равна n . Более подробно эти вопросы рассмотрены в [3]. Эффективность работы генератора на базе программно-реализуемого счетчика можно повысить путем динамического изменения основания системы счисления каждого из разрядов [4]. Можно разработать счетчики, позволяющие вводить ограничения на использование звеньев конкретной сложности и на использование кинематических пар определенных классов [5]. Аналогичным образом эмулируется работа генератора кинематических пар.

Программная реализация других узлов виртуального специализированного вычислительного устройства не представляет сложностей.

Заключение

Объектно-ориентированное мышление формируется не просто. Однако, элементы такой парадигмы, вводимые уже в первом семестре учебных программ студентов, обучающихся по направлению «Информатика и вычислительная техника», и развитие этого эффективного стиля мышления на всех последующих курсах позволяют дать бакалаврам еще один из эффективных инструментов решения практических задач.

Список литературы

1. Грэхем, Иан. Объектно-ориентированные методы. Принципы и практика / Иан Грэхем: Пер. с англ. – М.: Издательский дом Вильямс, 2004. – 880 с.:ил.
2. Вайсфельд М. Объектно-ориентированное мышление. – СПб.: Питер, 2014. – 304 с.
3. Степанов А.В. решение универсальной структурной системы профессора Л.Т. Дворникова / Вестник Кузбасского государственного технического университета. 2007. № 3. С. 43-46.
4. Степанов А.В. Программно-реализуемые счетчики с динамически изменяемыми основаниями систем счисления их разрядов / Вестник компьютерных и информационных технологий. 2017. № 10 (160). С. 49-56.
5. Степанов А.В. Счетчики с выключаемыми разрядами и изменяемым основанием системы счисления в компьютерных процедурах, реализующих метод простого перебора / Вестник Кемеровского государственного университета. 2014. № 3-1 (59). С. 46-50.

СЕКЦИЯ 2. СОЗДАНИЕ ЭФФЕКТИВНЫХ ПРОГРАММНЫХ ПРОДУКТОВ

УДК 004.422

В.Н. Касьянов

д.ф.-м.н., профессор, ИСИ СО РАН, НГУ, Новосибирск, Россия
kvn@iis.nsk.su

Е.В. Касьянова

к.ф.-м.н., доцент, ИСИ СО РАН, НГУ, Новосибирск, Россия
kev@iis.nsk.su

WEB-СИСТЕМЫ ПОДДЕРЖКИ ПРИМЕНЕНИЯ ТЕОРЕТИКО- ГРАФОВЫХ МЕТОДОВ ДЛЯ КОНСТРУИРОВАНИЯ ЭФФЕКТИВНЫХ И НАДЕЖНЫХ ПРОГРАММ

Аннотация: работа посвящена методам и средствам поддержки применения теоретико-графовых методов для конструирования эффективных и надежных программ. В ней также рассмотрены работы по созданию электронного словаря WikiGRAPP по теории графов и его применениям в информатике и программировании, а также электронной энциклопедии WEGA теоретико-графовых алгоритмов решения задач информатики и программирования.

Ключевые слова: теория графов, электронный словарь, электронная энциклопедия, графовые алгоритмы.

V.N. Kasyanov

Dr.Sc., Professor, Institute of Sociology, SB RAS, Novosibirsk State University, Novosibirsk, Russia, kvn@iis.nsk.su

E.V. Kasyanova

Ph.D., Associate Professor, Institute of Sociology, SB RAS, Novosibirsk State University, Novosibirsk, Russia
kev@iis.nsk.su

WEB-SYSTEMS FOR SUPPORTING THE APPLICATION OF THEORETICAL AND GRAPH METHODS FOR CONSTRUCTING EFFICIENT AND RELIABLE PROGRAMS

Abstract: The work is devoted to methods and means of supporting the use of graph-theoretic methods for constructing efficient and reliable programs. It also discusses the work on creating a WikiGRAPP electronic dictionary on graph theory and its applications in computer science and programming, as well as the WEGA electronic encyclopedia of graph-theoretic algorithms for solving computer science and programming problems.

Keywords: graph theory, electronic dictionary, electronic encyclopedia, graph algorithms.

Введение

Теория графов из академической дисциплины все более превращается в средство, владение которым становится решающим для успешного применения компьютеров во многих прикладных областях.

Проблема терминологии, без сомнения, является одной из основных проблем в применении теоретико-графовых методов в программировании и информатике. Терминология в прикладной теории графов далеко не устоялась и постоянно развивается, в том числе, в связи с активным развитием теоретико-графовых методов решения задач программирования, а также из-за постоянного расширения их применения на новые предметные области, такие как Web-графы, социальные сети, семантический Web, базы знаний, сети белок-белковых взаимодействий, библиографические сети и т. д.

Существует большое количество открытых веб-систем, аккумулирующих знания по различным предметным областям, связанным с теоретико-графовыми методами. Среди них всем известная Википедия, а также более специализированные системы, посвященные отдельным прикладным областям, такие как MathWorld или AlgoWiki. MathWorld – это наиболее обширный математический ресурс паутины, предоставляемый в качестве бесплатной услуги компанией Wolfram Research, создавшей хорошо известную систему Mathematica. С 1995 сайт MathWorld активно развивается и поддерживается. В настоящее время MathWorld – это открытая энциклопедия по математике, которая содержит почти 13600 статей и считается не только самым ярким и самым читаемым интернет-ресурсом по математике, но и одним из самых надежных. Её статьи широко упоминается в журналах и книгах. MathWorld продолжает расти и развиваться при поддержке тысяч активных пользователей. AlgoWiki – это открытая энциклопедия по свойствам алгоритмов и особенностям их реализации на различных программно-аппаратных платформах от мобильных платформ до экзафлопсных суперкомпьютерных систем с возможностью коллективной работы всего мирового вычислительного сообщества, которая разрабатывается с 2014 в НИВЦ МГУ как вики-система и в настоящее время содержит порядка 200 статей, посвященных в основном алгоритмам линейной алгебры.

Доклад посвящен методам и средствам поддержки применения теоретико-графовых методов для конструирования эффективных и надежных программ, работа над которыми ведется в Институте систем информатики им. А.П. Ершова СО РАН при частичной финансовой поддержке Российского фонда фундаментальных исследований (грант РФФИ № 18-07-00024). В нем рассматриваются работы по созданию электронного словаря WikiGRAPP по теории графов и её применениям в информатике и программировании и электронной энциклопедии WEGA теоретико-графовых алгоритмов решения задач информатики и программирования.

1. Словарь WikiGRAPP

В 1999 г. в издательстве «Наука» вышел в свет наш толковый словарь по теории графов и её применению [1], который охватывал основные связанные с графами термины из монографий, вышедших на русском языке. Это был первый словарь по графам в информатике, и он вызвал большой интерес среди читателей. Электронная версия словаря получила название GRAPP (GRaphs and their APPLications).

Новое исправленное и пополненное издание словаря [2], работа над которым была завершена в 2009 г., представляет собой расширение первого словаря и включает в себя более 1000 новых терминов из статей, рефераты которых публиковались в РЖ «Математика» в разделе «Теория графов», а также из томов ежегодных конференций «Graph-Theoretic Concepts in Computer Science» и книг серии «Graph Theory Notes of New York». На его базе была подготовлена новая версия электронного словаря WikiGRAPP, которая стала расширяемой.

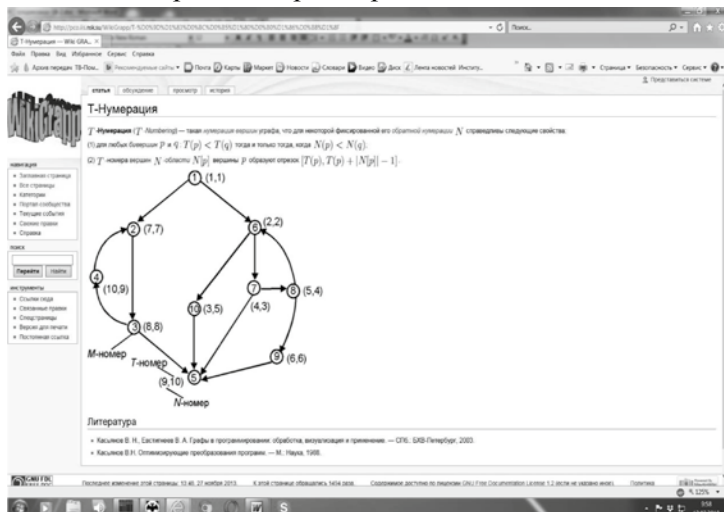


Рис. 1. Вики-словарь WikiGRAPP

Новый электронный словарь WikiGRAPP уже является вики-системой и поддерживает удобный поиск и интерактивное взаимодействие с пользователями по своему пополнению и развитию (Рис. 1). Для его создания мы использовали систему MediaWiki написанное на препроцессоре гипертекста (PHP) свободно распространяемое программное обеспечение, предназначенное для создания гипертекстовых «вики»-систем таких вебсайтов, структуру и содержимое которых пользователи могут сообщать изменять с помощью инструментов, предоставляемых самими сайтами. Описания терминов в словаре сопровождаются рисунками и гиперссылками. К настоящему времени завершена работа по наполнению словаря WikiGRAPP до уровня, покрывающего наши печатные издания.

Начальная версия вики-системы WikiGRAPP прошла государственную регистрацию, и сейчас ведется работа по совершенствованию словаря и его развитию.

2. Энциклопедия WEGA

Несмотря на наличие обширной специальной литературы по решению задач на графах, широкое применение в практике программирования полученных математических результатов затруднено в силу отсутствия систематического их описания, ориентированного на программистов. Поэтому значительный класс практических задач, по существу сводящихся к простому выбору подходящего способа решения и к построению конкретных формулировок абстрактных алгоритмов, для многих программистов все еще остается полем для интеллектуальной деятельности по «переоткрытию» методов.

Выполнен цикл работ по изучению и систематизации алгоритмов обработки, визуализации и применения графовых моделей в программировании. Впервые издана книга [3], которая содержит систематическое и полное изложение фундаментальных основ современных компьютерных технологий, связанных с применением теории графов. Ведётся работа по созданию на базе этой книги вики-системы WEGA, являющейся расширяемой интерактивной электронной энциклопедией теоретико-графовых алгоритмов решения задач информатики и программирования (Рис. 2).

The screenshot shows a web browser window displaying the WEGA wiki page for 'Коды Прюфера'. The page content includes:

статья обсуждение править история удалить переименовать защитить следовать

Коды Прюфера

Пусть T - дерево с множеством вершин $\{v_1, v_2, \dots, v_n\}$. Будем считать, что номер вершины v_i равен i . Сопоставим дереву T последовательность $\{a_1, a_2, \dots, a_{n-2}\}$ называемую кодом Прюфера, по следующему правилу:

```

функция код_пруфера( $T$ : дерево) =
1. Пусть  $n$  - число вершин в  $T$ ,
   а  $A$  - целочисленный вектор длины  $n-2$ ;
2.  $B := \{1..n\}$ ;
3. для  $i$  от 1 до  $n-2$  цикл
4.    $b_i := \min\{k \in B; k - \text{номер висячей вершины}\}$ ;
5.    $A[i] :=$  номер вершины, смежной вершине с номером  $b_i$ ;
6.    $B := B - \{b_i\}$ ;
7.   Удалить из  $T$  вершину с номером  $b_i$ 
8. все
возврат  $A$ 
все
  
```

Рассмотрим следующий пример. Для дерева T (рис.) код Прюфера имеет вид:

$P_2(T) = [4, 4, 5, 5, 7, 7]$.

The diagram shows a tree with 7 vertices. Vertex 4 is the root. Vertex 2 is a child of 4. Vertex 3 is a child of 4. Vertex 5 is a child of 4. Vertex 6 is a child of 5. Vertex 7 is a child of 5. Vertex 8 is a child of 7.

Рис. 2. Вики-энциклопедия WEGA.

В отличие от Д. Кнута, мы при создании данной энциклопедии, как и в книге [4], ориентируемся на абстрактную модель современных компьютеров (равнодоступная адресная машина – РАМ) и высокоуровневое описание алгоритмов в терминах специального языка высокого уровня – ВУ-язык. Этот язык является псевдоязыком (лексиконом) программирования и содержит в качестве базовых традиционные конструкции математики и языков программирования. Наряду с обычными для современных языков типами простых и составных данных он допускает такие более сложные структуры данных, как, например, деревья, графы и т.д. Для каждой базовой конструкции ВУ-языка фиксируется класс её допустимых реализаций на РАМ. Предполагается, что ВУ-язык позволяет наряду с базовыми использовать любые необходимые конструкции, если очевидны или заранее зафиксированы оценки их сложности, а также те реализации этих конструкций на РАМ, которые допускают такие оценки. Такой подход позволяет формулировать алгоритмы в естественной форме, допускающей прямой анализ их корректности и сложности, а также простой перенос алгоритмов на реальные языки программирования и компьютеры с сохранением полученных оценок сложности.

Еще одной важной особенностью создаваемой энциклопедии WEGA является её ориентация на поддержку анимационного исполнения представленных в ней графовых алгоритмов. На наш взгляд, анимация является удобным средством демонстрации работы любых алгоритмов, в том числе алгоритмов на графах. Она помогает человеку понять на конкретных примерах смысл и последовательность работы алгоритма, делая это гораздо более доходчиво, чем любые текстовые описания, пояснения и отдельные рисунки.

Разрабатываются методы и средства поддержки полноценной динамической визуализации графовых алгоритмов, представленных в энциклопедии. Визуализация информации – это процесс преобразования больших и сложных видов абстрактной информации в интуитивно понятную визуальную форму. Универсальным средством такого представления структурированной информации являются графы. Графы применяются для представления любой информации, которую можно промоделировать в виде объектов и связей между объектами. Поэтому визуализация графовых моделей является ключевой компонентой во многих приложениях в науке и технике, а методы визуализации графов представляют собой теоретическую основу методов визуализации абстрактной информации [4].

Начальная версия вики-системы WEGA прошла государственную регистрацию и в настоящее время содержит порядка 5000 статей.

Заключение

Наш доклад был посвящен методам и средствам поддержки применения теоретико-графовых методов для конструирования эффективных

и надежных программ, работа над которыми ведется в Институте систем информатики им. А.П. Ершова СО РАН при частичной финансовой поддержке Российского фонда фундаментальных исследований (грант РФФИ № 18-07-00024). В нем мы рассмотрели работы по созданию электронного словаря WikiGRAPP по теории графов и её применениям в информатике и программировании и электронной энциклопедии WEGA теоретико-графовых алгоритмов решения задач информатики и программирования. К настоящему времени подготовлены начальные версии указанных Web-систем, которые прошли государственную регистрацию.

Понятно, что обе системы нуждаются в развитии. Происходит постоянная работа по улучшению статей словаря и энциклопедии и по их пополнению новыми терминами и алгоритмами.

Разработаны методы и создан программный комплекс Wiki2TeX, поддерживающий работу со словарем и энциклопедией. Он автоматизирует процесс построения набора TeX-документов, образующих офлайн версию вики (базы данных, построенной с помощью MediaWiki), а также позволяет выполнять и обратную операцию (преобразовывать TeX документы в статьи MediaWiki и добавлять их к заданной вики в виде правок к её существующим статьям либо как её новые статьи).

Авторы благодарны всем коллегам, которые принимали участие в работах, рассмотренных в докладе.

Список литературы

1. Евстигнеев, В. А. Толковый словарь по теории графов в информатике и программировании / В. А. Евстигнеев, В. Н. Касьянов. – Новосибирск: Наука, 1999. – 288 с.
2. Евстигнеев, В. А. Словарь по графам в информатике / В. А. Евстигнеев, В. Н. Касьянов. – Новосибирск: Сибирское Научное Издательство, 2009. – 300 с.
3. Касьянов, В. Н. Графы в программировании: обработка, визуализация и применение / В. Н. Касьянов, В. А. Евстигнеев. – СПб.: БХВ-Петербург, 2003. – 1104 с.
4. Касьянов, В. Н. Визуализация информации на основе графовых моделей / В. Н. Касьянов, Е. В. Касьянова. – Новосибирск: Новосиб. гос. ун-т, 2014. – 149 с.

УДК 004.451

Назаров С.В.,

доктор технических наук, профессор, ЗАО «Московский научно-исследовательский телевизионный институт, Москва, Россия, s_nazarov@mail.ru, mniti.ru

ТЕХНОЛОГИИ ПРОЕКТИРОВАНИЯ ПРОГРАММНЫХ СИСТЕМ

Аннотация: в статье исследуются три технологии проектирования программных систем: структурное проектирование и программирование, объектно-ориентированное проектирование и программирование и компонентное проектирование. Каждая из этих технологий предоставляет определенные возможности в части декомпозиции архитектуры проектируемой программной системы. В данной работе рассматриваются достоинства и недостатки каждой технологии и целесообразность их использования при проектировании архитектуры программной системы и последующей ее реализации.

Ключевые слова: технология проектирования программных систем, программные средства, трансляция, структурное проектирование программных средств, объектно-ориентированное проектирование программных средств, компонентное проектирование программных средств.

S.V. Nazarov,

Doctor of Technical Sciences, Professor, CJSC Moscow Research Television Institute, Moscow, Russia, s_nazarov@mail.ru, mni-ti.ru

TECHNOLOGY OF DESIGN PROGRAM SYSTEMS

Abstract: The article explores three technologies for designing software systems: structural design and programming, object-oriented design and programming, and component design. Each of these technologies provides certain opportunities in terms of the decomposition of the architecture of the projected software system. This paper discusses the advantages and disadvantages of each technology and the feasibility of using them when designing the architecture of a software system and its subsequent implementation.

Keywords: software design technology, software, translation, structural software design, object-oriented software design, component software design.

Введение

До сих пор продолжают дискуссии о лучшей технологии и методах проектирования и программирования программных систем. Основных претендентов – три: структурное проектирование и программирование и связанный с ним модульно-интерфейсный подход, объектно-ориентированное проектирование и программирование и компонентное проектирование. Каждая из трех технологий, обсуждаемых в этой статье, предоставляет определенные возможности в части декомпозиции архитектуры проектируемой программной системы. В данной работе рассматриваются достоинства и недостатки каждой технологии и целесообразность их использования при проектировании архитектуры программной системы и последующей ее реализации. Выявляются положительные и отрицательные стороны каждой технологии и устанавливаются возможности сочетания полезных свойств этих технологий при проектировании архитектуры программной системы.

1. Общие принципы проектирования архитектуры ПС

В статье [1] рассмотрены архитектурные представления программных систем (ПС) инфокоммуникационных предприятий. Отмечено, что процесс создания архитектуры программной системы предполагает разработку системы с конкретными свойствами и функциональностью, причем каждое из этих свойств имеет заданный приоритет. Начиная с общей структуры, которая поддерживает всю требуемую функциональность, архитектор проводит декомпозицию функциональности и распределяет ее между элементами (модулями, объектами, компонентами и т.п.) структуры системы.

В целом сущность разработки ПС можно описать как ряд процессов перевода информации из одного представления в другое, начинающихся с постановки задачи и заканчивающихся набором подробных инструкций, управляющих компьютером при решении задачи. Создание ПС в этом случае представляется последовательной совокупностью процессов трансляции, т.е. перевода исходной задачи в различные промежуточные решения, пока, наконец, не будет получен подробный набор команд машинного языка. Такую последовательность можно представить схемой или макромоделью перевода следующего вида [2]:

$$\langle Z, L_z, P_z \rangle \rightarrow \bar{P}R_1 \langle TR_s, L_{tr} \rangle \rightarrow \overset{PR_2}{C_s, L_c} \rightarrow \overset{PR_3}{A_s, L_A} \rightarrow \dots \overset{PR_N}{S} \rightarrow \langle S, P_s, L_M \rangle.$$

Здесь приняты следующие обозначения:

Z – постановка задачи (возможно техническое задание) на разработку ПС;

P_z – параметры (характеристика) задачи;

L_z – язык описания задачи;

$\bar{P}R_i$ – процедура трансляции на i -ом этапе перевода, $i = 1, 2, \dots$;

N – количество различных этапов разработки системы;

TR_s – требования, предъявляемые к программной системе;

$$TR_s = TR_f \cup TR_{nf};$$

TR_f – функциональные требования, предъявляемые к системе;

TR_{nf} – нефункциональные требования, предъявляемые к системе;

L_{tr} – язык описания требований;

C_s – цели программной системы (программного продукта);

L_c – язык описания целей;

A_s – архитектура программной системы;

L_A – язык описания архитектуры программной системы;

S – код разработанной программной системы;

P_s – параметры кода программной системы;

L_M – машинный язык.

Заметим, что на каждом этапе процесса создания ПС предполагается использование определенного языка описания (постановки) задачи. Степень формализации этого языка растет по мере продвиже-

ния от одного этапа к следующему. Если на начальном этапе постановка задачи на разработку ПС заказчик использует чаще всего естественный разговорный язык, то с каждым следующим этапом язык становится все более формализованным, и наконец, созданный программный продукт на машинном языке исключает его неоднозначную трактовку. Практическая реализация схемы (1) связана с возможными ошибками на каждом ее этапе. Сложность – основная причина ошибок перевода и одна из главных причин ненадежности программных систем. В общем случае сложность объекта является функцией взаимодействия между его составными частями. Сложность архитектуры ПС определяется связями между ее подсистемами (компонентами, модулями и т. п.¹) системы. Сложность отдельного модуля – функция связи между его командами.

В борьбе со сложностью применимы концепции общей теории систем. Первая – независимость. В соответствии с этой концепцией для минимизации сложности необходимо усилить независимость компонентов системы. По существу, это означает такое разбиение системы, чтобы высокочастотная динамика (взаимодействие) системы была заключена в единых компонентах, а межкомпонентные взаимодействия представляли лишь низкочастотную динамику системы.

Вторая концепция – иерархическая структура и вытекающая из нее методология проектирования, основанная на понятии виртуальной системы. Иерархия позволяет стратифицировать систему по уровням понимания. Каждый уровень представляет собой совокупность структурных отношений между элементами нижних уровней. Концепция уровней позволяет понять систему, скрывая несущественные уровни детализации. Это, по сути, третья концепция декомпозиции – абстракция – отвлечение от несущественных подробностей с целью лучшего понимания стороны изучаемого явления. Декомпозиция наиболее эффективна тогда, когда осуществляется на базе абстракции.

И все же, существует ли наилучший метод проектирования программных систем? По словам Г.Буча [3], на этот вопрос нет однозначного ответа. По сути дела, это был бы ответ на вопрос: «Существует ли лучший способ декомпозиции сложной системы?» Лучшего способа пока никто не заявил, но достаточное количество различных методов декомпозиции и методов проектирования архитектуры программных систем известно и применяется на практике достаточно давно.

2. Структурно-функциональная декомпозиция и методы проектирования

2.1. Модульно-интерфейсный подход

В общем случае модульная программная система представляет собой

¹ Далее любую часть ПС будем называть модулем или компонентом.

древовидную структуру, в узлах которой размещаются программные модули, а направленные дуги показывают статическую подчиненность модулей. Если в тексте модуля имеется ссылка на другой модуль, то их на структурной схеме соединяет дуга, которая исходит из первого модуля и входит во второй модуль. При этом модульная структура программной системы, кроме структурной схемы, должна включать в себя совокупность спецификаций модулей, образующих эту систему.

Функции верхнего уровня обеспечиваются главным модулем. Он управляет выполнением нижестоящих функций, которым соответствуют подчиненные модули. При определении набора модулей, реализующих функции конкретного алгоритма, необходимо учитывать следующее:

1) модуль вызывается на выполнение вышестоящим по иерархии модулем и, закончив работу, возвращает ему управление;

2) принятие основных решений в алгоритме выносится на максимально высокий по иерархии уровень;

3) если в разных местах алгоритма используется одна и та же функция, то она оформляется в отдельный модуль, который будет вызываться по мере необходимости.

Структурное проектирование обычно подразумевает использование структур для разделения различных частей системы. Каждая часть при этом может проектироваться отдельно. Один из наиболее традиционных методов проектирования операционных систем в 60-е годы 20 века состоял в том, что проект каждого основного модуля системы выполнялся отдельно. Этот подход называется модульно-интерфейсным.

Анализируя существующие системы, можно заключить, например, что операционная система состоит из системы ввода-вывода, планировщика времени процессора, менеджера памяти, системы управления файлами и т. д. Каждый модуль выделяется и описывается, кроме того, определяется его интерфейс с другими модулями. Модули и их взаимные связи образуют абстракцию системы высокого уровня. После этого этапа для дальнейшего проектирования и реализации модули рассматриваются по отдельности. Такая же процедура может быть повторена в отношении какого-либо одного модуля системы, и последующую работу над модулем можно расчленить на индивидуальные задания по программированию. После того как все эти модули спроектированы и реализованы, они связываются воедино в соответствии с заранее определенными интерфейсами. Соблюдение правильного интерфейса представляет собой серьезную проблему в модульно-интерфейсном подходе.

Модули и их интерфейсы описываются весьма неформально и неточно. Поэтому, хотя отдельные модули и пишутся в соответствии с заданными спецификациями, часто оказывается, что в них заложено неверное представление об операционном окружении. И когда модули связываются вместе, они не взаимодействуют должным образом. После того, как

модули получены в коде компьютера, разрешение возникших конфликтов может оказаться чрезвычайно трудным делом.

В принципе, если интерфейс полностью специфицирован, то в этом отношении не должно возникать трудностей. К сожалению, первоначальный общий проект и планирование редко доводят до такого уровня детализации, когда используется язык программирования. В результате важные решения, которые должны были быть приняты опытными разработчиками в начале работы над проектом, отодвигаются внутрь модулей. В конечном счете, программисты принимают решения, имеющие глобальные последствия, базируясь на очень ограниченной и узкой точке зрения на систему. Решения могут приниматься и чрезвычайно квалифицированными людьми, но не в тот момент, когда нужно.

Правильное определение и выделение модулей представляет собой трудную задачу (заметим, что при объектно-ориентированном проектировании возникает не менее сложная задача определения объектов и выделения классов). Тесно связанные между собой части системы должны входить в один и тот же модуль. Было предложено считать мерой связности двух частей системы число предположений, которое одна часть должна делать относительно другой части. Слишком много или слишком мало информации об окружении, в котором работают соседние модули, может иметь плохие последствия при проектировании модуля.

Модульно-интерфейсный подход – один из методов структурного проектирования. Спецификации модулей и их интерфейсов дают структурную основу для проектирования каждого модуля и системы в целом. Спецификация программного модуля состоит из функциональной спецификации, описывающей семантику функций, выполняемых этим модулем по каждому из его входов, и синтаксической спецификации его входов, позволяющей построить на используемом языке программирования синтаксически правильное обращение к модулю.

Существуют различные методы разработки модульной структуры программной системы, в зависимости от которых определяется порядок разработки структуры системы, программирования и отладки модулей, указанных в этой структуре. Обычно в литературе выделяют два основных метода: метод нисходящей разработки (метод сверху-вниз или top-down) и метод восходящей разработки (метод снизу-вверх или bottom-up).

По мнению В. Турского, известного специалиста в области модульного программирования[4], формулировка аналитический и синтетический подходы к проектированию программ является более правильной, поскольку действительная суть проблемы не столько в выборе направления при проектировании, сколько в определении того, что именно преобладает: факторизация (аналитические шаги) или композиция (синтетические шаги). Ни один из этих подходов в их чистом виде не является жизнеспособной методологией проектирования: чисто аналитическое

проектирование сравнимо с построением “пирамиды с плавающей в воздухе вершиной”, а чисто синтетическое – с построением пирамиды” стоящей на голове”. Реальная стратегия проектирования почти всегда представляет собой разумное сочетание этих двух подходов.

2.2. Метод восходящей разработки (снизу-вверх)

Этот метод применялся в 60-х годах 20 века при разработке операционных систем, в частности операционной системы ТНЕ [5]. Он получил название метода восходящей разработки, или синтетического, или снизу-вверх. Применяя этот метод, разработчик начинает с основного аппаратного оборудования. “Чистая” аппаратура представляет собой для пользователя довольно неудобную машину, чтобы решать на ней задачу. Поэтому разработчик на первом шаге добавляет к компьютеру слой программного обеспечения ПО₁. Этот слой программ вместе с нижележащей аппаратурой обеспечивает выполнение некоторого множества команд, определяющих новую виртуальную машину VM₁ (рис. 1). На следующем шаге выделяется другое нужное свойство, добавляется новый слой программного обеспечения и получается очередная виртуальная машина. Процесс продолжается до тех пор, пока не будет получена виртуальная машина с требуемыми пользователю свойствами.

Аналогично проводится процесс конструирования программных систем, который назван архитектурным подходом. Модульная структура системы формируется в процессе программирования модулей, начиная с самого низшего уровня, затем следующего и т. д. При этом модули реализуются в таком порядке, чтобы для каждого программируемого модуля были уже запрограммированы все модули, к которым он может обращаться. Главной целью архитектурного подхода является повышение уровня используемого языка программирования, а не разработка конкретной программы. Это означает, что для заданной предметной области выделяют типичные функции, каждую из которых можно использовать при решении разных задач в этой области.

Основные преимущества метода проектирования снизу-вверх вытекают из способа структурирования, ограничивающего построение системы. Между различными уровнями можно организовать четкий интерфейс. При этом, поскольку любой процесс может требовать обслуживания только от процесса на более низком уровне, практически исключаются непредвиденные ситуации. Тестовые процедуры, используемые в этом методе проектирования, могут быть исчерпывающими, так как каждый уровень можно проверить по отдельности и достаточно полно. Когда некоторый уровень проверен до конца, можно добавлять части следующего уровня, реализующего более сложные функции и продолжать проверку.

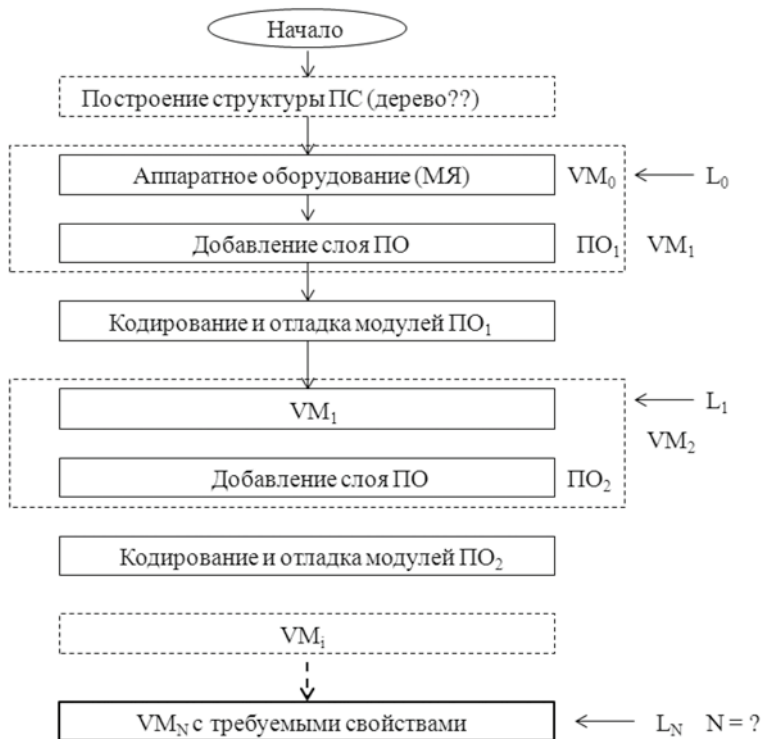


Рис. 1. Последовательность проектирования снизу-вверх

При восходящем программировании разработчик начинает с полного набора базовых средств, обеспечиваемых выбранным языковым уровнем (например, машинным языком МЯ). Все, что можно запрограммировать на этом языковом уровне, можно выразить в терминах таких средств. Для облегчения процесса программирования разработчик создает более высокие слои модулей таким образом, чтобы они облегчали использование доступных средств в форме, позволяющей абстрагироваться от обременительных деталей. Но на каждом следующем уровне (слое) разработчик должен быть уверен, что все средства, выразимые в терминах, установленных на этом слое понятий, доступны и что слой модулей представляет собой полное и логическое описание всей совокупности средств, обеспечиваемых базовым слоем.

Главная трудность в применении метода восходящей разработки заключается в выборе числа уровней и иерархическом упорядочении. Головной модуль проектируется и реализуется в последнюю очередь, это не дает возможности продемонстрировать его работу заказчику и про-

верить его соответствие спецификациям. Разработчики, использующие синтетический подход к проектированию программ, имеют значительные трудности на начальных этапах работы, поскольку они затрагивают самые базовые понятия и связаны с написанием текстов на машинном языке программирования. С другой стороны, нельзя их пропустить и сразу перейти к понятиям “среднего уровня”. Описание представляет собой как бы двойную загадку: первое – как догадаться, какие понятия являются полезными для решения конкретной задачи; второе – как получить эти промежуточные понятия из базовых.

На первый взгляд порядок разработки снизу-вверх кажется вполне естественным. Каждый модуль при программировании выражается через уже запрограммированные, непосредственно подчиненные модули, а при тестировании используются уже отлаженные модули. Однако современная технология не рекомендует такой порядок разработки программной системы. Во-первых, для программирования какого-либо модуля можно обойтись без текстов используемых им модулей, достаточно чтобы используемый модуль был специфицирован. Во-вторых, каждая программа в какой-то степени подчиняется некоторым внутренним для нее, но глобальной для ее модулей информацией, что определяет ее концептуальную целостность. Такая информация формируется в процессе разработки ПС. При восходящей разработке для модулей нижних уровней такая информация еще не ясна в полном объеме. поэтому эти модули часто приходится перепрограммировать. В-третьих, при восходящем тестировании для каждого модуля (кроме головного) приходится создавать ведущую программу, которая должна подготовить для тестируемого модуля необходимое состояние информационной среды и произвести требуемое обращение к нему.

Это приводит к большому объему отладочного программирования и не дает гарантии, что тестирование модулей производилось именно в тех условиях, в которых они будут выполняться в рабочей программной системе. Заметим, что последовательные слои программной системы, спроектированной снизу-вверх, не обладают свойством быть решением, но обеспечивают весь спектр возможностей, предоставляемых базовым слоем, хотя и выраженных через более общие понятия. Необходимо еще раз отметить, что слои возникают в синтетическом проекте тогда, когда разработчик получает набор модулей, реализующих понятия, в терминах которых можно выразить любую программу, выразимую в базовых понятиях. Модули, в которых детали, если они не выразимы в данном слое, упрятываются, не меняя при этом смысла программы (выраженного в понятиях, доступных в рассматриваемом слое).

2.3. Метод нисходящей разработки (сверху-вниз)

Альтернативный метод проектирования ПС получил название анали-

тического, или нисходящего проектирования, или сверху-вниз (рис. 2). В этом случае разработчик исходит из желаемых свойств виртуальной машины пользователя и последовательно разрабатывает уточнения в направлении аппаратуры. Реализация этой виртуальной машины полностью не специфицируется. Неопределенные части проектируются в дальнейшем как компоненты (модули) системы. Эта работа продолжается до тех пор, пока система не определена настолько, что ее основные функции реализуются аппаратурой или модулями, непосредственно выполняемыми на аппаратуре.

В результате проектирования получается множество вложенных друг в друга компонентов. На каждом шаге проектирования у разработчика имеется некоторая абстракция функционального описания некоторого компонента (алгоритм), и он должен уточнить эту абстракцию, разбив ее на более мелкие и более подробно проработанные части. На начальном шаге, используя внешний проект (спецификации системы), формируется перечень всех функций (алгоритмов) системы. Затем определяются их подфункции. Далее каждая подфункция может расчленяться до тех пор, пока ее составные части не будут окончательно уточнены. Метод нисходящего проектирования, иногда называемый функциональной декомпозицией, основан на двух стратегиях: пошаговом уточнении (детализации), разработанном Е. Дейкстрой, и анализе сообщений, базирующемся на работах Йордана, Константайна и Майерса[6]. Эти стратегии отличаются способами определения начальных спецификаций, методами, используемыми при разбиении задачи на части, и правилами записи.

Для проверки проекта в процессе его разработки можно с самого начала применять моделирование. Этот подход побуждает разработчика думать о том, какие функции должен выполнять некоторый компонент, а не как данный компонент будет их реализовать. Компоненты системы можно моделировать более детально, по мере того, как продвигается работа по проекту. Первоначальное представление какого-либо компонента может быть некоторым алгоритмом, который для данного входа вырабатывает с некоторой временной задержкой соответствующий выход, и таким образом моделирует работу данного компонента. По мере продвижения работ по проектированию данный алгоритм замещается последовательным набором обращений к следующему множеству проектируемых компонентов. На каждой стадии разработки проекта можно оценить, чтобы проверить, продолжает ли он соответствовать поставленным целям.

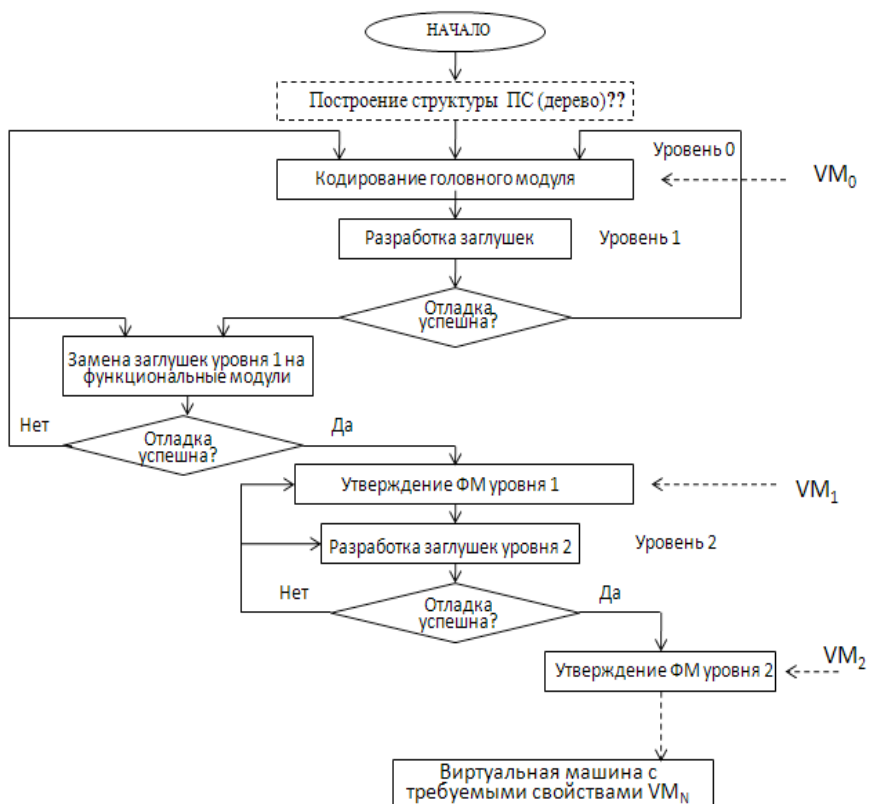


Рис. 2. Последовательность проектирования сверху-вниз

Каждая абстракция системы соответствует некоторой программе моделирования, которая построена из иерархии процедур. Эта программа определяет, как нужно действовать с теми переменными, которые влияют на состояние системы на этом уровне абстракции. Программа, назовем ее P , соответствующая некоторому уровню абстракции, управляется программой более высокого уровня абстракции, которую назовем Q . Программа Q принимает более глобальные решения, основываясь на значениях ее собственных переменных, которые на самом деле являются абстракциями переменных программы P . Изменения в переменных программы Q соответствуют изменениям в переменных программы P . Аналогично некоторые процедуры программы P делают запросы на выполнение работ к программам более низких уровней абстракции.

Когда система окончательно спроектирована, ее можно реализовать путем замены основных алгоритмов на самом низком уровне абстракции и средств, обеспечиваемых программой моделирования теми эле-

ментами, из которых должна быть построена система. Если эта замена сделана правильно, полученная система будет представлять в точности те средства, которые определены на самом высоком уровне.

При реализации программы нисходящим методом первым кодируется головной модуль – модуль верхнего уровня. При этом, поскольку этот вызывает модули соседнего низшего уровня, они представляются заглушками – фиктивными модулями (имитаторами). С помощью заглушек организуется непосредственный возврат в вызывающий модуль или осуществляться передача значений текстовых данных и печать результатов работы. Если заглушки имеют параметры, целесообразно организовать их печать. После того как отлажен головной модуль, Заглушки последовательно заменяются функциональными модулями, которые в свою очередь будут вызывающими для следующих заглушек (рис. 2).

В большинстве случаев аналитическое проектирование обеспечивает естественное расслоение программной системы: каждый следующий уровень понятий в расширяющемся дереве может рассматриваться как слой. Иногда выдвигаются возражения против методологии проектирования сверху-вниз, заключающиеся в том, что она не определяет, какое из двух направлений дерева программной структуры является более предпочтительным: от ветви к ветви или от уровня к уровню. Однако рекомендация выражать решение на каждом уровне исключительно в терминах этого уровня практически решает эту дилемму.

2.4. Замечания по структурному проектированию

Обе рассмотренные методологии естественным образом приводят к явно выраженной расслоенной структуре программной системы, обеспечивая равную доступность модулей одного слоя. В обеих методологиях слой представляет собой полное непротиворечивое множество модулей, даже если понятие полноты в каждом случае имеет различные оттенки. Сферы применения этих двух методов дополняют друг друга. Кратко это можно сказать так: при программировании решения отдельной задачи предпочтительнее аналитическое проектирование, а при создании программных систем для решения класса задач – синтетический метод. Поэтому синтетическое проектирование часто используется при написании операционных систем, а аналитическое – при написании специализированных программных систем.

Возможен другой путь использования методов структурного проектирования. Система может быть спроектирована методом сверху-вниз, а реализована методом снизу-вверх. В этом случае общий проект структуры программной системы выполняется методом сверху-вниз, а модули фактически объединяются в систему по методу снизу-вверх. При таком подходе налицо все преимущества реализации методом снизу-вверх и в тоже время не приносится в жертву возможность общего охвата системы, предоставляемая методом проектирования сверху-вниз.

Существует другая интерпретация взаимной дополняемости этих двух принципов проектирования, объясняющая, почему в большинстве важных задач используется оба подхода. Для трудных и больших задач обстановка, обеспечиваемая базовым уровнем, как правило, слишком удалена, чтобы к ней можно было прийти аналитическими методами проектирования. Поскольку при выборе абстрактных свойств, полезных для решения данной широкой задачи, можно учитывать предыдущий опыт, на этапах приведения системы к уровню этих предположительно полезных абстрактных понятий можно с успехом применять синтетическое проектирование, что значительно снижает объем разработок аналитического характера.

Между сторонниками методов проектирования снизу-вверх и сверху-вниз ведется много дискуссий. Многие полагают, что проектирование сверху-вниз больше соответствует идее, когда следует принимать решения на ранних стадиях проектирования и оставлять решения по детальной реализации на более поздние этапы. Однако проектирование сверху-вниз требует лучшего осмысливания на ранней стадии процесса проектирования. Проектирование снизу-вверх не требует вначале полного осмысливания системы, так как каждый слой программной системы добавляется к уже полностью определенному предыдущему слою. Вследствие этого, как указывается в [6], некоторые разработчики программ пропагандируют метод сверху-вниз, но на практике проектируют системы методом снизу-вверх; особенно это касается операционных систем.

Главный недостаток структурного проектирования заключается в следующем: процессы и данные существуют отдельно друг от друга, причем проектирование ведется от процессов к данным. Таким образом, помимо функциональной декомпозиции, существует также структура данных, находящаяся на втором плане.

3. Объектно-ориентированный поход (ООП)

В ООП основная категория объектной модели – класс – объединяет в себе на элементарном уровне как данные, так и операции, которые над ними выполняются. Именно с этой точки зрения изменения, связанные с переходом от структурного подхода к ООП, являются наиболее заметными. Разделение процессов и данных преодолено, однако остается проблема преодоления сложности системы, которая решается путем использования механизма пакетов. Существует две методики разработки объектно-ориентированных программных систем, которые можно условно называть “потребление” и “производство». Команды разработчиков могут работать совместно, используя либо одну из методик, либо обе. Но если руководитель проекта не уверен и не знает, хватит ли квалификации его сотрудников для того, чтобы использовать чужие объекты, или способны ли они разрабатывать собственные, то это может привести к серьезным проблемам.

Первая методика заключается в том, что команда разработчиков активно использует библиотеки и классы, разработанные другими. При такой методике разработчики понимают, что их опыт и квалификация пока недостаточны для разработки серьезных объектов. Вторая методика применяется тогда, когда команда хорошо осведомлена о процессе разработки шаблонов, рефакторинге и имеет удачный опыт разработки объектно-ориентированных проектов, включая разработку собственных классов. Обе техники приемлемы для использования, но важно знать, какая из них дает большие шансы добиться успеха.

Выявление классов, правильно описывающих предметную область проекта, является самой сложной задачей, которую предстоит решить разработчику. Выявление нужных классов намного сложнее процесса создания диаграмм, используемых в структурном подходе. Если не определены классы, точно описывающие задачу, то в конечном счете есть большая вероятность того, что созданные модели окажутся неработоспособными. Не каждый класс является классом предметной области. Массивы, наборы и классы графического интерфейса не являются классами предметной области. Классы предметной области – это сущности, описывающие задачу предметной области. Каждому классу предметной области соответствует объект предметной области.

Объект – это сущность, которая используется при выполнении некоторой функции или операции (преобразования, обработки, формирования и т.д.). Объекты могут иметь динамическую или статическую природу: динамические объекты используются в одном цикле воспроизводства, например, заказы на продукцию, счета на оплату, платежи; статические объекты используются во многих циклах воспроизводства, например, оборудование, персонал, запасы материалов. На концептуальном уровне построения модели предметной области уточняется состав классов объектов, определяются их атрибуты и взаимосвязи. Таким образом строится обобщенное представление структуры предметной области.

Рекомендации по объектно-ориентированному программированию предлагают выполнять поиск в предметной области существительных, а затем – привязку к ним глаголов. Существительные впоследствии переходят в классы, а глаголы – в методы. Это самый легкий способ, но он дает только 20 – 30 % всех классов, которые реально потребуется выявить. Если затем на этапе анализа обнаружили только классы и методы, выявленные сочетанием существительных и глаголов, то, вероятнее всего, в разработанных моделях будет наблюдаться нехватка классов и потребуется дополнительный углубленный анализ. Тем не менее, выявление существительных и глаголов предметной области является хорошим началом.

В дополнение к тем деталям по предметной области, которые могут сообщить эксперты заказчика, разработчику придется решить, как в системе представить эти данные в доступной форме для заказчика, и

в любом случае придется решать задачу того, как сохранять информацию, вводимую потенциальными пользователями программной системы. Для решения этих вопросов используются интерфейсные классы, классы-контроллеры и классы-сущности. Интерфейсный класс – это класс, предназначенный для связи элементов вне системы с элементами, входящими в состав системы. Поэтому такой класс часто называют граничным. Классы сущности представляют данные. Обычно сущности передают информацию, которая хранится в базе данных. Классы-контроллеры управляют другими классами.

Как правило, пользователи (заказчик) могут дать много полезной информации о классах-сущностях и помочь определить вид графического интерфейса пользователя, основываясь на том, как они выполняют задачи своей повседневной деятельности. Однако этого недостаточно, и разработчику нужно еще выявить классы-контроллеры и интерфейсные классы. Рекомендуется помнить тот факт, что эксперты по предметной области могут сообщить разработчику большое количество информации. Они могут рассказать многое о данных, которые им необходимо хранить, пояснить кое-какие детали о процессах, с помощью которых они получают эти данные, и немного о том, каким образом им хотелось бы вводить эти данные в компьютер.

Вторым важным моментом является то, что эксперты предметной области в рамках своей деятельности могут выполнять большое количество действий, которые будут просто непонятны постороннему наблюдателю. Это означает, что проектировщик, возможно, никогда и не пытался исследовать в целом то, что делает фирма-заказчик и как она это делает, и не пытался найти способ, как оптимизировать ее деятельность. В результате проектировщик может получить много информации от клиентов об их деятельности, которая, с его точки зрения, будет иметь слабое отношение к разрабатываемой системе. Но при этом эксперты по предметной области будут считать, что эта информация очень существенна.

После того, как разработчик выяснит у пользователей все о классах-сущностях, его задачей становится определение классов-контроллеров и интерфейсных классов. При разработке модели системы к сущностным классам добавляется стереотип “entity” и специальный символ класса-сущности, поддерживаемый во многих инструментах разработки, например, Rational Software Architect. Класс-контроллер является связующим кодом, который управляет другими классами и обеспечивает передачу данных между классами-сущностями и интерфейсными классами. Классы-контроллеры обозначаются при помощи стереотипа “control”, добавляемого к обычному классификатору, и специального символа. Интерфейсные или граничные классы моделируются, как показано на этом же рисунке и имеют стереотип “boundary”.

Заметим, что классы-сущности включают в себя логические сущно-

сти. Обычно их появление является результатом какого-либо смешанного запроса в базу данных. Выявление логических и простых сущностей является относительно простой задачей, так как теория реляционных баз данных легка для понимания. Классам-сущностям, определенным для отдельных таблиц и различных смешанных запросов, включающих обращения к нескольким таблицам, можно присваивать стереотип “table”. Вообще процесс выявления классов при работе над проектом системы весьма нетривиален. Существует много рекомендаций по организации такой работы. Одна из них – использование CRC (Classresponsibilityandcollaborator) – карты функциональности и сотрудников класса.

Поскольку данные по сравнению с процессами являются более стабильной и относительно редко изменяющейся частью системы, отсюда следует главное достоинство ООП. Г. Буч сформулировал его следующим образом [3]: объектно-ориентированные системы более открыты и легче поддаются внесению изменений, поскольку их конструкция базируется на устойчивых формах. Это дает возможность системе развиваться постепенно и не приводит к полной ее переработке даже в случае существенных изменений исходных требований.

Г. Буч отметил также ряд следующих преимуществ ООП:

- объектная декомпозиция дает возможность создавать программные системы меньшего размера путем использования общих механизмов, обеспечивающих необходимую экономию выразительных средств. Использование ООП существенно повышает уровень унификации разработки и пригодность для повторного использования не только ПО, но и проектов, что в конце концов ведет к сборочному созданию ПО;

- объектная декомпозиция уменьшает риск создания сложных систем ПО, так как она предполагает эволюционный путь развития системы на базе относительно небольших подсистем. Процесс интеграции системы растягивается на все время разработки, а не превращается в единовременное событие;

- объектная модель вполне естественна, поскольку в первую очередь ориентирована на человеческое восприятие мира, а не на компьютерную реализацию;

- объектная модель позволяет в полной мере использовать выразительные возможности объектных и объектно-ориентированных языков программирования.

К недостаткам ООП относятся некоторое снижение производительности функционирования ПО (которое, однако, по мере роста производительности компьютеров становится все менее заметным) и высокие начальные затраты. Объектная декомпозиция существенно отличается от функциональной, поэтому переход на новую технологию связан как с преодолением психологических трудностей, так и дополнительными финансовыми затратами. При переходе от структурного

подхода к объектному, как при всякой смене технологии, необходимо вкладывать деньги в приобретение новых инструментальных средств. Здесь следует учесть расходы на обучение методу, инструментальным средствам и языку программирования. Для некоторых организаций эти обстоятельства могут стать серьезными препятствиями.

Следует отметить, что многие специалисты (программисты и ученые) высказывали отрицательные суждения по вопросу перспективности объектно-ориентированной технологии программирования. Поводом к таким высказываниям послужила публикация известной в мире программирования статьи “Почему объектно-ориентированное программирование провалилось?”. Статья была написана доктором компьютерных наук Стэнфорда Ричардом Гэбриелом, старшим архитектором по разработке ПО сначала Sun, а потом и IBM, [7]. Он не скрывал своего скептического отношения к парадигме ООП. В 2002 году, по прошествии 2 лет после первоначальной публикации критической статьи, его пригласили выступить на ежегодной конференции OOPSLA (центральная конференция IT-специалистов по объектно-ориентированным языкам и методологиям разработки ПО) и изложить свои критические взгляды.

В качестве оппонента пригласили Г.Стила (GuySteele), разработчика языка Scheme, крупнейшего специалиста-теоретика по ООП, авторитет которого в американской академической среде непререкаем. В качестве “анти-объектника” дополнительно пригласили П.Грэма (PaulGraham), крупнейшего специалиста по Lisp, автора многочисленных книг и стандартизаций Lisp. В стан “объектников” пригласили Джеймса Ноубла (JamesNoble), автора одних из первых книг и работ по теории ООП. По свидетельству участников, конференция запомнилась им по тому уровню обсуждения, которое завязалось в этой публичной “интеллектуальной дуэли” фактически диаметрально разных школ программирования. Но факт остаётся фактом: сторона, представлявшая ООП, во время открытой дискуссии с противниками под смех зала даже запуталась в своих же концепциях.

П.Грэм утверждал, что половина всех концепций ООП являются скорее плохими, чем хорошими, в связи с чем он искренне сочувствует ООП-программистам, тогда как вторая половина от оставшихся концепций и вовсе не имеет никакого отношения к ООП, с которыми их почему-то постоянно ассоциируют. Например, он говорил [7]: “В восьмидесятых годах метод повторного использования каким-то неясным мне образом связали с объектно-ориентированным программированием, и сколь угодно многочисленные имеющиеся доказательства обратного, по-видимому, уже не избавят этот метод от клейма ООП. Хотя иногда объектно-ориентированный код годится для повторного использования, таким его делает вовсе не объектно-ориентированность, а программирование в стиле «снизу-вверх». Возьмём, например, библиотеки: их мож-

но подгружать и повторно использовать сколько угодно, потому что, по сути, они представляют собой отдельный язык. И при этом совсем неважно, написаны ли они в объектно-ориентированном стиле или нет”.

Читателям будет интересно познакомиться с полным текстом статьи [7]. Однако только время покажет, кто в итоге окажется правым, а кто искренне заблуждался. Тем не менее, критика основных концепций и соответствующих языков ООП с выяснением, какой язык, по-настоящему объектно-ориентированный не прекращается. Признанный эксперт в объектных методологиях программирования и гибких методологиях, таких как agile и scrum, Дэвид Уэст, автор книг *Object Thinking* и *Design Thinking* [8], утверждает, что идея объектов вообще никогда не была по-настоящему воплощена в каком-либо языке, даже в *Smalltalk*. По его мнению, понятие класса – не объектно-ориентированно. Классы не имеют никакого отношения к объектам. Они были лишь способом эффективно хранить код, позволяя вносить изменения в одном месте вместо многих сразу. Дэвид Уэст также считает, что наиболее действительно настоящий объектно-ориентированный язык – это *Self* (проект финансировался компанией *Sun Microsystems* и в 1994 году был закрыт).

В нем нет ничего, кроме объектов, имеется только один класс, или один тип сущностей, и эта сущность – объекты. Что касается *C++*, то это не объектно-ориентированный язык, и не задумывался таким. Его автор Страуструп писал, что он хотел сделать программистов на *C* дисциплинированнее. Так что *C++* был задуман для того, чтобы программисты писали более хороший и более дисциплинированный код.

В заключение отметим, объектно-ориентированный подход не дает немедленной отдачи. Эффект от его применения начинает сказываться после разработки двух-трех проектов и накопления повторно используемых компонентов, отражающих типовые проектные решения в данной области. Переход организации на объектно-ориентированную технологию – это смена мировоззрения, а не просто изучение новых CASE-средств и языков программирования.

4. Компонентный подход

Компоненты представляют собой автономный код, который может быть повторно использован за счет его независимого развертывания. Компоненты не обязательно должны быть большими по объему кода, но в основном они больше по объему отдельного класса или группы слабо связанных классов. Компоненты имеют множество предоставляемых и требуемых интерфейсов и используются в больших сложных приложениях с десятками и сотнями классов предметной области. В сложных случаях разработки больших систем, например, при разработке корпоративной информационной системы потребуются диаграммы компонентов [9].

Как отмечается автором книги [10], существует два основных способа выявления компонентов. Первым является метод движения сверху

вниз, а вторым – движение снизу-вверх. По сути это методы проектирования архитектуры программной системы, которые характерны и для структурного проектирования с модульно-интерфейсным подходом. Если смотреть глубже – это не что иное, как метод декомпозиция сложной системы в первом случае методом сверху вниз, а во втором случае – синтез сложной системы на основе метода проектирования снизу-вверх. При этом образуется многослойная архитектура программной системы (см. п. 2.3).

В ряде монографий и статей рекомендуется определять архитектуру системы на основе компонентов методом сверху вниз. В этом случае в первую очередь определяются компоненты (крупные модули системы), а потом интерфейсы этих компонентов. Как только компоненты и их интерфейсы определены, можно разделить задачи реализации системы среди участников процесса разработки, поручив проектирование каждого компонента отдельным группам. С того момента, как каждый согласится с разработкой интерфейсов, проектировщики будут свободны в реализации внутренних частей компонента и могут осуществлять ее любым способом, который они выберут.

По мнению автора работы [10], такой поход может быть эффективным в том случае, если команда проектировщиков применяет хорошо продуманные компоненты с понятными и известными им интерфейсами. Однако правильно выявить все новые компоненты путем проектирования сверху вниз достаточно сложно. Кроме того, это достаточно сложный стиль программирования, связанный с тем, что системы, основанные на компонентах, имеют от трех до пяти вспомогательных интерфейсов и промежуточных классов для каждого класса предметной области. Поэтому системы, основанные на компонентах, могут быть достаточно трудоемкими, дорогими и рискованными.

При разработке компонентной системы методом снизу-вверх в первую очередь определяются классы предметной области, т.е. те, которые служат непосредственно для решения задач предметной области, а не вопросов архитектуры приложения. В результате основные усилия концентрируются, прежде всего, на решении поставленной задачи вместо описания сложной архитектуры. Применяя метод разработки снизу-вверх и выявляя классы предметной области, проектировщики имеют больший потенциал для решения поставленной задачи. Кроме того, всегда можно объединить классы предметной области в компоненты, если сложность проекта растет, некоторые группы классов можно развернуть и повторно использовать, включив их в состав компонентов.

Оба метода проектирования компонентных систем имеют право на использование. Для небольших приложений и приложений среднего размера, вероятно, не потребуются компоненты и в этих случаях можно воспользоваться методом снизу-вверх. Для приложений корпоративного

масштаба потребуются более мощная методология, и в этом случае метод проектировании сверху вниз может оказаться более подходящим.

Способы декомпозиции могут быть различными. При модульно-интерфейсном подходе на верхних уровнях осуществляется структурная декомпозиция, на средних и нижних – функциональная. При объектно-ориентированном подходе декомпозиция связана с правильным подходом к определению классов предметной области. Каждый класс должен согласоваться с принципами хорошего объектно-ориентированного проектирования. К принципам хорошего объектно-ориентированного проектирования, в частности, относятся принцип единственной ответственности (ПЕО – SingleResponsibilityPrinciple (SRP)), принцип открытости-закрытости (ПОЗ – OpenClosedPrinciple (OCP)) и принцип подстановки Лискова (ППЛ – LiskovSubstitutionPrinciple (LSP)) [11].

Принцип единственной ответственности говорит о том, что у такого элемента, как класс, должна быть лишь одна, определяемая им ответственность. Если же класс отвечает и за представление данных и доступ к ним, это характерный пример нарушения принципа ПЕО. Принцип открытости-закрытости заключается в том, что класс должен быть закрыт для модификации, но открыт для расширения. При изменении класса всегда существует риск что-то нарушить. Но если вместо модификации класс расширяется подклассом, такое изменение менее рискованно.

Принцип подстановки Лискова (ППЛ) лучше пояснить следующим примером. Допустим, что существует иерархия наследования для классов Person (Лицо) и Student (Студент). При использовании класса Person должна быть возможность использовать класс Student, поскольку он является подклассом Person. На первый взгляд, это всегда происходит автоматически, хотя и не совсем очевидно в отношении рефлексии – метода, позволяющего проверять программными средствами тип экземпляра объекта, считывать и устанавливать его свойства и поля, а также вызывать его методы, ничего не зная заранее об этом типе. Так, в методе, использующем рефлексия для обращения к классу Person, может и не предполагаться подкласс Student.

Проблема рефлексии имеет отношение к синтаксису. Мартин Р. [12] приводит в большей степени семантический пример класса Square (Квадрат), который относится к классу Rectangle (Прямоугольник). Но когда используется метод задания ширины квадрата SetWidth() для класса Square, это не имеет смысла – по крайней мере, для этого нужно еще вызвать (внутренним образом) метод задания высоты квадрата SetHeight(). Такое поведение отличается оттого, что требуется для класса Rectangle.

Безусловно, все эти принципы оказываются спорными в определенных контекстах. Ими следует руководствоваться лишь как рекомендациями, но не “истиной в последней инстанции”. Например, применяя ПОЗ, нетрудно переусердствовать до такой степени, когда становится ясно, что для реал-

лизации метода лучше модифицировать класс, чем расширять его. И ввод метода в класс можно также рассматривать как расширение.

Заключение

Разработка архитектуры любой сложной системы связана с тем или иным подходом к ее декомпозиции. Очевидно, что в конкретном проекте сложной программной системы невозможно обойтись только одним способом декомпозиции. Можно начать декомпозицию каким-либо одним способом, а затем, используя полученные результаты, попытаться рассмотреть систему с другой точки зрения. Рассмотрение различных методов представления архитектуры сложных программных систем позволяет констатировать факт представления архитектуры программной системы в виде многослойной иерархической системы, которую удобно представлять системой вложенных виртуальных машин. Структурный подход по-прежнему сохраняет свою значимость и достаточно широко используется на практике. На примере языка UML хорошо видно, что его авторы заимствовали то рациональное, что можно было взять из структурного подхода: элементы функциональной декомпозиции в диаграммах вариантов использования, диаграммы состояний, диаграммы деятельности и др. Основой взаимосвязи между структурным и объектно-ориентированным подходами является общность ряда категорий и понятий обоих подходов (процесс и вариант использования, сущность и класс и др.). Эта взаимосвязь может проявляться в различных формах.

Объектно-ориентированный подход непросто использовать на начальных этапах проектирования системы в связи с неполнотой и сложностью выделения всех объектов, которые должны быть учтены в проекте и которые должны будут стать классами. Одним из возможных вариантов является использование структурного анализа и функциональной декомпозиции как основы для объектно-ориентированного проектирования. После выполнения структурного анализа можно различными способами приступить к определению классов и объектов. Так, если взять какую-либо отдельную диаграмму потоков данных, то кандидатами в классы могут быть элементы структур данных. Структурный подход к определению архитектуры программной системы на основе компонентной технологии напрямую связан со структурными методами проектирования сверху-вниз и снизу-вверх. Это основные способы выделения компонентов системы.

Критикуемый недостаток структурного проектирования заключается в том, что процессы и данные существуют отдельно друг от друга, причем проектирование ведется от процессов к данным. Таким образом, структура данных находится на втором плане. В этом отношении объектно-ориентированный подход имеет преимущество – возможность интеграции объектной и реляционной технологий, поскольку объектно-ориентированное проектирование имеет определенные точки со-

прикосновения с реляционным проектированием. Классы в объектной модели могут определенным образом соответствовать сущностям (которые кстати рассматриваются и в структурных подходах). Однако, такое соответствие имеет место только на ранней стадии разработки системы. В дальнейшем цели объектно-ориентированного проектирования и разработки реляционной БД расходятся. Преодоление данного разрыва возможно построением отображения между диаграммами классов и реляционной моделью. Концепция многоуровневой виртуальной машины в совокупности с пошаговой детализацией и объектно-ориентированной декомпозицией позволяют построить формальную методику разработки модульной архитектуры программной системы, которую предполагается рассмотреть в продолжение данной статьи.

Список использованной литературы

1. Назаров С.В., Вилкова Н.Н. Архитектурные представления программных систем инфокоммуникационных предприятий. Новосибирск, Изд. «Печатный дом-Новосибирск». Журнал «Инфосфера», № 71. С. 15 – 18.
2. Назаров С.В. Архитектура и проектирование программных систем: монография /С.В. Назаров. – 2-е изд., перераб. и доп. – М.: ИНФРА-М, 2016 – 374 с.
3. Буч Г. Объектно-ориентированный анализ и проектирование с примерами приложений на С++ (Второе изд.). Rational Санта-Клара, Калифорния:перевод с английского под редакцией И. Романовского и Ф. Андреева. Спб. Изд. Невский Диалект. 2001. 359 с.
4. Турский В. Методология программирования: пер. с англ. – М.: Мир, 1981. – 264 с.
5. Цикритзис Д.,Бернстайн Ф. Операционные системы. Пер с англ. М.: Мир, 1977. – 336 с.
6. Зиглер К. Методы проектирования программных систем: Пер. с англ. – М.: Мир, 1985. – 328 с.
7. Савчук И. Почему объектно-ориентированное программирование провалилось? [Электронный ресурс]. Режим доступа: http://bloggerator.ru/page/oor_why-objects-have-failed
8. Классы – это не объектно: интервью Егора Бугаенко с Дэвидом Уэстом. [Электронный ресурс]. Режим доступа: <https://jug.ru/2016/09/bugayenko-west/>
9. Кулямин В. В. Технологии программирования. Компонентный подход. [Электронный ресурс]. Режим доступа: <http://lib.mdpu.org.ua/e-book/vstup/L/Logolev.pdf>
10. Киммел П. UML. UML-Основы визуального анализа и проектирования UML- Универсальный язык программирования / Пол Киммел; пер. с англ. Кедрова Е.А. – М.: НТ Пресс, 2008. – 272 с.
11. Нильссон Дж. Применение DDD и шаблонов проектирования. Проблемно-ориентированное проектирование приложений с примерами на С# и .NET. Издательство: Вильямс. 2008. – 560 с.
12. Мартин Р., Ньюкирк Дж., Косс Р. Быстрая разработка программ: принципы, примеры, практика. Вильямс, 2004. – 752 с.

УДК 004.42

Соколова Е.А.

СКГМИ (ГТУ), 362021, РСО-Алания, г.Владикавказ, ул.Николаева,44

e-mail:katya_sea@mail.ru

<http://www.skgmi-gtu.ru/ru-ru/>

ЭКСПЕРИМЕНТАЛЬНОЕ ИССЛЕДОВАНИЕ ЭФФЕКТИВНОСТИ РАБОТЫ АЛГОРИТМОВ KIF

Анотация: в работе рассмотрены наиболее часто используемые форматы для 3D моделирования. Проанализированы их свойства, достоинства и недостатки. Предлагается новый формат для хранения 3D моделей – .DBE, который позволит уменьшить размер модели без потери качества.

Ключевые слова: 3Dмоделирование, форматы для хранения 3Dизображений, сжатие без потерь, оптимизация.

Sokolova E.A.

SKGMI (GTU), 362021, RNO-Alania, Vladikavkaz, Nikolayev St., 44

e-mail: katya_sea@mail.ru

EXPERIMENTAL STUDY OF THE EFFICIENCY OF THE KIF ALGORITHMS

Abstract: *The most frequently used formats for 3D modeling are considered. Analyzed their properties, advantages and disadvantages. A new format for storing 3D models – .DBE, which will reduce the size of the model without loss of quality.*

Keywords: *3D modeling, formats for storing 3D images, lossless compression, optimization.*

1. Введение

Одной из основных задач 3D форматов – это хранение различных моделей трехмерных объектов (в английской транскрипции – keepers of images files, сокращенно – kif), которые успешно используются в различных областях. Важную роль в сфере 3D моделирования для создания и обучения интеллектуальных систем играют типы расширений трехмерных математических моделей. При создании новых виртуальных тренажеров и обучающих систем, а также виртуальной реальности для проведения научных исследований и трехмерной визуализации большинство специалистов сталкиваются с проблемой выбора правильного формата для хранения 3D моделей и объектов.

Существует пять основных этапов для формирования информации: сбор, передача, накопление, обработка и представление информации. Именно на этапе представления информации используются технологии трехмерного моделирования. При использовании данных технологий, возникает вопрос о правильном хранении трехмерных объектов [1].

Если говорить о необходимости использования трехмерных моделей для дальнейшего развития интеллектуальных систем, то она есть

и способна намного упростить хранение данных. Например, при решении задач распознавания образов необходимо обрабатывать фотографии одного и того же объекта, полученные с разных ракурсов и в разных условиях. Хранить все возможные двухмерные фотографии подобных объектов физически невозможно. Следовательно, остается одно из двух – хранить в памяти компьютера трехмерные модели объектов и сравнивать их с полученными двухмерными фотографиями и другими изображениями или использовать нейронные сети.

В задачах первого типа важную роль играет выбор формата хранения трехмерных объектов. Следовательно, появляется необходимость исследования существующих 3D форматов, с целью выявления наиболее эффективного типа файлов для хранения 3D объектов [2].

Существует множество форматов для хранения 3D моделей и объектов, многие из них давно устарели, но продолжают являться актуальными.

2. Сравнение 3D Форматов

Таблица 1

Типы данных	Преимущества	Недостатки
.3DS	Открытый формат, возможность чтения другими программами	Ограниченное количество полигонов, устаревшая технология по сравнению с .max
.MAX	Хорошая степень сжатия	Закрытый формат, большой объем файла, открывается только с помощью программы Autodesk 3ds Max
.OBJ	Открытый формат, маленький объем файла, бинарный и ASCII, самый распространенный среди других форматов, легко воспринимается пользователями без изучения дополнительных языков программирования	Формат не хранит иерархию и связи объектов сцены, не поддерживает анимацию
.FBX	Поддерживает анимацию, настройку освещения, расположение камер, поддерживается и обновляется компанией Autodesk, обладает бесплатной и открытой SDK, поддерживает сценическую графику, формат ASCII имеет древовидную структуру с четкими обозначениями идентификаторов	Проприетарное программное обеспечение, закрытый исходный код
.WRL	Открытый формат, используется в качестве файлового формата для обмена 3D-моделями в САПР, используется в интернет-браузерах	Устаревшая технология по сравнению с .x3d
.STL	Кроссплатформенность, открытость исходного кода	Невысокая точность геометрии, большой объем файла для сложных моделей

Проведенное исследование выявило то, что каждый из 3D форматов имеет свои достоинства и недостатки, которые необходимо учитывать при решении различной сложности задач [3].

В решении вопросов хранения трёхмерных математических объектов, связанных с интеллектуальными системами и системами виртуальной реальности, которые широко применяются в различных сферах деятельности человека, наиболее эффективным форматом является OBJ. Применение формата OBJ позволит минимизировать затраты на распознавание и хранение 3D объектов[4].

Однако возможно данный формат улучшить, а именно, создать новый (.DBE) и взяв за основу формат .OBJ создать совершенно новую структуру файла, а также оптимизировать объем файла, что облегчит хранение 3D-модели.

В формате .OBJ каждый символ кодируется 1 байтом и при кодировании координаты из 8 цифр будет затрачено 8 байт памяти. В новом же формате, после конвертации создается пакет из пяти файлов (файл координат вершин, файл координат нормали, файл координат текстур, параметры одной поверхности объекта и solution), каждый из которых содержит числа типа float. Соответственно теперь, эта же координата в файле будет занимать 4 байта, что существенно уменьшает объем файла[5].

Таблица 2

Тип данных	Объем файла модели		Объем файла после сжатия модели		Степень сжатия модели в %		Хранение текстур в файле
	Низк., Кб	Высок., Кб	Низк., Кб	Высок., Кб	Низк., Кб	Высок., Кб	
.3DS	9,78	621	3,28	100	7,95	7,76	+
.MAX	176	176	13,5	14	7,67	7,95	+
.OBJ	15,04	1638	4,7	330	20,15	31,25	+
.FBX	21,2	671	10,6	402	59,91	17,95	+
.WRL	8,6	958	2,3	176	18,37	26,79	+
.STL	12,8	1251	4,4	445	35,57	34,38	-
.DBE	15,04	1638	2,7	290	17,7	15,04	+

Как видно из Таблицы 2, формат .DBE показывает лучший результат в сравнении с форматом .OBJ.

3. Постановка и результаты экспериментов.

Для проведения эксперимента, было выбрано пять 3D моделей.

На таблице 3 представлены 3D модели, а также, объем каждой из них в определенных форматах.

Таблица 3

Наименование объекта	Объем 3D модели в формате .OBJ (Кб)	Объем 3D модели в формате .MAX (Кб)	Объем 3D модели в формате .DBE (Кб)
1.OBJ	1069	2400	882
2.OBJ	569	1500	350
3.OBJ	160	990	52
4.OBJ	700	1530	298
5.OBJ	2950	4000	2110

Как видно из проведенного эксперимента, формат .DBE существенно сжимает размер файла .OBJ. И несмотря на то, что некоторые форматы имеют размер меньше, в случае с форматом .DBE отсутствует потеря качества 3D модели.

4. Заключение

С течением времени все терпит изменения, и такое быстроразвивающееся направление как 3D графика – не исключение.

В рядах графических форматов появилось пополнение. Насколько активно оно будет востребовано пользователями – покажет время, ясно одно: новейшие разработки могут значительно улучшить соотношение «сжатие/качество» для 3D изображений.

Несмотря на то, что разработанный формат не дает наилучшего результата компрессии 3D изображений, это окупается тем, что сжатие происходит без малейшей потери качества, что является немаловажным фактором для дальнейшей работы с этим изображением.

С учетом того, что размер файла является минимальным, в перспективе его можно приспособить для работы в геоинформационных системах. Так как данное направление развивается сейчас очень стремительно, а картография требует больших ресурсов от вычислительной машины, формат .DBE помог бы существенно сократить затраты на хранение 3D карт.

Литература

1. Соколова Е.А. КОМПРЕССИЯ ИЗОБРАЖЕНИЙ ВАРИАБЕЛЬНЫМИ ФРАГМЕНТАМИ. Вестник компьютерных и информационных технологий. 2008. № 10. С. 31-34.
2. Кумаритов А.М., Соколова Е.А. Разработка системы анализа и обработки информации по стратегическому управлению предприятий топливно-энергетического комплекса. Наука и бизнес: пути развития. 2014. № 5 (35). С. 113-116.
3. Соколова Е.А. Математическая модель компрессии статичных изображений переменными фрагментами с учетом погрешностей, депонированная рукопись № 748-B2007 19.07.2007

4. Соколова Е.А. Использование теоретико-множественного подхода для поиска необходимого контента по атрибутам и ключевым словам. Фундаментальные исследования. 2013. № 8-6. С. 1360-1363. Маров А.А. – 3ds max. Реальная анимация и виртуальная реальность. Электронное издание, 2010.
5. Киан Б.Н.– Цифровые эффекты в MAYA. Создание и анимация, Электронное издание, 2015.

УДК 681.343.001

Томаев М.Х.

Доцент кафедры Автоматизированной обработки информации
Северо-Кавказского государственного горно-металлургического института
(государственного технологического университета), Владикавказ.
murat@risk-modelling.com

ДИСКРЕТНЫЕ ПОДХОДЫ К ОПТИМИЗАЦИИ ПОЛЬЗОВАТЕЛЬСКИХ ПРОГРАММНЫХ КОДОВ

Аннотация: в статье предлагаются методы и средства для улучшения производительности пользовательских программных кодов. Предлагаются дискретные подходы – в частности, аппарат теории графов, – для анализа, выделения и ранжирования отдельных участков исходного кода по «важности». Приводятся описания программных средств, автоматизирующих процесс оптимизации пользовательских кодов на основе двух методов: 1) оптимизации работы с медленной памятью за счет минимизации числа DLL-подсистем на критически важных участках программной системы; 2) повышение производительности функций программы за счет оптимального размещения переменных в двухуровневой оперативной памяти ЭВМ.

Ключевые слова: программа, оптимизация, модель, качество, метод, экстремальные, автоматизация, память.

М.Н. Томаев

Associate Professor at the Department of Automated Information Processing
North Caucasus State Mining and Metallurgical Institute (State Technological University),
Vladikavkaz. murat@risk-modelling.com

DISCRETE APPROACHES TO OPTIMIZING CUSTOM PROGRAM CODES

Abstract: *The article suggests methods and tools for improving the performance of custom program codes. Discrete approaches are proposed – in particular, the apparatus of graph theory – for analyzing, isolating, and ranking individual sections of the original code of “importance”. Describes software tools that automate the process of optimizing user codes based on two methods: 1) optimizing work with slow memory by minimizing the number of DLL-subsystems of critical parts of the software system, 2) improving the performance of program functions by optimally placing variables in two-level RAM COMPUTER.*

Keywords: *program, optimization, model, quality, method, extreme, automation, memory.*

Одним из направлений научно-исследовательских работ на кафедре «Информатики и вычислительной техники» Северо-Кавказского госу-

дарственного горно-металлургического института (технологического университета) является поиск новых подходов и методов решения задачи оптимизации пользовательских программных кодов. Полученные результаты характеризуются новым подходом к анализу качества программных продуктов, а также рядом новых программных алгоритмов оптимизации, позволяющих улучшить производительность пользовательских программных кодов.

Актуальность направления исследований тесно связана с широким распространением средств быстрой разработки, благодаря которым производительность труда коллективов разработчиков резко возросла. Однако рост производительности не сопровождается аналогичным ростом качества. Клиентам приходится мириться с необходимостью своевременного обновления аппаратных средств вычислительных систем, так как минимальные требования к производительности ЭВМ постоянно возрастают. Разработанные оптимизационные подходы учитывают специфику функционирования продукта, т.е. является ли это программа для компактных систем либо масштабная система, состоящая из нескольких модулей. Кроме того, отличительной чертой является глобально-оптимальный подход к оценке и модификации кодов, когда распределение наиболее ценных вычислительных ресурсов (процессорного времени либо памяти) между отдельными участками кода осуществляется на основе объективных критериев. До настоящего времени препятствиями широкому распространению такого рода подходов является высокая вычислительная сложность, невозможность оптимизации в ручном режиме, без использования вспомогательных инструментов автоматизирующих процесс оптимизации.

В рамках исследований были предложены объективные критерии качества программного кода для различных условий функционирования, были предложены мат модели оптимизации программного кода для каждого из критериев, на основе которых были программно реализованы алгоритмы и инструментальные средства оптимизации пользовательских кодов.

Предлагаемый оптимизационный подход основан на дискретном анализе структуры алгоритма программы. Для этого, в частности, используется аппарат теории графов. Пользовательский программный код представляет собой последовательность операторов, которую можно представить виде дерева состояний массива данных программы. В зависимости от входных данных программы могут либо завершаться за конечное время, либо неопределенно долго заикливаться. В связи с этим, очевидно, что улучшение качества участков, образующих контура является задачей более приоритетной по сравнению с задачей оптимизации конечных участков.

На следующем рисунке (Рис.1) наглядно демонстрируется подход

к ранжированию операторных участков по важности для программы, алгоритм которой изображен в левой части рисунка. Вначале следуют участки заикливания затем конечные участки. Внутри каждой группы участки ранжируются по времени выполнения участка (время однократного заикливания если это контур).

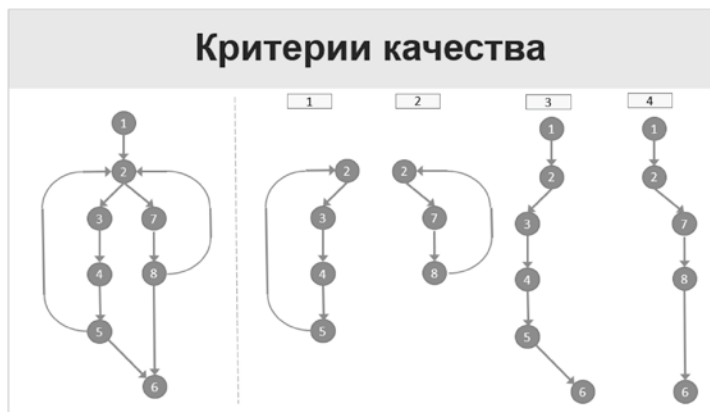


Рис.1. Графовое представление алгоритма (слева) и ранжирование участков (справа).

Полученные на текущий момент результаты можно разбить по специфике условий функционирования целевого программного обеспечения на 2 группы: первая – это модели оптимизации работы с медленной памятью для систем, работающих в условиях, когда размер кода системы значительно превышает размер быстрой памяти ЭВМ. Вторая группа – это модели так называемого экстремального программирования. Для улучшения производительности они используют имеющиеся избыточные объемы быстрой памяти.

Первая группа моделей включает в себя задачи оптимального кэширования пользовательских файлов, а также задачу оптимального декомпозиции пользовательской программной системы, функционирующей в условиях недостатка оперативной памяти. Следует отметить, что аналогов этим двум технологиям не существует.

Модели экстремального программирования представлены целой группой задач в том числе:

а) Перемещение переменных между различными областями памяти с целью сокращения времени на обслуживание временных локальных переменных;

б) Макрозамены функций;

с) Оптимизация передачи аргументов в функцию;

д) Замена одних операторов на более производительные аналоги и другие.

На основе теоретического задела были созданы ряд программных продуктов, демонстрирующих описанные выше подходы. Одним из них

является САПР оптимальных программных продуктов. Это демонстра- тор технологий, включающий основные алгоритмы оптимизации рабо- ты с медленной памятью: декомпозицию кода и оптимальное кэширова- ние файлов.

Еще один продукт, выполняющий оптимизацию медленной памяти, реализован в виде дополнительного инструмента к популярной среде разработки MSVisualStudio 2015. Отличительной особенностью про- дукта является многокритериальный подход – рассматриваются одно- временно конечные и цикляющиеся участки с учетом их приоритетов. Кроме того, алгоритмы решения не используют переборных подходов, в их основе быстрые алгоритмы на основе методов теории графов. На рис.2 демонстрируется модифицированный интерфейс VisualStudio состав стандартного меню включены пункты для вызова оптимизаци- онного модуля, а также возможность вызова справки.

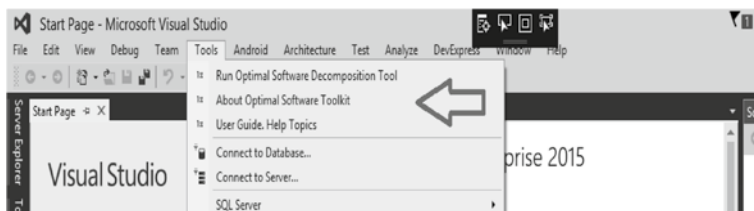


Рис.2. Модифицированный интерфейс VisualStudio.

Непосредственно процесс оптимизации проиллюстрирован на рис. 3. Слева исходный код пользователя, а справа выделенный программой критерии качества и оптимальная стратегия декомпози- ции.

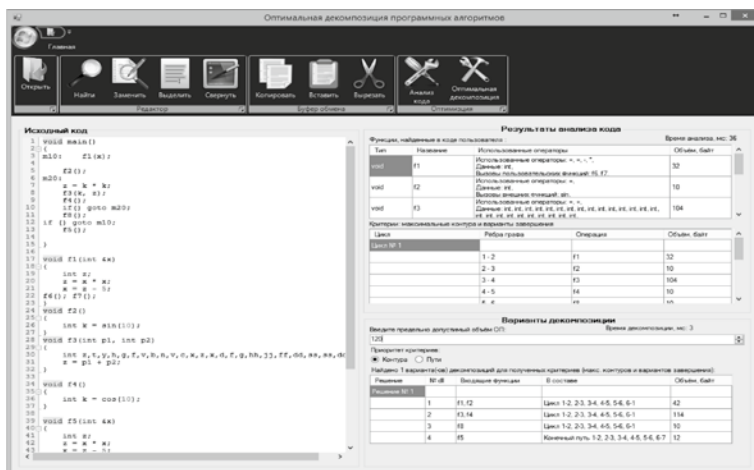


Рис.3. Результаты оптимизации.

Алгоритмы экстремального программирования характеризуются экстенсивным подходом к расходованию ресурсов при достижении даже небольшого прироста производительности. Однако для циклящихся систем, в которых большую часть времени система может проводить лишь в нескольких участках кода, использование таких подходов по отношению к подобным участкам может серьезно улучшить производительность системы в целом.

Подобный подход лежит в основе метода оптимального размещения переменных в двухуровневой памяти ЭВМ, который основан на том, что время работы функции сокращается по мере возможности на величину процессорного времени необходимого на размещение временных, так называемых локальных переменных. Однако замена временной переменной на постоянную приводит к увеличению объема оперативной памяти требуемой программе, что не всегда может являться допустимым. Вэтом заключается оптимизационный характер задачи. На рис. 4. демонстрируется программа, реализующая данный подход. Слева исходный код, а справа результат – оптимизированный в автоматическом режиме эффективный код.



Рис. 4. Оптимизация размещения переменных в двухуровневой памяти.

Заключение

Эффект от внедрения технологий оптимизации ожидается по нескольким направлениям: 1) Создание автоматизированных средств позволит снизить требования к уровню квалификации разработчиков систем, для которых производительность является критическим фактором качества. 2) Автоматизация трудоемких процессов оптимизации кода сократит время, а, следовательно, и ресурсы, выделяемые на решение прикладных задач.

Список использованной литературы

1. Гроппен В.О. Принципы принятия решений с помощью эталонов. Автоматика и телемеханика. 2006. № 4. С. 167-184.
2. Гроппен В.О., Томаев М.Х. Модели, алгоритмы и средства программной поддержки проектирования оптимальных программных продуктов //Автоматика и телемеханика. 2000.
3. Томаев М.Х. Модели экстремального программирования. Информационные технологии и системы. Наука и практика 2009. С. 150-152.
4. Томаев М.Х. Технологии глобальной оптимизации пользовательских программных кодов. Автоматизация и управление в технических системах. 2015. № 3 (15). С. 16-30.

СЕКЦИЯ 3. ПАРАЛЛЕЛЬНЫЕ ВЫЧИСЛЕНИЯ

УДК 004.272

Мирошников А. С.

к.т.н., каф. ИВТ СКГМИ(ГТУ)

mirandrey@mail.ru

МЕТОД ОПТИМИЗАЦИИ ПАРАЛЛЕЛЬНЫХ АЛГОРИТМОВ ДЛЯ SMP-СИСТЕМ

Аннотация: в статье описана параллельная реализация алгоритма решения уравнения Навье-Стокса в переменных скорость-давление с использованием высокопроизводительных вычислительных систем. Рассмотрены математические модели минимизации времени работы параллельного алгоритма.

Ключевые слова: уравнения Навье-Стокса, параллельный алгоритм, высокопроизводительные вычислительные системы, SMP-система, оптимизация программного обеспечения.

Miroshnikov A.S.

Ph.D. ICT SKGMI (GTU)

mirandrey@mail.ru

METHOD OF OPTIMIZATION OF PARALLEL ALGORITHMS FOR SMP SYSTEMS

Abstract: *The article describes a parallel implementation of the algorithm for solving the Navier-Stokes equation in variable speed-pressure using high-performance computing systems. Mathematical models for minimizing the operation time of a parallel algorithm are considered.*

Keywords: *Navier-Stokes equations, parallel algorithm, high-performance computing systems, SMP-system, software optimization.*

Введение

Сегодня, когда масштаб вычислений требует существенных вычислительных мощностей, для исследовательских, инженерных и других задач, необходим рациональный подход к использованию аппаратного обеспечения. Таким подходом стало использование в вычислениях параллельных алгоритмов и высокопроизводительных вычислительных систем.

Широкий круг задач решается с помощью сеточных методов, которые в свою очередь хорошо поддаются распараллеливанию. Одной из таких задач является численное решение системы уравнений Навье-Стокса – системы дифференциальных уравнений в частных производных, описывающих движение вязкой ньютоновской жидкости[1].

Уравнение Навье-Стокса в переменных скорость давление и метод его решения

Одним из основных компонентом любого моделирования жидкости является численное решение уравнений Навье-Стокса, представленное на равномерной сетке пространственных узлов или ячеек сетки.

Предположим, что скорость и давление известны в некоторый момент времени $t=0$, то их изменение с течением времени описывается уравнениями Навье-Стокса[2]:

$$\nabla \cdot \mathbf{u} = 0 \quad (1)$$

$$\frac{\partial \mathbf{u}}{\partial t} = -(\mathbf{u} \cdot \nabla) \mathbf{u} - \frac{1}{\rho} \nabla p + \nu \nabla^2 \mathbf{u} + \mathbf{f}, \quad (2)$$

где ν – кинематическая вязкость жидкости, ρ – плотность жидкости, \mathbf{f} – внешняя сила, $\nabla = \left(\frac{\partial}{\partial x}, \frac{\partial}{\partial y} \right)$.

Уравнение (1) здесь является уравнением неразрывности, а (2) – уравнением движения.

Уравнения Навье-Стокса можно решить аналитически только для нескольких простых физических конфигураций. Однако, чтобы решить их можно использовать численные методы. Для этого, как и любой другой алгоритм, необходимо разбить решение уравнений Навье-Стокса на несколько простых шагов. Для решения уравнений Навье-Стокса используется метод, основанный на технике «стабильных жидкостей» описанный в [3].

Решение уравнений Навье-Стокса включает три вычисления для обновления скорости на каждом временном шаге: вычисление промежуточных значений скоростей без учета давления, решение уравнения Пуассона для давления и обновление скоростей с учетом давления[4]. Выполнение этих трех шагов позволяет нам в результате получить новое поле скорости.

Метод решения.

Уравнения решаются с начальным состоянием $u_0=0$ и $p=0$ с шагом по времени Δt . Предположим, что поле было решено в момент времени t и что мы хотим вычислить поле в следующий момент времени $t+\Delta t$.

Решение происходит итеративно, то есть на каждом итерационном или временном шаге для небольших групп ячеек в сетке применяется набор дискретных дифференциальных операторов. Эти операторы обновляют скорость жидкости, основываясь на значениях в узлах сетки на предыдущем шаге по времени[5]. Таким образом, одна итерация состоит из следующих шагов:

Шаг1: Скорость жидкости рассчитывается без учета давления и уравнения неразрывности.

$$\mathbf{u}' = \mathbf{u}^n + \delta t \left[-\frac{\partial u u}{\partial x} - \frac{\partial v u}{\partial y} + \frac{1}{Re} \left(\frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2} \right) \right]$$

$$v' = v^n + \delta t \left[-\frac{\partial v u}{\partial x} - \frac{\partial v v}{\partial y} + \frac{1}{Re} \left(\frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2} \right) \right]$$

Шаг2: Чтобы получить корректное поле скоростей жидкости, на этом шаге необходимо найти давление, решив уравнение Пуассона.

Коррекция давления осуществлялась по формуле:

$$p_{ij}^{np} = p_{ij}^{np-1} - \omega \left(\frac{u_{i+1j} - u_{ij}}{h_x} + \frac{v_{ij+1} - v_{ij}}{h_y} \right)$$

Шаг3: Проверка сходимости состоит в вычислении погрешности разностного аналога уравнения неразрывности. Критерий останова итераций по давлению задан в виде

$$\max_{ij} \left| \frac{u_{i+1j} - u_{ij}}{h_x} + \frac{v_{ij+1} - v_{ij}}{h_y} \right| < \varepsilon$$

Если условие не выполняется повторяем шаг 2.

Шаг4: Обновление полей скоростей с учетом давления и уравнения неразрывности.

$$u^{(n+1)} = u' - \delta t \frac{\partial p^{(n+1)}}{\partial x}$$

$$v^{(n+1)} = v' - \delta t \frac{\partial p^{(n+1)}}{\partial y}$$

Построение параллельной версии алгоритма осуществлялось на основе принципа геометрической декомпозиции сеточной области, когда вся область исследования делится на равные по площади (или по количеству сеточных узлов) части, вычисления в которых следует проводить одновременно и независимо. Вся расчетная область делится на двумерные блоки одинаковых размеров (Рис. 1), в каждом из которых значения сеточных функций вычисляются одновременно и независимо.

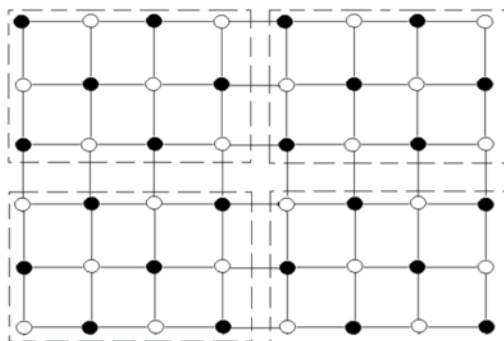


Рис. 1. Геометрическая декомпозиция данных

Однако для обеспечения таких идеально параллельных вычислений необходимо обеспечить каждую подобласть дополнительными значениями сеточных функций, которые принадлежат узлам из соседних по декомпозиции подобластей, но однако же необходимы для вычислений в соответствии с выбранным шаблоном «крест». Для получения эффективного параллельного алгоритма необходимо, чтобы затраты времени на передачу данных между соседними подобластями были существенно меньше времени вычислений на каждой итерации.

Математические модели

При реализации описанного выше алгоритма для высокопроизводительных SMP-систем процедуру расчета отдельной области выполняет отдельный поток. Для оптимизации использования центральных процессоров необходимо выбрать количество потоков, используемых при параллельном алгоритме решения уравнения, при котором время счета будет минимально.

С увеличением количества потоков скорость решения уменьшается, однако возрастает и время на организацию потоков. Учитывая эти факторы, была выбрана следующая математическая модель [10]:

$$\begin{cases} T = c + f \cdot x + \frac{d}{x} \rightarrow \min \\ 1 \leq x \leq P, \text{целое} \end{cases} \quad (3)$$

где:

c – среднее время на организацию вычислений;

f – среднее время на организацию одного потока;

x – количество потоков, задействованных в параллельном алгоритме расчета поля скоростей;

d – время расчета поля скоростей одним потоком;

P – максимальное количество потоков, поддерживаемых SMP-системой;

Целевая функция (3) отражает время работы алгоритма и зависит от числа потоков x задействованных для решения задачи.

Решение поставленной задачи (3) доставляет оптимальное значение количества потоков, необходимое для того чтобы время работы программного комплекса было минимальным.

Параметры c, f, P из (3) индивидуальны для каждой SMP-системы и рассчитываются экспериментальным путем. Время расчета поля скоростей d зависит от размера области и программной реализации алгоритма. Для его определения воспользуемся подсчетом количества выполняемых операций в алгоритме и времени их выполнения. В результате получим зависимость вида:

$$d(n,m) = k_1 \cdot n + k_2(n \cdot m) + k_3.$$

где:

m – количество точек сетки по оси OX;

n – количество точек сетки по оси ОУ;
 k_1, k_2, k_3 – коэффициенты, индивидуальны для каждой SMP-системы и рассчитываются экспериментальным путем.

Для вычислительной системы на базе процессора IntelCorei3-2350M, 2.3 ГГц параметры системы (3) следующие:

$c = 0.048$ сек., $f = 1.793 \cdot 10^{-4}$ сек., $\kappa_1 = 0,6260$, $k_2 = 0,0019$, $k_3 = 0,00001$.

Рассмотренную целевую функцию системы(3) можно считать частным случаем выражения вида:

$$t = a_0 + \sum_{i=1}^k a_i x^i + \sum_{i=1}^k b_i / x^i \quad \text{при } k=1.$$

Рассмотрим случай при $k=2$.

$$t = a_0 + a_1 * x + a_2 * x^2 + \frac{b_1}{x} + \frac{b_2}{x^2} \quad (4).$$

Используя экспериментальные данные, методом МНК получены следующие значения искомым коэффициентов (4):

a_0	0.029
a_1	-0.023
a_2	$2.412 \cdot 10^{-3}$
b_1	0.318
b_2	0.219

Сумма квадратов отклонений экспериментальных данных от найденных аналитически составила $\Delta = 1,17 \cdot 10^{-4}$.

Зависимость, построенная по данным, для выражения (4), полученным экспериментально и аналитически, показана на рисунке 2.

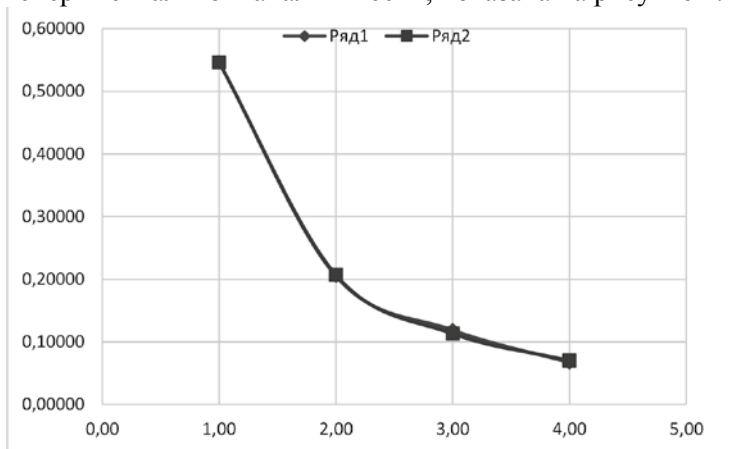


Рис. 2. Зависимость времени расчета одной итерации от количества используемых потоков. Ряд 1 – аналитически (4), ряд 2 –экспериментально.

Рассмотрим случай при $k=3$.

$$t = a_0 + a_1 * x + a_2 * x^2 + a_3 * x^3 + \frac{b_1}{x} + \frac{b_2}{x^2} + \frac{b_3}{x^3} \quad (5)$$

Используя экспериментальные данные, методом МНК получены следующие значения искомых коэффициентов (5):

a_0	-0.642
a_1	-0.028
a_2	0.018
a_3	$-1.188 * 10^{-3}$
b_1	3.528
b_2	-5.062
b_3	2.733

Сумма квадратов отклонений экспериментальных данных от найденных аналитически $\Delta = 9,233 * 10^{-6}$.

Зависимость, построенная по данным полученным экспериментально и аналитически, показана на рисунке 2.3

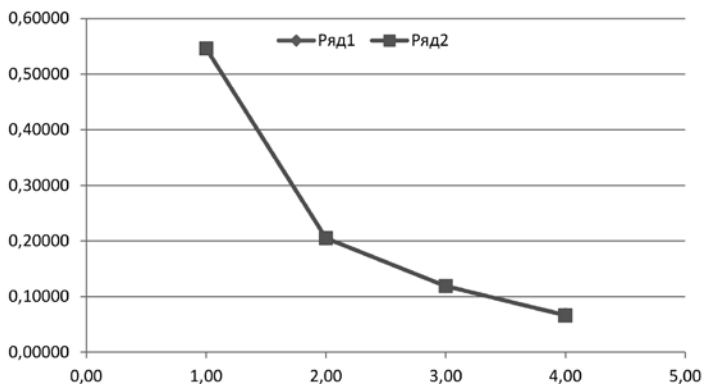


Рис. 3. Зависимость времени расчета одной итерации от количества используемых потоков.
Ряд 1 – аналитически (5), ряд 2 – экспериментально.

Как видно из графиков, с увеличением k точность данных полученных аналитически растет, однако также и увеличивается время I , затрачиваемое на проведение расчетов.

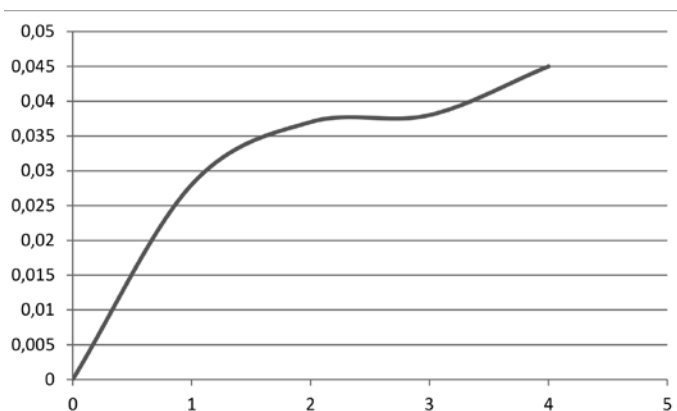


Рис. 4. Зависимость времени расчета целевой функции I от числа k

Для выбора оптимального числа k проведен анализ зависимости отношения времени I к сумме квадратов отклонений Δ от k .

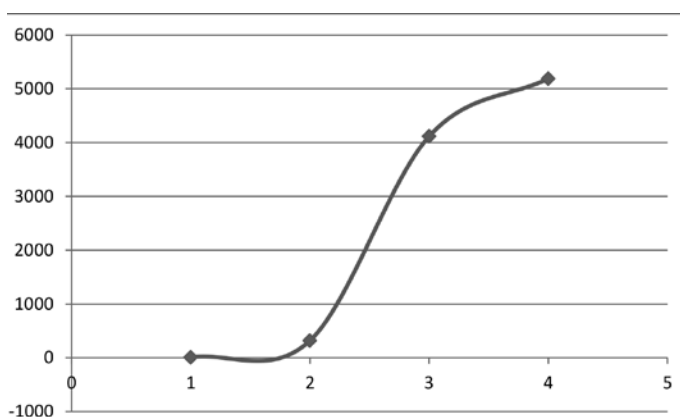


Рис. 5. Зависимость отношения $\frac{I}{\Delta}$ от числа k

Как видно из графиков существенный прирост точности достигается при $k=3$, при дальнейшем увеличении k в значительной степени увеличивается время, поэтому для программной реализации выбрано выражение (5).

Заключение

Решение уравнений Навье-Стокса – сложная математическая задача, существование аналитического решения для которой доказано лишь для частных, притом наиболее простых случаев, а численные решения настолько трудоемки, что не имеют смысла без использования достаточно производительной вычислительной техники. Решение же реальных или приближенных к реальным задачам требует больших

вычислительных ресурсов. Оптимальным для получения решения в этом случае является применение параллельных алгоритмов и параллельных систем, которые позволяют получить решение с заданной точностью за минимальное время.

Список литературы

1. Биркгоф Г. Гидродинамика. Москва: ИИЛ, 1963. стр. 244.
2. Патанкар С.В. Численное решение задач теплопроводности и конвективного теплообмена при течении в каналах. [перев.] Калабина Е.В. Москва: МЭИ, 2003. стр. 312.
3. Griebel M Dornseifer T Neunhoeffler. Numerical simulation in fluid dynamics T SIAM 1998.
4. Stam J. Stable Fluids. NY : SIGGRAPH 99 Conference Proceedings, 1999. стр. 121-128.
5. Chorin A., Marsden J. A., Mathematical Introduction to Fluid Mechanics. NY: Springer, 1993. стр. 168.
6. Мирошников А. С. Система управления параллельной обработки данных в локальных вычислительных сетях. Дисс. канд. техн. наук. Владикавказ, 2000.

УДК 004.45

Небаев И.А.

к.т.н., доцент

Санкт-Петербургский государственный университет
промышленных технологий и дизайна,
Институт информационных технологий и автоматизации,
Кафедра информационных технологий
191186 РФ, Санкт-Петербург, Большая Морская, 18
mail@radiocoder.spb.ru
http://www.sutd.ru

АНАЛИЗ ПОДДЕРЖКИ ПАРАЛЛЕЛЬНЫХ ВЫЧИСЛЕНИЙ И МНОГОЯДЕРНЫХ СИСТЕМ В СРЕДЕ ПАКЕТА ПРОГРАММНОГО ОБЕСПЕЧЕНИЯ MATLAB ДЛЯ ЗАДАЧ ПОЛИНОМИАЛЬНЫХ ВЫЧИСЛЕНИЙ В РАСШИРЕННЫХ ПОЛЯХ ГАЛУА

Аннотация: В работе приведены результаты моделирования распараллеливания полиномиальных вычислений в расширенных полях Галуа с помощью пакета MATLAB, которые показывают достаточно существенное сокращение времени вычислений с увеличением количества используемых ядер на многоядерном процессоре с поддержкой параллельных вычислений. При этом выделены некоторые особенности реализации и поддержки встроенными функциями MATLAB режима параллельной обработки данных и программной многопоточности.

Ключевые слова: параллельная обработка, MATLAB, многоядерный процессор, многопоточность, сокращение времени вычислений.

Nebaev I.A.

Ph.D., Associate Professor

St. Petersburg State University

industrial technology and design

Institute of Information Technology and Automation,

Department of Information Technology

191186 of the Russian Federation, St. Petersburg, Bolshaya Morskaya, 18

mail@radiocoder.spb.ru

http://www.sutd.ru

ANALYSIS OF SUPPORT OF PARALLEL COMPUTING AND MULTI-NUCLEAR SYSTEMS IN THE ENVIRONMENT OF MATLAB SOFTWARE FOR THE TASKS OF POLYNOMIAL COMPUTING IN EXTENDED GALOU FIELDS

Abstract: *The paper presents simulation results of parallelization of polynomial calculations in extended Galois fields using the MATLAB package, which show a rather significant reduction in computation time with an increase in the number of cores used on a multi-core processor with parallel computing support. At the same time, some features of the implementation and support of the built-in MATLAB functions of the parallel data processing mode and software multithreading are highlighted.*

Keywords: *parallel processing, MATLAB, multi-core processor, multithreading, reducing computation time.*

Введение

Поля Галуа (*Galois Field*) вида $GF(q^m)$, где $q=2$, имеют важное прикладное значение в области вычислительных систем и защиты информации от ошибок. Некоторые математические свойства полей позволяют конструировать специализированные процедуры кодирования информации, предназначенные для контроля, обнаружения и исправления ошибок, возникающих при хранении и передаче информации [1]. На математических свойствах поля Галуа основываются методы CRC кодирования, БЧХ, РС и др., широко используемые в компьютерных системах резервного хранения данных (RAID), DVD, цифровых системах радиосвязи (WCDMA, LTE и пр.) и различных компьютерных сетях (Ethernet, GE и т.д.). Для реализации алгоритмов контроля и защиты информации от ошибок, основанных на базе теории конечных полей на помощь приходят различные пакеты прикладного программного обеспечения с заранее подготовленными математическими модулями. Однако не все компьютерные системы численных вычислений, распространенные на практике (GNU Octave, Scilab, LabVIEW и пр.) обладают достаточной функциональностью программных модулей для прикладных вычислений в поле $GF(q^m)$. Одной из наиболее развитых реализаций математического аппарата и пользовательских функций полей Галуа обладает прикладной пакет MATLAB, в котором помимо базовых процедур (построение поля, полиномиальные вычисления и т.д.) реализованы востре-

бованные на практике функции (вычисление минимальных полиномов, расчет корней и т. д.). Однако, например в области обработки больших массивов информации или инфотелекоммуникационных приложениях реального времени, на первый план выходят вопросы вычислительной производительностью и полноты поддержки параллельной обработки данных. Т.о. наблюдающееся в последние несколько лет активное развитие методов параллельных вычислений со стороны модулей пакета MATLAB и значительный рост количества многоядерных систем (4-х, 8-ми ядерных ЦП и т.д.), поддерживающих параллельные вычисления, позволяют провести практическое исследование эффективности использования многоядерных систем с поддержкой режима параллельных вычислений и выявить оптимальную конфигурацию программных средств и аппаратного обеспечения для задач прикладных расчетов в конечных полях $GF(q^m)$.

Программно-аппаратное обеспечение

Для анализа производительности расчетов в конечном поле пакета MATLAB автором реализован ряд специальных программных сценариев на встроенном языке MATLAB включающих в себя операции построения поля $GF(2^m)$, поиска минимальных полиномов (для ненулевых элементов) поля, определение сопряженных элементов поля, создание массива полиномов из элементов принадлежащих данному полю, выполнение стандартных математических операций в поле (сложение, умножение, деление полиномов), вычисление корней заданных полиномов и т.д. Программа разработанного пакета тестовых сценариев отличается прикладной направленностью вычислений. Подобные вычисления характерны для задач моделирования и подбора свойств циклических кодов (БЧХ, РС и пр.) удовлетворяющих своими свойствами (длина, корректирующая способность и т.д.) заданным характеристикам прикладной системы (хранение данных, передача по беспроводным каналам и т. д.).

Для создания вычислительной нагрузки на программно-аппаратный комплекс используется 2048 примитивных неприводимых полиномов (максимально допустимой для ПО MATLAB степени $m=16$), каждый из которых образует собственное поле с 65536 элементами (16 битные символы). Для каждого из сопряженных элементов поля определяется свой циклокласс. Массив полиномов (для статистических испытаний) формируется из 100 случайных полиномов. Стандартные математические операции над полиномами одного поля выполняются последовательно, т. е. случайный полином минимальной степени $m-1$ последовательно складывается, умножается и делится на каждый из полиномов тестового массива. Итоговые результаты экспериментальных данных усреднены и получены на основе статистических испытаний по методу Монте-Карло.

Набор многоядерных процессоров представлен в табл.1. В качестве операционной системы использовалась ОС Astra Linux (разработки АО

«НПО РусБИТех»), для архитектуры x86-64. Версия пакета MATLAB – 2016a, 64-х разрядная (версия для UNIX-подобных систем). Выбор представленных ЦП обусловлен достаточно широкой доступностью в условиях коммерческого рынка, а также архитектурой ЦП, принятой промышленностью как стандарт по-умолчанию.

Аппаратное обеспечение (ЦП)

Табл. 1.

Модель ЦП	Кол-во ядер ЦП	Поддержка многопоточности
Intel Core i3 (3245, 3.4 ГГц)	2	4 поточная обработка, поддержка «HyperThreading»
Intel Core i5 (6500, 3.2 ГГц)	4	4 поточная обработка
Intel Core i7 (7820X, 3.6 ГГц)	8	16 поточная обработка, поддержка «HyperThreading»

Измерения и анализ результатов

Время выполнения каждой из операций представленного ранее сценария приведено в табл.2.

Время исполнения отдельных этапов программного сценария

Табл. 2

Кол-во ядер	Построение полей GF(2 ¹⁶)	Поиск мин. полиномов, опред. циклотомических классов	Создание массива полиномов поля	Операции над полиномами	Вычисление корней	Суммарное время
2 ядра (4 потока)	17,5	86,1	22,4	188,2	386,4	700,6
4 ядра (4 потока)	14,8	58,8	18	113,8	310,5	515,9
8 ядер (16 потоков)	13,5	46,3	17,2	53,3	78,1	208,4

Первое, что необходимо отметить это суммарное время выполнения сценария для каждого из многоядерных процессоров. Полученное значение показывает практически трехкратное сокращение времени исполнения сценария между двухядерной и восьмиядерной вычислительной системой. Исходя из технических особенностей рассматриваемых ЦП, можно отметить, что данный фактор получен не только за счет увеличенного количества ядер, но и поддержки восьмиядерным процессором режима многопоточности (16 программных потоков), т.н. функция «HyperThreading». При этом необходимо отметить, что поддержка многопоточного режима (функция *maxNumCompThreads*) в ПО MATLAB появилась только в версии R2007b, и до версии R2014a поддерживала работу не более 12 потоков. Т.о. образом важно подчеркнуть, что ис-

пользование мощного восьмиядерного ЦП с режимом многопоточности имеет смысл только при наличии актуальных (поддерживаемых) версий программного обеспечения. Разница во времени выполнения полного сценария для двух и четырехядерных ЦП не столь существенна, хотя и показывает прирост производительности на более мощном процессоре, эквивалентном выигрышу по времени порядка 3.07 минут (184.7 с). При этом следует отметить, что двухядерный процессор поддерживает многопоточный режим (4 потока) и в этом плане аналогичен более мощному ЦП, в котором режим HyperThreading отключен. Однако, в результате проведения экспериментов были выявлены некоторые особенности в поддержке функциями MATLAB режимов параллельной обработки данных. Результаты времени исполнения отдельных операций из табл.2 использованы для построения диаграмм, изображенных на рис.1

Диаграмма под литерой (а) рис.1 демонстрируют измерение времени, затраченное на построение полей $GF(2^{16})$. Время для различных ЦП отличается несущественно – максимальное отклонение между вычислениями составляет не более 4 с.. Это позволяет сделать вывод, что реализация функций построения поля в пакете MATLAB (функция $gf()$) достаточно слабо использует возможности многоядерной (параллельной) обработки и многопоточности. Однако, операция поиска минимальных полиномов (функции $minpol()$ и $cosets()$) для сопряженных элементов (диаграмма рис.1(б)) на восьмиядерном процессоре выполняется почти вдвое быстрее по сравнению с двухядерной системой. При этом отличие между временем моделирования на двух и четырехядерных ЦП отличается не так существенно. Типичная операция создания массива полиномов, для каждого из сгенерированных полей, вновь указывает на слабую поддержку параллельной обработки данных, поскольку разница между наиболее медленным и наиболее быстрым ЦП не превышает 5.2 с. Последующие диаграммы (рис.1 литеры (г) и (д)) демонстрируют полное доминирование наиболее мощного ЦП (Core i7) в плане производительности и скорости моделирования в конечных полях.

Полиномиальные операции (сложение, вычитание, умножение, деление) над полиномами заданного поля выполняются на восьмиядерном процессоре практически втрое быстрее чем на двухядерном ЦП, и более чем вдвое, чем на четырехядерном. Результаты моделирования функции $polyval()$, реализующей вычисление корней заданного полинома показывает подавляющее преимущество наиболее мощного процессора – в данном случае время вычисления последнего этапа сценария сокращается в 5 раз, по сравнению с двухядерным процессором. Аналогичное преимущество наблюдается и над четырехядерным ЦП. Исходя из этого, можно говорить об очень высоком уровне эффективности и широкой поддержки многопоточности в реализации функции $polyval()$ пакета MATLAB.

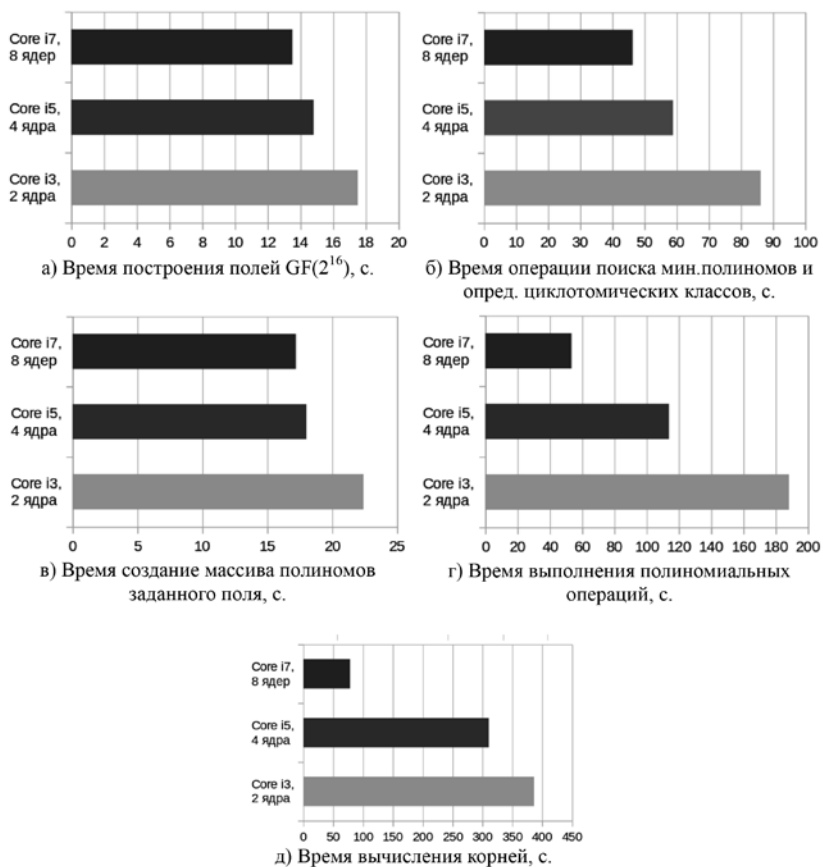


Рис.1 – Диаграммы времени исполнения отдельных этапов программного сценария.

Заключение.

Анализ результатов, полученных в ходе моделирования показывает достаточно существенное сокращение времени вычислений с увеличением количества используемых ядер на многоядерном процессоре с поддержкой параллельных вычислений. Однако существуют некоторые особенности реализации и поддержки встроенными функциями MATLAB режима параллельной обработки данных и программной многопоточности. Так наибольшей эффективностью в режиме параллельных вычислений обладают функции нахождения корней и арифметических операций над полиномами поля $GF(q^m)$. Очевидно, что в данном случае встроенные функции MATLAB (модуль Communication Toolbox) имеют специальную программную оптимизацию для выполнения параллельных вычислений. С другой стороны сокращение времени исполнения не столь очевидно для базовых операций над элементами поля – построение поля,

нахождение циклотомических классов и т. д. Как и отмечено ранее, это, скорее всего связано с недостаточным уровнем оптимизации диспетчера параллельных вычислений в указанных функция. В результате проведения эксперимента с версиями MATLAB до R2007b наличие функций программной многопоточности и многоядерная архитектура ЦП позволяет в целом увеличить производительность вычислений лишь на 3-6%, что практически не сказывается на общем времени моделирования. Т.о. можно сделать заключение, что для значительного повышения эффективности (т.е. сокращения времени расчетов) прикладных вычислений в конечном поле MATLAB следует учитывать не только факторы аппаратной поддержки ЦП многопоточного режима и наличие программного обеспечения поддерживающего обработку данных в режиме параллельных вычислений, но и степень программной поддержки многоядерных ЦП в реализациях конкретных функций. Наконец, важно еще раз подчеркнуть, что использование многоядерных ЦП с поддержкой многопоточного режима имеет смысл только при наличии программной поддержки параллельных вычислений в ПО MATLAB.

Список литературы

1. Блейхут Р. Теория и практика кодов, контролирующих ошибки: Пер. С англ. – М. : Мир, 1986. – 576 с.
2. Он-лайн руководство пользователя MATLAB <https://www.mathworks.com/help/comm/ug/error-detection-and-correction.html#fp5602> (дата обращения 09.02.2018).

УДК 515.1

Пекунов В.В.

д.т.н., ОАО «Информатика», Иваново, Россия
pekunov@mail.ru

СОКРАЩЕННАЯ АЛГЕБРАИЧЕСКАЯ НОТАЦИЯ ЗАПИСИ ПАРАЛЛЕЛЬНЫХ ВЫЧИСЛИТЕЛЬНЫХ ТОПОЛОГИЙ В PLANNING C

Аннотация: в работе предложена более удобная и наглядная по сравнению с иными средствами алгебраическая нотация записи вычислительных топологий небольшой и средней сложности. Описан синтаксис и сформулированы основные алгебраические законы для трансформации и упрощения топологий. Для получения описания топологий по алгебраической записи предложено применить специальный дедуктивный макро модуль.

Ключевые слова: параллельные вычисления, вычислительные топологии, алгебраическая нотация, упрощения топологий, дедуктивный макро модуль.

Pekunov V.V.

Doctor of Technical Sciences, OJSC «Informatics», Ivanovo, Russia
pekunov@mail.ru

REDUCED ALGEBRAIC NOTATION FOR WRITING PARALLEL COMPUTATIONAL TOPOLOGIES IN PLANNING C

Abstract: *In the paper, the algebraic notation of computational topologies of small and medium complexity, which is more convenient and more vivid than other means, is proposed. The syntax is described and the basic algebraic laws for the transformation and simplification of topologies are formulated. To obtain a description of topologies for algebraic writing, it was proposed to use a special deductive macro module.*

Keywords: *parallel computing, computational topologies, algebraic notation, simplifications of topologies, deductive macromodule.*

Введение

Параллельные вычислительные топологии являются удобными абстракциями, описывающими допустимые направления передачи данных при решении различных вычислительных задач (см., например, [1, 2]). Однако базовые средства описания таких топологий в различных языках/расширениях не всегда удобны и не очень наглядны. Это либо массивы (MPI), либо списки (Planning C [3]), либо вызовы специальных функций (ТВВ). В сложных случаях данная проблема решается путем составления специального алгоритма, который генерирует описание топологии (в Planning C, в частности, используются дедуктивные макро-модули). Однако в более простых случаях решать задачу алгоритмически не вполне целесообразно, а составление описания вручную небезопасно по причине возможного внесения трудноуловимых ошибок.

Актуальна задача разработки новых, интуитивно понятных средств описания топологий малой и средней сложности, которые можно было бы легко транслировать в классическое списковое или алгоритмическое описание.

Основные теоретические положения

В языке Planning C топология описывается набором прямых цепей и обратных дуг. При этом порядок следования цепей/дуг не имеет значения, но имеет значений порядок следования элементов в каждой отдельной цепи/дуге. Здесь можно провести *аналогию с алгебраическими операциями*, которые могут быть, соответственно, коммутативными и некоммутативными. Если обозначить узлы некоторыми идентификаторами, коммутативную операцию соединения цепей/дуг – знаком «+», а некоммутативную операцию соединения узлов в цепи/дуге – знаком «*», то получаем некоторую *алгебраическую нотацию записи топологий*. Например, запись вида « $A*B*C+B*A$ » будет означать наличие трех дуг, « $A \rightarrow B$ », « $B \rightarrow C$ » и « $B \rightarrow A$ ».

Для ее упрощения можно, по аналогии с обычной алгеброй, ввести возможность *указания скобок* «(» и «)», разрешить действие сочетательных законов сложения/умножения и распределительного закона умно-

жения относительно сложения. Далее, можно ввести *операцию возведения в константную степень* «**» (эквивалентно умножению основания степени на само себя указанное константой число раз) и *умножения на константу* (эквивалентно сложению первого множителя с самим собой заданное вторым множителем-константой число раз).

Также имеет смысл разрешить *вычитание* (некоммутативную операцию, обозначаемую знаком «-»), для которого будут справедливы все классические законы алгебры. Очевидно, что такая операция может использоваться *лишь для упрощения алгебраической записи топологии*, поскольку дуга/цепь с обратным знаком *не имеет какого-либо алгоритмического смысла*. Например, описание « $(A+B)^*2-2*A*B$ » преобразуется в « $A*A+B*B$ ».

Далее, обратим внимание на тот факт, что появление нескольких узлов с одним идентификатором в описании топологии должно, вообще говоря, сопровождаться указанием *уточняющего индекса*. Введем следующие *правила*:

а) если за идентификатором следует индекс в квадратных скобках «[» и «]», то узел с соответствующим идентификатором имеет *указанный индекс*;

б) если за идентификатором следует указание «[]», то узел с соответствующим идентификатором в конечной записи получит *уникальный индекс, больший или равный единице*;

в) если за идентификатором нет никаких указаний на индекс, то узел с соответствующим идентификатором в конечной записи получит индекс, *равный единице*.

Например, описание « $(A[]+B)^*2-2*A[]*B$ » сначала преобразуется в « $A[]*A[]+B*B$ », а затем, после расстановки индексов, превращается в « $A[1]*A[2]+B[1]*B[1]$ ».

Суммируя вышесказанное, дадим формальное описание синтаксиса предложенной нотации и дополнительно предложим перечень основных алгебраических законов, по которым возможна трансформация полученных выражений. Итак, *синтаксис описывается следующими правилами в расширенной форме Бэкуса-Наура*:

выражение = сумма

сумма = сумма «+» унарное | сумма «-» унарное | унарное

унарное = «+» с_константой | «-» с_константой | с_константой

с_константой = число «*» произведение | произведение

произведение = степенное «*» произведение | степенное

степенное = с_переменными «**» число | с_переменными

с_переменными = «(» выражение «)» | элемент
 элемент = идентификатор «[]» | идентификатор «{» число «}» | идентификатор

Сформулируем *основные алгебраические законы* для трансформации описаний топологий в предложенной нотации, в которых выражения, не содержащие констант, обозначены прописными буквами, а числовые константы – строчными:

1. $A+B = B+A$.
2. $(A+B)+C = A+(B+C) = A+B+C$.
3. $-(-A) = A$.
4. $A+(-B) = A-B$.
5. $+A = A$.
6. $A-A = 0$.
7. $A+0 = A$.
8. $(A*B)*C = A*(B*C) = A*B*C$.
9. $k*A = (A+A+...+A)$ {k раз, $k > 0$ }.
10. $(A+B)*C = A*C+B*C$.
11. $A*(B+C) = A*B+A*C$.
12. $A+B*C = A+(B*C)$.
13. $(-A)*(-B) = A*B$.
14. $(-A)*B = -(A*B)$.
15. $A*(-B) = -(A*B)$.
16. $A**k = (A*A*...*A)$ {k раз, $k > 0$ }.
17. $A*B**k = A*(B**k)$.
18. $A**k*B = (A**k)*B$.
19. $A+B**k = A+(B**k)$.
20. $A-B**k = A-(B**k)$.

Отметим, что законы 12-й и с 17-го по 20-й введены исключительно для установления приоритетов операций.

Применение специализированного макромодуля для реализации нотации

Поскольку синтаксические правила разбора могут быть достаточно хорошо описаны декларативно, наиболее простой и уместной реализацией транслятора алгебраических нотаций в стандартные для Planning С описания топологий представляется дедуктивный макромодуль (правила синтаксического разбора и все преобразования записываются в форме GNU Prolog-предикатов). Такой модуль принимает в качестве параметров, *во-первых*, список соответствий между идентификаторами процедур с планированием повторного входа (ПППВ), входящими в запись,

и их списками параметров и, *во-вторых*, собственно алгебраическую запись топологии.

Соответствующий макромодуль **altopo** включен в язык Planning C в формате подключаемого файла **altopo.h**. Далее приведен *пример применения макромодуля*: работа с топологией «бинарное трехуровневое дерево». В примере число **22** передается от корня всем узлам дерева, далее листовые узлы выводят полученные ими числа на экран, после чего топология завершает работу. Из примера достаточно очевидно, что примененная алгебраическая форма записи топологии существенно короче и нагляднее простого спискового описания топологии базовыми средствами.

```
#include "altopo.h"
#include <iostream>
using namespace std;
chain A(bool init, input_proc Ref, int Result) {
    static int leafs = 0;
    input_proc incoming[128];
    int n_in;
    input_proc outgoing[128];
    int n_out;
    plan_neighbours(true, A[plan_linear_num()], &n_in, incoming);
    plan_neighbours(false, A[plan_linear_num()], &n_out, outgoing);
    if (init) {
        if (n_in == 0)
            plan_first(false, empty_proc, Result);
        if (n_out == 0) {
            plan_critical(outCount) {
                leafs++;
            }
        }
    } else if (n_out == 0) {
        plan_critical(outResult) {
            cout<<Result<<" ";
            leafs--;
        }
        if (leafs == 0) plan_topology_quit();
    } else {
        for (int i = 0; i < n_out; i++) {
            throw_first(false, outgoing[i], Result);
        }
    }
}
```

```

    }
  }
}
int main() {
  const int N = 22;
  altopo([[ 'A', '(true, empty_proc, N)']], 'A[1]*(A[2]*(A[]+A[])
+A[3]*(A[]+A[]))');
  cout<<endl;
}

```

В качестве еще одного, заключительного примера алгебраической нотации приведем запись топологии «труба»: « $A^*B^*C^*D+D^*C^*B^*A$ »

Выводы

В данной работе предложена более удобная и наглядная по сравнению с иными средствами алгебраическая нотация записи вычислительных топологий небольшой и средней сложности. Описан синтаксис. Сформулированы основные алгебраические законы для трансформации и упрощения топологий. Для получения описания топологий (в языковом формате Planning C) по алгебраической записи предложено применить специальный дедуктивный макромодуль. Поскольку синтаксические правила и законы алгебраических трансформаций могут быть достаточно легко записаны в предикативной форме, такой подход является вполне естественным и эффективным.

Список литературы

1. Пекунов В.В. Модель образования и распространения твердых, жидких и газообразных загрязнителей. Оптимальное распараллеливание // Математическое моделирование.- 2009. – Т.21. – №3. – С.69-82.
2. Пекунов В.В. Новые методы параллельного моделирования распространения загрязнений в окрестности промышленных и муниципальных объектов // Дис. докт. тех. наук. – Иваново, 2009. – 274 с.
3. Пекунов В.В. Язык параллельного программирования Planning C. Применение при обучении глубоких нейронных сетей на гибридных системах с OpenCL-видеоускорителями // Мат. Междунар. науч.-техн. конф. «XIX Бернардовские чтения». – Иваново, 2017. – Т.3. – С.44-47.

УДК 002.6

Посохов А.С.,

ИФП СО РАН, Новосибирск, Россия, pas@isp.nsc.ru

Головина И.П.,

ИФП СО РАН, Новосибирск, Россия, golovina@isp.nsc.ru

ПОДХОД К ОРГАНИЗАЦИИ МУЛЬТИПРОГРАММНОГО РЕЖИМА ФУНКЦИОНИРОВАНИЯ ВЫЧИСЛИТЕЛЬНЫХ СИСТЕМ

Аннотация: В работе предлагается подход, который основан на разбиении набора задач, поступающего для решения в многомашинную вычислительную систему, на не-сколько частей. Распределением каждой части занимается отдельно взятая элементарная машина. В этом случае можно сократить время, затраченное системой на распределения набора, а значит и общее время работы вычислительной системы. Приведены результаты машинного моделирования предлагаемого подхода.

Ключевые слова: многопроцессорные вычислительные комплексы, мультипрограммный режим, распределение задач, минимизация времени решения пакета задач.

Posokhov A.S.,

Institute of Physics and Technology, Siberian Branch of the Russian Academy of Sciences, Novosibirsk, Russia, pas@isp.nsc.ru

Golovina I.P.,

Institute of Physics and Technology, Siberian Branch of the Russian Academy of Sciences, Novosibirsk, Russia, golovina@isp.nsc.ru

APPROACH TO THE ORGANIZATION OF MULTIPROGRAM MODE OF FUNCTIONING OF COMPUTATIONAL SYSTEMS

Abstract: The paper proposes an approach that is based on dividing the set of tasks that comes in for solving in a multi-machine computing system into several parts. The distribution of each part is carried out separately taken elementary machine. In this case, you can reduce the time spent by the system on the distribution of dialing, and hence the total time of the computing system. The results of machine modeling of the proposed approach are given.

Keywords: multiprocessor computing systems, multiprogramming mode, task distribution, minimization of solution time for a package of tasks

Современные вычислительные системы являются распределенными и больше масштабными [1]. Число элементарных машин в них неуклонно растет. Например, лидер списка TOP500 [2] редакции ноября 2017 года, суперкомпьютер Sunway TaihuLight состоит из 10,649,600 процессоров. Такие вычислительные системы могут работать не только в монопрограммных режимах, когда решается одна сложная задача, но и в мультипрограммных – когда на системе решается набор параллельных задач. Таким образом, в очереди на решение может оказаться набор, состоящий из сотен тысяч и миллионов задач. Необходимость эффективного использования вычислительных мощностей в такой ситуации становится актуальной. Точные алгоритмы, построенные на основе методов математического программирования, при росте числа задач

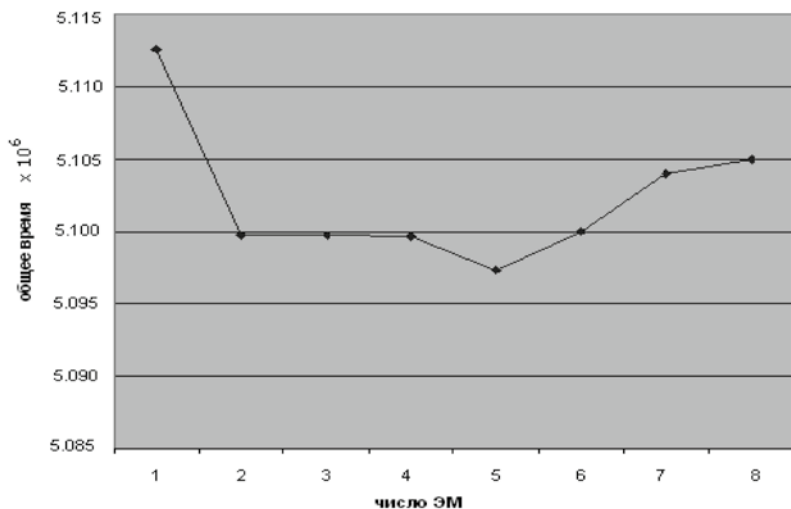


Рисунок 1 – Результаты моделирования

в наборе имеют экспоненциальную трудоемкость, и, следовательно, их применение ограничено. Достаточно эффективными оказываются эвристические алгоритмы [3].

Подход основан на разбиении набора задач, поступающего для решения на вычислительную систему, на несколько частей. Распределением каждой части занимается отдельно взятая элементарная машина. В этом случае можно сократить время, затраченное системой на распределения набора, а значит и общее время работы вычислительной системы.

Каждая машина распределяет свою часть набора по последовательному алгоритму [4]. Наборы, сформированные каждой элементарной машиной, объединяются в общий, который в свою очередь и является оптимальным распределением задач по вычислительной системе.

В результате моделирования алгоритма, были получены следующие значения общего времени (затраченного на распределение и последующее решение набора) для числа элементарных машин, участвующих в распределении набора из 2×10^6 задач (рис. 1).

При этом общее время решения задач $GT=10,089 \times 10^6$, нижняя граница времени решения $LB=5,038 \times 10^6$. В данном случае рекомендуется разбивать набор на пять частей, при этом время решения набора плюс время, затраченное на распределение, $RT=5,097 \times 10^6$, что достаточно близко к нижней границе.

Список литературы

1. Хорошевский В.Г. Архитектура вычислительных систем. – М.: МГТУ им. Н.Э. Баумана, 2005. – 512 С.
2. TOP 500 Supercomputer sites [Электронный ресурс]: November 2017 / – Режим доступа: <http://www.top500.org>
3. Евреинов Э.В., Хорошевский В.Г. Однородные вычислительные системы, Наука, 1978
4. Посохов А.С. Алгоритм функционирования вычислительных систем при решении набора сложных задач. // Материалы Российской научно-технической конференции «Информатика и проблемы телекоммуникаций – 2005», Новосибирск, СибГУТИ, 2005 г., том 1, стр. 223.

СЕКЦИЯ 4. ЗАЩИТА ИНФОРМАЦИИ

УДК 621.39

Анисимов А.В.,

кандидат технических наук, доцент, Microsoft Certified Professional, профессор кафедры Информационных Систем факультета Информационных Технологий Смольного Института РАО, 195197 Россия, Санкт-Петербург, Полустровский пр., 59, тел. 5402037, www.smun.spb.ru, e-mail: avanisimov@yandex.ru

ОБЕСПЕЧЕНИЕ БЕЗОПАСНОСТИ DNS

Аннотация: Описываются различные виды атак, целью которых является удаленный несанкционированный доступ к ресурсам сервера, а также методы противодействия им. К последним в работе отнесены организационные методы и специализированные аппаратно-программные комплексы, оснащенные межсетевыми экранами, брандмауэрами или системами Firewall.

Ключевые слова: DNS – сервер, режим работы, противодействие атакам, классификация атак, процессы в сетях ЭВМ.

Anisimov A.V.,

Candidate of Technical Sciences, Associate Professor, Microsoft Certified Professional, Professor, Department of Information Systems, Faculty of Information Technologies, Smolny Institute of the Russian Academy of Education, 195197 Russia, St. Petersburg, Polyustrovsky Ave., 59, tel. 5402037, www.smun.spb.ru, e-mail: avanisimov@yandex.ru

SECURING DNS

Abstract. Describes the various types of attacks, the purpose of which is remote unauthorised access to server resources, as well as methods to counter them. The latter include organizational methods and specialized hardware and software systems equipped with firewalls, firewalls, or Firewall systems.

Keywords: DNS – server, mode of operation, counteracting attacks, attack classification, processes in computer networks.

Введение

Взаимосвязь между именем компьютера и IP-адресом DNS – алгоритм удаленного поиска IP – адреса по имени в сети Internet сводится к отправке запроса. Результат запроса достигается путем поиска по базе данных службы доменных имен -DNS (DomainNameSystem). База данных DNS содержит отдельные записи, которые называются Записи Ресурса (ResourceRecords). Отдельные части базы данных DNS, называемые зонами, размещены на специфических серверах имен. Запрос на IP – адрес DNS – сервера, и на запросивший хост DNS – сервер инициирует DNS ответ, в котором указывает искомый IP – адрес. Если имя DNS – сервер не обнаружил, то DNS – запрос отправляется на один из корневых DNS.

Безопасность DNS

С точки зрения безопасности уязвимость этой схемы удаленного поиска состоит в возможности осуществления в сети удаленной атаки "Ложный объект". Служба доменных имён является системной службой, и DNS – преобразование символьных имен в IP адреса оперирует с распределённой иерархической базой данных имеющей вид дерева имен. Внедрение в сеть ложного DNS – сервера путем перехвата DNS – запроса может быть реализовано как удаленная атака, связанная с ожиданием поискового DNS – запроса. Алгоритм атаки на службу DNS отталкивается от того, что по умолчанию служба DNS функционирует как базе протокола UDP так и протокола TCP, что делает ее менее защищенной, так как протокол UDP в отличии от TCP вообще не предусматривает средств идентификации сообщений. Для реализации атаки путем перехвата DNS – запроса атакующему необходимо перехватить DNS – запрос и извлечь из него номер UDP – порта отправителя запроса. Благодаря этому можно воздействовать по схеме "Ложный объект " на трафик между хостом и сервером. Если отталкиваться от расположения доменов верхнего уровня (TLD), таких как org, com, gov, edu, то выделенный домен имеет двухсимвольное имя в соответствии со стандартом ISO. У домена есть свои правила, в соответствии с которыми предоставляется имя хоста. Обобщенная схема работы ложного DNS- сервера имеет следующий вид. Сначала – ожидание DNS – запроса; потом получение DNS – запроса, извлечение из него сведений и передача по сети ложного DNS – ответа. Необходимым условием осуществления данной атаки является перехват DNS – запроса [1]. Поскольку атакующий не имеет возможности перехватить DNS – запрос, то основную проблему представляет номер UDP – порта, с которого был послан запрос. Однако, номер порта отправителя принимает ограниченный набор значений, и DNS – преобразование символьных имен в IP адреса оперирует с распределённой иерархической базой данных имеющей вид дерева имен. Корректность базы данных в каждом узле дерева – прерогатива администратора домена. Внедрение в сеть ложного DNS – сервера путем перехвата DNS – запроса может быть реализовано как удаленная атака, связанная с ожиданием поискового DNS – запроса.

Режимы работы сервера

С точки зрения общей реализации взаимодействия в сети существуют следующие режимы работы сервера DNS:

master, в котором файлы баз данных ведутся вручную;

slave, в котором автоматически осуществляется копирование с master сервера;

hint – режим кеширования всех запросов.

Если атакующий находится либо на пути основного трафика, либо в сегменте настоящего DNS – сервера, то выполнение условий местонахождения атакующего делает удаленную атаку затрудненной. Однако,

если правильно выбрать режим, то возможно осуществление межсегментной удаленной атаки. Реализация удаленной атаки связана с особенностями работы протокола FTP и процедурой идентификации TCP – пакетов. При подключении FTP – клиента к удаленному FTP – серверу через ложный DNS – сервер после выдачи пользователем прикладной команды FTP следует ряд действий FTP – клиента. Например – команда PORT по передаче на FTP – сервер. При этом на ложном DNS – сервере пакет воспринимается как нормальный пакет. С точки зрения рассматриваемой проблемы существенным является организация записи ресурсов в базе данных домена. При этом следует учитывать особенности размещения и конфигурирования серверов для корпоративной сети. Сервис DNS является критическим для функционирования сети, и корпоративной в сети создают два сервера. На внутреннюю сеть обращены оба сервера, а на внешнюю – один.

Существует алгоритм, в котором атакующий выполняет передачу на атакуемый хост ложного DNS – ответа от имени настоящего DNS – сервера без приема DNS –запроса. Достигается такой подход при использовании протокола UDP. Поскольку в этом протоколе не предусмотрена идентификации пакетов, то такая атака реализуема. При этом DNS – ответ должен быть направлен на тот же UDP – порт, с которого был послан DNS – запрос. Поскольку атакующий не имеет возможности перехватить DNS – запрос, то для преодоления проблемы используется номер UDP – порта, с которого был послан запрос. Нужно иметь в виду, что номер порта отправителя принимает ограниченный набор значений и атакующий может осуществить простой перебор, направляя ложные ответы на выбранный перечень портов.

Противодействие атакам

Для противодействия атакам в корневых DNS – серверах, адреса которых содержатся в файле настроек сервера root.cache используются средства обеспечения безопасности. Технология работы DNS – сервера, при которой каждый DNS – сервер кэширует в области памяти свою таблицу соответствия имен и IP – адресов хостов является одним из таких средств. В кэш заносится динамически изменяемая информация о именах и IP – адресах хостов, поскольку он находится посередине между сайтом и сервером, тосоответствующая настройка снижает нагрузку на сервер. Под настройкой подразумевается не настройка кэширующего прокси, а настройка веб-сервера. Если в ответ на запрос от DNS – сервера атакующий направит ложный DNS, то в кэш-таблице сервера появится соответствующая запись с ложными сведениями. Эта запись может быть использована с целью искажения сведений при обращении к хосту.

Тестирование работы сервера производится командой dig, которая позволяет производить произвольные запросы к указанному в командной строке серверу. Поскольку в кэш заносится динамически изменя-

емая информация о именах и IP – адресах, то при попадании запроса кеширующему серверу информация может быть устаревшей, в связи с чем запросы для проверки обратной зоны целесообразно осуществлять с использованием полного имени записи PTR.

Многие функции безопасности, могут быть применены для ненадлежащего использования. Перехват данных, передаваемых по сети, может быть осуществлен под видом мероприятий, связанных с информационной безопасностью. Протоколы сетевого анализа следует удалять после их использования. Несанкционированный перехват данных представляет собой одно из нарушений должностных инструкций системных администраторов. Сетевые фильтры имеют ряд недостатков с позиций безопасности-аутентификация отправителя производится только на основании IP-адреса. С помощью подмены IP-адресов, что носит название IP-spoofing можно обойти преграду. Сетевой фильтр не может отслеживать работу сетевых приложений, и не контролирует содержимое пакетов транспортного, сеансового и прикладного уровня.

Классификация атак

Если сопоставить осуществление в сети удаленной атаки "Ложный объект" с другими атаками, то уместно воспользоваться моделью STRIDE. Термин STRIDE является обобщением следующих атак:

- Spoofing (подмена данных),
- Tampering (изменение),
- Repudiation (аннулирование),
- InformationDisclosure (разглашение сведений),
- DenialofService (отказ в обслуживании),
- Elevation of Privilege (повышение уровня доступа).

Дадим краткую характеристику этих атак:

SpoofingofIdentity – представляет собой способ искажения параметров идентификации участника сетевого обмена.

TamperingwithData – предусматривает подмену данных, которые предполагается передавать между сервером и клиентом.

Repudiation – представляет собой атаку в форме отрицания действий, выполненным участником транзакции.

InformationDisclosure – при сопоставлении с DNS – алгоритм удаленного поиска IP – адреса по имени в сети Internet сводится к перенаправлению информации.

DenialofService – отказ в обслуживании.

ElevationofPrivilege – это изменение уровня привилегий.

Процессы в сети

Внедрение в сеть ложного DNS – сервера путем перехвата DNS сопоставим с процессами в локальной сети. При трансляции Интернет-пакетов по локальной сети, локальная сеть маршрутизирует локальные IP-адреса, а также еще и IP-адреса Интернета. В итоге такая процедура

создает реальные интернет-адреса. Описанная последовательность действий порождает отрицательные эффекты:

- Маршрутизация локальной сети должна обеспечить, чтобы все IP-адреса, не принадлежащие локальной сети, были перенаправлены на брандмауэр. Указанные настройки предусматривают создание записи в таблицах маршрутизации каждого компьютера и каждого маршрутизатора локальной сети.

- Сетевые администраторы, должны таким образом настроить, её безопасность, чтобы защитить сеть от внешних атак. Возможность атак трансляции Интернет-пакетов по локальной сети зависит от настройки брандмауэра.

Передача адресов Интернета по локальной сети используется в случае запуска прозрачного прокси-сервера на брандмауэре. Прозрачный прокси приводит к тому, что пользователи сети не подозревают о его существовании и он незаметен для них. При подключении любой закрытой компьютерной сети к открытым сетям, проблему составят угрозы несанкционированного вторжения в закрытую сеть из открытой. При этом возникает риск несанкционированного доступа из закрытой сети к ресурсам открытой. Для случая объединения отдельных сетей, ориентированных на обработку конфиденциальной информации, возникают угрозы разного уровня, а также категорий секретности. Ограничение взаимного обмена данными при осуществлении доступа сетей друг к другу возможны нарушения установленных ограничений.

Внешняя сеть должна подчиняться установленным правилам межсетевого взаимодействия:

- Недопустимости неправомерного вторжения во внутреннюю сеть из внешней.

- Недопустимости неправомерного вторжения во внешнюю сеть из внутренней.

Если в качестве внешней сети используется открытая, либо любая другая потенциально враждебная сеть, то неправомерное вторжение во внутреннюю сеть из внешней может выполняться как с целью несанкционированного использования ресурсов внутренней сети, так и с целью нарушения работоспособности сети. Без использования средств защиты передачи адресов Интернета по локальной сети высока вероятность успешной реализации подобных угроз. Порождает нарушение информационной безопасности недостатки набора протоколов TCP/IP. Даже версия протокола IPv4 не устраняет вероятность подобных несанкционированных действий.

Корпоративные сети

Для корпоративных сетей предусматриваются ограничения доступа, связанные с особенностями компании, использующей корпоративную сеть. Причины установления указанных ограничений следующие:

- Борьба с утечкой секретной информации.
- Поддержка установленных регламентов в корпоративной сети.
- Сегментация информации по профилю использования.
- Контроль за использованием информационных ресурсов в интересах производства.

Решение задачи установления указанных ограничений с позиций угроз безопасности межсетевого взаимодействия связано с характеристиками системного программного обеспечения компьютера. Администратор операционных систем не всегда в состоянии эффективно сконфигурировать системное программное обеспечение компьютера. Кроме того, защита от нежелательных действий пользователей также зачастую оставляет желать лучшего. Сеть зачастую формируется из различных компьютеров, неоднородных по настройкам и связанных между собой сетевым оборудованием. В качестве возможной рекомендации решения поставленных проблем можно указать специализированные аппаратно-программные комплексы. Подобный подход обеспечит целостность в защите корпоративной сети.

Аппаратно-программные комплексы укомплектовываются межсетевыми экранами, брандмауэрами или системами Firewall. Межсетевой экран размещается между внутренней и внешней сетями. Существуют решения, основанные на применении описанной технологии внутри корпоративных сетей. При этом решаются вопросы внутриотраслевой этики, а также ограничения доступа сотрудников к закрытым ресурсам сети.

Заключение

С точки зрения правовой базы подобной организации использования информационных ресурсов следует применять «Руководящий документ. «Межсетевые экраны» Гостехкомиссии при Президенте РФ». Согласно данному документу «межсетевым экраном называется локальное (однокомпонентное) или функционально-распределенное средство (комплекс), которое реализует контроль за информацией, поступающей в автоматизированную систему и/или выходящей из нее, и обеспечивает защиту автоматизированной системы посредством фильтрации информации, т. е. анализа по совокупности критериев и принятия решения об ее распространении в (из) автоматизированной системе».

Библиографический список

1. Крикет Ли, Пол Альбитц DNS и BIND М.: Изд. Символ-Плюс, 2008. 712 с.

УДК 004.942

Козунова С.С.,

аспирант, ФГБОУ «Волгоградский государственный технический университет»,
г. Волгоград, Россия, one1100on@gmail.com, vstu.ru

Бабенко А.А.,

к.п.н, доцент, ФГАОУ ВПО «Волгоградский государственный университет»,
г. Волгоград, Россия, ba_benko@mail.ru, volsu.ru

ОБЕСПЕЧЕНИЕ ИНФОРМАЦИОННОЙ БЕЗОПАСНОСТИ В СЕГМЕНТЕ КОРПОРАТИВНОЙ ИНФОРМАЦИОННОЙ СИСТЕМЫ

Аннотация: Процесс обеспечения информационной безопасности (ИБ) корпоративной информационной системы (КИС) является набором сложных процедур, представляющим собой решение частных проблем, которые обладают рядом специфических особенностей. Проблема обеспечения ИБ в КИС не является новой, однако вопросы обеспечения ИБ в сегментах КИС не достаточно исследованы. В данной статье предложена модель безопасности информации в сегменте КИС.

Ключевые слова: корпоративная информационная система, сегмент, типовая архитектура, модель безопасности информации, зона безопасности, иерархическая структура.

Kozunova S.S.,

graduate student, Volgograd State Technical University, Volgograd, Russia,
one1100on@gmail.com, vstu.ru

Babenko A.A.,

Ph.D., Associate Professor, Volgograd State University, Volgograd, Russia,
ba_benko@mail.ru, volsu.ru

ENSURING INFORMATION SECURITY IN THE SEGMENT OF CORPORATE INFORMATION SYSTEM

Abstract. The process of ensuring information security (IS) of a corporate information system (CIS) is a set of complex procedures, representing the solution of particular problems, which have a number of specific features. The problem of providing information security in a CIS is not new, but the issues of providing information security in segments of a CIS are not sufficiently studied. This article proposes an information security model in the EIS segment.

Keywords: corporate information system, segment, typical architecture, information security model, security zone, hierarchical structure.

Введение

Актуальность проблемы обеспечения ИБ в сегменте КИС обоснована, прежде всего, масштабируемостью и высоким спросом на системы такого класса. Это подтверждается следующими факторами: увеличение числа КИС и их сегментов в организациях, разнообразие сегментов КИС, проблема эффективного управления сегментами КИС, реорганизация и рост производственных работ на предприятиях [1]. Результаты исследований [1-2] показывают, что основным вопросом при обеспечении ИБ КИС является классификация информационных активов, обрабатываемых в КИС, учёт специфики предприятия, подбор и внедрения средств

защиты информации (СЗИ), которые уникальны для каждого сегмента КИС. При низкоуровневой реализации вопросов, перечисленных выше, значительно снижается качество эксплуатации ИС [3]. Для обеспечения безопасности информации, циркулирующей в КИС, необходимо использовать специализированную управленческую концепцию, основанную на модели безопасности [1, 4].

Исходя из этого, сформулируем цель данного исследования: разработка модели обеспечения ИБ в сегменте КИС.

Сегментирование корпоративной информационной системы

Независимо от типа обработки данных: централизованная, распределённая, файл-серверная, клиент-серверная, многоуровневая или иерархическая, присущие современным КИС, такие системы содержат несколько сегментов [1, 2]. В работах [1, 5] дано определение «сегмент КИС» как часть КИС, полученная при делении КИС на составные части, описывающие базовые и функциональные характеристики всей системы. Основываясь на исследованиях [1-6], спроектируем типовую архитектуру КИС, представленную на рис. 1.

Типовая архитектура КИС (рис. 1) состоит из интегрированной ИС и шести сегментов: информационное обеспечение, управление, автоматизированные рабочие места пользователей, принятие решений, обработка данных, сетевой сегмент. Такое выделение сегментов КИС позволяет отобразить их функционал и взаимосвязи, что даёт возможность учесть воздействия, которые будут оказывать на них СЗИ. С практической точки зрения при эксплуатации и управлении КИС выделены в отдельную сеть, чаще всего в подсеть сети общего пользования организации. В связи с этим, сетевой сегмент и сегмент управления имеют наиболее сложную структуру.

Описание модели безопасности информации в сегменте корпоративной информационной системы

В качестве решения исследуемой проблемы, мы предлагаем модель безопасности информации в сегменте КИС, которая представлена на рис. 2. Центральным понятием такой модели является «зона безопасности». Под «зоной безопасности» в общем случае понимается совокупность механизмов защиты, объектов защиты и сегментов КИС, которые способны подвергнуться угрозам КИС [7]. В [1] предложена модель защищённой ИС корпоративного типа, которая поддерживает пять уровней ИБ. В связи с тем, что в [1, 2] данные уровни распространяются на всю КИС, то при разработке модели (рис. 2) мы будем руководствоваться стратегическим и управленческим уровнями.

Данная модель поддерживает иерархическую структуру. Единицами иерархии являются зоны безопасности. К каждому из сегментов А-Е необходимо применять соответствующие механизмы защиты. Максимальное число мер по обеспечению безопасности приходится

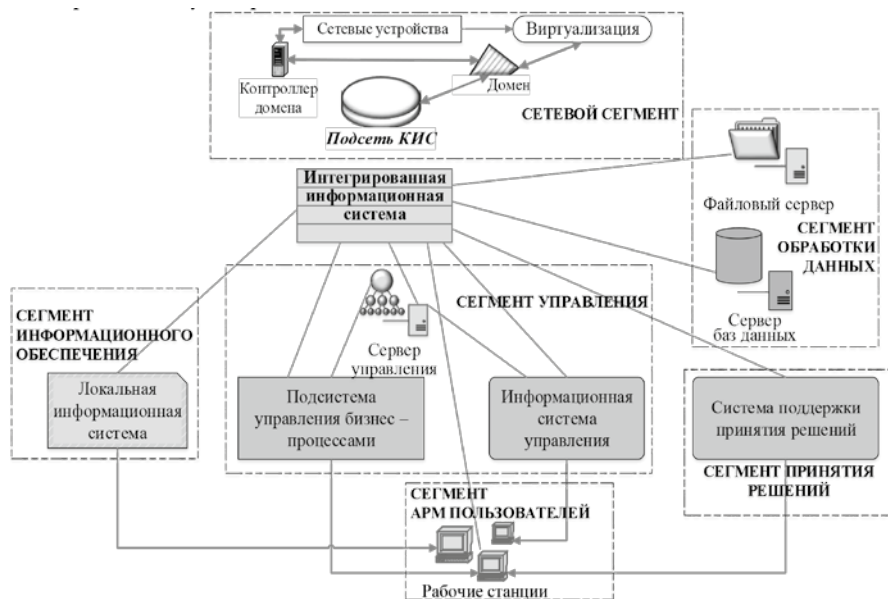


Рисунок 1 – Типовая архитектура корпоративной информационной системы

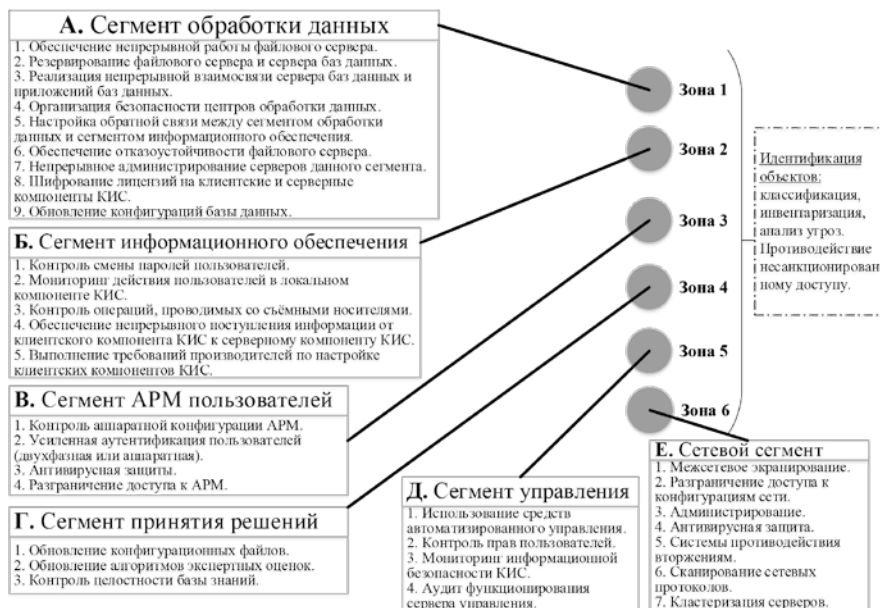


Рисунок 2 – Модель безопасности информации в сегменте корпоративной информационной системы

на сегмент обработки данных, минимальное число мер – на сегмент принятия решений. Самым трудоёмким является обеспечение безопасности сетевого сегмента, а именно: межсетевое экранирование, кластеризация серверов, системы противодействия вторжений (СПВ). При организации защиты с применением межсетевого экранирования присутствует риск получения избыточной защиты и конфликта межсетевых экранов, которые могут привести к недоступности сервисов КИС.

СПВ должны быть подсистемой КИС, что обуславливает сложность их использования. Для качественного внедрения СПВ, необходимо учесть их наличие и особенности функционирования в техническом проекте. Организация кластера серверов представляет собой операцию объединения всех серверов в единую систему. Эта процедура повышает отказоустойчивость ИС в целом. Модель (см. рис. 2) даёт возможность визуально-аналитического проектирования [8] ИС в защищённом исполнении.

Заключение

Обеспечение ИБ в сегменте КИС исследовано как сложный процесс. Обоснована актуальность проблематики обеспечения ИБ в сегменте КИС. Разработана модель безопасности информации в сегменте КИС, которая основана на «зоне безопасности». Данная модель ориентирована на сегментную структуру КИС.

Список литературы

1. Козунова С.С., Бабенко А.А. Модель построения защищённой информационной системы корпоративного типа // Информационные системы и технологии, 2016. – №3(95). – с. 112-120.
2. Гнеушев В.А., Кравец А.Г., Козунова С.С., Бабенко А.А. Моделирование сетевых атак злоумышленников в корпоративной информационной системе // Промышленные АСУ и контроллеры. 2017. №6. С. 51-60.
3. Ажмухамедов И.М., Князева О.М. Комплексная оценка качества информационных систем на основе нечёткого когнитивного моделирования // Научные тенденции: Вопросы точных и технических наук сб. научных трудов по матер. X межд. научн. конф. (г. Санкт-Петербург, 12 октября 2017). С. 10-12.
4. Кравец, А.Г. The Risk Management Model of Design Department's PDM Information System / А.Г. Кравец, С.С. Козунова // Creativity in Intelligent Technologies and Data Science. Second Conference, CIT&DS 2017 (Volgograd, Russia, September 12-14, 2017) : Proceedings / ed. by A. Kravets, M. Shcherbakov, M. Kultsova, Peter Groumpos ; Volgograd State Technical University [et al.]. – [Germany] : Springer International Publishing AG, 2017. – P. 490-500. – (Ser. Communications in Computer and Information Science ; Vol. 754).
5. Богданов Н.Г., Бочков П.В., Нечаенко Н.Д. Администрирование безопасности корпоративных информационных систем на основе ролевого управления доступом // Информационные системы и технологии, 2015. – №2(88). – с.124-130.

6. Суркова Н.Е. Методология структурного проектирования информационных систем // Монография / Н.Е. Суркова, А.В. Остроух. Красноярск: Научно-инновационный центр. 2014. 190 с. – [Электронный ресурс.] – URL: <http://lib.madi.ru/fel/fel1/fel16S061.pdf> (дата обращения: 06.12.2017).
7. Бабенко А.А., Козунова С.С. Information security model in the segment of corporate information system // Информационные системы и технологии, 2017. – №1(99). – с. 87-91.
8. Гордеев Д.С. Визуализация алгоритмов на графах: интерпретация алгоритма в качестве программы // XIII Междунар. научно-технич. Конференция «ИТ-технологии: развитие и приложения», Владикавказ, 14-15 дек. 2012 г., с. 220-227.

УДК 004.056

Ибрагимов Байрам,

д.т.н., профессор, Азербайджанский Технический
Университет, Баку, E-mail: i.bayram@mail.ru

Гаджиев Тамерлан,

магистрант, Азербайджанский Технический
Университет, Баку.

ЭФФЕКТИВНОСТЬ КРИПТОГРАФИЧЕСКИХ МЕТОДОВ АКТИВНОЙ ЗАЩИТЫ ИНФОРМАЦИИ В СИСТЕМЕ ТЕЛЕКОММУНИКАЦИЙ

Аннотация: проанализированы методы передачи и криптозащиты информации от несанкционированного доступа к абонентской линии связи. На основе исследования предложен криптографический метод активной защиты информации в системах телекоммуникации, учитывающий показатели методов шифрования и алгоритмов кодирования сообщений.

Ключевые слова: активная защита информации, кодирование, алгоритм, криптоатака, криптостойкость, несанкционированный доступ.

Ibrahimov Bayram,

m, Doctor of Technical Sciences, Professor, Azerbaijan Technical
University, Baku, E-mail: i.bayram@mail.ru

Tamerlan Hajiyev,

Undergraduate, Azerbaijan Technical
University, Baku.

EFFICIENCY OF CRYPTOGRAPHIC METHODS OF ACTIVE PROTECTION OF INFORMATION IN THE SYSTEM OF TELECOMMUNICATIONS

Abstract. The methods of transmitting and encrypting information from unauthorized access to the subscriber line of communication are analyzed. Based on the research, a cryptographic method of active information protection in telecommunication systems was proposed, taking into account the indicators of encryption methods and message coding algorithms.

Keywords: *active information protection, coding, algorithm, crypto-attack, cryptograph-ic resistance, unauthorized access.*

Введение

Бурное развитие мультисервисных сетей связи на базе современных информационных и компьютерных технологий, рост передаваемого мультимедийного трафика по каналам связи, требуют создания эффективных методов активной защиты передаваемой информации, предъявляемые к качеству обслуживания трафика (QoS – Quality of Service).

В системе телекоммуникаций под активной защитой информации подразумевается система, с помощью которой обеспечивается защита информации, передаваемой от источника информации, шифратора и кодера канала связи к дешифратору, декодеру и получателю сообщения. Предполагается, что данная система является комплексной системой активной защиты от непрерывных попыток несанкционированного доступа к защищаемой информации, и криптозащита производится от начала до конца тракта (End to end) при передаче, обработке и приеме сообщения.

В работах [1, 2] рассмотрены средства и алгоритмы активной защиты информации в распределенных системах связи, которые выполнены с моделированием процесса разработки и исследования системы активной защиты информации.

Среди последних публикаций по активной защите информации внимания заслуживают исследования, посвященные синтезу операций на основе криптографических методов преобразования [1-3], среди которых особое место занимают способы замены, перестановки, гаммирования и комбинированные. При этом центральное место занимает метод шифрования с помощью датчика псевдослучайных чисел, алгоритма прямого криптопреобразования и эффективного метода кодирования.

Таким образом, разработка криптографического метода и средства активной защиты информации с использованием способов шифрования и сетевого кодирования сообщений безусловно является актуальной задачей в системе телекоммуникаций.

Рассматривается решение задачи исследования криптографических методов активной защиты передаваемой информации, использующей сетевое кодирование, совмещенное с шифрованием передаваемых сообщений в системе телекоммуникаций.

Общая постановка задачи

На основе системного анализа установлено [1, 4-6], что среди современных криптографических методов активной защиты передаваемой информации особое место занимает метод шифрования с использованием алгоритмов прямого криптопреобразования и сетевого кодирования.

Проведенные исследования показали [3, 7, 8], что концепция сетей NGN требует разработки адекватных методов шифрования и сетевого

кодирования, обеспечивающих простую аппаратную и программную реализацию, а также высокий уровень криптографической стойкости за счет расширения спектра используемых операций криптографического преобразования информации. И это является одной из актуальных задач информационной безопасности в системе телекоммуникаций.

В работе [4, 5, 8] вопросу применения группы операций замены, перестановок, гаммирования и сетевого кодирования управляемых информацией для реализации криптографических методов активной защиты информации достаточное внимание не уделено.

Для решения поставленных выше задач предлагаются эффективные криптографические методы активной защиты информации от нелегальных пользователей в системе телекоммуникаций, базирующейся на алгоритмах и средствах методов шифрования с помощью датчика псевдослучайных чисел и сетевого кодирования.

На основе системно-технического анализа математическая задача предлагаемого криптографического метода активной защиты передаваемой информации в системе телекоммуникаций может быть сформулирована следующими целевыми функциями:

$$E_{эф} = Arg \max_i T_{i,max} [I(X)], \quad i = \overline{1, n}, \quad (1)$$

при следующих ограничениях

$$P_{ou} \leq P_{ou,дон}, \quad \eta_i \leq \eta_{i,дон}, \quad T_{i,сп.з} \leq T_{i,сп.з,дон}, \quad (2)$$

где P_{ou} – вероятность ошибки при выполнении криптографического преобразования информации от источника до получателя сообщения;

$T_{i,max} [I(max)]$ – величина периода времени, необходимого для выполнения операции декодирования злоумышленником информации i -го типа при передаче $I(X)$ количество информации по каналу связи, $i = \overline{1, n}$; $T_{i,сп.з}$ – среднее время задержки при передаче i -го потока сообщения, $i = \overline{1, n}$; η_i – коэффициент эффективного использования сетевых ресурсов системы и средства активной защиты информации при передаче i -го потока сообщения, $i = \overline{1, n}$.

Выражения (1) и (2) определяют сущность рассматриваемого криптографического метода активной защиты информации в системе телекоммуникации и являются одним из важных показателей криптостойкости системы при передаче информации.

Таким образом, вышеперечисленные варианты злоумышленников и вредоносные приложения нелегальным пользователям могут представлять собой угрозу целостности или доступности передаваемых данных по каналам связи, а нежелательные – угрозу конфиденциальности.

Целью данной работы является исследование эффективности криптографических методов и алгоритмов активной защиты передаваемой ин-

формации, классификации нежелательных приложений трафика для обнаружения угроз безопасности в системе телекоммуникации с использованием сетевое кодирование, совмещенное с шифрованием передаваемых сообщений.

Схема функционирования исследуемого метода и средства активной защиты информации

На основе системно-технического анализа [2, 3] установлено, что для построения системы телекоммуникаций на базе концепции сетей связи NGN будущего поколения необходимо создание криптостойких методов и средств активной защиты передаваемой информации от несанкционированного доступа.

Выявлено [4, 5], что методы и средства активной защиты информации в основном зависят от типа передаваемой информации, способов передачи сообщения, формы ее хранения и обработки и носителя сигнала, а также модели и методов несанкционированного доступа к абонентской и сетевой линиям связи. При этом важное место занимают предполагаемые методы нападения – криптоатака, и последствия ее влияния на информацию – копирование, искажение и уничтожение.

С целью реализации поставленной выше задачи и целевых функций (1) и (2) была предложена приведенная на рис. 1 структурно-функциональная схема активной защиты передаваемой информации от несанкционированного доступа к абонентской и сетевой линии связи, т.е. от нелегального пользователя или злоумышленника.

Исследуемый рис. 1 характеризует криптографический метод и средства активной защиты информации и в основном состоит из системы передачи, канала связи и системы приема сообщения, использующей симметричные алгоритмы с открытой парой ключей ke и kd , $K \in (ke, kd)$ при воздействии криптоатакой. В данном случае под криптоатакой подразумевается попытка несанкционированного доступа к защищаемому передаваемому сообщению. Определено, что успешная криптоатака производится методом взлома и методом вскрытия шифра [3, 4].

Из схемы следует, что активная защита передаваемой информации от нелегального пользователя или злоумышленника производится от абонента-А до абонента-В, т.е. по всему тракту системы передачи сообщения – от начала до конца.



Рис. 1. Структурно-функциональная схема активной защиты информации от несанкционированного доступа к абонентской и сетевой линиям связи

Исследование и создание криптографического метода активной защиты информации

Известно [4-6], что в основе всех криптографических способов преобразования информации и алгоритмов активной защиты информации лежат две базовых операции:

1. Способ подстановки, при котором производится управляемая (в зависимости от используемого ключа) замена каждого символа или группы символов исходного сообщения другим символом или группой символов. Данный способ может быть реализован по различным принципам [3, 4]: простой способ подстановки, множественный способ подстановки, динамический способ подстановки, многоэлементный способ подстановки.

В современных средствах активной криптозащиты способ подстановки реализуется с помощью операций сложения по модулю 2 (mod2) над защищаемой информацией и некоторой псевдослучайной последовательностью, задаваемой ключом $K[ke, kd]$, ke – ключ шифрования, а kd – ключ дешифрования (иногда mod 2 называется логической операцией «исключающее ИЛИ»). Здесь, ключ ke называется открытым ключом, а ключ kd – индивидуальным ключом или закрытой частью ключа.

Следует отметить, что алгоритмы с открытым ключом называются также двухключевыми [4, 6]. В отличие от симметричных, использование двухключевых алгоритмов активной криптозащиты не требует решения задачи распространения ключей. При этом существенно снижается время поиска перехвата ключей.

На основе исследования [4], определено, что многие двухключевые алгоритмы активной защиты информации обладают свойством линейности, которое описывается следующим образом:

$$X = D_{kd}[E_{ke}(X)] = E[D_{kd}(X)], \quad (3)$$

Выражение (3) определяет свойства линейности и способствует изменению порядка операций шифрования и дешифрования. При этом любое сообщение X , зашифрованное его владельцем с помощью закрытого ключа kd , может быть дешифровано любой персоной, но только с помощью открытого ключа ke того же персоны, подтверждая авторства сообщения.

Из алгоритма видно, что данная технология открывает новые функции криптозащиты информации – аутентификацию и идентификацию, которые называются электронной подписью.

Учитывая алгоритмы защиты информации вышеназванными способами и средствами подстановки, установлено [4], что двукратное применение операции сложения по модулю 2 – один при шифровании и один раз при дешифровании – с одним и тем же ключом позволяет восстановить исходную информацию:

$$X \oplus K = Y, \quad Y \oplus K = X \oplus K \oplus K = X \quad (4)$$

Выражение (4) выполнения логической операции сложения mod2 над шифруемым сообщением и псевдослучайной последовательностью часто называют гаммированием, а саму псевдослучайную последовательность – гаммой шифра.

Кроме того, в данном способе операция «сложения mod 2» с учетом (4) только реализует механизм подстановки и не обеспечивает секретности активно защищаемой информации. Здесь секретность заключается в длине и статических свойствах ключевой псевдослучайной последовательности и, в частности, ее предсказуемости. С этой точки зрения вопрос генерации псевдослучайных последовательностей и вопрос генерации ключей, предназначенных для использования в системах криптографической защиты информации, является важным направлением криптологии [4].

2. Способ перестановки, которым предполагается зависимое от ключа изменение порядка следования символов в сообщении и алгоритма перестановки, в современных системах криптозащиты обычно выполняется в оперативной памяти данных процессора, реализующего алгоритм активной защиты передаваемой информации.

Определение способности криптографического метода к активной защите информации

Проведенные исследования показали [4-6], что способность крипто-системы с использованием алгоритма криптозащиты, способов шифрования и сетевого кодирования сообщения противостоять попыткам несанкционированного доступа к защищаемой информации с последующим ее искажением и уничтожением, называется криптостойкостью метода защиты и преобразования информации.

В данном случае одним из важных направлений в системе криптографических методов активной защиты информации является метод криптоанализа, алгоритмы и средства которого дают возможность оценить криптостойкость системы активной защиты передаваемой информации $T_{i,max}[I(max)]$ – в системе телекоммуникаций, и который способствует проверке ее функциональности в различных условиях.

Учитывая вышеизложенное, криптостойкость защищаемой информации от несанкционированного доступа на основе предложенного алгоритма и средства криптографического метода активной защиты передаваемой информации оптимизируется следующими целевыми функциями:

$$T_{i,max}[I(X)] = |T_{i+1}[I(X)] - T_i[I(X)]| \rightarrow \max, \quad i = \overline{1, n} \quad (5)$$

Выражения (5) характеризует криптостойкость предложенного алгоритма и средства криптографического метода активной защиты передаваемой информации и определяют надежность защиты сообщения в системе телекоммуникаций.

Оценка криптостойкости методов активной защиты информации

На основе предложенного метода предполагаем, что источник информации генерирует сообщение m из общего множества X , представляющего собой последовательность символов над полем

$$GF^{(N)}(q), N = (N_k - N_k^{mn}) > 0,$$

где N_k – общее число каналов в системе телекоммуникаций; N_k^{mn} – число каналов, доступных нелегальному пользователю (злоумышленнику).

Считаем, что в данной подсистеме источник генерирует сеансовый ключ ke и kd из множества K представляющего собой последовательность символов над полем $GF^{(N)}(q)$.

Из рис. 1 следует, что в исследуемой системе телекоммуникаций имеется несколько подсистем, которые состоят из N_k – каналов связи, и предполагается, что нелегальный пользователь во время одной из передач сообщений имеет доступ к одной из этих подсистем – каналам связи N_k^{mn} .

Допустим, система телекоммуникаций может быть подвержена вмешательству злоумышленников в ее работу. Поставленные условия могут быть выполнены для данного варианта злоумышленника и канала с несколькими получателями.

Учитывая информационные характеристики источника сообщения X и алгоритма предлагаемого метода защиты передаваемой информации, можно обеспечить требуемый уровень защиты сообщений при выполнении следующего условия:

$$H(K) \geq H(A) \cdot \frac{N_k^{mn}}{N_k - N_k^{mn}}, N_k > N_k^{mn}, \quad (6)$$

где $H(K)$ – энтропия открытого и закрытого ключа, является длиной ключа, измеренной в битах; $H(A)$ – энтропия источника информации определяется следующим образом:

$$H(A) = - \sum_{i=1}^N P(a_i) \log P(a_i), \quad (7)$$

где N – объем алфавита источника, который вырабатывает сообщения в виде последовательности дискретных элементов, выбираемых из конечного множества

$$A = (a_1, a_2, \dots, a_n) = \{a_i\}^N$$

Из (6) следует, что выполнение условия, обеспечивающего требуемый уровень активной защиты передаваемой информации по каналам связи, определяет безусловную криптостойкость системы телекоммуникации.

Выводы

В результате исследования предложен криптографический метод активной защиты передаваемой информации от несанкционированного

доступа в систему телекоммуникаций, учитывающий способ шифрования и сетевого кодирования сообщения. Разработана структурно-функциональная схема активной защиты информации с открытым ключом.

Получены условия на основе информационных показателей открытого ключа и источника сообщения, выполнение которых определяет безусловную криптостойкость системы активной защиты информации.

Список литературы

1. Заседателева П.С., Маслов О.Н. Моделирование процесса разработки и исследования системы активной защиты информации // Инфокоммуникационные технологии, Том 12, № 3. ПГУТИ. 2014. – с.51-55.
2. Рябко Б. Я., Фионов А. Н. Криптографические методы защиты информации М.: Горячая линия–Телеком, 2005. 229 с.
3. Ибрагимов Б.Г., Ибрагимов Г.Г. Исследование криптографического метода и средств защиты передаваемой информации в системах телекоммуникаций // Ученые записки, АзТУ. №1, 2018. – С.40 – 45.
4. Шаврин С.С. Защита информации в многоканальных телекоммуникационных системах. Ч.1.,М.: МТУСИ. 2002. – 62 с.
5. Костин Д. В., Шелухин О. И. Сравнительный анализ алгоритмов машинного обучения для проведения классификации сетевого зашифрованного трафика // Т- Comm: Телекоммуникации и транспорт. № 9, 2016. – С. 46-52
6. Криптографическое кодирование/Под ред. В.Н. Рудницкого, В.Я. Мильевича.– Харьков : Изд-во «Щедрая усадьба плюс», 2014. – 240 с.
7. Schneyer B. Applied Cryptography: Protocols, algorithms and source code. C.J.Willer&Sons, inc. New York, 1994. – 618 p.
8. Blake I., Seroussi G., Smart N. Elliptic Curves in Cryptography. Cambridge University Press, 2002. 204 p.

СЕКЦИЯ 5. МОДЕЛИРОВАНИЕ СИСТЕМ

УДК 518.61

С. Д. Алгазин,

д.ф. -м.н., вед. н. сотр.

Федеральное государственное бюджетное учреждение науки Институт проблем механики им. А. Ю. Ишлинского РАН,
119526, Москва, проспект Вернадского 101-1, Россия.
e-mail: algszinsd@mail.ru

МНОГОСЛОЙНЫЙ, НЕЯВНЫЙ АЛГОРИТМ ДЛЯ НЕОДНОРОДНОГО ДВУМЕРНОГО УРАВНЕНИЯ ТЕПЛОПРОВОДНОСТИ

Аннотация: рассматривается двумерное уравнение теплопроводности в плоской области с неоднородными краевыми условиями и правой частью обеспечивающими гладкость решения. Для приближенного нахождения решения этого уравнения построен численный алгоритм без насыщения. Указан эффективный способ решения соответствующей дискретной задачи.

Ключевые слова: уравнение теплопроводности, краевое условие Дирихле, формула Бабенко, численный пример.

S.D. Algazin,

Doctor of Physics and Mathematics, Leading Researcher

Federal state budgetary institution of science Institute of mechanics problems. A. Ju RAS,
119526, Moscow, Prospekt Vernadskogo 101-1, Russia.
e-mail: algszinsd@mail.ru

MULTI-LAYER, IMPLICIT ALGORITHM FOR INHOMOGENEOUS TWO-DIMENSIONAL HEAT EQUATION

Abstract: A two-dimensional heat conduction equation is considered in a flat region with inhomogeneous boundary conditions and the right-hand side ensuring smoothness of the solution. For an approximate finding of the solution of this equation, a numerical algorithm without saturation is constructed. The effective method of solving the corresponding discrete problem is indicated.

Keywords: heat equation, Dirichlet boundary condition, equation of Babenko, numerical example.

Рассматривается двумерное уравнение теплопроводности в плоской области с неоднородными краевыми условиями и правой частью обеспечивающими гладкость решения. Для приближенного нахождения решения этого уравнения построен численный алгоритм без насыщения. Указан эффективный способ решения соответствующей дискретной задачи.

Введение.

В [1-9] описаны дискретизации уравнения теплопроводности с однородным краевым условиями. В этой методике требование однородности краевых условий существенно, поскольку идея дискретизации состоит в построении конечномерной задачи со спектральными свойствами, аналогичными дифференциальной задаче. Вместе с тем практические задачи часто приводят к уравнению теплопроводности с неоднородными краевыми условиями. В качестве примера можно привести задачу о термо-упругих деформациях рельса, вызванного трением от колес проходящего поезда.

Ниже построен численный алгоритм без насыщения для уравнения теплопроводности в плоской двумерной области с неоднородными краевыми условиями. Особенности численных алгоритмов без насыщения является то, что они в отличие от разностных методов приводят к дискретной задаче с полностью заполненной матрицей. Эффективное решение которой и представляет основную трудность.

I. Постановка задачи и дискретизация.

Рассматривается уравнение теплопроводности с краевым условием Дирихле

$$\frac{\partial u(\zeta, t)}{\partial t} = \Delta u(\zeta, t) + f(\zeta, t), \quad (1)$$

$$u|_{\partial G} = g \quad (2)$$

$$u|_{t=0} = 0. \quad (3)$$

Решение ищется в круговом цилиндре единичного радиуса, т.е. в области $G = \{\rho \leq 1; 0 \leq t \leq 1\}$, где $\zeta = \rho \cdot \exp(i\varphi)$, (ρ, φ) – полярные координаты. Для примера рассмотрено краевое условие Дирихле. Проведённые ниже рассуждения без труда обобщаются на другие типы краевых условий.

Будем обозначать: $g = g(t, \varphi)$ – граничное условие на границе круга.

$$u(\zeta, z) = - \int_{|\xi| \leq 1} K(\zeta, \xi) (f(\xi, t) - \frac{\partial u(\xi, t)}{\partial t}) d\xi + \int_0^{2\pi} K_0(\zeta, \varphi) g(t, \varphi) d\vartheta. \quad (4)$$

Здесь

$$K(\zeta, \xi) = -\frac{1}{2\pi} \ln|(1 - \zeta\bar{\xi})/(\zeta - \xi)|, \quad \zeta = \rho \exp(i\varphi), \quad \xi = r \exp(i\theta)$$

- функция Грина оператора Лапласа с краевым условием Дирихле (однородным);

$$K_0(\zeta, \vartheta) = \frac{1}{2\pi} \frac{1 - \rho^2}{1 + \rho^2 - 2\rho \cos(\vartheta - \varphi)}, \quad \zeta = \rho \exp(i\varphi)$$

- ядро Пуассона

Выберем в круге сетку из m окружностей и $N=2n+1$ точек на каждой окружности. На границе круга также выберем $N=2n+1$ точек. Причём радиус ν -ой окружности $r_\nu = \cos((2\nu-1)\pi/4m)$, $\nu=1,2,\dots,m$.

По окружностям узлы располагаются через равные углы $\vartheta_l = 2\pi l/N$, $l=0,1,\dots,2n$.

Применим для функции $F(\xi, t) = f(\xi, t) - \frac{\partial u(\xi, t)}{\partial t}$ интерполяционную формулу К. И. Бабенко для функции двух переменных в круге [10, стр. 40], а для функции $g(\vartheta, t)$ применим интерполяцию тригонометрическим многочленом (здесь t – фиксировано), тогда из (1.4) получаем приближённое значение $u(\zeta, t)$ для функции $u(\zeta, t)$

$$\tilde{u}(\zeta, t) = \sum_i H_i(\zeta) F(\xi_i, t) + \sum_{j=0}^{2n} H_j^0(\zeta) g(\vartheta_j, t) \quad (5)$$

где

$$H_i(\zeta) = \int_{|\xi| \leq 1} K(\zeta, \xi) l_i(\xi) d\xi, \quad H_j^0 = \frac{2}{N} (0.5 + \sum_{l=1}^n \rho^l \cos l(\varphi - \vartheta_j)), \quad \zeta = \rho \exp(i\varphi).$$

Здесь $l_i(\xi)$, $i=1, 2, \dots, R$, $R = mN$ – фундаментальные функции интерполяционной формулы К. И. Бабенко в круге. Конкретный вид этих функций приведен в [10, стр. 40]. Для дальнейших рассуждений он не важен.

Проведём теперь дискретизацию $\frac{\partial u(\zeta, t)}{\partial t}$ по t . Выберем для $u(\zeta, t)$ интерполяционную формулу по k узлам

По t выберем сетку, состоящую из k узлов:

$$t_\nu = \frac{1}{2}(z_\nu + 1) \quad z_\nu = \cos \chi_\nu, \quad \chi_\nu = \frac{(2\nu-1)\pi}{2k}, \quad \nu = 1, 2, \dots, k,$$

и применим интерполяцию многочленом:

$$q(t) = \sum_{\nu=1}^k \frac{T_m(t) t q_\nu}{m \frac{(-1)^{\nu-1}}{\sin \chi_\nu} t_\nu (z - z_\nu)}, \quad T_m(t) = \cos(m \arccos t). \quad (6)$$

Значения первой производной от $u(x, t)$ по t , входящие в правую часть соотношений (5), получим дифференцированием интерполяционной формулы (6).

Исследование структуры конечномерной задачи.

Перенумеруем узлы в цилиндре сначала по t , а затем по ζ , т.е. быстрее всего меняется индекс i , потом q . Тогда имеем из соотношений (5) в матричной форме

$$u = (I_k \otimes H)f - H \otimes Du + \sum_{j=0}^{2n} H_j^0(\zeta)g(\theta_j, k) \quad (7)$$

где u – вектор столбец, содержащий приближённые значения искомого решения в узлах сетки; f – вектор столбец, содержащий значения соответствующей функции (см. 8) в узлах сетки; знаком \otimes обозначено кронекеровское произведение матриц H и D . Размерность этих векторов $N_g = Rk$ равна числу внутренних узлов сетки. Решив конечномерную задачу (7) получим приближённое значение решения в узлах сетки. В остальных точках цилиндра решение может быть восстановлено по используемым выше интерполяционным формулам.

Таким образом, для решения системы линейных уравнений (I.7) требуется обратить матрицу $I + H \otimes D$ большого размера $N_g \times N_g$.

Обсудим теперь изменения, которые следует внести в методику, чтобы рассмотреть цилиндр с произвольным основанием в виде области G , с гладкой границей ∂G . Пусть $\varphi(z)$, $|z| \leq 1$ конформное отображение круга на область G . Обозначим $Z = \text{diag}(|\varphi'(\zeta_1)|^2, \dots, |\varphi'(\zeta_R)|^2)$. Тогда аналогично получаем, что в формулы этого параграфа вместо H нужно подставить матрицу HZ .

II. Обсуждение методики и численный пример.

Рассмотрим класс областей, получающийся из круга конформным отображением

$$\zeta = \xi(1 + \varepsilon \xi^{2n}), \quad \zeta = \rho \exp(i\varphi), \quad \xi = r \exp(i\theta).$$

Если $u = u(\rho, \varphi, t) = t^2 \rho^2 \cos(\varphi) = t^2 |\zeta| \text{Re} \zeta$, то $f(\rho, \varphi, t) = 2t^2 \rho^2 \cos \varphi - 3t^2 \cos \varphi = 2t |\zeta| \text{Re} \zeta - 3t^2 \text{Re} \zeta / |\zeta|$, $g = g(\rho, \varphi, t) = t^2 \rho^2 \cos(\varphi) = t^2 |\zeta| \text{Re} \zeta$, $|\zeta| = 1$.

1) Единичный круг.

$$K = \quad 2 \quad M = \quad 3 \quad N = \quad 3$$

Норма матрицы дискретной задачи
 BNORM R2^-1 = 1.66959631005056

Точное решение

$$0.67974956743D+00 \quad 0.36427669530D+00 \quad 0.48803823163D-01$$

Приближённое решение

$$0.67825965095D+00 \quad 0.38599513248D+00 \quad 0.28475956684D-01$$

ОТНОСИТЕЛЬНАЯ ПОГРЕШНОСТЬ % 0.85

$$\text{RNORM} = 5.770771892073479E-003$$

$$\text{NORM} = 2.171843718125033E-002$$

2) Эпитрохоида

$$\text{EPS} = 1/6 \quad \text{NP} = \quad 4 \quad K = \quad 2 \quad M = 3 \quad N = \quad 3$$

Норма матрицы дискретной задачи

$$\text{BNORM R2}^-1 = 1.71519000078133$$

Точное решение

0.89130169583D+00 0.39526551139D+00 0.48876849562D-01

Точное решение

0.89130169583D+00 0.39526551139D+00 0.48876849562D-01

Приближённое решение

0.90023831744D+00 0.56147540729D+00 0.61891747216D-01

ОТНОСИТЕЛЬНАЯ ПОГРЕШНОСТЬ % 1.00

RNORM = 8.936621610843742E-003

NORM = 0.166209895899030

Выведено решение на действительной оси на втором слое по времени в узлах сетки по радиусу.

Библиография

- 1) Алгазин С. Д. Неявный численный алгоритм без насыщения для решения двумерного уравнения теплопроводности // *Proceedings of the International Conference "Education, science and economics at universities. Integration to international education area"* Plock, Poland, 9 -14 September 2008.
- 2) Алгазин С. Д. Non-saturating numerical algorithm for solution of heat conduction equation // *Proceeding of the International Conference "Numerical geometry, grid generation and high performance computing (NUMGRID2008" and the International Workshop VORONOI-2008, honoring 140th anniversary of G. F. Voronoi A. A. Dorodnicyn Computing Center RAS Moscow, 10-13 June, 2008*
- 3) Алгазин С. Д. Численный алгоритм без насыщения для двумерного уравнения теплопроводности // *Аналитические и численные методы моделирования естественнонаучных и социальных проблем: сборник статей III Международной научно-технической конференции. – Пенза: Приволжский Дом знаний, 2008. – 316 с.*
- 4) Алгазин С. Д. Численные алгоритмы классической матфизики. XXII. Двумерное уравнение теплопроводности. Новая программа. М., 2008. 25 с. (Препр. ИПМех; № 883).
- 5) Алгазин С. Д. Численный алгоритм без насыщения для решения нестационарных задач. *Инженерно-физический журнал. 2009. Т. 82, № 5. С. 950–960.* На англ. языке: *Numerical algorithm without saturation for solving nonstationary problems. J. Eng. Physics and Thermophysics. 2009. Vol. 82, No.*
- 6) Рр. 956-Алгазин С. Д. Численный алгоритм без насыщения для решения нестационарных задач // *Труды XV Международного симпозиума «Методы дискретных особенностей в задачах математической физики» (МДОЗМФ-2011). Харьков-Херсон, 2011. – С. 25 – 28: ил. 0. – табл. 0. – Библиогр.: 7 назв. ISBN 978-966-623-764-7*
- 7) Алгазин С. Д. Многослойный, неявный, параллельный алгоритм для уравнения теплопроводности в параллелепипеде // *Математика, информатика, естествознание в экономике и обществе (МИЕСЭКО-2015). Труды Всероссийской научной конференции / Отв. Байков А. Ю.- Москва, МФЮА, 2015–145 с. Т.1, ISBN 978-5-94811-203-9, с. 26-29.*
- 8) Алгазин С. Д. Многослойный, неявный, параллельный алгоритм для уравнения теплопроводности в параллелепипеде // *Инженерно-физический журнал. Минск, 2015, т. 88, № 5, стр. 1096-1099.* На англ. языке:

Multilayer, Implicit, Parallel Algorithm for the Equation of Heat Conduction in a Parallelepiped. J. Eng. Physics and Thermophysics. September 2015, Volume 88, Issue 5, pp 1131–1134.

- 9) Сергей Д. Алгазин. Численные алгоритмы без насыщения в классических задачах математической физики. – М: «Агентство Интеллектуальной собственности на Транспорте», Издание 3-е, переработанное и дополненное. 2016, 390 с.
- 10) Бабенко К. И. Основы численного анализа. М.: Наука, 1986, с.744.

УДК 624.07

Альшакова Е.Л.,

к.т.н., доцент, Финансовый университет при Правительстве Российской Федерации, Колледж информатики и программирования, г. Москва, Российская Федерация, alshakova_el@mail.ru, fa.ru

Альшакова Е.А.,

Юго-Западный государственный университет, г. Курск, Российская Федерация, katya.alshakova@mail.ru, swsu.ru

ПРИМЕНЕНИЕ ИНФОРМАЦИОННЫХ ТЕХНОЛОГИЙ ДЛЯ ИССЛЕДОВАНИЯ КОНСТРУКТИВНЫХ И ОБЪЕМНО-ПЛАНИРОВОЧНЫХ РЕШЕНИЙ ЖИЛЫХ ЗДАНИЙ

Аннотация: в работе с использованием программных продуктов информационного моделирования и расчета зданий выполнена оценка конструктивной безопасности многоэтажных жилых зданий нового поколения из железобетонных панельно-рамных элементов индустриального производства. В расчете использовалась конечно-элементная модель, построенная для всего каркаса здания, при применении специального программного комплекса.

Расчет по несущей способности, проведенный на максимальные нагрузки и воздействия, возможные в период эксплуатации, использован для подбора размеров сечения элементов и диаметра арматуры.

Созданы информационные модели многоэтажных жилых зданий нового поколения из железобетонных панельно-рамных элементов индустриального производства с использованием специального программного комплекса.

Созданы расчетные схемы многоэтажных жилых зданий нового поколения из железобетонных панельно-рамных элементов индустриального производства.

Ключевые слова: расчет зданий, оценка конструктивной безопасности, информационные модели, несущая способность, расчетная схема.

Alshakova E.L.,

Candidate of Technical Sciences, Associate Professor, Financial University under the Government of the Russian Federation, College of Computer Science and Programming, Moscow, Russian Federation, alshakova_el@mail.ru, fa.ru

Alshakova E.A.,

Southwestern State University, Kursk, Russian Federation, katya.alshakova@mail.ru, swsu.ru

APPLICATION OF INFORMATION TECHNOLOGIES FOR THE INVESTIGATION OF CONSTRUCTIVE AND SURROUND-PLANNING SOLUTIONS OF RESIDENTIAL BUILDINGS

Abstract: *In the work with the use of software products of information modeling and calculation of buildings, the constructive safety of multi-storey residential buildings of new generation from reinforced concrete panel-frame elements of industrial production was evaluated. In the calculation, a finite element model built for the entire building frame was used when using a special software package.*

The calculation of the bearing capacity, carried out on the maximum loads and impacts, possible during the operation period, was used to select the cross sectional dimensions of the elements and the diameter of the reinforcement.

Created information models of multi-storey residential buildings of a new generation of reinforced concrete panel-frame elements of industrial production using a special software package.

Created design schemes of high-rise residential buildings of a new generation of reinforced concrete panel-frame elements of industrial production.

Keywords: *calculation of buildings, assessment of constructive safety, information models, bearing capacity, design scheme.*

Введение

При выполнении выпускной квалификационной работы магистра направления подготовки Строительство решаются задачи исследования, анализа моделей, конструктивных систем, объемно-планировочных решений зданий, обеспечения определенного уровня конструктивной безопасности зданий при эксплуатационных и возможных запроектных воздействиях. В таких задачах применяются различные методы моделирования и расчета зданий.

Применение информационных технологий, цифрового компьютерного моделирования в данной области и, следовательно, в учебном процессе направления подготовки Строительство, является актуальным направлением, особенно при создании сценариев возможных запроектных воздействий, которые приводят к прогрессирующему разрушению здания [1, 2, 3, 4]. В данном случае возникает необходимость работать с большими объемами информации. Здание – это сложная система различного вида элементов и связей, которые при выключении из системы принимают, передают и перераспределяют внешние воздействия. Современные компьютеры и программы расчета имеют возможность производить необходимые объемы вычислений, в отличие от человека, работая по заданной программе, не пропуская возможные варианты воздействий.

Возникает актуальная задача разработки компьютерных моделей зданий [5], расчетных схем и расчетных моделей, объективно отражающих реальные процессы, возникающие в элементах конструкции и в здании в целом, симулирующих реальные объекты строительства, их свойства и поведение.

Кроме того, программы компьютерного проектирования позволяют автоматизировать выполнение рутинных задач: на основе созданной компьютерной трехмерной модели выполняют расчеты на прочность конструкций зданий, их элементов, формируют рабочую документацию проекта, составляют сметы, следят за графиком работ на строительной площадке и т.д.

Применение информационных технологий для расчета и проектирования железобетонных зданий

В ходе выполнения выпускной квалификационной работы маги-

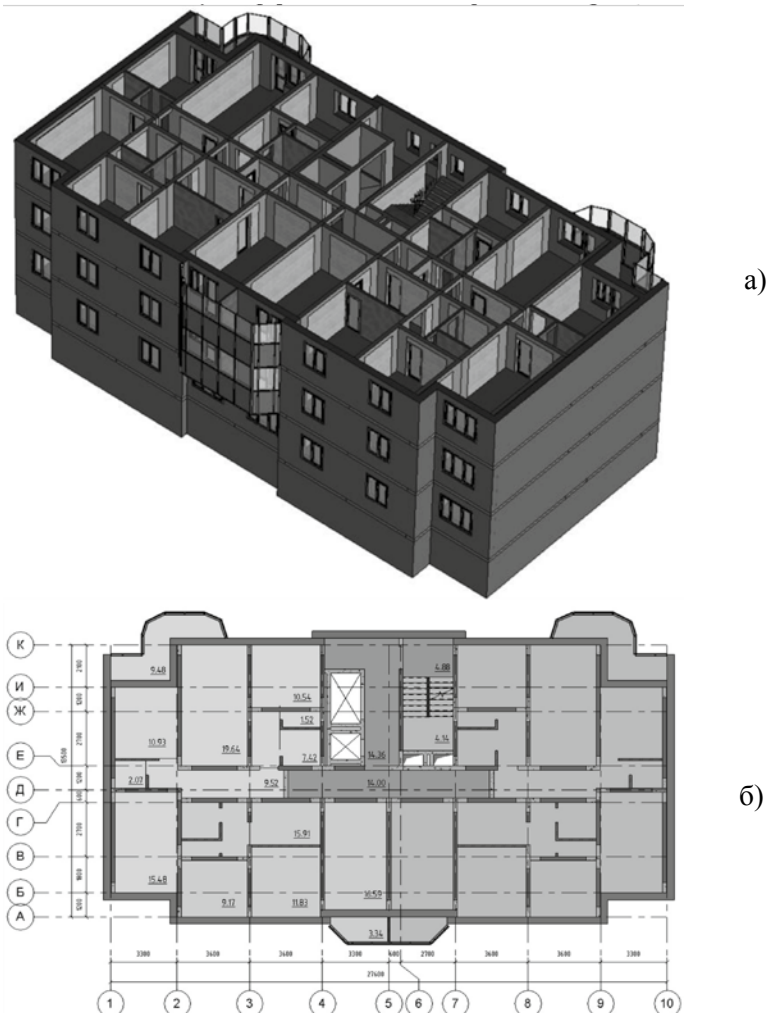


Рис. 1. Информационная модель здания: а – модель каркаса здания в изометрии; б – план типового этажа

стра направления подготовки Строительство был проведен анализ конструктивных и объемно-планировочных решений жилых зданий нового поколения из железобетонных панельно-рамных элементов индустриального производства [6] на основе компьютерного моделирования и исследования информационных моделей, содержащих всю необходимую информацию об объекте строительства (рис. 1).

Индустриальное строительство, относящееся к области исследования данной работы, обладает рядом достоинств: быстрое возведение зданий, достаточно высокое качество конструкций, всесезонность строительства, низкая себестоимость [7]. Развивается направление крупнопанельного домостроения, использующее комбинированные конструктивные системы, обеспечивающие свободные объемно-планировочные решения здания, снижение материалоемкости, повышение энергоэффективности и экологичности здания [8].

Разработка методики выполнения расчетов конструкций зданий на проектные и возможные запроектные воздействия с использованием программных продуктов Autodesk Revit и Autodesk Robot Structural Analysis, реализующих технологию информационного моделирования зданий, является актуальной задачей, которая решается в данной работе [9, 10]. Предложена методика создания расчетной схемы конструкций каркаса здания из панельно-рамных элементов индустриального производства, используемая для расчета напряженно-деформированного состояния и устойчивости конструктивной системы (рис. 2).

Расчет здания из панельно-рамных элементов индустриального производства выполняется в два этапа: первый этап – расчет напряженно-деформированного состояния и устойчивости конструктивной системы здания в целом; второй этап – конструктивный расчет элементов системы здания. Исследования в данной работе осуществлялись в

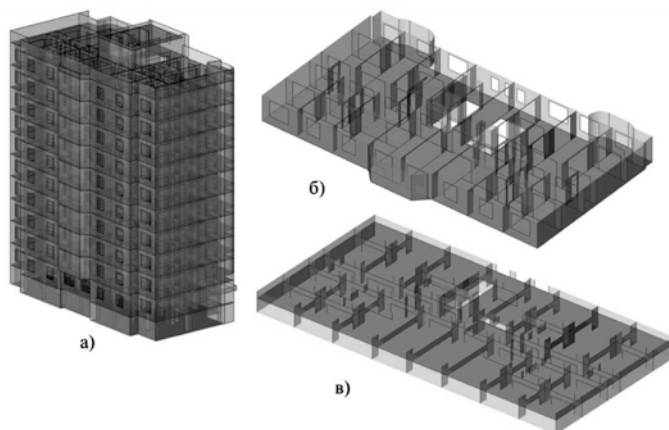


Рис. 2. Аналитическая модель здания: а – модель в изометрии; б – план типового этажа в изометрии; в – план первого этажа в изометрии

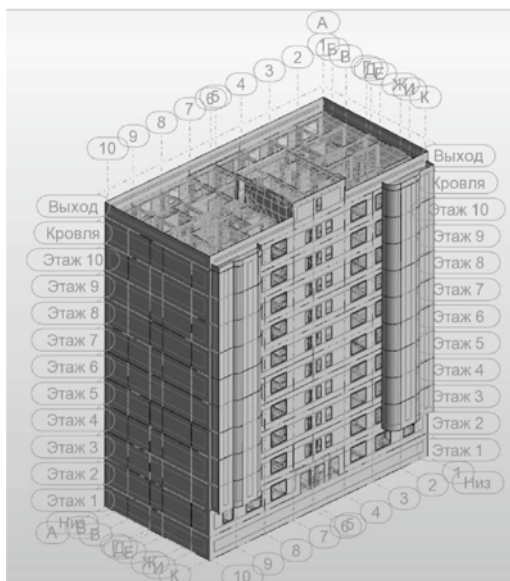


Рис. 3. Расчетная схема – конечно-элементная модель

следующих направлениях:

- назначить расчетную схему;
- определить постоянные и временные нагрузки;
- выполнить статический расчет с использованием специализированных программных средств;
- выполнить расчет прочности панели-рамы.

Конструктивная безопасность системы определялась путем расчета каркаса здания с использованием метода конечных элементов в программе Autodesk Robot Structural Analysis (рис. 3). Задавались постоянные и временные нагрузки, выполнялся статический расчет конструкции по предельным состояниям. Проанализированы результаты расчета. Кроме того, необходимо провести расчет каркаса здания на запроектные воздействия, вызванные выключением из расчетной схемы некоторых несущих элементов при возможных ЧС.

В программе расчета создано нагружение – Ветер 1-10, действующие по направлению цифровых осей для наветренной стороны, как временное кратковременное с коэффициентом надежности 1,4. В качестве элементов, к которым приложена ветровая нагрузка выбраны фасадные элементы – самонесущие наружные стены и несущее наружное ограждение. Аналогично создана нагрузка с подветренной стороны (в осях 10-1) для этого же нагружения.

Создано нагружение Ветер А-К, как временное кратковременное с коэффициентом надежности по нагрузке 1,4. Так же, как и в предыдущем случае, нагрузка создана для наветренной стороны. Аналогично

создана нагрузка с подветренной стороны (в осях К-А) для этого же нагружения.

Таблица нагрузок (с именами нагружений) представлена на рисунке 4.

Нагружение	Тип нагружения	Спецификация								
1 Собственный вес	собственный вес	1до42 50до5	Все констр	-Z	Коэффициент	MEMO				
1 Собственный вес	(К3) равномерно распрд	3 1013	RX=0,0	PY=0,0	PZ=-1.13	общий	не запроекти	абсолютно	Ограничения	MEMO:
2 Эксплуатация	(К3) равномерно распрд	102 106 194	RX=0,0	PY=0,0	PZ=-1.50	общий	не запроекти	абсолютно	Ограничения	MEMO:
3 Перегородки	(К3) равномерно распрд	102 106 194	RX=0,0	PY=0,0	PZ=-1.50	общий	не запроекти	абсолютно	Ограничения	MEMO:
4 Снег	(К3) равномерно распрд	3 970 971 10	RX=0,0	PY=0,0	PZ=-1.50	общий	не запроекти	абсолютно	Ограничения	MEMO:
5 Лоджии	(К3) равномерно распрд	968 969 1019	RX=0,0	PY=0,0	PZ=-2.00	общий	не запроекти	абсолютно	Ограничения	MEMO:
6 Ветровая 1-10	(К3) плоский	2 4др6 60 15	RX1=0,0	PY1=0,40	PZ1=0,0	RX2=0,0	PY2=0,40	PZ2=0,0	RX3=0,0	PY3=0,38
6 Ветровая 1-10	(К3) плоский	7 10 50 62 63	RX1=0,0	PY1=0,25	PZ1=0,0	RX2=0,0	PY2=0,25	PZ2=0,0	RX3=0,0	PY3=0,24
7 Ветровая А-К	(К3) плоский	11 95 974до9	RX1=-0,50	PY1=0,0	PZ1=0,0	RX2=-0,25	PY2=0,0	PZ2=0,0	RX3=-0,24	PY3=0,0
7 Ветровая А-К	(К3) плоский	1 51 64 973д	RX1=-0,25	PY1=0,0	PZ1=0,0	RX2=-0,50	PY2=0,0	PZ2=0,0	RX3=-0,47	PY3=0,0

Рис. 4. Таблица нагрузок

В программе Autodesk Robot Structural Analysis для определения нагрузок установлены нормы проектирования СП 20.13330.2016 «Нагрузки и воздействия» [11], в том числе для сочетаний нагрузок (рис. 5).

Nature		Subnature	γmax	γmin	γs	γa	Ψ _{0,1}	Ψ _{0,2}	Ψ _{0,3}	Ψ _{0,n}	Ψ ₁	Ψ _{2,1}	Ψ _{2,n}	Ψ _K	ξ _i	ξ _c
1	Dead	пост_1_0	1	1	1	1									0,9	0,9
2	Dead	пост_1_05	1,05	1,05	1	1									0,9	0,9
3	Dead	пост_1_1	1,1	1,1	1	1									0,9	0,9
4	Dead	пост_1_15	1,15	1,15	1	1									0,9	0,9
5	Dead	пост_1_2	1,2	1,2	1	1									0,9	0,9
6	Dead	пост_1_3	1,3	1,3	1	1									0,9	0,9
7	Live	1 врем_длит_1			1		0,95				0,95	0,95	0,8	0,8	1	
8	Live	1 врем_длит_1,1			1		0,95				0,95	0,95	0,8	0,8	1	
9	Live	1 врем_длит_1,2			1		0,95				0,95	0,95	0,8	0,8	1	
10	Live	1 врем_длит_1,3			1		0,95				0,95	0,95	0,8	0,8	1	
11	Live	1 врем_длит_1,4			1		0,95				0,95	0,95	0,8	0,8	1	
12	Live	1 врем_кратк_1			1			0,9	0,7		0,9	0,8	0,8	0,5	1	
13	Live	1 врем_кратк_1,1			1			0,9	0,7		0,9	0,8	0,8	0,5	1	
14	Live	1 врем_кратк_1,2			1			0,9	0,7		0,9	0,8	0,8	0,5	1	

Combination type	User-defined type	Loads		
		Dead	Live	Acc
1	ULS USR ShortTerm <=2	(2) $\sum_{i \geq 1} G_i \cdot \gamma_{max}^{(i)}$	(73) [49] + [71]	(0)
2	ULS USR ShortTerm >=3	(2) $\sum_{i \geq 1} G_i \cdot \gamma_{max}^{(i)}$	(74) [49] + [72]	(0)
3	SLS USR ShortTerm <=2	(1) $\sum_{i \geq 1} G_i \cdot \gamma_s^{(i)}$	(77) [53] + [75]	(0)

Рис. 5. Сочетания нагружений по нормативам

Программа Robot Structural Analysis создает возможные сочетания на основе созданных нагружений, которые отображаются в Таблице сочетаний. По умолчанию, создаются сочетания для предельных состояний первой группы (ПС1) и второй (ПС2). В строках таблицы указываются номера нагружений умноженные на коэффициенты надежности и с учетом коэффициентов сочетаний.

Проектирование элементов конструкции здания

После вычисления усилий, действующих в конструкции, переходят к проектированию отдельных элементов конструкции. В данной работе

для многоэтажного здания из панельно-рамных элементов выполняется проектирование железобетонной панели-рамы – несущего элемента конструкции, железобетонной панельной стены и плиты перекрытия.

Перед проведением расчетов по проектированию элементов конструкции необходимо задать (уточнить) используемые нормы проектирования железобетонных конструкций. В работе заданы действующие в настоящее время СП 63.13330.2012 [12].

На стадии проектирования элементов конструкции выбираются железобетонные элементы или вся модель для переноса в модули проектирования железобетонных конструкций. Расчетная программа Autodesk Robot Structural Analysis анализирует выбранные элементы и на основе типа железобетонного элемента и его свойств, таких как геометрические параметры, способ опирания, тип армирования оболочки: сжатие и растяжение, изгиб, открывается (с выбранным элементом) соответствующий железобетонный модуль. Железобетонные элементы относятся к следующим модулям: колонны; ленточный фундамент; балки; балки-стенки; стены; плиты; фундамент.

Проектирование балки с повышенной строительной высотой – балки-стенки предусмотрено в программе Autodesk Robot Structural Analysis.

Выбранная панель-рама, расположенная на первом этаже, перенесена программой в модуль стен (рис. 6). Панель-рама имеет опирание на плиту перекрытия этажа. Тип арматуры оболочки, используемой в панели-раме, создан под сжатие и растяжение, что соответствует условиям перехода в модуль расчета железобетонной балки-стенки, тем не менее программа выбрала для панели-рамы модуль стены.

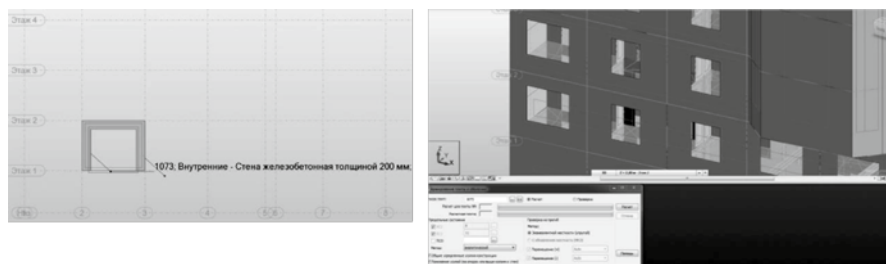


Рис. 6. Выбор проектируемого элемента в конструкции здания

Для стен и плит перед фактическим армированием – раскладкой арматуры, необходимо выполнить расчет теоретической площади арматуры (рис.7).

В программе Autodesk Robot Structural Analysis выполнена проверка на прогиб и на трещинообразование при расчете теоретической площади арматуры.

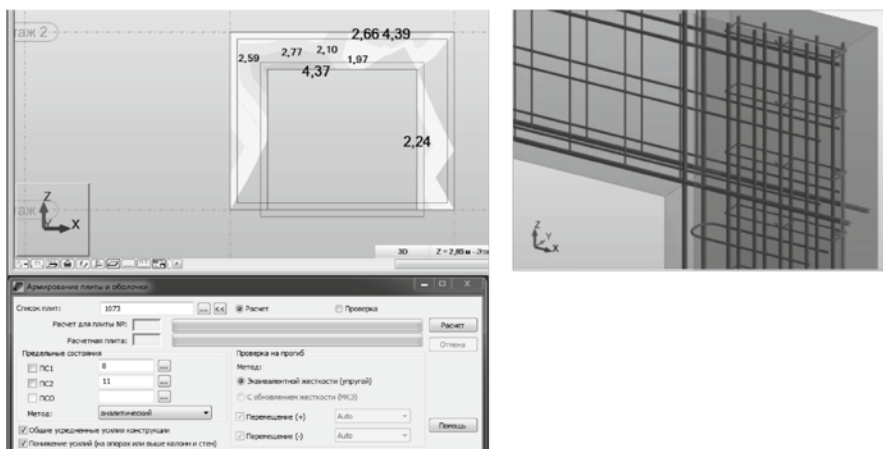


Рис. 7. Проектирование железобетонного панельно-рамного элемента

На чертежах армирования панели-рамы необходимо предусмотреть выпуски арматуры, которые служат для соединения бетоном замоноличивания с плитами перекрытия [6].

Заключение

С использованием программных продуктов информационного моделирования и расчета зданий выполнена оценка конструктивной безопасности многоэтажных жилых зданий нового поколения из железобетонных панельно-рамных элементов индустриального производства: проектный расчет каркаса здания по предельным состояниям на постоянные и временные нагрузки, на возможное запроектное воздействие, вызванное внезапным выключением из расчетной схемы вертикального несущего элемента. В расчете использовалась конечно-элементная модель, построенная для всего каркаса здания, при применении специального программного комплекса.

Расчет по несущей способности, проведенный на максимальные нагрузки и воздействия, возможные в период эксплуатации, использован для подбора размеров сечения элементов и диаметра арматуры.

Созданы информационные модели многоэтажных жилых зданий нового поколения из железобетонных панельно-рамных элементов индустриального производства с использованием специального программного комплекса.

Созданы расчетные схемы многоэтажных жилых зданий нового поколения из железобетонных панельно-рамных элементов индустриального производства.

Список литературы

1. Альшакова Е.Л. Создание моделей объемно-планировочных решений жилых и общественных зданий в технологии информационного моделирова-

- ния [Текст] // Гагаринские чтения – 2016: XLII Международная молодежная научная конференция: Сборник тезисов докладов: В 4 т. М.: Московский авиационный институт (национальный исследовательский университет), 2016. – Т. 3. – С. 53 – 54.
2. Альшакова Е.Л. Информационные технологии в образовательном процессе – модели конструктивных решений жилых зданий из железобетонных панельно-рамных элементов [Текст] // Преподавание информационных технологий в Российской Федерации: материалы Пятнадцатой открытой Всероссийской конференции, Архангельск: Северный (Арктический) федеральный университет, 2017, – С. 20 – 24.
 3. Юрин В.Н. Компьютерный инжиниринг в инженерном образовании: эволюция [Текст] // Труды XXII Международной научно-технической конференции «Информационные средства и технологии». В 3-х томах. М.: Издательский дом МЭИ, – 2014. – Т. 2. – С. 102 – 108.
 4. Альшакова Е.Л. Содержание дисциплины «Начертательная геометрия и инженерная графика» специальности «Строительство уникальных зданий и сооружений» // Геометрия и графика. – 2013. – Т. 1. – № 2. – С. 24 – 26.
 5. Альшакова Е.Л. Применение 3D моделирования на стадии эскизного проектирования объектов строительства и архитектуры // Информационные технологии в образовании XXI века. Сборник научных трудов Международной научно-практической конференции. – М.: НИЯУ МИФИ. – 2015. – С. 150 – 156.
 6. Ключева Н.В., Колчунов В.И., Рыпаков Д.А., Бухтиярова А.С. Жилые и общественные здания из железобетонных панельно-рамных элементов индустриального производства [Текст] // Жилищное строительство. 2015. № 5. С. 69 – 75.
 7. Николаев С.В. Архитектурно-градостроительная система панельно-каркасного домостроения [Текст] // Жилищное строительство. 2016. № 6. С. 15 – 25.
 8. Патент РФ 2506385. Здание из панельных элементов / Ильичев В.А., Колчунов В.И., Ключева Н.В., Бухтиярова А.С. Заявл. 01.08.2012. Оpubл. 10.02.2014. Бюл. №4.
 9. Альшакова Е.Л. Особенности проектирования, расчета и анализа железобетонных панельно-рамных элементов индустриального производства [Текст] // Проектирование и строительство: сборник тезисов докладов I Международной (III региональной) научно-практической конференции молодых ученых, аспирантов, магистрантов и бакалавров (Курск, 17–18 марта 2016 г.) / ред. кол.: Н.В. Федорова (отв. ред.) [и др.]; Юго-Зап. гос. ун-т. – Курск, 2017. – С. 17 – 20.
 10. Альшакова Е.Л. САПР в исследовании моделей объемно-планировочных решений жилых зданий нового поколения из железобетонных панельно-рамных элементов [Текст] // Гагаринские чтения – 2017: XLIII Международная молодежная научная конференция: Сборник тезисов докладов: М.; Московский авиационный институт (национальный исследовательский университет), 2017, – С. 668 – 669.
 11. СП 20.13330.2016. Нагрузки и воздействия. Актуализированная редакция СНиП 2.01.07-85* [Текст]. – М.: Минстрой России, 2016, – 104 с.
 12. СП 63.13330.2012. Бетонные и железобетонные конструкции. Основные положения [Текст]. – М.: Аналитик, 2012, – 158 с.

УДК 004.942.

Крутиков Д.И.,

студент, ФГАОУ ВПО «Волгоградский государственный университет», г. Волгоград,
Россия, krutikov1996@yandex.ru, volsu.ru

Бабенко А.А.,

к.п.н, доцент, ФГАОУ ВПО «Волгоградский государственный университет»,
г. Волгоград, Россия, ba_benko@mail.ru, volsu.ru

ВЕРОЯТНОСТНАЯ МОДЕЛЬ ОЦЕНКИ НАДЁЖНОСТИ ИНФОРМАЦИОННЫХ СИСТЕМ

Аннотация: в статье предлагается объединять однотипные компоненты аппаратно-программных комплексов в группы и оценивать надежность группы в целом. Приводится порядок такого рода расчета надежности.

Ключевые слова: надежность, среднее время наработки на отказ, среднее время восстановления, коэффициент готовности.

Krutikov D.I.,

student, Volgograd State University, Volgograd, Russia, krutikov1996@yandex.ru, volsu.ru

Babenko A.A.,

Ph.D., Associate Professor, Volgograd State University, Volgograd, Russia, ba_benko@mail.ru, volsu.ru

A PROBABLE MODEL FOR ASSESSING THE RELIABILITY OF INFORMATION SYSTEMS

Abstract: The article proposes to combine the same type of components of hardware and software systems into groups and evaluate the reliability of the group as a whole. An order of this kind of reliability calculation is given.

Keywords: reliability, average time between failures, average recovery time, availability factor.

Введение

Основными показателями качества информационных систем (ИС) являются надежность, достоверность и безопасность. В последние десятилетия проблема повышения надежности информационных систем не только не ослабела, но и стала еще более значимой. Это обусловлено бурным ростом в области информационных технологий и вычислительной техники, который привел к массовой информатизации общества. От надежности информационных систем зависят достоверность и безопасность обрабатываемых в них информации. Современные информационные системы имеют сложную структуру. Они различаются по назначению, принципу построения и взаимодействию составляющих их компонентов. Все это накладывает определённые сложности при оценивании надежности ИС.

Согласно [1], «надёжность – свойство объекта сохранять во времени в установленных пределах значения всех параметров, характеризующих способность выполнять требуемые функции в заданных режимах и условиях применения, технического обслуживания, хранения

и транспортирования» [1]. Надежность – составное свойство. В него могут входить сохраняемость, долговечность, безотказность, ремонтно-пригодность или какая-либо комбинация данных свойств [1].

Показатели надежности ИС

Надежность ИС можно оценить количественно через показатели надежности. Они отражают количественные оценки одного или нескольких свойств ИС, определяющих надежность.

Надежность ИС можно выразить через интенсивность отказов, вероятность безотказной работы, среднее время восстановления, среднюю наработку на отказ, и коэффициент готовности, коэффициент оперативной готовности [2].

1. «Интенсивность отказов – соотношение числа отказов компонентов системы в единицу времени к среднему числу наблюдений, работающих без отказов в данный отрезок времени при условии, что отказавшие наблюдения не восстанавливаются и не заменяются исправными» [2].

2. «Среднее время наработки на отказ – отношение суммарной наработки восстанавливаемого объекта к математическому ожиданию числа его отказов в течение этой наработки» [2].

3. «Среднее время восстановления – показатель, характеризующий ремонтпригодность технического устройства; равно математическому ожиданию времени восстановления его работоспособности» [2].

4. «Коэффициент готовности – вероятность того, что компонент ИС окажется в работоспособном состоянии в произвольный момент времени, кроме планируемых периодов, в течение которых применение объекта по назначению не предусматривается» [2].

5. «Вероятность безотказной работы – это вероятность того, что в пределах заданной наработки или заданном интервале времени отказ компонента ИС не возникает» [2].

6. «Коэффициент оперативной готовности – это вероятность того, что компонент ИС окажется в работоспособном состоянии в произвольный момент времени, кроме планируемых периодов, в течение которых применение компонента по назначению не предусматривается, и, начиная с этого момента, будет работать безотказно в течение заданного интервала времени» [2].

После того, как определен состав показателей надежности необходимо получить статистические данные о вероятности отказов компонентов ИС во времени. Для этого сначала нужно определить, основные компоненты ИС. Такими компонентами являются

- информация;
- каналы и средства связи (КС);
- маршрутизаторы;
- автоматизированные рабочие места (АРМ) пользователей ИС, в том числе и удалённые;

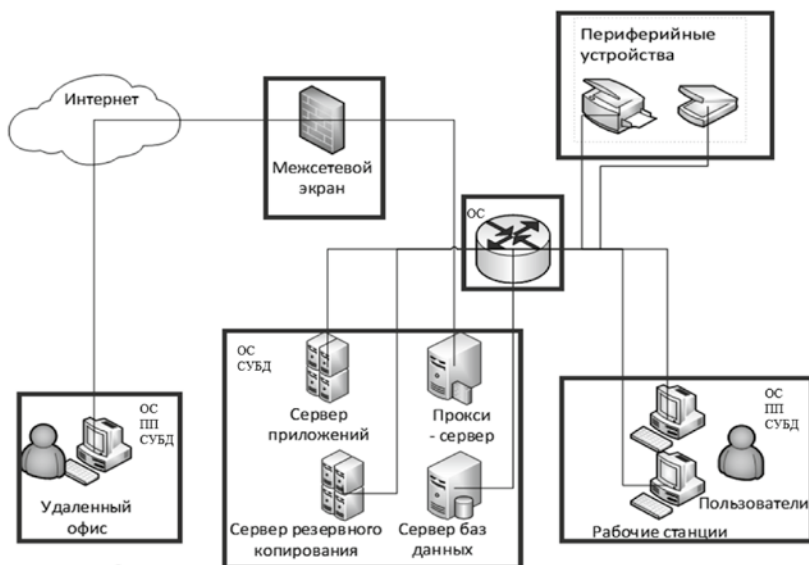


Рисунок 1 – Обобщенная схема типовой архитектуры ИС, отражающая особенности, влияющие на ее надежность

- рабочие станции;
- сервера, которые разделяются на: сервера баз данных, сервера приложений, резервного копирования);
- периферийные устройства;
- пользователи ИС;
- администраторы ИС.

В результате анализа ИС составлена типовая архитектура, которая отражает основные компоненты, влияющие на ее надежность (рисунок 1).

При этом важно обеспечить надежность серверного оборудования, сетевого оборудования, каналов связи, рабочих станций, периферийных устройств, программного обеспечения, поэтому необходимо собрать в первую очередь статистику их отказов.

В зависимости от выполняемых элементами ИС функций можно выделить следующие виды связи:

- связь вида «включает/хранит». При данном виде связи информационный ресурс содержит другой информационный ресурс. Например, при выходе из строя сервера баз данных, нарушится доступность или произойдет уничтожение зависимого ресурса – носителя информации и данных;

- связь вида «обрабатывает/передает». При данном типе связи один ресурс участвует при передаче или обработке другого ресурса. Например, маршрутизатор передает данные от одного узла сети ИС к другому, поэтому при отказе данного элемента, нарушится доступ-

ность узлов сети подключенных через данный маршрутизатор.

На объекты уязвимости воздействуют различные дестабилизирующие факторы, которые являются угрозами:

1) нарушение работы серверного и сетевого оборудования вследствие: физического износа; низкого качества технических средств; перебоев с энергоснабжением; обрушения помещения, в котором расположены компоненты ИС; перегрев оборудования по причине выхода из строя системы кондиционирования; возгорание вследствие неисправности электрической проводки и заземления; затопление помещения с оборудованием в результате прорыва систем водоснабжения и отопления.

2) нарушение работы каналов связи вследствие: возгорание вследствие неисправности электрической проводки и заземления; повреждение линий связи в результате строительных или ремонтных работ вблизи прохождения каналов связи;

3) нарушение работы рабочих станций вследствие: перебоев с энергоснабжением; возгорание вследствие неисправности электрической проводки, заземления или технических средств, составляющих рабочую станцию; нарушение правил противопожарной безопасности пользователями рабочих станций или администраторами; физическое устаревание технических компонентов рабочих станций;

4) нарушение работы программного обеспечения вследствие: алгоритмических ошибок разработки программных средств; некорректная обработка ошибок, возникающих в процессе работы программного обеспечения; использование некачественных программных средств, не обеспечивающих при сбоях сохранения целостности информации.

Математическая модель оценки надежности ИС

На следующем этапе необходимо составить математическую модель, по которой рассчитываются показатели надежности [3]. При составлении математической модели можно опираться на определение каждого из показателей надежности ИС, а также использовать аппарат теории вероятности.

Исходными данными являются:

N_r – количество серверов;

N_o – среднее количество единиц оборудования, приходящегося на один сервер;

m – количество разновидностей оборудования;

n_i – количество оборудования i -ой разновидности, $i = \overline{1, m}$;

λ_i – интенсивность отказов i -ой разновидности, ($i = \overline{1, m}$) (1/год);

τ_{vi} – среднее время восстановления элемента i -ой разновидности, ($i = \overline{1, m}$) (часы);

t – планируемое время безотказной работы ИС (годы).

Далее выполняется расчет показателей надежности по формулам 1-8:

1) рассчитывается количество оборудования ИС:

$$N = N_c + N_o N_c, \quad (1)$$

2) определяется усреднённая интенсивность отказов оборудования $\bar{\lambda}$:

$$\bar{\lambda} = \frac{\sum_{i=1}^m \lambda_i}{m}, \quad (2)$$

где m – число видов оборудования; λ_i – интенсивность отказов i -ого вида оборудования;

3) рассчитывается интенсивность отказов ИС по формуле:

$$\lambda_{ис} = N * \bar{\lambda}, \quad (3)$$

4) рассчитывается коэффициент средней наработки на отказ ($T_{0ис}$) (формула 4), а также рассчитывается вероятность безотказной работы за требуемое время t ($R_{ис}(t)$) (формула 5):

$$T_{0ис} = \frac{1}{\lambda_{ис}}, \quad (4)$$

$$R_{ис}(t) = e^{-\lambda_{ис} * t}, \quad (5)$$

5) рассчитывается коэффициент готовности ИС:

$$A_{ис} = \frac{T_{0ис}}{T_{0ис} + T_в}, \quad (6)$$

где $T_в$ – время восстановления ИС, которое рассчитывается по формуле 7:

$$T_в = \frac{1}{\lambda_{ис}} \sum_{i=1}^m \tau_{ви} * 0,000114 * \lambda_i * n_i * N_c, \quad (7)$$

где τ_0 – время необходимое для старта ремонтных работ; $\tau_{ви}$ – среднее время восстановления i -го элемента; n_i – количество оборудования i -ой разновидности.

6) определяется коэффициент оперативной готовности ИС:

$$K_{ог} = R(t)_{ис} * A_{ис}. \quad (8)$$

Стоит отметить, что для программного обеспечения необходимо считать интенсивность отказов постоянной величиной, т.к. оно физически не стареет.

На последнем этапе необходимо осуществить расчет по формулам, определённым в математической модели.

Заключение

Современные информационные системы являются аппаратно-программными комплексами, надежность которых сложно оценить из-за трудностей сбора статистических данных о вероятности отказов компонентов ИС. Кроме того, программное обеспечение, для различных компонентов отличается, что усложняет задачу оценки надежности. Поэтому необходимо объединять однотипные компоненты в группы и оценивать группу в целом. Именно такой подход предложен в данной статье.

Список литературы

1. ГОСТ Р 27.002–2009. Надежность в технике. Термины и определения. М.: Изд-во стандартов, 2009.
2. Рябинин И. А. Надежность и безопасность структурно-сложных систем: монография. СПб.: Изд-во Санкт-Петербургского ун-та, 2007. 248 с.
3. Воронин А.А., Морозов Б.И. Надежность информационных систем: Учебное пособие. СПб.: Изд-во СПбГТУ, 2001. 89 с.

УДК 004.422

Алексеева А.В., Клячкин В.Н.,

д-р техн. наук, профессор

Ульяновский государственный технический университет, Ульяновск, Россия

v_kl@mail.ru, www.ulstu.ru

ПРОГРАММНОЕ ОБЕСПЕЧЕНИЕ ДЛЯ КОНТРОЛЯ СТАБИЛЬНОСТИ ВИБРАЦИЙ ГИДРОАГРЕГАТА

Аннотация: В работе описаны принципы, заложенные в программное обеспечение, которое предназначено для оперативного выявления нарушения стабильности вибраций гидроагрегата по результатам мониторинга. Показания датчиков визуализируются, усредняются, оцениваются корреляции между показателями с целью разбиения на подмножества независимых и коррелированных показателей. Для оценки стабильности вибраций по группам коррелированных показателей используются методы многомерного статистического контроля, по результатам которого строится карта Хотеллинга для мониторинга среднего уровня процесса и карта обобщенной дисперсии для мониторинга многомерного рассеивания.

Ключевые слова: программное обеспечение, нарушение стабильности вибраций гидроагрегата, коррелированные показатели, карта Хотеллинга.

Alexeyeva A.V., Klyachkin V.N.,

Dr. Tech. sciences, professor

Ulyanovsk State Technical University, Ulyanovsk, Russia

v_kl@mail.ru, www.ulstu.ru

SOFTWARE FOR MONITORING THE STABILITY OF VIBRATIONS OF A HYDRO-AGREEDER

Abstract: The paper describes the principles laid down in the software, which is designed

to quickly detect violations of the stability of vibrations of a hydraulic unit as a result of monitoring. Sensor readings are visualized, averaged, correlations between indicators are evaluated in order to divide them into subsets of independent and correlated indicators. To assess the stability of vibrations by groups of correlated indicators, multidimensional statistical control methods are used, the results of which are used to construct a Hotelling map to monitor the average level of the process and a generalized dispersion map to monitor multidimensional dispersion.

Keywords: software, violation of the stability of vibrations of a hydraulic unit, correlated indicators, Hotelling map.

Введение

Система вибромониторинга гидроагрегата включает показания распределенной сети датчиков относительной и абсолютной вибрации различных элементов агрегата. Замер вибраций проводился в заранее определенных точках с заданной частотой. Для повышения эффективности управления гидроагрегатом необходимо иметь возможность предупреждения об аварийной ситуации по множеству его характеристик, в частности по критерию нарушения стабильности вибраций [1]. В режиме стационарной работы агрегата для оценки стабильности процесса можно применять контрольные карты, аналогичные используемым при статистическом управлении технологическими процессами [2-4].

При этом для контроля независимых показателей используют стандартные карты Шухарта, а для анализа коррелированных показателей могут быть применены алгоритм Хотеллинга при контроле среднего уровня процесса и алгоритм обобщенной дисперсии при контроле многомерного рассеивания [5].

Для оперативного выявления возможных нарушений стабильности процесса при вибромониторинге необходима разработка программного обеспечения, включающего предварительный анализ данных о вибрациях с целью оценки коррелированности показателей и разбиения на соответствующие подмножества, а также программы построения контрольных карт.

Предварительный анализ данных о вибрациях гидроагрегата

Соответствующая программа обеспечивает ввод исходных данных, визуализацию вибраций (рис. 1), их усреднение за заданный период (рис. 2), а также оценку степени коррелированности показателей вибраций с проверкой их значимости по критерию Стьюдента [6] (рис.3).

Апробация программы проводилась на примере данных по вибрациям системы управления гидроагрегатом на Краснополянской ГЭС (Краснодарский край) как в режиме запуска, так и в режиме работы в сети. Контролировались биение вала и вибрация гидроагрегата. Процесс определяется десятью показателями [5,7]: X_1 – вибрация нижнего генераторного подшипника верхнего бьефа, X_2 – вибрация нижнего генераторного подшипника, правый берег, X_3 – вибрация верхнего ге-

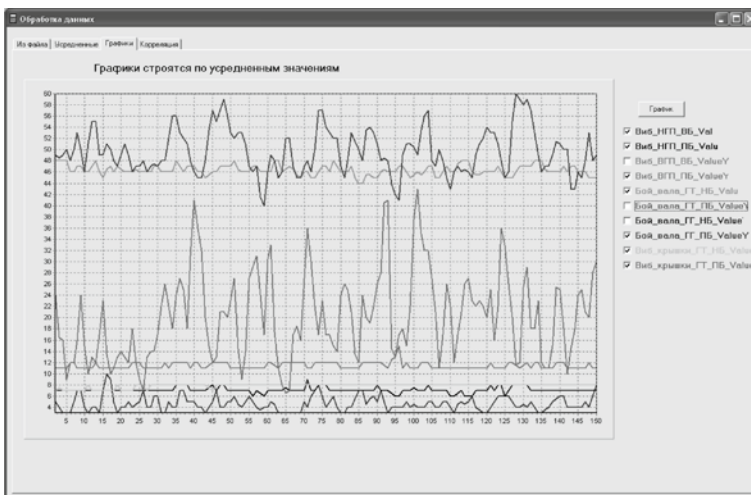


Рис. 1. Визуализация вибраций

7. Нормализация данных. Карта Хотеллвича. Обобщенные дисперсии.

Загрузка из файла | Нормализация | Карта Хотеллвича | График | Обобщенные дисперсии | График Обобщенной дисперсии

Нормализация

	X1	X2	X3	X4	X5	X6	X7	X8	X9
1	2.19307417661433	2.27212589593034	2.5649433574154	3.32939515762342	4.662604493297619	2.95537548071022	4.16704575473034	2.07526617038936	2.65930039637958
2	2.1953710006389	2.2210134300023	2.5776883623897	3.52461012033362	4.65368970195607	2.9196712274175	4.22122192134210	2.07944154167934	2.31482288374
3	2.179443836362	2.27959642006613	2.577006445912831	3.91567891636434	4.67229601503563	2.9132505172975	4.2463500570297	2.07944154167934	2.5481422394515
4	2.1930532979288	2.21920348495499	2.57261223002711	3.9307900309836	4.6791101038981	2.91596730508512	4.19636951344075	2.07944154167934	2.61251777438064
5	2.17953244432407	2.22315728797159	2.57642175862277	3.91833636748474	4.6662652863479	2.93474300152756	4.1964529597827	2.08152270789358	2.47092040781326
6	2.20460484652384	2.2824619172853	2.577006445912831	3.92756821676361	4.65268970195607	2.9400433678394	4.19128648295991	2.07944154167934	2.55986663156647
7	2.19165383228676	2.2867797467786	2.56238196199629	3.9289116179219	4.67995782652809	2.94443897996644	4.21015033924147	2.07944154167934	2.63300876263876
8	2.1860517673809	2.263487704236	2.57261223002711	3.91667891636434	4.67799569571374	2.94388305246416	4.22048761792426	2.07917192868634	2.45500518007171
9	2.17953244432407	2.27726728950976	2.58021602950233	3.92633316454016	4.67842356477256	2.9432040378989	4.167045754634	2.07395603518881	2.63687089059408
10	2.20423998150433	2.2564237302048	2.5649433574154	3.93096521321262	4.67198195150735	2.9492040378989	4.19167296616256	2.07317192868634	2.59993448648532
11	2.20092143921755	2.24424951598100	2.57261223002711	3.93098274679962	4.6510931787643	2.93030448637153	4.17104368638032	2.07317192868634	2.23421684322353
12	2.19165383228676	2.20460484652384	2.5649433574154	3.91886754818881	4.67314021242736	2.9338568969269	4.1920979470191	2.06886275947239	2.39949252772959
13	2.1897898984987	2.20052143921755	2.57261223002711	3.92197338628131	4.6794959416234	2.93088224647554	4.219073889129	2.07944154167934	2.49044882016362
14	2.21193072472591	2.2828024013142	2.56781917832321	3.91701054693919	4.6599397802876	2.95143201465744	4.22146596968144	2.07526617038936	2.3953006950378
15	2.20052143921755	2.2137838792543	2.57518392763879	3.923131305232624	4.68522388688083	2.93296389181767	4.20641707849195	2.06886275947239	2.32537491581703
16	2.19907417661433	2.239592219209	2.59036218544562	3.9399523073936	4.67646976769895	2.92673940047074	4.172079191076245	2.07526617038936	2.37024741467486
17	2.21395962920743	2.21953791600442	2.57421795862377	3.92131305232624	4.65014395163082	2.95230271562666	4.17590315157626	2.07317192868634	2.28916260258191
18	2.179443836362	2.29253475741054	2.5649433574154	3.91526079553882	4.66188252866463	2.9177733020426	4.1971262967307	2.07944154167934	2.79458925923095
19	2.1972457733622	2.28916260258191	2.57388388619531	3.9158826620963	4.6625020070259	2.94706710157271	4.19887389906455	2.07944154167934	2.428225965366

Рис. 2. Усреднение данных

нераторного подшипника верхнего бьефа, X4 – вибрация верхнего генераторного подшипника, правый берег, X5 – бой вала гидротурбины нижнего бьефа, X6 – бой вала гидротурбины, правый берег, X7 – бой вала гидрогенератора нижнего бьефа, X8 – бой вала гидрогенератора, правый берег, X9 – вибрация крышки гидротурбины нижнего бьефа, X10 – вибрация крышки гидротурбины, правый берег.

Из корреляционной матрицы (рис. 3) видно, что значимая корреляция в рассматриваемых условиях (по критерию Стьюдента на уровне значимости 0,05) имеет место между показателями X2-X6-X8, а также

Обработка данных

Из файла | Массивы | Графики | Корреляция

Корреляция

Корреляция рассчитывается по усредненным значениям

Специальный коэффициент корреляции

	Виб_НПТ_БВ_ValueY	Виб_НПТ_ПВ_ValueY	Виб_ВПТ_БВ_ValueY	Виб_ВПТ_ПВ_ValueY	Виб_вект_ПТ_БВ_ValueY	Виб_вект_ПТ_ПВ_ValueY
Виб_НПТ_БВ_ValueY						
Виб_НПТ_ПВ_ValueY	0,08195035	1				
Виб_ВПТ_БВ_ValueY	-0,19822343	-0,07902641	1			
Виб_ВПТ_ПВ_ValueY	0,00219193	0,11452627	0,16227822	1		
Виб_вект_ПТ_БВ_ValueY	-0,12530098	0,01962857	-0,12198026	-0,02265890	1	
Виб_вект_ПТ_ПВ_ValueY	0,08850060	0,49667689	-0,18945178	0,05427430	0,21342889	1
Виб_вект_ПТ_БВ_ValueY	0,15979672	0,03325203	0,10595007	0,20306005	-0,77220020	0,220447753
Виб_вект_ПТ_ПВ_ValueY	0,14952246	0,01120612	0,14600054	0,15240258	0,08641099	0,61910225
Виб_вект_ПТ_БВ_ValueY	0,09512525	0,03094452	-0,07308915	0,14370074	-0,12619548	0,01292110
Виб_вект_ПТ_ПВ_ValueY	0,11120434	0,09738430	-0,06712270	0,00381069	0,09648144	0,07396797

Рис. 3. Расчет корреляции (выделены значимые корреляции на уровне 0,05)

X5-X7. Таким образом, стабильность вибраций по этим двум подмножествам показателей должна контролироваться с помощью многомерный методов, для оставшихся показателей X1, X3, X4, X9 и X10 могут быть использованы карты Шухарта [2,8-9], например, по средним значениям и стандартным отклонениям.

Контроль стабильности вибраций

Для контроля стабильности вибраций по независимым показателям с применением карт Шухарта могут использоваться различные программные средства, например, система Statistica [10] (Statistics / Industrial Statistics & Six Sigma / Quality Control Chart). В качестве примера на рис. 4 показаны карты Шухарта для средних значений и

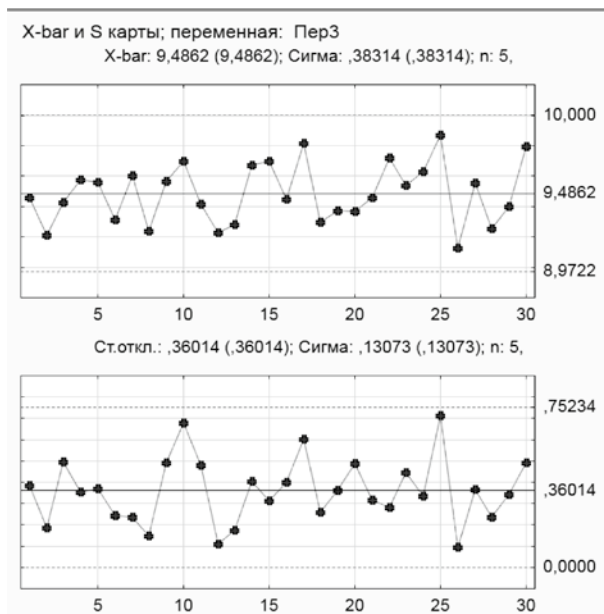


Рис. 4. Карты Шухарта для параметра X₃

стандартных отклонений по параметру X_3 , из которых следует, что вибрации стабильны: нет выхода точек за контрольные границы.

Многомерный контроль по подмножествам коррелированных показателей X_2 - X_6 - X_8 , а также X_5 - X_7 в принципе также можно проводить с применением этой же системы, однако возникает ряд проблем. Во-первых, в системе Statistica есть средства для построения карты Хотеллинга, но нет инструментов для построения карты обобщенной дисперсии (обобщенная дисперсия – определитель ковариационной матрицы – характеризует многомерное рассеяние). Кроме того, карта Хотеллинга дает корректные результаты лишь для многомерного нормального распределения, что не всегда имеет место, поэтому часто бывает необходимо преобразовывать данные, обеспечивая нормализацию. Для повышения эффективности многомерного контроля полезно выявление неслучайных структур на карте Хотеллинга, свидетельствующих о возможных нарушениях процесса.

Перечисленные обстоятельства обусловили необходимость разработки специальной программы для многомерного контроля. Программа обеспечивает:

- проверку нормальности распределения по каждому показателю и, при необходимости, – нормализацию данных; используются нормализующие преобразования Джонсона трех типов [2] (в простейшем случае – логарифмирование),

- настройку параметров алгоритма Хотеллинга и построение контрольной карты с поиском неслучайных структур,

- настройку параметров алгоритма обобщенной дисперсии и построение соответствующей карты для контроля стабильности многомерного рассеяния; кроме того, в программе заложены средства для проведения статистических испытаний [11] – моделирования изменения рассеивания с целью оценки чувствительности карты обобщенной дисперсии.

На рис. 5 показано окно программы с нормализованными данными.

Карта Хотеллинга по подмножеству коррелированных показателей X_5 - X_7 показана на рис. 6: карта свидетельствует о стабильности процесса. По этой же группе показателей на рис. 7 представлена карта обобщенной дисперсии, свидетельствующая о нарушении стабильности вибраций в выборке 2: обобщенная дисперсия вышла за контрольную границу.

Многомерный контроль по подмножествам коррелированных показателей X_2 - X_6 - X_8 с применением карт Хотеллинга и обобщенной дисперсии не выявил нарушений стабильности процесса за исследованный период мониторинга вибраций.

Заключение

Разработанное программное обеспечение предназначено для опера-

Нормализация данных. Карта Хотеллинг. Обобщенные дисперсии

Загрузка из файла | Нормализация | Карта Хотеллинг | График | Обобщенная дисперсия | График: Обобщенная дисперсия

Нормализация

	X1	X2	X3	X4	X5	X6	X7	X8	X9
1	2.199741166243	2.271258989263	2.5645432674154	3.9239815782342	4.66260482927615	2.9633744871022	4.1987045774634	2.0753617102636	2.8590323667358
2	2.195371008689	2.22101343930023	2.5766883832897	3.5246101203362	4.6536897019569	2.91867122741753	4.22122192154218	2.07544154167984	2.3141842983374
3	2.1756438936362	2.279564200613	2.57006449812031	3.9166781638434	4.67229615032663	2.91326569172975	4.24635008570297	2.07544154167984	2.54814223149515
4	2.19351395793969	2.23130384945493	2.5726122302711	3.5331903093836	4.67111081038981	2.91598720569513	4.19636951344075	2.07544154167984	2.61251777438064
5	2.17952444324027	2.2315728761758	2.57542179862377	3.9183368746474	4.6862852853479	2.93474300192256	4.19845429597823	2.08152270788366	2.47092403791326
6	2.20469584663384	2.29824619123853	2.57026445812801	3.52796821673651	4.65526897019569	2.94004338253394	4.19828482295981	2.07544154167984	2.65556663156647
7	2.19165352236676	2.2867797467768	2.56238196196429	3.8289116173219	4.67098732852809	2.94443897916844	4.21015032957417	2.07544154167984	2.633087142628976
8	2.1860517673809	2.283487704256	2.5726122302711	3.9166781638434	4.67795989571374	2.94288305246416	4.22048761736246	2.07317192086624	2.45530818021171
9	2.17953244434027	2.27726778903976	2.58021652995233	3.52636318454015	4.67942064772765	2.94952403705899	4.1987045774634	2.07736603918881	2.63680780050408
10	2.22462395154333	2.22642373200384	2.5645432674154	3.5209523321362	4.67198195150735	2.94952403705899	4.1987045774634	2.07736603918881	2.63680780050408
11	2.20052142921795	2.24424951930188	2.5726122302711	3.5299824679962	4.65109911797643	2.93008446823753	4.17104036863012	2.07317192086624	2.29421084323293
12	2.1919535220876	2.20469484846384	2.5645432674154	3.9186874814681	4.67314031242738	2.9339868898969	4.1972019479191	2.06886279847298	2.39949227762959
13	2.1897895898487	2.20092143921795	2.5726122302711	3.5219733828131	4.67934938418234	2.93088224647754	4.2190173889123	2.07544154167984	2.4304483016382
14	2.21153072472591	2.3289240131142	2.56791017832321	3.91701054633919	4.6593737802876	2.95143201485741	4.22146850986144	2.07526617028306	2.33653085030378
15	2.20052142921795	2.21379307903743	2.57915952763878	3.5213130523924	4.69522938989893	2.932968998191767	4.20841701948198	2.06686279847298	2.33527491981703
16	2.199741166243	2.23389220150703	2.58021652995233	3.5209523321362	4.67624937676908	2.92673340267034	4.17207810392645	2.07526617028306	2.31024374146796
17	2.21379307920743	2.21957371600442	2.57042179862377	3.5213130523924	4.65014395150002	2.952380271502666	4.15980015197905	2.07317192086624	2.28916202676191
18	2.1764389836362	2.2503457514054	2.5645432674154	3.91529079950082	4.6618852956463	2.9177703203428	4.19712629567307	2.07544154167984	2.79349882930095
19	2.1922245773922	2.28916202676191	2.57388388619531	3.91588298629863	4.66292020702059	2.94706710157221	4.19887309936455	2.07544154167984	2.4262795683369

Рис. 5. Нормализация исходных данных

тивного выявления нарушения стабильности вибраций гидроагрегата по результатам вибромониторинга. Показания датчиков визуализируются, усредняются, оцениваются корреляции между показателями с целью разбиения на подмножества независимых и коррелированных показателей.

Для оценки стабильности вибраций по группам параметризованных показателей используются методы многомерного статистического контроля, по результатам которого строится карта Хотеллинга для мониторинга среднего уровня процесса и карта обобщенной дисперсии для мониторинга многомерного рассеивания.

При обнаружении нарушения стабильности вибраций с помощью

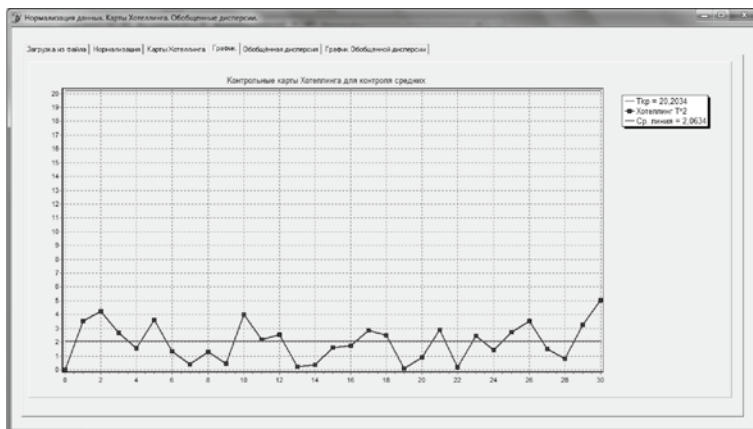


Рис. 6. Карта Хотеллинга

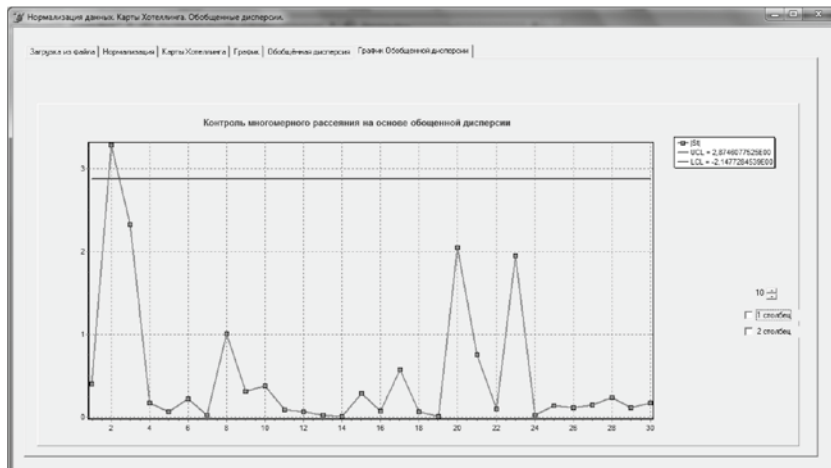


Рис. 7. Карта обобщенной дисперсии

системы управления гидроагрегатом производятся соответствующие корректирующие воздействия, обеспечивающие стабилизацию вибраций.

Список литературы

1. Стандарт организации ОАО РАО «ЕЭС России». Методика оценки технического состояния основного оборудования гидроэлектростанций. СТО 17330282.27.140.001-2006.
2. Клячкин В.Н. Статистические методы в управлении качеством: компьютерные технологии. – М.: Финансы и статистика, ИНФРА-М, 2009. – 304 с.
3. Зенцова Е.А., Клячкин В.Н. Адаптивный многомерный статистический контроль технологического процесса // Радиотехника. – 2017. – № 6. – С. 45-48.
4. Клячкин В.Н., Кравцов Ю.А. Обнаружение нарушений при многомерном статистическом контроле технологического процесса // Программные продукты и системы. – 2016. – № 3. – С. 192-197.
5. Иванова А.В., Карпунина И.Н., Клячкин В.Н. Статистическая обработка результатов вибромониторинга гидроагрегата // Научный вестник УВАУГА. – 2017. – № 9. – С. 144-150.
6. Клячкин В.Н., Кувайскова Ю.Е., Алексеева В.А. Статистические методы анализа данных. – М.: Финансы и статистика, 2016. – 240 с.
7. Иванова А.В., Клячкин В.Н., Кувайскова Ю.Е. Предварительная обработка данных при контроле стабильности вибраций // Радиоэлектронная техника. – 2017. – № 1 (10). – С. 174-177.
8. Уиллер Д., Чамберс Д. Статистическое управление процессами / Оптимизация бизнеса с использованием контрольных карт Шухарта; пер. с англ. – М.: Альпина Бизнес Букс, 2009. – 409 с.
9. Montgomery D. C. Introduction to statistical quality control. – New York, John Wiley and Sons, 2009. – 754 p.
10. Боровиков В. Statistica: искусство анализа данных на компьютере. Для профессионалов. – СПб.: Питер, 2001. – 656 с.

10. Иванова А.В., Клячкин В.Н. Оценка эффективности алгоритма обобщенной дисперсии по результатам статистических испытаний // Современные проблемы проектирования, производства и эксплуатации радиотехнических систем. – 2017. – № 1-2 (10). – С. 186-188.

УДК 621.382.33-025.13(042.3)

Lagunovich N.L.,

the Affiliate Research & Design Centre “Belmicrosystems” of Openly Join-Stock Company «INTEGRAL»-“INTEGRAL” Holding Managing Company, Minsk, Republic of Belarus, e-mail: n_dudarby@tut.by

THE NEW PROCESS FLOW OF MAKING AND DEVICE-PROCESS SIMULATION OF BIPOLAR STATIC INDUCTION

Abstract: Presented results shows possibility to receive required silicon device structure with short channel working on field effect parameters, which was made by improved process flow proposed in this paper. This paper also includes results of BSIT device simulation and their comparison with experimental data.

Keywords: transistor, simulation, improved process flow, experimental data, experimental samples.

1. Introduction

The bipolar static induction transistors [1, 2] are device structures suitable for work in conditions of relative high power, temperature and frequency (it is up to 85 Celsius degrees and 100 kHz in this case) and which can be used both as discrete devices and in high-voltage integrated circuits (IC) composition of different assignment. At present a large quantity of software products [3, 4] which gives the possibility to perform numerical computer simulation of semiconductor structures exist. Such modeling allows to reduce cost required for experimental samples making and to optimize discrete devices and elements of IC parameters [5, 6] up to stage of its manufacture. The aim of this work was to improve previously used forming BSIT process flow [1] for receiving one with collector-emitter breakdown voltage of more than 100 V at collector current of 50 mA, collector-emitter saturation voltage of less or equal to 0.3 V at collector current of 5 A and base current of 2.5 mA as well as to define electrophysical parameters and characteristics of searched transistor with help of its device-process simulation [3, 4] and to compare with experimental data.

2. The BSIT structure

BSIT is the device structure with short channel working on field effect. The construction of researched BSIT is shown on Figure 1. The amplification factor of such transistor (it is at switching scheme with common emitter) must be not less or equal to 200, the collector-emitter breakdown voltage must be more than 100 V, the collector-emitter saturation voltage must be less than 0.3 V. The structure of this device makes on silicon wafer with epitaxial layer. The base region of p-type forms

by means of boron impurity distillation at temperature of $(1200 \pm 2)^\circ$ Celsius degrees after ion-implantation doping. Base of p-type and ohmic contacts of p⁺-type regions form by boron ion-implantation doping through silicon nitride mask but emitter region of n⁺-type is created by means of arsenic ion-implantation doping into silicon area free from local oxide. The local oxide provides the reliable isolation regions of n⁻ and p⁺-types each from other. The polysilicon layer is deposited after

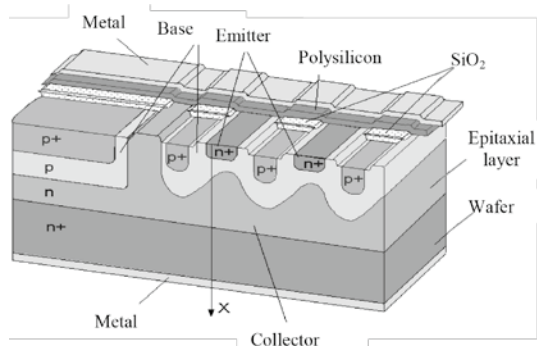


Figure 1. The BSIT structure

forming base of p-type and emitter of n⁺-type regions. The epitaxial layer and silicon wafer itself perform the role of collector in described structure.

3. The developing new BSIT making process flow

It is to receive the structure shown on Fig. 1 the BSIT making process flow was developed [1]. The one of the method of researched

BSIT making process flow improving is using thick (local) oxide in the capacity of masking cover at forming emitter of n⁺-type region.

The main researched BSIT structure forming stages as well previously used as used in suggested new process flow are presented in Fig. 2. The method of manufacturing investigated transistor involves the following major stages: 1) the epitaxial film of n-type 2 is grown on silicon wafer of n⁺-type 1 (Fig. 2, a, b); 2) the guard ring of p⁺-type 3 is formed in epitaxial layer by means of boron ion-implantation doping and its distillation, the base of p-type 6 and ohmic contacts of p⁺-type to base 7 are formed through silicon nitride mask 5 with under nitride oxide 4 (Fig. 2, c, d); 3) the photoresist layer 9 was applied and developed (Fig. 2, e) after growing local oxide 8 for creating emitter of n⁺-type 10 by arsenic impurity ion-implantation doping and its next distillation (Fig. 2, f) in previously used making BSIT process flow; 4) the transistor emitter 9 is formed after growing local oxide 8 by means of arsenic impurity ion-implantation doping into silicon region free from local oxide and its next distillation in new BSIT making process flow as follows from Figure 2, g; 5) the polysilicon layer (it is 11 on Fig. 2, h or 10 on Fig. 2, i) was applied for the purpose of creating contact to emitter, then metal contacts to base and emitter (it is 12 on Fig. 2, h or 11 on Fig. 2, i) are formed and metal contact to collector (it is 13 on Fig. 2, h or 12 on Fig. 2, i) is created.

Experimental samples of BSIT were produced by using proposed method of fabrication in this paper and it was established that new process flow makes possible to receive transistors with fixed parameters both in discrete version

and in composition of integrated circuits (IC) with smaller time and material costs. The quantity of photolithographies but therefore used metal intermediate subject copies (MISC) is equal to seven in previously used BSIT making process flow but the number of ones is equal to six in new process flow that led to reduction of MISC set cost on 16,7 % and to decrease of timetable and material costs which necessary at researched BSIT making. Moreover the reduction of photolithography number leads to lowering general accumulated imperfection and therefore permits to improve discrete transistors yield percent more than on 20 % as well as to raise microchips (the investigate BSIT is in composition of such microchips) yield percent more than on 15 %.

The previously used making BSIT process flow

The new making BSIT process flow

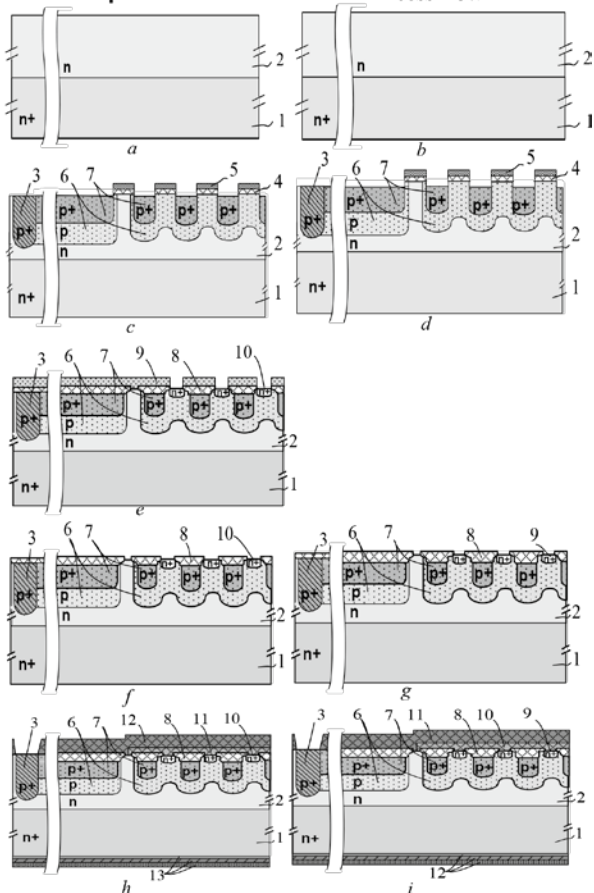


Figure 2. The BSIT forming stages

4. The results of BSIT process simulation

The results of BSIT process simulation and of researched transistor making flow declared in previous section are presented in this one. This process simulation was performed with help of program package of SUPREM3 of Silvaco company [3] and results of such simulation served by input data for further device modeling transistor. Dopant profiles were received with help of program package of SUPREM3 in direction of X (see Fig. 1). These profiles were used as initial data for device simulation execution. The calculation emitter-base junction depth value is approximate to 2.5 μm , the base-collector junction depth value is approximate to 4.8 μm . The impurity peak concentra-

tion of p-type is approximate to $8.65 \cdot 10^{15} \text{ sm}^{-3}$ on the depth of $2.84 \text{ }\mu\text{m}$ in base region 7 (see Fig. 2); the impurity peak concentration of n-type is approximate to $1.38 \cdot 10^{20} \text{ sm}^{-3}$ on the depth of $0.64 \text{ }\mu\text{m}$ in emitter region 10 (see Fig. 2). The doping concentration of n-type is significantly more in emitter region than doping concentration of p-type in base region that gives an opportunity to receive the transistor amplification factor more than 200 at device switching on scheme with common emitter. Thereby the investigated BSIT structure geometrics were defined yet to performance of direct measurements on its experimental samples.

5. The BSIT model

The device modeling investigated transistor was realized in this paper with help of software system of MOD-1D [4]. The BSIT one-dimensional model proposed by author is in foundation of MOD-1D one of the program. This model is in diffusion-drift approximation and hases the next composition:

the Poisson equation

$$\frac{\partial^2 \varphi}{\partial x^2} = -\frac{\rho}{\varepsilon \varepsilon_0} = -\frac{q}{\varepsilon \varepsilon_0} (p - n + N_d - N_a), \quad (1)$$

where φ is electrostatic potential; x is coordinate value along direction X ; p , n are concentrations of holes and electrons, respectively; q is electron charge; ε is relative permittivity; ε_0 is permittivity of free space; N_d , N_a are donor and acceptor concentrations, respectively; and continuity equations for electrons (eq. (2)) and holes (eq. (3))

$$\frac{\partial J_n}{\partial x} - qR - q \frac{\partial n}{\partial t} = 0, \quad (2)$$

$$\frac{\partial J_p}{\partial x} + qR + q \frac{\partial p}{\partial t} = 0, \quad (3)$$

where J_n и J_p are electron and hole current densities; t is time; R is recombination rate.

The J_n и J_p were assigned in expressions (2) and (3) with a glance of diffusion and drift constituents by next equations

$$J_n = qn\mu_n E + q\varphi_T \mu_n \frac{\partial n}{\partial x}, \quad (4)$$

$$J_p = qp\mu_p E - q\varphi_T \mu_p \frac{\partial p}{\partial x}, \quad (5)$$

where μ_n and μ_p are electron and hole mobilities; E is electric intensity; $\varphi_T = k_B T/q$ is temperature voltage, but k_B is Boltzman constant and T is chip thermodynamic temperature.

It was assumed that recombination mechanism is described by Shockley-Read-Hall equation with a glance of the expression which describes Auger recombination process. Then total recombination rate R can be written in the next view

$$R = (pn - n_i^2) \cdot \left(\frac{1}{\tau_n (p + n_i) + \tau_p (n + n_i)} + c_n \cdot n - c_p \cdot p \right), \quad (6)$$

where n_i is intrinsic density of charge carriers in semiconductor; τ_n , τ_p are electron and hole lifetimes, respectively; c_n , and c_p are Auger recombination coefficients.

The solution of equations set (1) – (5) with a glance of equations (6) and (14) can be performed by numerical methods only. Therefore this solution was realized by Gummel method [5]. Equations (1) – (5) were solved in discrete form by Gauss method [7] on each iteration of Gummel method numerical implementation.

6. The results of BSIT device simulation and their comparison with experimental data

The current distributions along direction X and device volt-ampere characteristic (VACH) calculations were carried out with help of researched BSIT device simulation (Fig. 3). It is completed the modeling only that structure part which work as transistor immediately. The epitaxial film thickness value was $30 \mu\text{m}$ but resistivity of epitaxial layer was $6 \text{ Ohm}\cdot\text{sm}$ where the BSIT structure was formed during the simulation. The epitaxial film was grown on substrate which have resistivity $\rho_v = 0.01 \text{ Ohm}\cdot\text{sm}$ and thickness of $460 \mu\text{m}$. Thereby it is possible to neglect by resistance which included by wafer in total BSIT one at conditions of low injection level but that part of structure which was considered at simulation can be considered as the resistor with very low resistance series-connected with transistor.

It was supposed that simulated transistor was connected by common-base circuit but currents leaked through device will be very small (Fig. 3) at small bias voltage of base-to-emitter junction ($0 \div 0.5 \text{ V}$). The emitter voltage is lowered with respect to base potential at raising base-to-emitter voltage U_{be} . Consequently the potential barrier is reduced for charge mobile car-

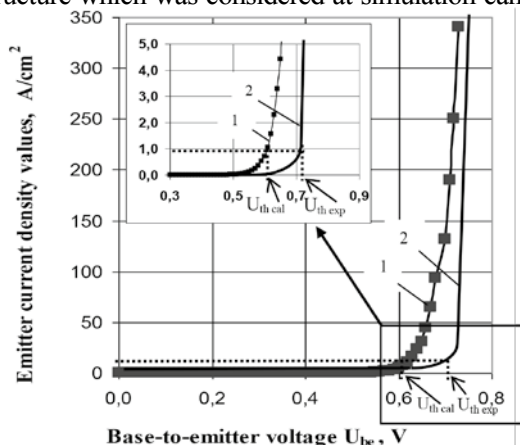


Figure 3. Calculate (1) and experimental (2) dependences of BSIT emitter current density on base-to-emitter voltage at base-to-collector voltage $U_{bc} = 3 \text{ V}$

riers on base-to-emitter junction that puts to rise of hole concentration in the n-type emitter and to growth of electron concentration in the p-type base. In whole transistor is opened gradually and current grows exponentially (Fig. 3) with rise of U_{be} at inversely biased base-to-collector junction ($U_{bk}=3$ V).

The BSIT experimental samples were received by making process flow proposed in this paper and device VACH was measured on mentioned samples for case of transistor diode connection (it is collector shorted with base). The BSIT threshold voltage $U_{th\ exp}$, which was received consequently direct (experimental) measurements, was equal to 0.72 V at emitter current of 100 mA. The active area of explored transistor was equal $3 \times 3 \text{ mm}^2$, that is current density will equal 1.11 A/sm^2 at emitter current of 100 mA. While the calculated threshold voltage $U_{th\ cal}$ value will equal 0.61 V as shown from Figure 3, but difference between calculated and experimental data will be 0.11 V or 18 %.

7. Conclusion

Presented results shows possibility to receive required silicon BSIT parameters and VACH which was made by improved process flow proposed in this paper and in which six photolithographies but therefore six MISC used instead of seven theirs utilized previously that led to decrease of timetable and material costs and to reduction of MISC set cost on 16,7 %. The reduction of photolithography number leads to lowering imperfection that is to improve discrete transistors yield percent more than on 20 % as well as to raise microchips (the investigate BSIT is in composition of such microchips) yield percent more than on 15 %. The BSIT process modeling performed with help of suite SUPREM3 allowed to define device design parameters yet before the stage of manufacturing its experimental samples. Impurities and potential distributions received consequently of process simulation were used as initial data for device modeling transistor which was carried out by software system MOD-1D developed by author. The device modeling enabled to calculate transistor threshold voltage $U_{th\ cal}$. BSIT experimental samples were produced by process flow suggested in this paper and transistor VACH was measured. It was established that variance between the estimated value of $U_{th\ cal} = 0.61$ V and its experimental one of $U_{th\ exp} = 0.72$ V is 0.11 V and not exceeds 18 %.

References

- [1] Dudar N.L. Simulation of Si based static induction transistor//Doklady of the Belarusian State University of Informatics and Radioelectronics. 2005, N 2 (10). P. 79-85.
- [2] Ismailov T.A., Shharmaeva A.R., Zaharova P.R. Technological solution for bipolar static induction transistor chip parameters improvement//Reporter of Dagestan State Technical University. 2011, Vol. 20, N 1. P. 6-11.
- [3] <http://www.silvaco.com>.
- [4] Lagunovich N.L. The computer program MOD-1D: deposit-receipt N 742 Republic of Belarus; the declarant is Openly Join-Stock Company «IN-

- TEGRAL» – № t 20140041; it was written in List of computer programs registered in Nat. center of intellectual property 10.03.15.
- [5] H.K. Gummel H.K. A self consistent iterative scheme for one-dimensional steady state transistor calculations//IEEE Trans. Electron. Dev, 1964. Vol. ED-11, N 10. P. 455-465.
- [6] Polskiy B.S. Numerical simulation of semiconductor devices. Riga: Zinatne, 1986 – 168 p.
- [7] Samarskiy A.A., Nikolaev E.S. Methods of solving net equations. M.: Science, 1978 – 326 p.

УДК 519.6:532.5

Леонтьев А.Л.¹; Никитина А.В.²,

д-р техн. наук,

Южный федеральный университет, Таганрог, Россия;

Email:Leontyev_Anton@mail.ru1; Nikitina.v@gmail.com

РАЗРАБОТКА МОДЕЛИ ГИДРОДИНАМИКИ ГЛУБОКОВОДНОГО ВОДОЕМА НА ОСНОВЕ МОДЕЛИ ПЕРЕМЕННОЙ ПЛОТНОСТИ^{1*}

Аннотация: Работа посвящена разработке и исследованию математической модели гидродинамики водоема переменной глубины. Модель описывает движение водного потока, учитывает силу Кориолиса, горизонтальную и вертикальную турбулентность водной среды, сложную геометрию береговой линии и дна, трение о дно и ветровые течения, испарение, скорость ветра и плотность атмосферы, отклонение значения поля давления от гидростатического приближения. В работе проведено исследование влияния пространственного распределения температуры, солености, ионного состава и гидростатического давления на плотность воды. В качестве объекта моделирования был выбран водоём (озеро) – Каспийское море. В качестве входных данных используются изменения распределений температуры и солености, измеренные по электропроводности воды с учетом ионного состава вод Каспийского моря за многолетний период. Для проведения численных экспериментов, калибровки и верификации моделей гидродинамических наблюдений используются базы данных «ЕСИ-МО» и «БМЭ Черного, Азовского и Каспийского морей». Осуществлена локализация точки проведения измерений для проверки адекватности модели: «Тюлений остров». Приведен анализ и графическая интерпретация погрешности расчета плотности воды за 177 месяцев. Моделирование движения водного потока позволят изучать механизмы формирования природных катастроф, таких как наводнения, абразия берегов, эвтрофикация, заморные явления в ускоренном режиме времени. Представленные в работе модели гидродинамики, могут быть адаптированы для других водоемов при использовании соответствующих геоинформационных систем.

Ключевые слова: Каспийское море; математическое моделирование; поле скоростей водного потока; переменная плотность; соленость; ионный состав; температура; попеременно-треугольный метод.

¹ Работа выполнена при поддержке РФФ, проект № 17-11-01286

Leontyev A.L.¹; Nikitina A.V.²,

Dr. Tech. sciences,

Southern Federal University, Taganrog, Russia;

Email: Leontyev_Anton@mail.ru¹; Nikitina.vm@gmail.com

DEVELOPMENT OF THE MODEL OF HYDRODYNAMICS OF DEEP WATER BASED ON THE BASIS OF THE VARIABLE DENSITY MODEL *

Abstract: *The work is devoted to the development and study of a mathematical model of hydrodynamics of a reservoir of variable depth. The model describes the movement of water flow, takes into account the Coriolis force, horizontal and vertical turbulence of the aquatic environment, complex geometry of the coastline and bottom, friction about the bottom and wind currents, evaporation, wind speed and density of the atmosphere, the deviation of the pressure field from the hydrostatic approximation. In this paper, we studied the influence of the spatial distribution of temperature, salinity, ion composition, and hydrostatic pressure on the density of water. The reservoir (lake) – the Caspian Sea was chosen as the modeling object. The input data are changes in temperature and salinity distributions measured by the electrical conductivity of water, taking into account the ionic composition of the waters of the Caspian Sea over a long-term period. For conducting numerical experiments, calibration and verification of hydrodynamic observation models, the databases ESIMO and BME of the Black, Azov and Caspian Seas are used. The measurement point was localized to verify the adequacy of the model: "Seal Island". The analysis and graphical interpretation of the error in calculating the density of water for 177 months is given. Modeling the movement of water flow will study the mechanisms for the formation of natural disasters, such as floods, abrasion of the coast, eutrophication, and observational phenomena in an accelerated mode of time. The hydrodynamics models presented in the paper can be adapted for other water bodies using appropriate geographic information systems.*

Keywords: *aspian Sea; math modeling; water flow velocity field; variable density; salinity; ion composition; temperature; alternately-triangular method.*

Введение

В качестве объекта моделирования был выбран глубоководный водоём – Каспийское море. Каспийское море – глубоководный водоем с сильно развитой шельфовой зоной. Протяжённость Каспийского моря с севера на юг – примерно 1200 км, с запада на восток – от 195 до 435 км, в среднем 310 – 320 км. Каспийское море условно делится по физико-географическим условиям на 3 части – Северный Каспий, Средний Каспий и Южный Каспий. Мангышлакский порог отделяет Северный Каспий от Среднего. Апшеронский порог разделяет Средний и Южный Каспий. Площадь Северного, Среднего и Южного Каспия составляет соответственно 25, 36, 39 % от общей площади моря. Максимальная глубина Каспийского моря (рис. 1) – в Южно-Каспийской впадине, составляет 1025 м. По величине максимальной глубины Каспийское море уступает лишь Байкалу (1620 м) и Танганьике (1435 м). Средняя глубина Каспийского моря, рассчитанная по батиметрической кривой, составляет 208 м. В то же время северная часть Каспия – мелководная, её максимальная глубина не превышает 25 м, а средняя глубина – 4 м [1].

Математическим моделированием процессов гидрофизики и биологической кинетики Каспийского моря занимались многие ученые. Работы ученых Р.А. Ибраева, Ю.Б. Силантьева, В.А. Калмыкова, О.В. Олейник, Р.П. Ходоревской, А.А. Жилкина, И.С. Гулиева, А. Naderi Beni, Н. Lahjani, R. Mousavi Harami, K. Apre, S.A.G. Leroy, N. Marriner, M. Berberian, V. Andrieu-Ponel, M. Djamali, A. Mahboubi, P.J. Reimer посвящены математическому моделированию процессов термогидродинамики, анализу проблем сохранения водных ресурсов и влиянию деятельности человека на экологическую безопасность, а также изучению состояния запасов ценных промысловых рыб. Но, несмотря на большое количество работ в данной области, многие эффекты, существенные для повышения точности и надежности прогнозов изменения экологической обстановки глубоководных водоемов, связанные с пространственной неоднородностью и движением водной среды, температурным и кислородным режимами, соленостью, не учитывались ранее при построении гидрофизических моделей [2 – 8]. В качестве объекта моделирования было выбрано Каспийское море, в виду того, что это единственный водоем, располагающийся в нескольких климатических зонах и имеющий как глубоководные, так и мелководные зоны. Изучение распределения поля плотности морской воды необходимо для точного определения осадки судов в разных точках моря, позволяет совершенствовать средства мониторинга экологической ситуации в водоеме, разрабатывать оптимальные сценарии развития биоты глубоководного водоема – Каспийское море. Разработка и численная реализация гидродинамических моделей является одним из этапов создания проектов устойчивого развития водоемов, а также одной из важных и первоочередных задач экологической безопасности и экономики стран.

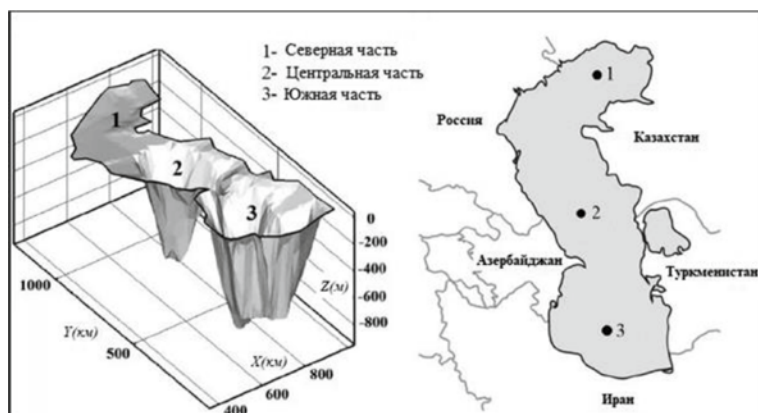


Рис. 1. Батиметрия Каспийского моря

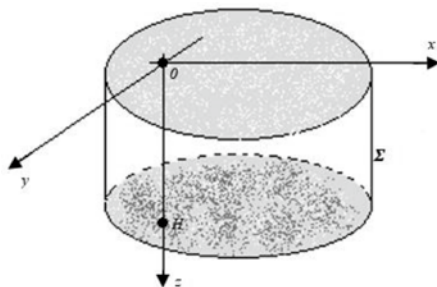


Рис. 2. Расчетная область

$$\frac{\partial u}{\partial t} + u \frac{\partial u}{\partial x} + v \frac{\partial u}{\partial y} + w \frac{\partial u}{\partial z} - 2\Omega(v \sin\theta - w \cos\theta) + \frac{1}{\rho} \frac{\partial p}{\partial x} + \frac{\partial \varphi}{\partial x} = \mu \Delta u + \frac{\partial}{\partial z} \left(\nu \frac{\partial u}{\partial z} \right), \quad (1)$$

$$\frac{\partial v}{\partial t} + u \frac{\partial v}{\partial x} + v \frac{\partial v}{\partial y} + w \frac{\partial v}{\partial z} + 2\Omega u \sin\theta + \frac{1}{\rho} \frac{\partial p}{\partial y} + \frac{\partial \varphi}{\partial y} = \mu \Delta v + \frac{\partial}{\partial z} \left(\nu \frac{\partial v}{\partial z} \right), \quad (2)$$

$$\frac{\partial w}{\partial t} + u \frac{\partial w}{\partial x} + v \frac{\partial w}{\partial y} + w \frac{\partial w}{\partial z} - 2\Omega u \cos\theta + \frac{1}{\rho} \frac{\partial p}{\partial z} + \frac{\partial \varphi}{\partial z} = \mu \Delta w + \frac{\partial}{\partial z} \left(\nu \frac{\partial w}{\partial z} \right), \quad (3)$$

$$\frac{\partial \rho}{\partial t} + \frac{\partial(\rho u)}{\partial x} + \frac{\partial(\rho v)}{\partial y} + \frac{\partial(\rho w)}{\partial z} = 0, \quad (4)$$

$$\frac{\partial T}{\partial t} + u \frac{\partial T}{\partial x} + v \frac{\partial T}{\partial y} + w \frac{\partial T}{\partial z} + \gamma_T w = \mu_T \Delta T + \frac{\partial}{\partial z} \left(\nu_T \frac{\partial T}{\partial z} \right), \quad (5)$$

$$\frac{\partial S}{\partial t} + u \frac{\partial S}{\partial x} + v \frac{\partial S}{\partial y} + w \frac{\partial S}{\partial z} + \gamma_S w = \mu_S \Delta S + \frac{\partial}{\partial z} \left(\nu_S \frac{\partial S}{\partial z} \right), \quad (6)$$

где u, v, w – компоненты вектора скорости течения; $\gamma_{T,S}(z)$ – заданные параметры стратификации: $\gamma_T = d\bar{T}/dz, \gamma_S = d\bar{S}/dz$; T, S – отклонения температуры и солёности воды от средних значений \bar{T}, \bar{S} ; $\Delta = \partial^2/\partial x^2 + \partial^2/\partial y^2$ – оператор Лапласа; ρ – превышение давления над гидростатическим давлением невозмущенной жидкости; φ – значение гравитационного потенциала; θ – угол между угловой скоростью и вертикалью; $\Omega = 7,2921 \cdot 10^{-5} \text{c}^{-1}$ – угловая скорость вращения Земли; ρ – плотность воды; μ, ν – горизонтальные и вертикальные коэффициенты диффузии соответственно [8].

К системе (1) – (6) добавим граничные:

верхняя граница:

$$\rho_v \mu u_n(x, y, z, t) = -\tau_x(t), \quad \rho_v \mu v_n(x, y, z, t) = -\tau_y(t), \quad \rho_v \mu w_n(x, y, z, t) = -\tau_z(t),$$

$$w(x, y, t) = -\psi - p'_t / \rho g, \quad p'_n(x, y, t) = 0, \quad T'_z = f_3, \quad S'_z = f_4; \quad (7)$$

берег и дно:

$$\rho_v \mu u_n(x, y, z, t) = -\tau_x(t), \quad \rho_v \mu v_n(x, y, z, t) = -\tau_y(t), \quad \rho_v \mu w_n(x, y, z, t) = -\tau_z(t),$$

$$V_n(x, y, z, t) = 0, \quad pn'(x, y, z, t) = 0, \quad Tn' = 0, \quad Sn' = 0; \quad (8)$$

и начальные условия:

$$t = t_0, u = u_0, v = v_0, T = T_0, S = S_0 \quad (9)$$

Здесь ψ – интенсивность испарения жидкости; ρ_v – плотность взвеси; $u_0, v_0, S_0, T_0, f_3, f_4$ – заданные функции; n – вектор нормали; g – ускорение свободного падения.

Составляющие тангенциального ускорения определяются как:

$$\tau_x = \rho_a C_p(\vec{\omega}) \omega_x |\vec{\omega}|, \quad \tau_y = \rho_a C_p(\vec{\omega}) \omega_y |\vec{\omega}|, \quad \tau_z = \rho_a C_p(\vec{\omega}) \omega_z |\vec{\omega}|. \quad (10)$$

Здесь $\vec{\omega}$ – скорость ветра относительно воды; ρ_a – плотность атмосферы; безразмерный коэффициент $C_p(x)$ определяется как:

$$C_p(x) = \begin{cases} 0,0088, & \text{при } x < 6,6 \text{ м/с} \\ 0,0026, & \text{при } x \geq 6,6 \text{ м/с} \end{cases} \quad (11)$$

Модель учитывает движение водного потока, микротурбулентную диффузию, силу Кориолиса, пространственное распределение солёности и температуры, а также скорость ветра и плотность атмосферы. Часть работ по океанологии и термогидродинамике рассматривают распределение плотности воды как заранее заданный параметр, другая часть использует линейный вариант распределения, зависящий только от изменения температуры:

$$\rho_L = 1028(1 - \beta T), \quad (12)$$

где β – коэффициент, выражающий зависимость изменения плотности воды от изменения температуры (табл. 1).

1. Значение коэффициента β

T, °C	$\beta \cdot 10^4, \text{ } ^\circ\text{C}^{-1}$
0	-0,6300
5	0,0350
7,5	0,3675
10	0,7000
12,5	0,9800
15	1,2600
17,5	1,5400
20	1,8200
22,5	2,1675
25	2,5150

Данная зависимость, как и предположение о том, что распределение плотности может быть задано фиксированным значением, является максимально допустимым упрощением, что в свою очередь существенно понижает точность моделей, построенных на данных предположениях. Так как плотность воды зависит от температуры и солёности, которая в свою очередь зависит от ионного состава воды, то необходима модель наблюдений, учитывающая все нюансы. Предлагаемая модель наблюдений, с помощью которой будет проводиться прогнозирование

изменения гидрофизических процессов в Каспийском море, основывается на уравнении состояния, используемым для определения плотности воды во всех водоемах, находящихся на планете и одобренном ЮНЕСКО [9]:

$$\rho = \rho(T, S, 0) / [1 - p / (k(T, S, p) \cdot 10^3)], \quad (13)$$

$$\rho(T, S, 0) = \rho_0 + AS - BS^{3/2} + CS^2, \quad (14)$$

$$\rho_n = i_0 + i_1 T - i_2 T^2 + i_3 T^3 - i_4 T^4 + i_5 T^5, \quad (15)$$

$$A = A_0 - A_1 T + A_2 T^2 - A_3 T^3 + A_4 T^4, \quad (16)$$

$$B = B_0 - B_1 T + B_2 T^2, \quad (17)$$

$$C = 4,8314 \cdot 10^{-4}, \quad (18)$$

$$k(T, S, p) = k(T, S, 0) + Z_1 \cdot 10^{-5} p + X_1 \cdot 10^{-10} p^2, \quad (19)$$

$$k(T, S, 0) = k_n + aS + bS^{3/2}, \quad (20)$$

$$Z_1 = z_w + (z_0 - z_1 T - z_2 T^2)S + z_3 S^{3/2}, \quad (21)$$

$$X_1 = x_w - (x_0 - x_1 T - x_2 T^2)S, \quad (22)$$

$$z_w = c_0 + c_1 T + c_2 T^2 - c_3 T^3, \quad (23)$$

$$x_w = d_0 - d_1 T + d_2 T^2, \quad (24)$$

$$k_n = k_{n0} + k_{n1} T - k_{n2} T^2 + k_{n3} T^3 - k_{n4} T^4, \quad (25)$$

$$a = a_0 - a_1 T + a_2 T^2 - a_3 T^3, \quad (26)$$

$$b = b_0 + b_1 T - b_2 T^2, \quad (27)$$

$$p = l_0 h + l_1 h^2, \quad (28)$$

где $k(T, S, p)$ – средний модуль упругости; T, S – отклонения температуры и солёности от средних значений \bar{T}, \bar{S} ; h – глубина; ρ_n – плотность стандартной среднеокеанической воды;

$i_{0-5}, A_{0-4}, B_{0-2}, z_{w,0-3}, x_{w,0-2}, k_{n0-n4}, a_{0-3}, b_{0-2}, l_{0,1}, c_{0-3}, d_{0-2}$ –

заданные коэффициенты.

Данная модель наблюдений применима в диапазонах:

- для солёности $S = 0 \div 42 \text{ ‰}$;

- для температуры $T = -2 \div 40^{\circ}\text{C}$;
- для давления $p = 0 \div 1000$ бар.

Как было сказано ранее, соленость напрямую зависит от ионного состава воды. Ионный состав замкнутых водоемов сильно отличается от состава океанических вод (табл. 2). На гидрологических станциях «ЕСИМО» соленость измеряется с учетом содержания хлора в воде для Северного, Среднего и Южного Каспия по отдельности, что позволяет учесть ионный состав вод Каспийского моря в целом, и не вносить дополнительные правки касающиеся ионного состава воды.

2. Сравнительная таблица ионного состава океана и Каспия

Состав солей	Общая соленость	
	Океан (35‰)	Каспийское море (13‰)
Na	30,593	24,82
K	1,106	0,66
Ca	1,197	2,70
Mg	3,725	5,70
Cl	55,292	41,73
Br	0,188	0,06
SO ₄	7,692	23,49
CO ₃	0,207	0,84

Соотношение между соленостью и хлорностью в водах Каспийского моря сильно отличается от установленного для океана:

$$S_{\text{‰}} = 2,386 \text{ Cl}_{\text{‰}} \quad (29)$$

$$S_{\text{‰}} = 2,396 \text{ Cl}_{\text{‰}} \quad (30)$$

$$S_{\text{‰}} = 0,14 + 1,36 \text{ Cl}_{\text{‰}} \quad (31)$$

Зависимость вида (29) получена А.А. Лебединцевым в 1901 году, для вод Среднего Каспия. Соотношение вида (30) применимо для вод Южного Каспия и получено в 1937 году С.В. Бруевичем. В 1939 году А.В. Трофимов установил зависимость вида (31) для вод Северного Каспия [10].

Входные данные

Для решения поставленной задачи вида (1) – (11) в расчетной области сложной формы – глубоководное озеро Каспийское море, используются входные данные с портала единой государственной системы информации об обстановке в мировом океане «ЕСИМО» (рис. 3) [11].

На береговой линии Каспийского моря располагается 38 гидрологических станций (рис. 4), формирующих постоянно обновляемую базу данных «ЕСИМО», использование которых позволяет произвести сравнительный анализ модели вида (12) с моделью вида (13).

По результатам историко-сравнительного анализа, охватывающего 50-летний период, установлено, что соленость воды в Каспийском море достигает своего максимального значения (13,5‰) в районе Южного Каспия. Соленость вод в Северном Каспии близка к нулю. Средний Каспий является переходной зоной, значение солёности здесь колеблется в интервале от 10‰ до 13‰. На рисунке 5 представлены средние распределения солёности вод Каспийского моря в весенне-летний период за 50 лет.

Мониторинг температурного режима Каспийского моря проводится организацией «ЕСИМО» начиная с 1960 года. Весенний период характеризуется равномерным возрастанием температуры вод Каспийского



Рис. 3. Главная навигационная панель портала «ЕСИМО»



Рис. 4. Расположение гидрологических станций

моря с севера на юг, но к середине летнего периода температурный режим претерпевает значительные изменения. Во время выравнивания температуры в районах Северного и Южного Каспия, температура в районе Среднего Каспия, а именно мысов Песчаный и Саржа, расположенных западнее села Курык, Каракиянского района, Мангистауской области Казахстана, понижается на 5 – 10 °С [11]. На рисунке 6 представлены средние распределения температуры в весенне-летний период за последние 50 лет.

Каспийское море омывает берега пяти государств (Россия, Казахстан, Туркмения, Иран, Азербайджан). В период с 1991 по 2010 года координация работы «ЕСИМО» была нарушена. Начиная с 2010 года и вплоть до нынешнего времени обмен данными между странами осложняется геополитическими факторами, такими как борьба за разработку нефтяных месторождений и неутрачивающим спором о перераспределении нормы вылова осетровых. В связи с этим сравнительный анализ данных был проведен для периода с января 1977 года по октябрь 1991 года. За выбранный период накоплено большое количество информации о гидрологических показателях Каспийского моря. При использовании неспециализированных средств моделирования такие массивы данных делают невозможным сравнительный анализ показателей плотности воды во всех точках моря.

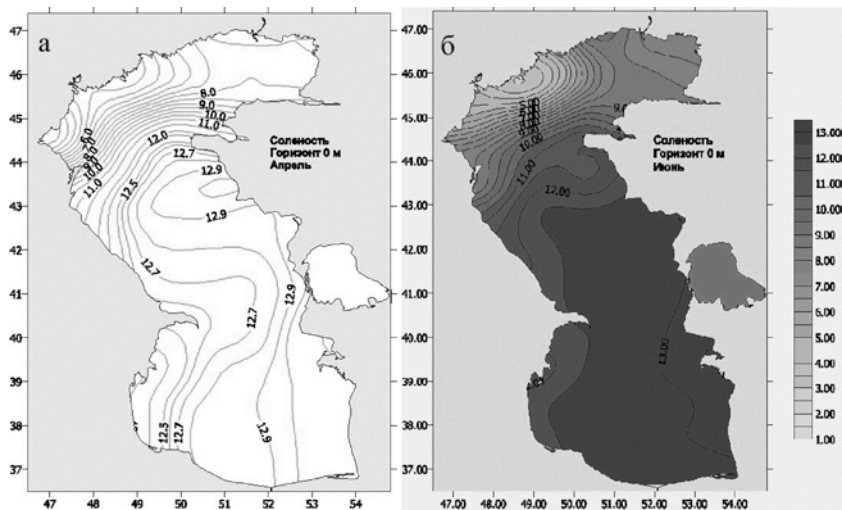


Рис. 5. Среднее распределение солёности в Каспийском море:
(а) – с апреля по июнь; (б) – с июня по август

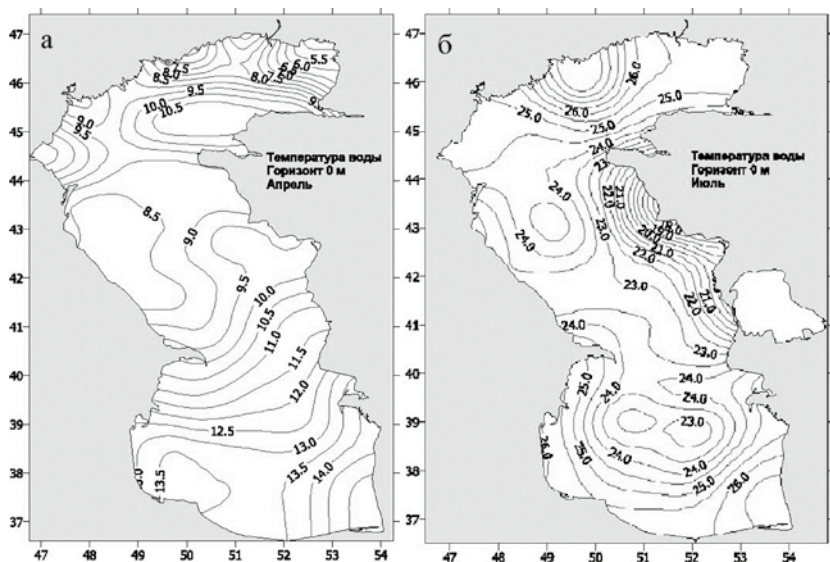


Рис. 6. Среднее распределение температуры в Каспийском море:
(а) – с апреля по июль; (б) – с июля по август

Сравнение моделей наблюдений вида (12) и (13) проведено по данным, собранным на гидрологической станции «Тюлений остров» (Даг. ЦГМС) (рис. 7)

Станция расположена на Мангышлакском пороге, разделяющим Северный Каспий – мелководный и Средний Каспий – глубоководный, что делает моделирование в данной области особо важным как для дальнейшего использования полученных данных в моделировании водных потоков, так и для судоходства в регионе (осадка судов рассчитывается по плотности воды).

Анализ моделей наблюдений

Для моделирования значений плотности воды за выбранный временной период в заданной точке моря используются массивы данных для температуры, солёности и давления по $n = 177$ элементов каждый (одно измерение в месяц).

Так как модель (12) не является универсальной, то β зависимости от значений температуры будет меняться и коэффициент β , выражающий зависимость изменения плотности воды от изменения температуры.

Графики зависимостей вида (12) и (13) для расчета плотности воды представлены на рисунках 8 и 9 соответственно.



Рис. 7. Станция «Тюлений Остров»

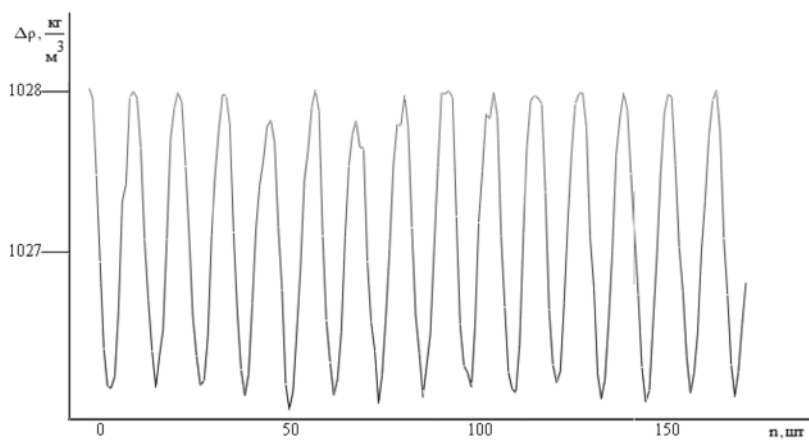


Рис. 8. График плотности воды для линейной зависимости

Критерием проверки адекватности рассмотренных моделей наблюдений вида (12) и (13) служит оценка погрешности моделирования с одновременным учетом натуральных данных по имеющимся n размерам, которая вычисляется по формуле:

$$\delta = \sqrt{\sum_{k=1}^n (\rho_{k \text{ nat}} - \rho_k)^2} / \sqrt{\sum_{k=1}^n \rho_{k \text{ nat}}^2}, \quad (32)$$

где $\rho_{k \text{ nat}}$ – значение плотности воды, взятое из океанографической базы данных «Большие морские экосистемы Чёрного, Азовского и Каспийского морей (Большие морские экосистемы (БМЭ) Чёрного (включая подрайон Азовского моря) и Каспийского морей)», ρ_k – значение плотности, рассчитанное с помощью моделей (12), (13).

Обозначим $\delta_{(i)}$ – значение погрешности, рассчитанное по формуле (32) с использованием i -ой зависимости, $i \in \{12, 13\}$ в качестве модели наблюдений. Для модели (12) получим:

$$\delta_{(12)} = 0,03;$$

для модели (13):

$$\delta_{(13)} = 0,0005394.$$

Таким образом, установлено, что универсальная модель вида (13) в 60 раз точнее модели (12) описывает изменение плотности воды в Каспийском море. На рисунке 10 представлен график универсальной зависимости (формула (13)) и натурные измерения из базы данных «БМЭ Чёрного, Азовского и Каспийского морей».

Заключение

В работе предложена математическая модель гидродинамики водоема переменной глубины Каспийское море, учитывающая сложную геометрию береговой линии и дна, трение о дно и ветровые течения, испарение, скорость ветра относительно воды, плотность атмосферы, отклонение значения поля давлений от гидростатического приближения, плотность воды как функцию распределения температуры и со-

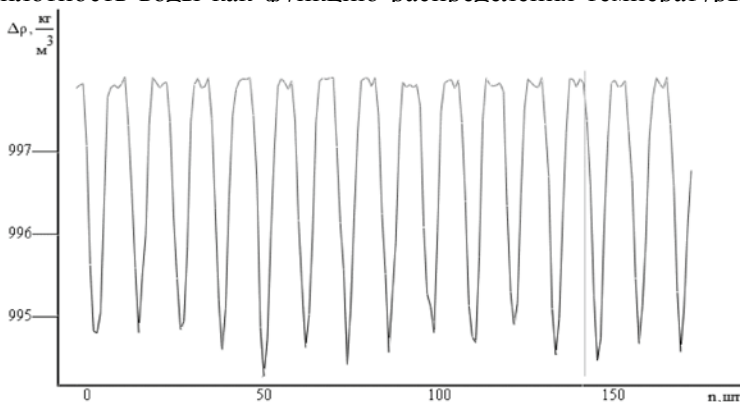


Рис. 9. График плотности воды для универсальной зависимости

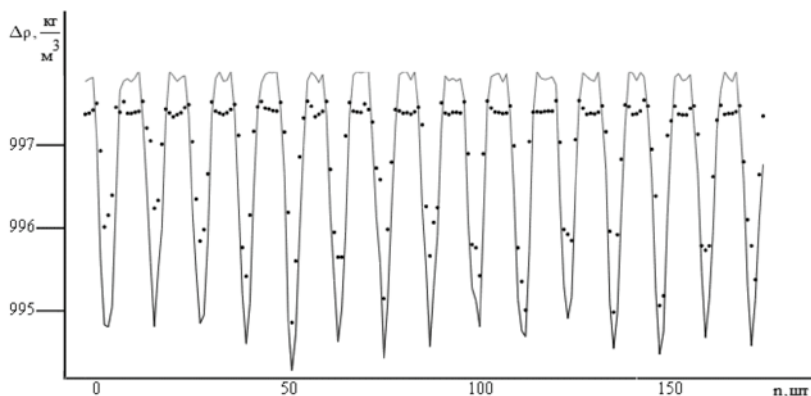


Рис. 10. Значения плотности воды:
 (•••) – натурные измерения; (—) – модельные данные

лености. Проведён историко-сравнительный анализ распределений температуры и солёности в весенне-летние периоды за 50-летний период. Произведено моделирование функций, отражающих изменение солёности в выбранной точке моря, с целью определения более жизнеспособной модели. Проведен сравнительный анализ результатов моделирования, линейной и универсальной зависимостей, по итогу которого доказано, что универсальная модель значительно точнее линейной. При моделировании учтено, что измерения солёности проводились по значениям электропроводности. Электропроводность является параметром, по которому находится хлорность воды, что в свою очередь позволяет учесть ионный состав вод Каспийского моря. Для решения возникающих при дискретизации гидродинамической модельной задачи сеточных уравнений в области сложной формы – Каспийское море будет разработан алгоритм использования модифицированного попеременно-треугольного метода (МПТМ), обладающий высокой скоростью сходимости при решении систем сеточных уравнений большой размерности, в самосопряженном и несамосопряженном случае [12]. С целью повышения эффективности вычислительного процесса для численной реализации гидродинамической модели Каспийского моря в дальнейшем будут разработаны параллельные алгоритмы, ориентированные на МВС [13-16].

Исследование гидрофизических процессов Каспийского моря, с учетом зависимости плотности от пространственного распределения температуры, солёности и ионного состава воды, повышает точность математического моделирования. Моделирование движения водного потока в глубоководном озере позволят совершенствовать точность расчетов перегрузок на суда, рассчитывать моменты формирования природных катастроф, таких как наводнения, абразия берегов, эвтро-

фикация, заморы, и в режиме реального времени производить расчеты для выделения зон, загрязненных в следствие катастроф техногенного характера, например, в случае аварийных разливов нефти.

Список литературы

1. http://esimo.oceanography.ru/esp1/index.php?sea_code=2§ion=2&menu_code=777
2. Ибраев Р.А. Математическое моделирование термогидродинамики Каспийского моря. Диссертация. 2002. 209 с.
3. Силантьев Ю.Б. Особенности геологического строения и нефтегазоносности северо-западной части Каспийского моря. // Вести газовой науки. 2014. № 3 (19). С. 49 – 55.
4. Калмыков В.А. Миграции, распределение, структура популяции и запасы стерляди Нижней Волги. Диссертация. 2005. 162 с.
5. Олейников Е.П. Исследование краниологических и молекулярно-генетических маркеров разнообразия популяции тюленя в Каспийском море. Диссертация. 2015. 116 с.
6. Павлова М.В. Современное экологическое состояние и перспективы устойчивого развития Каспийского региона. Диссертация. 2007. 150 с.
7. Гулиев И.С., Кулаков С.И., Федоров Д.Л. Нефтегазоносность Каспийского региона. Nafta-Press, Баку, 2009. 409 с.
8. Марчук Г.И. Математическое моделирование в проблеме окружающей среды. М.: Наука, 1982. 319 с.
<http://docs.cntd.ru/document/1200105587>
9. Алекин О.А., Ляхин Ю.И. Химия океана. Гидрометеиздат, 1984. 344с.
<http://esimo.ru/portal/auth/portal/arm-csmonitor/Расчетно-модельный+комплекс>
10. Сухинов А.И., Чистяков А.Е. Адаптивный модифицированный попеременно-треугольный итерационный метод для решения сеточных уравнений с несамосопряженным оператором. // Математическое моделирование. 2012. С. 398–409.
11. Petrov I.B. Application of grid-characteristic method for numerical solution of deformable solid mechanics dynamical problems // Computational Mathematics and Information Technologies. – 2017. Vol.1, No 1. – Pp. 1-20.
12. Sukhinov A.I., Sidoryakina V.V., Sukhinov A.A. Sufficient convergence conditions for positive solutions of linearized two-dimensional sediment transport problem // Computational Mathematics and Information Technologies. – 2017. Vol.1, No 1. – Pp. 21-35.
13. Nikitina A.V., Semenyakina A.A. Mathematical modeling of eutrophication processes in Azov Sea on supercomputers // Computational Mathematics and Information Technologies. – 2017. Vol.1, No 1. – Pp 82-101.
14. Chistyakov A.E., Protzenko E.A., Timofeeva E.F. Mathematical modeling of oscillatory processes with a free boundary // Computational Mathematics and Information Technologies. – 2017. Vol.1, No 1. – Pp. 102-112.

УДК 519.6

Никитина А.В.,

д.т.н., доцент, ООО «Научно-исследовательский центр супер-ЭВМ и нейрокомпьютеров», nikitina.vm@gmail.com, г. Таганрог, Россия

Семенякина А.А.,

к.т.н., ООО «Научно-исследовательский центр супер-ЭВМ и нейрокомпьютеров», j.a.s.s.y@mail.ru, г. Таганрог, Россия

МАТЕМАТИЧЕСКОЕ МОДЕЛИРОВАНИЕ ИЗМЕНЕНИЯ ГИДРОХИМИЧЕСКИХ ПОКАЗАТЕЛЕЙ КАЧЕСТВА ВОД ПРИБРЕЖНЫХ СИСТЕМ НА СУПЕР-ЭВМ*

Аннотация: В работе проводится построение и исследование многовидовых пространственно-неоднородных математических 3D моделей взаимодействия планктона и рыб, а также моделей изменения нефтяного фракционного состава вод при аварийных разливах, предназначенные для моделирования экологического состояния прибрежных систем. Верификация и калибровка разработанных моделей выполнена на основе данных спутникового мониторинга, экспедиционных данных. На базе многопроцессорных вычислительных систем разработано программное обеспечение, предназначенное для контроля гидрохимических показателей состава воды прибрежной зоны водоемов, анализа и дальнейшей разработки рекомендаций по устойчивому развитию и улучшению качества вод прибрежных экосистем.

Ключевые слова: математическое моделирование, гидрохимические параметры, нефтяное пятно, прибрежная система, супер-ЭВМ.

A.V. Nikitina,

Doctor of Technical Sciences, Associate Professor, Research Center for Supercomputers and Neurocomputers, LLC, nikitina.vm@gmail.com, Taganrog, Russia

A.A. Semenyakina,

Candidate of Technical Sciences, Research Center for Supercomputers and Neurocomputers, j.a.s.s.y@mail.ru, Taganrog, Russia

MATHEMATICAL MODELING OF CHANGES IN HYDROCHEMICAL INDICATORS OF WATER QUALITY COASTAL SYSTEMS ON A SUPER-COMPUTER

Abstract. In the work is carried out the construction and study of multiple species spatially inhomogeneous 3D mathematical models of interaction between plankton and fish, as well as models of changes in the oil fractional composition of waters during accidental spills, designed to simulate the ecological state of coastal systems. The verification and calibration of the models developed are based on satellite monitoring data and expedition data. On the basis of multiprocessor computing systems, software has been developed for monitoring the hydrochemical indicators of the composition of the water of the coastal zone of water bodies, analyzing and further developing recommendations for the sustainable development and improvement of water quality of coastal ecosystems.

Key words: mathematical modeling, hydrochemical parameters, oil slick, coastal system, super-computer.

* Работа выполнена при поддержке РФФ, проект №17-11-01286.

Введение

Проблемы обеспечения стабильного функционирования водных экосистем, предотвращения уменьшения биологического разнообразия, предупреждения негативных последствий антропогенного воздействия, обеспечения безопасного водопользования в процессе социально-экономического развития территории приобретают в настоящее время особую актуальность. Возрастающая в настоящее время антропогенная нагрузка в виде увеличения объемов производства, многочисленных сбросов сточных вод, производственных отходов и загрязняющих веществ способствуют ухудшению экологи-химического состояния и качества вод прибрежных систем, эвтрофикации водоема, снижению уровня растворенного в воде кислорода, изменению в структуре сообществ планктона и влиянию на воспроизводство рыб.

Построение моделей гидробиологии мелководных водоемов требует развития, в свою очередь, пространственно-трехмерных моделей биогеохимических циклов с высоким разрешением. Важнейшими техногенными факторами, оказывающими значительное воздействие на экосистему водоема, являются: металлургические и химические отходы производственной деятельности и коммунальные загрязнённые сточные воды; нефть и нефтепродукты; траление дна, уничтожающее донные биоценозы; сложно регулируемый лов рыбы браконьерами; строительство водохранилищ; насыщенная химизация, загрязнение почвы и воды, засоление водоёма; увеличение неконтролируемости слива пестицидов в водоем, влекущее за собой цветение воды; усиленное строительство объектов вдоль побережья, не соответствующее экологическим нормам; дампинг и др.

Особую опасность для экосистемы водоемов, особенно прибрежной зоны, представляют нефтяные загрязнения. Нефть является продуктом длительного распада; она очень быстро покрывает поверхность воды тонким слоем; образовавшаяся нефтяная плёнка ограничивает доступ света и воздуха. В результате происходит нарушение естественных процессов и взаимосвязей живых организмов, негативно влияя на устойчивое развитие прибрежных систем.

В этой связи возникает необходимость в разработке эффективных методов построения прогнозов изменения гидрохимических показателей качества вод прибрежных экосистем, особенно при аварийных загрязнениях, в условиях ограниченного времени.

Экспедиционные исследования

Калибровка и верификация разработанных моделей проводились на основе экологических данных по Азовскому морю, полученных в ходе научно-исследовательских экспедиций, проводимых учеными Южного федерального университета (ЮФУ), начиная с 2000 года [1]. В июле 2017 г. сотрудниками Донского государственного технического университета,

ЮФУ, Южного научного центра РАН на акватории Азовского моря были проведены экспедиционные исследования на научно-исследовательском судне «Денеб»[2], основная задача которых заключалась в комплексном изучении текущего состояния и пространственно-временных изменений гидробиологического, гидрологического и гидрохимического режимов Азовского моря и Таганрогского залива (рис. 1).

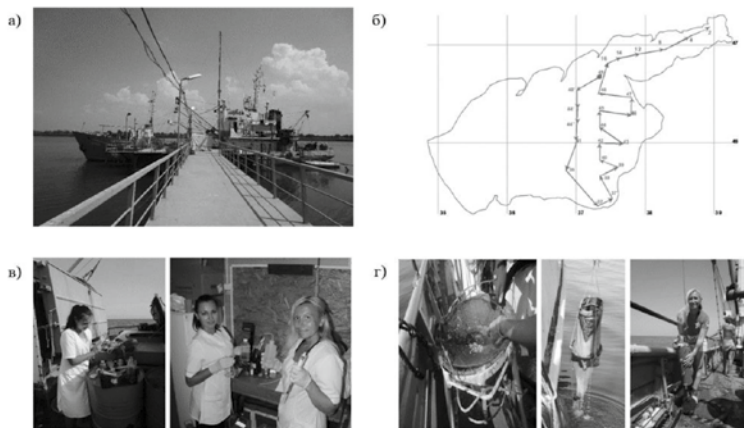
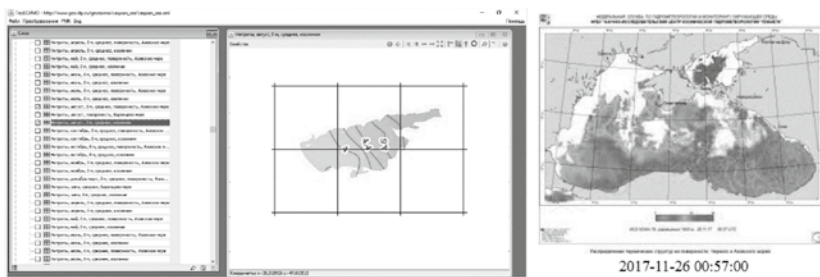


Рис. 1. Экспедиционные исследования 2017 г.: а) НИС «Денеб»
 б) маршрут движения судна; в) гидрохимические исследования;
 г) биологические исследования

Проведенная морская экспедиция была условно разделена по районам исследования на две группы: Таганрогский залив, Азовское море. В Таганрогском заливе изучались термогалинная структура и гидрохимические параметры в комплексе с изучением динамики планктонного и бентосного сообществ Таганрогского залива. Использовался типовой состав океанологических наблюдений: измерение метеорологических величин (температура воздуха, скорость и направления ветра) и регистрация гидрометеорологических явлений; местоположение станции (координаты, время по GPS), глубина (эхолот Furuno FCV 1200L), измерение прозрачности воды (диск Секки); отбор проб воды для определения гидрохимических параметров на борту судна (температура воды, рН (экспресс метод), кислород (титрование), биогенные элементы) и для определения общей взвеси и взвешенного органического вещества; отбор проб зообентоса модифицированным дно черпателем Петерсена в 3-х кратной повторности.

В качестве входных данных для моделирования гидрофизических процессов помимо экспедиционных данных, литературных источников использовались системы портала «Аналитические ГИС», разработанного Институтом проблем передачи информации РАН (ИППИ РАН, г. Москва) [3] и данные НИЦ «Планета» [4] (рис. 2).



а) б)
 Рис. 2. Экологические данные Азово-Черноморского бассейна.

а) Данные портала «Аналитические ГИС»;
 б) спутниковый снимок Азовского моря базы данных НИЦ «Планета».

Постановка задачи

Модель биохимической трансформации биогенных веществ (форм фосфора, азота и кремния) поступающих в Азовское море со стоками рек, включает три вида фитопланктонных водорослей: зеленой – *Chlorellavulgaris*, синезеленой – *Aphanizomenonflos-aquae*, диатомовой – *Skeletonemacostatum*, основана на системе уравнений транспорта биогенных веществ для каждого модельного блока q_i имеет вид [5-7, 8]:

$$\frac{\partial q_i}{\partial t} + \text{div}(\mathbf{U}, q_i) = \text{div}(\mathbf{k}_i, \text{grad } q_i) + R_i, \mathbf{k}_i = \{\mu_i, \nu_i, v_i\}, i = \overline{1, 10}, \quad (1)$$

где q_i – концентрация i -ой компоненты, u, v, w – компоненты вектора скорости водного потока, $\mathbf{u} = \{u, v, w\}$, $\mathbf{U} = \mathbf{u} + \mathbf{u}_{0i}$, $\mathbf{U} = \{U_1, U_2, U_3\}$ – скорость конвективного переноса вещества; u_{0i} – скорость осаднения i -й субстанции; R_i – химико-биологический источник, индекс i указывает на вид субстанции: 1 – 3 – индексы для концентраций водорослей *Chlorellavulgaris*, *Aphanizomenonflos-aquae* и *Skeletonemacostatum* соответственно; 4 – PO_4 , 5 – POP , 6 – DOP , 7 – NO_3 , 8 – NO_2 , 9 – NH_4 , 10 – Si ; PO_4 – фосфаты; POP – взвешенный органический фосфор; DOP – растворенный органический фосфор; NH_4 – аммоний; NO_2 – нитриты; NO_3 – нитраты; Si – растворенный неорганический кремний; μ_i, ν_i – коэффициенты диффузии в горизонтальном и вертикальном направлениях.

Модели наблюдений для описания химико-биологических источников [22]:

$$R_5 = \sum_{i=1}^3 s_p K_{D,i} q_i - K_{PD} q_5 - K_{PN} q_5, \quad R_6 = \sum_{i=1}^3 s_p K_{E,i} q_i + K_{PD} q_5 - K_{DN} q_6,$$

$$R_7 = \sum_{i=1}^3 s_N C_i (K_{R,i} - 1) \frac{f_N^{(1)}(q_7, q_8)}{f_N(q_7, q_8, q_9)} \cdot \frac{q_8}{q_7 + q_8} q_i + K_{23} q_8,$$

$$R_8 = \sum_{i=1}^3 s_N C_i (K_{R,i} - 1) \frac{f_N^{(1)}(q_7, q_8)}{f_N(q_7, q_8, q_9)} \cdot \frac{q_7}{q_7 + q_8} q_i + K_{42} q_9 - K_{23} q_8,$$

$$R_8 = \sum_{i=1}^3 s_N C_i (K_{R,i} - 1) \frac{f_N^{(1)}(q_7, q_8)}{f_N(q_7, q_8, q_9)} \cdot \frac{q_7}{q_7 + q_8} q_i + K_{42} q_9 - K_{23} q_8,$$

$$R_9 = \sum_{i=1}^3 s_N C_i (K_{R,i} - 1) \frac{f_N^{(2)}(q_9)}{f_N(q_7, q_8, q_9)} q_i - K_{42} q_9, R_{10} = s_{Si} K_{D,3} q_3,$$

где $K_{R,i}$ – удельная скорость дыхания фитопланктона; $K_{E,i}$ – удельная скорость отмирания фитопланктона; $K_{E,i}$ – удельная скорость экскреции фитопланктона; K_{PD} – удельная скорость автолиза *POP*; K_{DN} – коэффициент фосфатофикации *POP*; K_{DN} – коэффициент фосфатофикации *DOP*; K_{42} – удельная скорость окисления аммония до нитритов в процессе нитрификации; K_{23} – удельная скорость окисления нитритов до нитратов в процессе нитрификации, s_P, s_{Si}, s_{Si} – нормировочные коэффициенты между содержанием N, P, Si в органическом веществе.

Скорость роста фитопланктонных водорослей определяется выражениями:

$$C_i = K_{N,i} \min \{f_P(q_4), f_N(q_7, q_8, q_9)\}, i = \overline{1,2}, C_3 = K_{N,3} \min \{f_P(q_4), f_N(q_7, q_8, q_9), f_{Si}(q_{10})\}$$

где $K_{N,i}$ – максимальная удельная скорость роста фитопланктона.

Опишем зависимости содержания биогенных веществ:

– для фосфора $f_{Si}(q_{10}) = \frac{q_{10}}{q_{10} + K_{10}}$, где K_{10} – константа полунасыщения фосфатами;

– для кремния $f_{Si}(q_{10}) = \frac{q_{10}}{q_{10} + K_{10}}$, где K_{10} – константа полунасыщения кремнием;

– для азота $f_N(q_7, q_8, q_9) = f_N^{(1)}(q_7, q_8) + f_N^{(2)}(q_9) = \frac{(q_7 + q_8) \exp(-K_{psi} q_9)}{K_7 + (q_7 + q_8)} + \frac{q_9}{K_9 + q_9}$

где K_7 – константа полунасыщения нитратами, K_9 – константа полунасыщения аммонием, K_{psi} – коэффициент ингибирования аммония.

Для системы (1) необходимо в любой момент времени задавать вектор скоростей водного потока, а также начальные значения функций q_i :

$$q_i(x, y, z, 0) = q_i^0(x, y, z), (x, y, z) \in \overline{G}, t = 0, i = \overline{1,10}. \quad (2)$$

Пусть граница \sum цилиндрической области G является кусочно-гладкой и $\sum = \sum_H \cup \sum_0 \cup \sigma$, где \sum_H – поверхность дна водоема, \sum_0 – невозмущенная поверхность водной среды, σ – боковая (цилиндрическая) поверхность. Пусть u_n – нормальная по отношению к \sum составляющая вектора скорости водного потока, n – вектор внешней нормали к \sum . Для концентраций q_i будем предполагать:

– на боковой границе $q_i = 0$, если $u_n < 0, i \in M$; $\frac{\partial q_i}{\partial n} = 0$, если $u_n \geq 0, i = \overline{1,10}$;

– на Σ_o : $\frac{\partial q_i}{\partial z} = \varphi_i, i = \overline{1,10}$; на дне: $\frac{\partial q_i}{\partial z} = \varepsilon_{1,i} q_i, i = \overline{1,3}, \frac{\partial q_i}{\partial z} = \varepsilon_{2,i} q_i; i = \overline{4,10}$,

где φ – заданные функции; $\varepsilon_{1,i}, \varepsilon_{2,i}$ – неотрицательные постоянные, $\varepsilon_{1,i}, i = \overline{1,3}$ учитывают опускание водородней на дно и их затопление; $\varepsilon_{2,i}, i = \overline{4,10}$ учитывают поглощение питательных веществ донными отложениями.

Для моделирования изменения нефтяного фракционного состава в прибрежных водоемах в результате аварийного разлива разработана пространственно-трехмерная непрерывная модель транспорта нефти и продуктов ее переработки вида [9]:

$$(c_i)'_t + u(c_i)'_x + v(c_i)'_y = \left(\mu^*(c_i)'_x \right)'_x + \left(\mu^*(c_i)'_y \right)'_y - \alpha_i X_i M_i^m - \beta_i M,$$

$$M'_t + uM'_x + vM'_y = (\mu M'_x)'_x + (\mu M'_y)'_y + \gamma M - \lambda M,$$

$$(\varphi_i)'_t + u(\varphi_i)'_x + v(\varphi_i)'_y + w(\varphi_i)'_z = \left(\mu(\varphi_i)'_x \right)'_x + \left(\mu(\varphi_i)'_y \right)'_y + \left(\mu(\varphi_i)'_z \right)'_z + f, \quad (4)$$

$$(c_i)'_n \Big|_{(x,y) \in \Gamma} = 0, \quad M'_n \Big|_{(x,y) \in \Gamma} = 0, \quad (\varphi_i)'_n \Big|_{(x,y,z) \in \Gamma(z=0)} = 0, \quad (\varphi_i)'_z \Big|_{z=0} = \varepsilon_i M_i^m, \quad (5)$$

где c_i – концентрация i -ой фракции нефти, находящейся в поверхностном слое, $i = \overline{1,k}$; ρ_o, ρ – плотности воды и нефти соответственно; $\mu^* = \mu + (\rho - \rho_o) gh^3/\mu$ – функция, описывающая процесс разложения загрязняющего вещества, μ – коэффициент диффузии; f – функция источника ЗВ; g – ускорение свободного падения; $h = \sum c_i$ – толщина пленки, $i = \overline{1,k}$; S_i – растворимость в воде i -й фракции нефти, $i = \overline{k+1,l}$; $K_E = 2,5 \cdot 10^3 U^{0,78}$ – коэффициент массопереноса для углеводорода; $\alpha_i = \alpha_i(P_i, S_i) = K_E P_i / (RT) + K_D S_i$, $R = 8,314$ – универсальная газовая постоянная; $\beta_i = \beta_i(c_i) = \mu_m c_i / \{q(c_i + K_s)\}$, $\gamma_i = \gamma_i(c_i) = \mu_m c_i / (c_i + K_s)$, $\varepsilon_i = \varepsilon_i(S_i, X_i) = K_D S_i X_i$; M_i^m – значение молярной массы i -ой фракции; U – скорость ветра относительно воды; u, v, w – компоненты вектора скорости водного потока; P_i – давление паров i -го компонента; X_i – молярная доля i -ой фракции нефти, равная $v_i / \sum v_i, i \in \overline{1,l}$; v_i – количество вещества i -ой фракции нефти; T – температура окружающей среды над поверхностью пятна; $K_D = kK_{Do}$ – коэффициент массопереноса растворения; K_{Do} – начальное значение коэффициента массопереноса растворения; K_s – значение коэффициент насыщения; k – коэффициент, который зависит от волнения исследуемой прибрежной системы; M – концентрация микроорганизмов; q – значение коэффициента пропорциональности между количеством бактерий и поглощенным субстратом; μ_m – значение максимальной скорости роста микроорганизмов; λ – скорость отмирания клеток; φ_i – концентрация i -ой фракции нефти, находящейся в растворенном состоянии, $i = \overline{k+1,l}$; γ – область, описывающая поверхностные слои водоема; Γ – граница расчетной области, описы-

вающей водоем.

Для решения вышеописанной системы уравнений были определены начальные условия в случае одномоментного выброса нефти:

$$c_i \Big|_{t=0, (x,y) \in \sigma} = c_{i0}, \quad c_i \Big|_{t=0, (x,y) \notin \sigma} = 0, \quad (6)$$

здесь c_0 – концентрация нефти в исследуемой области; σ – часть поверхности водоема, покрытая нефтяным пятном, $\sigma \subset \gamma$.

Модель позволяет выявлять динамику распространения нефтяных пятен в акватории прибрежных систем, а также изменение гидрохимического состава вода под влиянием различных гидродинамических и гидрофизических факторов. Благодаря учету значительного количества факторов, влияющих на распределение нефтяных загрязнений в акватории прибрежных систем, а также методам математического моделирования, их реализующих, данная модель () – () позволяет повысить точность прогнозного моделирования на 10 – 30% [10].

Поля скоростей водного потока, рассчитанные на основе математической модели гидродинамики [11, 12], относятся к входным данным для разработанных моделей транспорта биогенных, взвешенных, загрязняющих веществ, включая нефть и нефтепродукты, а также донных материалов.

Модифицированный попеременно-треугольный метод (МПТМ)

Сеточные уравнения, полученные в результате конечно-разностных аппроксимаций поставленных задач с помощью схемы с весами [13] можно записать в матричном виде [14]:

$$Ax = f, \quad (7)$$

где A – линейный, положительно определенный оператор ($A > 0$). Для нахождения решения задачи (8) будем использовать неявный итерационный процесс:

$$B \frac{x^{m+1} - x^m}{\tau_{m+1}} + Ax^m = f \quad (8)$$

В уравнении (9) m – номер итерации, $\tau > 0$ – итерационный параметр, а B – некоторый обратимый оператор, который называется преобуславливателем или стабилизатором. Обращение оператора B в (9) должно быть существенно проще, чем непосредственное обращение исходного оператора A в (8). При построении B исходили из аддитивного представления оператора A_0 – симметричной части оператора A :

$$A_0 = R_1 + R_2, \quad R_1 = R_2^*, \quad (9)$$

где $A = A_0 + A_1$, $A_0 = A_0^*$, $A_1 = -A_1^*$.

Оператор-преобуславливатель запишется в следующем виде:

$$B = (D + \omega R_1)D^{-1}(D + \omega R_2), \quad D = D^* > 0, \quad \omega > 0 \quad (10)$$

где D – некоторый оператор.

Соотношения (9), (10) задают модифицированный попеременно-треугольный метод (МПТМ) решения задач, если определены операторы R_1, R_2 и указаны способы определения параметров τ_{m+1}, ω и оператора D .

Алгоритм адаптивного модифицированного попеременно-треугольного метода минимальных поправок для расчета сеточных уравнений с несамосопряженным оператором имеет вид:

$$r^m = Ax^m - f, \quad B(\omega_m)w^m = r^m, \quad \tilde{\omega}_m = \sqrt{\frac{(Dw^m, w^m)}{(D^{-1}R_2w^m, R_2w^m)}}, \quad (11)$$

$$s_m^2 = 1 - \frac{(A_0w^m, w^m)^2}{(B^{-1}A_0w^m, A_0w^m)(Bw^m, w^m)}, \quad k_m = \frac{(B^{-1}A_1w^m, A_1w^m)}{(B^{-1}A_0w^m, A_0w^m)}, \quad \theta_m = \frac{1 - \sqrt{\frac{s_m^2 k_m}{(1+k_m)}}}{1 + k_m (1 - s_m^2)},$$

$$\tau_{m+1} = \theta_m \frac{(A_0w^m, w^m)}{(B^{-1}A_0w^m, A_0w^m)}, \quad x^{m+1} = x^m - \tau_{m+1}w^m, \quad \omega_{m+1} = \tilde{\omega}_m,$$

где r^m – вектор невязки, w^m – вектор поправки, в качестве оператора D используется диагональная часть оператора A .

Оценка скорости сходимости разработанного метода имеет вид:

$$\rho \leq \frac{v^* - 1}{v^* + 1}, \quad v^* = v \left(\sqrt{1+k} + \sqrt{k} \right)^2, \quad k = \frac{(B^{-1}A_1\omega^m, A_1\omega^m)}{(B^{-1}A_0\omega^m, A_0\omega^m)}, \quad (12)$$

где v – число обусловленности оператора C_0 , $C_0 = B^{-1/2}A_0B^{-1/2}$.

В данном исследовании были разработаны и численно реализованы параллельные алгоритмы методов решения поставленных задач [15, 16].

Программный комплекс

На базе многопроцессорной вычислительной системы (МВС)ЮФУ разработан программный комплекс [17, 18], предназначенный для математического моделирования сценариев развития прибрежных экосистем на примере Азово-Черноморского бассейна (рис. 3). Комплекс позволяет выполнять оперативные прогнозы движения турбулентных потоков водной среды – поля скоростей водной среды на сетках с высокой разрешающей способностью; учитывает физические параметры: силу Кориолиса, турбулентный обмен, сложную геометрию дна и береговой линии, испарение, стоки рек, сгонно-нагонные явления, ветровые течения и трение о дно; обеспечивает выполнение функций: расчет поля скорости без учета давления; расчет гидростатического давления (используется в качестве начального приближения для гидродинамического давления); расчет гидродинамического давления; расчет трехмерного поля скорости водного потока.

Разработанный для МВС программный комплекс включает: блок управления, базы океанологических и метеорологических данных, системы интерфейсов, системы ввода – вывода и визуализации. Программный продукт обладает удобным пользовательским интерфейсом и имеет универсальный характер: привязка его к условиям конкретных объектов и районов осуществляется на уровне входной информации, т.е. для практического использования модулей требуется создание специальной информационной базы, содержащей сведения о физико-географических и климатических условиях исследуемых объектов, о параметрах, определяющих источники примесей.

Выходными параметрами программного комплекса являются: шаги по пространственным координатам, погрешность решения сеточных уравнений, размеры расчетной сетки, временной интервал, интенсивность испарения, начальные распределения компонент вектора скорости движения водной среды и давления. Разработанный комплекс программ допускает внедрение новых расчетных функций, в частности, в данный комплекс были встроены программные блоки, предназначенные для моделирования транспорта нефтепродуктов с учетом испарения легкой, нейтральной и неиспаряющейся псевдофракции нефтяного пятна, растворения нефтяного пятна и биоразложения [19].

Разработанный программный комплекс позволяет осуществлять:

- совершенствование и внедрение системы комплексного мониторинга в водоемах (наблюдение, оценка и прогноз состояния режима экосистем);
- разработку, согласование предложений и мероприятий по обеспе-

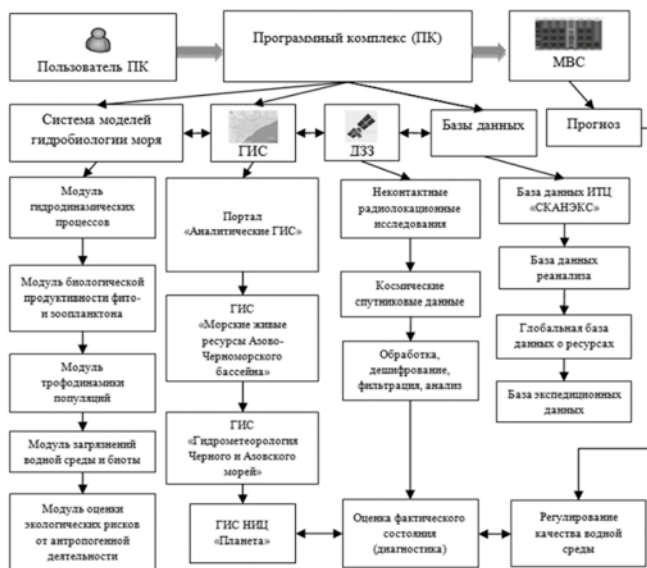


Рис. 3. Программный комплекс

чению оптимального режима, сохранения биоразнообразия промышленных ресурсов, прибрежных экосистем;

– совершенствование методологии природоохранных исследований, разработка новых, апробация и внедрение перспективных методов изучения состояния прибрежных экосистем и отдельных компонентов;

– разработку и совершенствование методов диагностики токсического воздействия биогенных веществ на гидробионты, в том числе ранней и дифференциальной диагностики токсикоза, а также поиск средств антидотной защиты водных экосистем;

– организацию и проведение исследований по выявлению тенденций и закономерностей изменения состояния прибрежных экосистем под воздействием антропогенных факторов, разработку предложений и мероприятий по снижению и предупреждению таких воздействий.

Результаты экспериментальных исследований

Приведенные ниже рис. 4, 5 отражают влияние структур течений водного потока в Азовском море на распределение загрязняющего биогенного вещества и фитопланктона [20, 21]. На рис. 4. показаны результаты расчета концентрации загрязняющего биогенного вещества для модели (1) – (3) (начальное распределение полей течений при северном ветре).

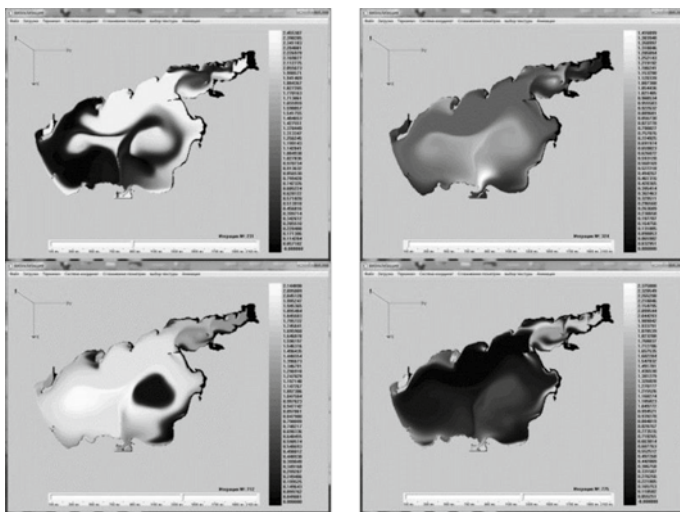


Рис. 4. Распределение концентрации загрязняющих биогенных веществ

Результаты моделирования динамики фитопланктона в Азовском море представлены на рис. 5. (N – номер итерации, начальное распределение полей течений водного потока при северном направлении ветра). Белым цветом выделены максимальные значения концентраций биогенного вещества (азота) и фитопланктона черным – минимальные.

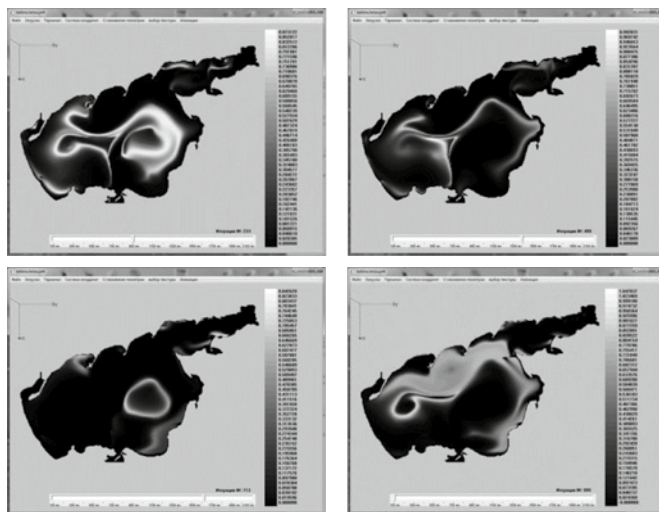
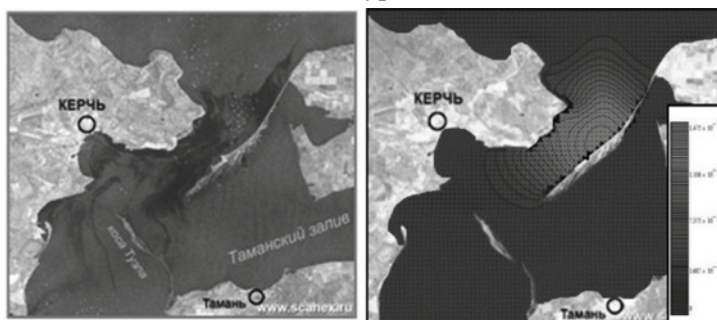


Рис. 5. Распределение концентрации фитопланктона в различные моменты времени

С помощью разработанного программного комплекса также была проведена реконструкция аварийного разлива нефти, произошедшего в Керченском проливе в ноябре 2007 г. [22]. При анализе полученных результатов моделирования получено соответствие имеющимся радиолокационным снимкам и натурным данным участка катастрофы (рис. 6) на качественном и количественном уровнях.



а) б)

Рис. 6. Сравнение результатов моделирования с данными дистанционного зондирования Земли: а) спутниковый снимок ИСЗ Radarset-1 участка разлива нефти на 16.11.2007; б) результат работы программного комплекса «Azov3d»

Важнейшим этапом в проведении любого полноценного прогнозного исследования является оценка качества построенного предсказания, выполненная на основе оценки качества информационной пригодности разработанных моделей и их прогностической ценности. В данном исследовании критерием проверки адекватности предложенных моделей

служила оценка погрешности моделирования с одновременным учетом натуральных данных по имеющимся n замерам, которая вычислялась по формуле: $\delta = \sqrt{\sum_{k=1}^n (S_{k \text{ nat}} - S_k)^2} / \sqrt{\sum_{k=1}^n S_{k \text{ nat}}^2}$, где $S_{k \text{ nat}}$ – значение концентрации, полученное с помощью натуральных экспедиционных измерений; S_k – значение, рассчитанное с помощью разработанной модели. Рассчитанные при различных ветровых ситуациях концентрации загрязняющих веществ и планктона принимались к рассмотрению, если относительная погрешность не превышала 30%.

Заключение

Для разработки комплекса мероприятий по обеспечению контроля гидрохимического состава прибрежных систем, а также рекомендаций по принятию мер и ликвидации негативных последствий загрязнения водоемов, в том числе, для предотвращения чрезвычайных ситуаций, был разработан программный комплекс, объединяющий в себе многовидовые пространственно-неоднородные математические 3D модели и многолетние постоянно пополняемые базы экологических данных.

На основе проведенных экспедиционных исследований, а также данных спутникового мониторинга земли была проведена верификация разработанных моделей развития экосистемы Азовского моря при различных видах загрязнений.

Отличительными особенностями разработанных алгоритмов решения поставленных задач является оптимальный раздел объема вычислительной работы, высокая производительность, достигаемая за счет использования эффективных численных методов решения сеточных уравнений, ориентированных для применения на параллельных вычислительных системах в реальном и ускоренном масштабах времени достоверность, обусловленная учетом определяющих физических факторов (сила Кориолиса, турбулентный обмен, сложная геометрия дна и береговой линии, испарение, стоки рек, динамическое перестроение расчетной области, ветровые напряжения и трение о дно, отклонение значения поля давления от гидростатического приближения) и точность получаемых результатов моделирования, достигаемая использованием подробных расчетных сеток, учитывающих частичную заполненность расчетных ячеек, а также отсутствием неконсервативных диссипативных слагаемых и нефизичных источников (стоков) [23–26].

Сравнение работы созданного программного комплекса для МВС, реализующего сценарии развития экологической обстановки в Азовском море с использованием численной реализации разработанных модельных задач биологической кинетики и транспорта загрязняющих веществ, с подобными работами показало, что в результате точность прогнозов изменения концентраций загрязняющих веществ и, как следствие, гидрохимических показателей качества вод прибрежных систем удалось повысить на 10 – 30% в зависимости от решаемой модельной задачи водной экологии [27].

Таким образом, разработанный программный комплекс может эффективно применяться для моделирования изменения гидрохимических показателей вследствие аварийных загрязнений и дальнейшей разработки рекомендаций по улучшению и сохранению экологического равновесия прибрежных систем.

Список литературы

1. Сухинов А.И., Никитина А.В. Математическое моделирование и экспедиционные исследования качества вод в Азовском море // Известия ЮФУ. Технические науки. – 2011. – № 8(121). – С. 62–73.
2. Никитина А.В., Сухинова Т.Г., Проценко С.В., Семенякина А.А., Бедная Т.А. Эколого-гидрофизическое обоснование влияния коэффициента вертикального турбулентного обмена на содержание растворенного кислорода в придонном слое мелководного водоема // Успехи современного естествознания. – 2018. – № 1. – С. 115-119.
3. Аналитические ГИС Online. URL: <http://geo.iitp.ru/index.php>.
4. Научно-исследовательский центр космической гидрометеорологии "Планета". URL: <http://planet.iitp.ru/index1.html>.
5. Сухинов А.И., Никитина А.В., Чистяков А.Е., Семенов И.С., Семенякина А.А., Хачунц Д.С. Математическое моделирование процессов эвтрофикации в мелководных водоемах на многопроцессорной вычислительной системе // Вестник Южно-Уральского государственного университета. Серия: Вычислительная математика и информатика. – 2016. – Т. 5, № 3. – С. 36-53.
6. Sukhinov A., Chistyakov A., Isayev A., Nikitina A., Sumbaev V., Semenyakina A. Complex of Models, High-Resolution Schemes and Programs for the Predictive Modeling of Suffocation in Shallow Waters // Communications in Computer and Information Science. – 2017. – Т. 753. – С. 169-185.
7. Sukhinov A., Nikitina A., Chistyakov A., Sumbaev V., Abramov M., Semenyakina A. Predictive modeling of suffocation in shallow waters on a multiprocessor computer system // 14th International Conference on Parallel Computing Technologies, PaCT 2017; Nizhny Novgorod; Russian Federation; 4 September 2017 to 8 September 2017. Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics). Vol. 10421 LNCS, 2017. –P. 172-180. DOI10.1007/978-3-319-62932-2_16
8. Сухинов А.И., Белова Ю.В. Математическая модель трансформации форм фосфора, азота и кремния в движущейся турбулентной водной среде в задачах динамики планктонных популяций // Инженерный вестник Дона. – 2015. – Т. 37, № 3. – С. 50.
9. Сухинов А.И., Никитина А.В., Семенякина А.А., Чистяков А.Е. Комплекс моделей явных регуляризованных схем повышенного порядка точности и программ для предсказательного моделирования транспорта нефтепродуктов // Вестник Уфимского государственного авиационного технического университета. – 2016. – Т. 20, № 3(73). – С. 143-152.
10. Sukhinov A., Nikitina A., Semenyakina A., Chistyakov A. Complex of models, explicit regularized schemes of high-order of accuracy and applications for predictive modeling of after-math of emergency oil spill // Proceedings of the 10th Annual International Scientific Conference on Parallel Computing Technologies, 2016. – Vol. 1576. – P. 308-319. ISSN: 16130073

11. Сухинов А.И., Чистяков А.Е., Алексеенко Е.В. Численная реализация трехмерной модели гидродинамики для мелководных водоемов на супервычислительной системе // Математическое моделирование. – 2011. – Т. 23, № 3. – С. 3-21.
12. Сухинов А.И., Чистяков А.Е. Параллельная реализация трехмерной модели гидродинамики мелководных водоемов на супервычислительной системе // Выч. мет. Программирование. 2012. –Т. 13, № 1. – С. 290–297.
13. Сухинов А.И., Чистяков А.Е., Шишениа А.В. Оценка погрешности решения уравнения диффузии на основе схем с весами // Математическое моделирование. 2013. – Т. 25, № 11. – С. 53-64.
14. Сухинов А.И., Чистяков А.Е. Адаптивный модифицированный попеременно-треугольный итерационный метод для решения сеточных уравнений с несамосопряженным оператором// Математическое моделирование. – 2012. – Т.24, №1. – С. 3-20.
15. Сухинов А.И., Чистяков А.Е., Семенякина А.А., Никитина А.В. Параллельная реализация задач транспорта веществ и восстановления донной поверхности на основе схем повышенного порядка точности // Вычислительные методы и программирование: новые вычислительные технологии. – 2015. – Т.16. – С. 256-267.
16. Никитина А.В., Семенякина А.А., Чистяков А.Е. Параллельная реализация задачи диффузии-конвекции на основе схем повышенного порядка точности // Вестник компьютерных и информационных технологий. – 2016. – № 7 (146). С. 3-7.
17. Sukhinov, A.I., Nikitina, A.V., Semenyakina, A.A., Protsenko, E.A. Complex programs and algorithms to calculate sediment transport and multi-component suspensions on a multiprocessor computer system // Engineering journal of Don, vol. 38, no. 4 (38), p. 52 (2015).
18. Sukhinov A., Nikitina A., Semenyakina A., Chistyakov A. Complex of models, explicit regularized schemes of high-order of accuracy and applications for predictive modeling of after-math of emergency oil spill // Proceedings of the 10th Annual International Scientific Conference on Parallel Computing Technologies, 2016. – Vol. 1576. – P. 308-319. ISSN: 16130073.
19. Sukhinov A.I., Chistyakov A.E., Nikitina A.V., Semenyakina A.A., Korovin I.S., Schaefer G. Modelling of the Oil Spill Spread // Proceedings of the 5th International Conference on Informatics, Electronics and Vision (ICIEV). Dhaka, Bangladesh, 2016. – P. 1134-1139. – ISBN: 978-150901269-5. – DOI: 10.1109/ICIEV.2016.7760176.
20. Sukhinov A.I., Chistyakov A.E., Semenyakina A.A., Nikitina A.V. Numerical modeling of an ecological condition of the Sea of Azov with application of schemes of the raised accuracy order on the multiprocessor computing system // Computer researches and modeling. – 2016. – Т. 8, No. 1. – P. 151-168.
21. Nikitina A.V., Semenyakina A.A. Mathematical modeling of eutrophication processes in Azov Sea on supercomputers // Computational Mathematics and Information Technologies. – 2017. – Vol.1, No 1. – P. 82-101.
22. Semenyakina, A., Protsenko, S. Complex of parallel programs for modeling oil products transport in coastal systems // MATEC Web of Conferences, 2017. 132, 04016 (2017). 6 pp. DOI: 10.1051/mateconf/201713204016.

23. Сухинов А. И., Чистяков А. Е., Якобовский М. В. Точность численного решения уравнения диффузии-конвекции на основе разностных схем второго и четвертого порядков погрешности аппроксимации// Вестн. ЮУрГУ. Сер. Выч.матем. информ., 2016. – Т. 5, № 1. – С. 47-62.
24. Никитина А.В., Семенякина А.А., Чистяков А.Е., Проценко Е.А., Яковенко И.В. Применение схем повышенного порядка точности для решения задач биологической кинетики на многопроцессорной вычислительной системе // Фундаментальные исследования. – 2015. – № 12-3. – С. 500-504.
25. Korovin I.S., Sukhinov A.I., Chistyakov A.E., Semenyakina A.A., Protsenko E.A., Schaefer G. Parallel Implementation of Spatially Three-Dimensional Model of Two-Phase Fluid Filtration Based on the Improved Alternating Triangular Method // Proceedings of the 5th International Conference on Informatics, Electronics and Vision (ICIEV). Dhaka, Bangladesh, 2016. – P. 1143-1148. – ISBN: 978-150901269-5. – DOI: 10.1109/ICIEV.2016.7760178.
26. Sukhinov A.I., Khachunts D.S., Chistyakov A.E. A mathematical model of pollutant propagation in near-ground atmospheric layer of a coastal region and its software implementation // Computational Mathematics and Mathematical Physics. – 2015. – Т. 55, № 7. – P. 1216-1231.
27. Никитина А.В., Сухинов А.И., Угольницкий Г.А., Усов А.Б., Чистяков А.Е., Пучкин М.В., Семенов И.С. Оптимальное управление устойчивым развитием при биологической реабилитации Азовского моря // Математическое моделирование. – 2016. Т. 28, №7. – С. 96-106.

УДК 519.87:66.096.5

Орлова Н.С.,

кандидат технических наук, ЮМИ ВНЦ РАН, Финансовый университет при Правительстве РФ, Владикавказ, Россия, norlova.umi.vnc@gmail.com, www.smath.ru, <http://www.old.fa.ru>

МОДЕЛИРОВАНИЕ ВИБРОКИПЕНИЯ ПОЛИДИСПЕРСНОЙ СРЕДЫ

Аннотация: В работе для моделирования динамики полидисперсного виброкипящего слоя использовался метод дискретного элемента. Первые результаты такого моделирования показали, что полидисперсность заметно изменяет характеристики виброкипящего слоя. В связи с этим для полидисперсной среды необходимо дополнительное исследование влияния амплитуды и частоты колебаний полки, а также толщины слоя частиц на режимы виброкипания.

Ключевые слова: виброкипение, полидисперсность, моделирование, метод дискретного элемента.

N.S. Orlova,

Candidate of Technical Sciences, YMI VSC RAS, Financial University under the Government of the Russian Federation, Vladikavkaz, Russia, norlova.umi.vnc@gmail.com, www.smath.ru, <http://www.old.fa.ru>

MODELING OF VIBRATION OF A POLYSPISSIVE MEDIUM

Abstract: *In the work, the discrete element method was used to model the dynamics of a polydisperse vibro-boiling layer. The first results of such a simulation showed that polydispersity noticeably changes the characteristics of a vibro-boiling layer. In this regard, for a polydisperse medium, additional research is needed on the influence of the amplitude and oscillation frequency of the shelf, as well as the thickness of the particle layer on the vibro-boiling regimes.*

Keywords: *vibro-boiling, polydispersity, modeling, discrete element method.*

Введение. Процесс виброкипения широко используется в различных областях, например в процессах каталитического синтеза порошковых материалов, которые могут применяться в химических технологиях, фармацевтической и косметической промышленности, энергетической промышленности и при изготовлении биоматериалов. Одним из способов конструирования аппаратов, использующих процесс виброкипения, является математическое и компьютерное моделирование. В связи с этим важное значение имеют разработка и исследование математических моделей виброкипения.

В процессе виброкипения очень часто используется полидисперсный гранулированный материал. Но во многих работах, как правило, представлены результаты моделирования виброкипения монодисперсной среды [1-5]. Имеются также работы, посвященные исследованию моделей, описывающих процесс виброкипения полидисперсного слоя [6,7]. Но режимы виброкипения полидисперсной среды остаются не изученными до конца.

В данной работе для моделирования динамики полидисперсного виброкипящего слоя использовался дискретный подход, в частности метод дискретного элемента. При таком подходе движение частиц описывается как детерминированное движение их достаточно представительного дискретного набора. Дискретный подход непосредственным образом связан с реальными движениями отдельных частиц, поэтому его можно считать физически более естественным, чем континуальный подход. Кроме того в рамках дискретного подхода гораздо проще моделировать движение полидисперсных сред, но при этом требуются достаточно мощные вычислительные ресурсы, так как для имитации движения слоя частиц требуется проведение большого числа расчетов движения отдельных частиц.

Постановка задачи. Метод дискретного описания движения каждой отдельной частицы с учетом ее взаимодействия с соседними частицами получил название «Метод дискретного элемента» (МДЭ) (Discrete element method) [8]. В основе МДЭ лежат следующие уравнения [8]:

$$\begin{cases} m_i \frac{d\bar{V}_i}{dt} = m_i \bar{b} + \sum_{j=1}^k \bar{F}_{ij}; \\ I_i \frac{d\bar{\omega}_i}{dt} = \sum_{j=1}^k (\bar{T}_{ij} + \bar{M}_{ij}), \end{cases} \quad (1)$$

где m_i – масса частицы; i – индекс частицы, относительно которой рассматривается система уравнений (1); I_i – вектор линейной скорости центра масс; I_i – момент инерции; \bar{b} – вектор угловой скорости; \bar{b} – вектор массовой силы; \bar{F}_{ij} – внешняя сила, действующая на частицу i через контакт с частицей j ; j – индекс частицы, находящейся в контакте с частицей i ; k – число частиц в контакте с частицей i ; \bar{M}_{ij} – внешний крутящий момент, связанный с контактным взаимодействием частиц i и j ; \bar{M}_{ij} – момент сопротивления качению. Более подробное описание модели представлено в работе [8].

Для проведения вычислений по виброкипению использовался свободный открытый программный код LIGGGHTS [9], в котором реализован МДЭ. При проведении параллельных вычислений на четырехъядерном процессоре с частотой 3,1 Ггерц время расчета одного варианта составляло примерно 20 часов.

Рассматривалось виброкипение как монодисперсной [8,10], так и полидисперсной среды. В случае виброкипения монодисперсной среды использовались частицы силикагеля, средний диаметр которых равен 4 мм. Толщина слоя составляла 15 мм. Результаты расчетов были получены при амплитуде колебаний 2 мм, частоте 30 Гц и 36 Гц, а также при амплитуде 2,5 мм и частоте 33 Гц. Следует отметить, что в работе [8] представлены результаты тестирования модели на основе МДЭ. Рассматривалось виброкипение относительно тонких слоев. Получено удовлетворительное совпадение результатов расчетов с экспериментальными данными.

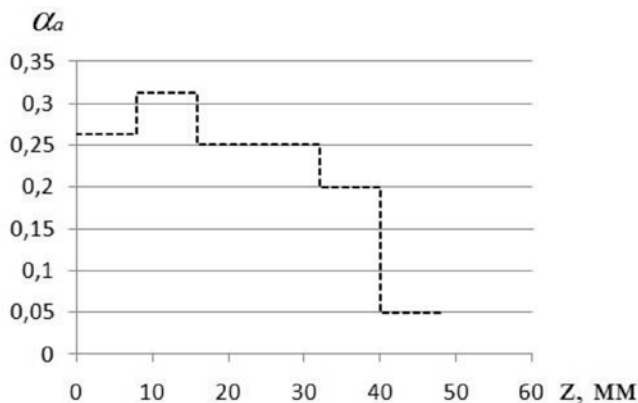


Рис. 1. Распределение объемной доли частиц по высоте

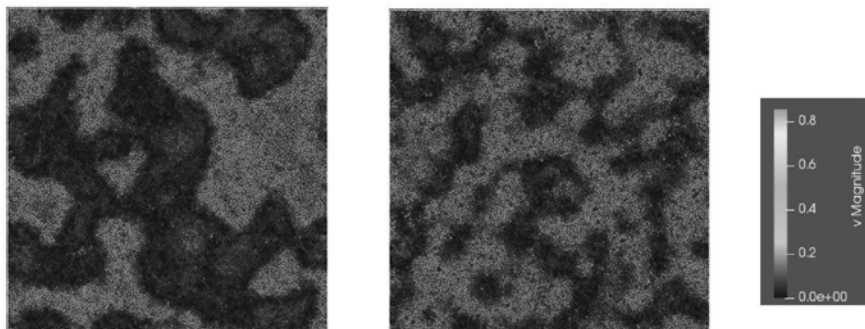


Рис. 2. Виброкипение монодисперсного слоя. Вид сверху.

Результаты моделирования. На рис. 1 представлены результаты моделирования среднего по времени распределения объемной доли частиц (α_v). Результаты получены при $A = 2,5$ мм и $f = 33$ Гц. Следует отметить, что в качестве значения коэффициента восстановления в случае столкновения частиц друг с другом использовалось значение $e = 0,15$, полученное по результатам экспериментов [8]. Методика обработки результатов моделирования подробно представлена в работе [8]. На рис. 2 представлено распределение скорости частиц (вид сверху) для случая, когда амплитуда колебаний равна 2 мм, частота 30 Гц (рис. 2 а) и 36 Гц (рис.2 б).

Результаты, представленные на рис. 1, показывают неравномерное распределение объемной доли частиц в слое. Следует отметить, что в процессе виброкипения монодисперсных слоев, наблюдается волнообразная поверхность и отдельные всплески над поверхностью слоя [11,12], как это видно из рис. 2. В результате предыдущих исследований [11] было установлено, что на степень расширения виброкипящего слоя (отношение максимальной высоты виброкипящего слоя к его на-

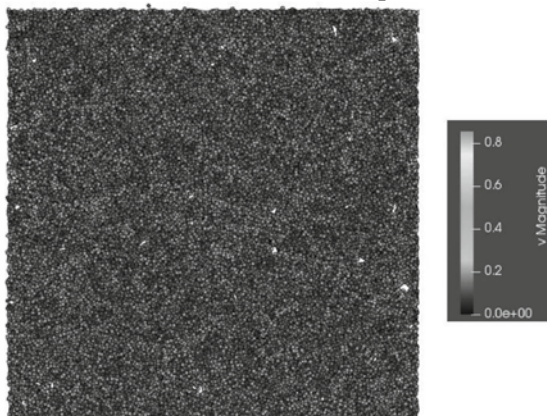


Рис. 3. Виброкипение полидисперсного слоя. Распределение частиц по скорости. Вид сверху.

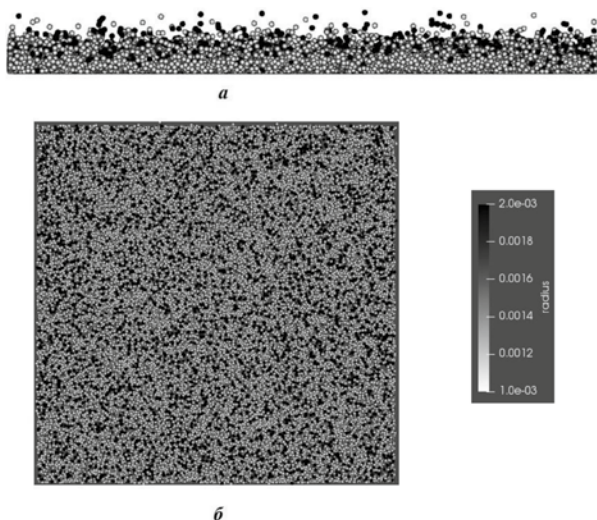


Рис. 4. Виброкипение полидисперсного слоя. Распределение частиц по размеру.

чальной толщине) большее влияние оказывает амплитуда колебаний, частота колебаний влияет на количество всплесков в виброкипящем слое. С увеличением частоты колебаний количество всплесков растет (рис. 2 б).

На рис. 3 представлены результаты моделирования виброкипания полидисперсной среды при амплитуде колебаний 2 мм и частоте 36 Гц. Приводится распределение скорости частиц. Использовались частицы диаметром 2 мм и 4 мм в соотношении 3:7. Общее количество частиц равно примерно 50000 (начальная толщина слоя примерно равна 15 мм).

На рис. 4 представлено распределение частиц в слое по их размеру. На рис. 4 а представлен вид сбоку, на рис. 4 б – вид сверху. Из рис. 4 видно, что более мелкие частицы расположены, в основном, в нижней части слоя, а более крупные – в верхней. При этом высоко над поверхностью слоя наблюдаются единичные частицы как мелких размеров, так и крупных. Слой можно поделить на две области, в первой нижней области расположены только мелкие частицы, диаметром 2 мм, во второй области (средняя и верхняя части слоя) расположены как мелкие, так и крупные частицы. Средняя высота первой области примерно равна 6 мм. Средняя высота слоя в целом не превышает 25 мм, как и в случае моделирования динамики монодисперсной среды, состоящей из частиц диаметром 4 мм.

Важно отметить, что структура слоя в случае полидисперсной среды значительно отличается от структуры монодисперсной среды в процессе виброкипания. Над поверхностью слоя полидисперсной среды всплески практически не визуализируются. Это видно из рис. 3 – 4.

Полученные результаты показывают, что виброкипение полидисперсной среды несколько отличается от виброкипания монодисперсной среды. Поэтому режимы виброкипания полидисперсной среды требуют более подробного изучения.

Заключение. Таким образом, использование метода дискретного элемента позволяет моделировать динамику виброкипящего слоя с учетом полидисперсности материала. Первые результаты такого моделирования показали, что полидисперсность заметно изменяет характеристики виброкипящего слоя. В связи с этим для полидисперсной среды необходимо дополнительное исследование влияния амплитуды и частоты колебаний полки, а также толщины слоя частиц на режимы виброкипания.

Список литературы

1. Tatemoto Y., Mawatari Y., Yasukawa T., Noda K. Numerical simulation of particle motion in vibrated fluidized bed. // Chem. Eng. Science. 2004. V. 59. P. 437 – 447.
2. Kolehmainen J., Sippola P., Raitanen O., Ozel A., Boyce Ch.M., Saarenrinne P., Sundaresan S. Effect of humidity on triboelectric charging in a vertically vibrated granular bed: experiments and modeling // Chemical Engineering Science. 2017. V. 173. P. 363-373.
3. Chung Y.C., Liao H.H., Hsiao S.S. Convection behavior of non-spherical particles in a vibrating bed: Discrete element method modeling and experimental validation // Powder Technology, 2013. V. 237. P. 53-66.
4. Laurentie J.C., Traoré P., Dascalescu L. Discrete element modeling of triboelectric charging of insulating materials in vibrated granular beds // Journal of Electrostatics, 2013. V. 71. P. 951-957.
5. Xiang L., Shuyan W., Huilin L., Goudong L., Juhui Ch., Yikun L. Numerical simulation of particle motion in vibrated fluidized beds // Powder Technology, 2010. V. 197. P. 25-35.
6. Lu L., Xu Yu., Li T., Benyahia S. Assesment of different coarse graining strategies to simulate polydisperse gas-solids flow // Chemical Engineering Science. 2018. V. 179. P. 53-63.
7. Li T., Zhang Y., Hernandez-Jimenez F. Investigation of particle–wall interaction in a pseudo-2D fluidized bed using CFD-DEM simulations // Particology. 2016. V. 25 P. 10–22.
8. Каменецкий Е. С., Орлова Н. С., Волик М.В., Минасян Д.Г. Тестирование модели виброкипящего слоя, использующей метод дискретного элемента // Известия вузов. Северо-Кавказский регион. Естественные науки. 2017 г. № 4-1. С. 18-23.
9. LIGGGHTS Open Source Discrete Element Method Particle Simulation Code. <http://www.liggghts.com>
10. Орлова Н.С. Математическое моделирование виброкипящего слоя на основе метода дискретных элементов // Тезисы докладов XII Всероссийской школы-семинара «Математическое моделирование и биомеханика в современном университете» (пос. Дивноморское, 29 мая – 3 июня 2017 г.) Таганрог: Издательство Южного федерального университета, 2017. С. 113.

11. Орлова Н.С., Качалкина Я.Н. Исследование режимов виброкипящего слоя с использованием пакета OpenFOAM // Труды Института системного программирования РАН, 2014. Выпуск 5, Том 26. С. 143-154.
12. Каменецкий Е. С., Орлова Н. С., Тагиров А.М., Волик М.В. Трехмерное моделирование виброкипящего слоя с использованием двухжидкостной модели гранулярного газа. Инженерно- физический журнал. 2016. Т. 89, № 6. С. 1480-1486.

УДК 519.85:621.396

Valentins Popovs¹, Vladimirs Skudnovs², Aleksejs Vasiljevs³

Riga Technical University, Institute of Transport

Azenes street 12a, Riga, LV-1048, Latvia

¹popovs@latnet.lv; ²vladimir.skudnov@sotus.net; ³aleksejs.vasiljevs@ldz.lv

MATHEMATICAL MODELS AND ALGORITHMS OF RADIO WAVE PROPAGATION IN CELLULAR NETWORKS OF MOBILE COMMUNICATION

Abstract: *The issue of ensuring uniform radio coverage of service areas and the associated reliability of mobile radio systems in the 3G and 4G standards requires a lot of attention upon designing cellular mobile networks. One of the main tasks at the first stage of the designing is the mathematical modeling of radio wave propagation under the influence of environment. This work provides classification of mathematical models of radio wave propagation in cellular mobile communication systems, the algorithm of calculating radio signal strength at the receiving point depending on the distance between the mobile and the base stations as well as factors and conditions affecting radio wave propagation are specified.*

Keywords: *radio wave propagation, cellular networks, mobile communication, mathematical models*

Introduction

Currently problems of radio wave propagation (RWP) in cellular networks of mobile communications (CNMC) are dealt with in many works. In them different methods of classification of mathematical models of RWP in the decimeter and centimeter wave band applied to CNMC are considered.

Depending on environmental influences, both deterministic and statistical models can be used for modeling RWP; often semi-deterministic, semi-empirical and empirical models are applied.

Deterministic mathematical models are based on the use of the physical laws of RWP. The calculation is based on a single-ray, double-ray or multi-beam models of RWP. They take account of influencing factors of obstacles in the path of radio wave propagation and local conditions in the radio channel (Fig. 1): attenuation in free space, reflection from objects, diffraction due to obstacles, absorption, refraction and cross-polarization of electromagnetic waves (EMW).

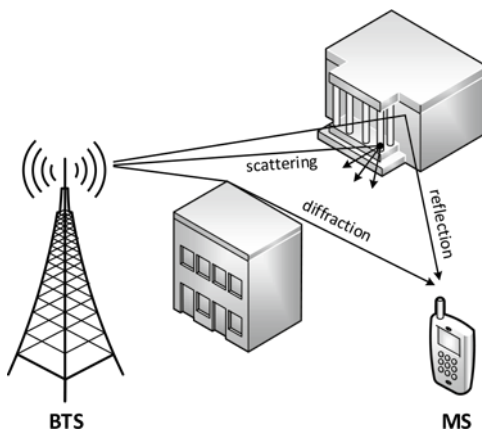


Figure 1. Impact of Obstacles on Radio Wave Propagation

Semi-empirical and empirical models are based on path loss formulas, including a number of empirical correction factors picked up by experiments. Accuracy of calculated results in such case depends on the preciseness of the correction factor values and restrictions having significance within the chosen frequency range and for the specific environmental conditions only.

It is important to point out, that results of experimental research for the real environmental conditions are often used to determine path loss of EMW propagating along the terrain. Mathematical models of RWP in heterogeneous environment, for example, have their own laws of attenuation for different surroundings and allow calculating median values of the received power depending on the distance to the receiver (considering ITU-R, CEPT recommendations). Combined methods represent a blend of deterministic and statistical

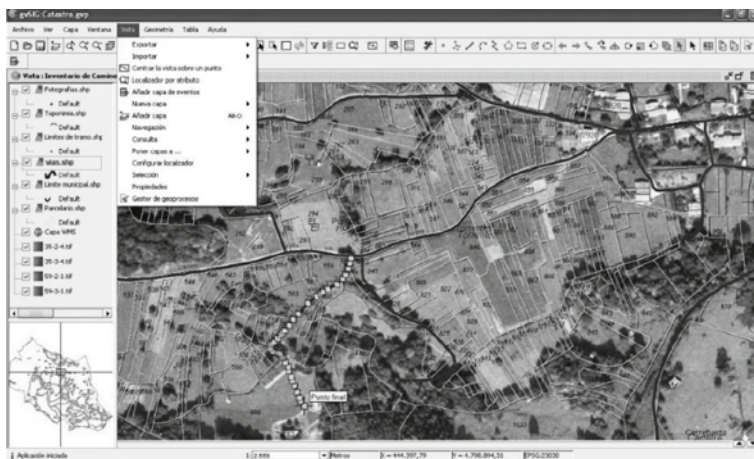


Figure 2. Example of the Geographic Information System–Open Source GIS gvSIG [5]

models. RWP models, based on these methods, may be applied for more accurate measurements in specific situations.

Nowadays majority of CNMC operators are using mobile network development and monitoring software products that are based on models listed above. Sadly, imperfections of digitized maps, approximation and averaging of path loss values lead to relatively large calculation and planning errors. Improvements in model simulation accuracy have been made possible lately by introducing geographic information systems (GIS) (Fig. 2) which use high-precision maps and geographical databases.


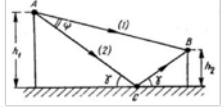




1. Classification of RWP Models in CNMC

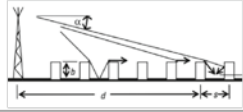
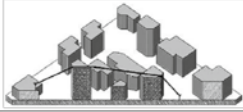



Let's consider the mathematical models of radio wave propagation in CNMC assuming standard conditions and propagation regimes are in place. Most characteristic regimes of wave propagation between base transceiver station (BTS) and mobile station (MS) in real world conditions are generalized in Table 1 and serve as a basis for further mathematical model classification.

The following abbreviations are used in the table:

- R – reflection;
- EMW – electromagnetic wave;

Table 1. Distinctive RWP Regimes in CNMC

No	Conditions	RWP Model, Regime, Cell Type
1		Free space <i>LOS</i> <i>macrocell</i>
2		Two-ray model <i>LOS+R/EMW</i> <i>macrocell</i>
3		Multiple ray model Log-distance path loss model $LOS+NLOS+D/EMW+\Sigma R_i/EMW_i$ <i>macrocell, microcell</i>
4		Multiple ray model for urban conditions Lee model $LOS+NLOS+\Sigma R_i/EMW_i$ <i>macrocell</i>
5		Multiple ray model for urban conditions Okumura model $LOS+NLOS+\Sigma R_i/EMW_i$ <i>microcell</i>
6		Multiple ray model for urban conditions Okumura-Hata model $LOS+NLOS+\Sigma R_i/EMW_i$ <i>microcell, picocell</i>

No	Conditions	RWP Model, Regime, Cell Type
7		Multiple ray model for urban conditions COST 231-Hata model $LOS+NLOS+\Sigma R_d/EMW_i$ microcell, picocell
8		Multiple ray model for urban conditions Walfish-Ikegami model $LOS+NLOS+\Sigma D_d/EMW_i$ macrocell, microcell, picocell
9		Multiple ray model for urban conditions Hata-Davidson model $LOS+NLOS+\Sigma R_d/EMW_i$ microcell, picocell
10		Radio wave propagation through the forest Semi-deterministic model $LOS_1+A/LOS_f+Lateral/EMW+LOS_2$ microcell, picocell
11		Radio wave propagation in free space with the reflection from the forest Deterministic model $NLOS+\Sigma R_d/EMW_i$ microcell, picocell

- LOS – line-of-sight;
- NLOS – non-line-of-sight;
- D/EMW – diffraction of electromagnetic wave;
- A/EMW – attenuation of electromagnetic wave.

This paper, as it follows from the Table 1, examines 11 path loss mathemat-

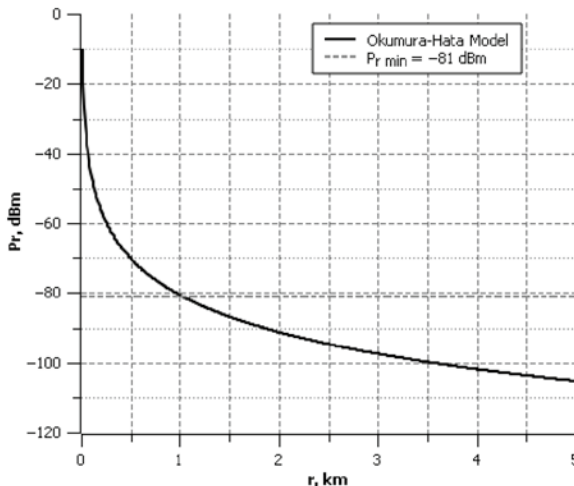


Figure 3. Finding Cell Tower Maximum Coverage Radius Using Okumura-Hata Model ($P_{r\ min} = -81\ dBm, R_{max} = 1.021\ km$) [2]

ical models, which represent the approximation of the received signal strength in relationship to the distance between BTS and MS, taking into account environmental conditions, transmitter and receiver antennas parameters. The determining of received signal strength for an array of distances allows finding the maximum radius of BTS cell coverage area R_{max} (Fig. 3) for a downlink transmission when P_r is equal to the MS receiver sensitivity $P_{r,min}$, which is one of the main tasks in mobile cellular network area coverage planning.

2. RWP Models in CNMC

Mathematical models, which determine the power of harmonic electromagnetic wave (EMW), aired by BTS antenna and received by MS, are shown in the Table 2 with the following set of parameters:

- P_t – BTS transmitter power, dBm;
- f^t – operational frequency, MHz;
- λ – operational wavelength, m;
- G_t, G_r – BTS, MS antenna gain;
- h_t, h_r – BTS, MS antenna height, m;
- r – distance between BTS transmitter and MS receiver, km;
- f_0 – norming frequency, $f_0 = 1$ MHz;
- λ_0 – norming wavelength, $\lambda_0 = 1$ m;
- h_0 – norming height, $h_0 = 1$ m;
- r_0 – norming distance between BTS and MS, $r_0 = 1$ km.

Table 2. RWP Mathematical Models

No.	Model	Formulas
1	Line-of-sight LOS	$P_{r,LOS} = P_t + 20\lg\left(\frac{\lambda}{\lambda_0}\right) - 20\lg\left(\frac{r \cdot 10^3}{r_0}\right) - 20\lg(4\pi) +$ $+ 10\lg G_t + 10\lg G_r, \quad (1)$ <p>where $P_{r,LOS}$ – received signal strength in free space (LOS) environment, $P_{r,LOS} < 0$, dBm;</p>
2	Two-ray model $LOS_t + R_2/$ EMW	$P_r = P_{r,LOS} + 10\lg \psi, \text{ dBm} \quad (2)$ $\psi \approx 1 + R^2 + 2R \cdot \cos\left(\varphi + 4\pi \cdot \frac{h_t \cdot h_r}{\lambda \cdot r \cdot 10^3}\right), \quad (3)$ $0.8r > r > 18 \frac{h_t h_r}{\lambda}, \quad (4)$ $P_r \approx P_t + 10\lg G_t + 10\lg G_r + 20\lg\left(\frac{h_t}{h_0}\right) + 20\lg\left(\frac{h_r}{h_0}\right) - 40\lg\left(\frac{r \cdot 10^3}{r_0}\right), \quad (5)$ <p>where ψ – value determined as a function of the terrain parameters; R – coefficient of reflection, $R \leq 1$; φ – reflection angle, usually $\varphi = 3.142$ rad.</p>
3	Log-distance path loss model	$P_r = P_{r,LOS} - 10 \cdot n \cdot \lg\left(\frac{r}{r_0}\right) - (\pm X_g), \text{ dBm} \quad (6)$

<p><i>LOS+NLO</i> <i>S+D/EM</i> <i>W+ΣR_f/</i> <i>EMW_i</i></p>	<p>where n – attenuation exponent (Table 3); X_g – random value for normal (Gaussian) distribution with a zero mean, representing signal attenuation condition, which was caused by its fading (in case of no fading $X_g = 0$).</p> <p>Table 3. Attenuation Exponent Values</p> <table border="1"> <thead> <tr> <th>Environment</th> <th>Attenuation Exponent n</th> </tr> </thead> <tbody> <tr> <td>Free space (LOS)</td> <td>2</td> </tr> <tr> <td>Suburban area (LOS/NLOS)</td> <td>2.7–3.5</td> </tr> <tr> <td>Urban area (NLOS)</td> <td>3–5</td> </tr> </tbody> </table>	Environment	Attenuation Exponent n	Free space (LOS)	2	Suburban area (LOS/NLOS)	2.7–3.5	Urban area (NLOS)	3–5
Environment	Attenuation Exponent n								
Free space (LOS)	2								
Suburban area (LOS/NLOS)	2.7–3.5								
Urban area (NLOS)	3–5								

<p>4 Lee model <i>LOS+ΣR_f/</i> <i>EMW_i</i></p>	$P_r = P_{r0} - 10\gamma \cdot \lg\left(\frac{r}{r_{n0}}\right) - 10n \cdot \lg\left(\frac{f}{f_{n0}}\right) + 10\lg(\alpha_0), \text{ dBm} \quad (7)$ $\alpha_0 = \left(\frac{h_t}{h_{m0}}\right)^2 \cdot \left(\frac{h_r}{h_{rn0}}\right)^v \cdot \left(\frac{P_{tmW} \cdot 10^3}{P_{m0}}\right) \cdot \frac{G_t \cdot G_r}{4}, \text{ dBm} \quad (8)$ $v = \begin{cases} 2, & \text{if } h_r > 10 \text{ m} \\ 1, & \text{if } h_r < 3 \text{ m} \end{cases}, \quad (9)$ <p>where P_r – median power level at the receiving point for point-to-point RWP, dBm; α_0 – adjustment factor, dB; P_{tmW} – BTS transmitter power, mW; P_{r0} – power at the r_{n0} point (Table 4), dBm; γ – path loss correction due to distance (Table 4); n – frequency correction factor, recommended values are given in Table 5, $2 \leq n \leq 3$; v – MS antenna height correction factor; f_{n0} – nominal frequency, $f_{n0} = 900$ MHz; P_{m0} – nominal BTS transmitter power, $P_{m0} = 10$ W; h_{m0} – nominal BTS antenna height, $h_{m0} = 30.48$ m; h_{rn0} – nominal MS antenna height, $h_{rn0} = 3$ m; r_{n0} – nominal distance between BTS and MS, $r_{n0} = 1.6$ km.</p> <p>Table 4. Parameters for Lee’s Model</p> <table border="1"> <thead> <tr> <th>Environment</th> <th>P_{r0}, dBm</th> <th>γ</th> </tr> </thead> <tbody> <tr> <td>Free space</td> <td>-45</td> <td>2</td> </tr> <tr> <td>Rural area</td> <td>-49</td> <td>4.35</td> </tr> <tr> <td>Suburban area</td> <td>-61.7</td> <td>3.84</td> </tr> <tr> <td>Urban area</td> <td>-70</td> <td>3.68</td> </tr> <tr> <td>Metropolitan center</td> <td>-84</td> <td>3.05</td> </tr> </tbody> </table> <p>Table 5. Parameters for Lee’s Model</p> <table border="1"> <thead> <tr> <th>Environment, Frequency f</th> <th>n</th> </tr> </thead> <tbody> <tr> <td>Open, rural, suburban area, $f < 450$ MHz</td> <td>2</td> </tr> </tbody> </table>	Environment	P_{r0} , dBm	γ	Free space	-45	2	Rural area	-49	4.35	Suburban area	-61.7	3.84	Urban area	-70	3.68	Metropolitan center	-84	3.05	Environment, Frequency f	n	Open, rural, suburban area, $f < 450$ MHz	2
Environment	P_{r0} , dBm	γ																					
Free space	-45	2																					
Rural area	-49	4.35																					
Suburban area	-61.7	3.84																					
Urban area	-70	3.68																					
Metropolitan center	-84	3.05																					
Environment, Frequency f	n																						
Open, rural, suburban area, $f < 450$ MHz	2																						

		Urban, metropolitan area, $f > 450$ MHz	3
5	Okumura model $LOS + \Sigma R_i / EMW_i$	Limitations: <ul style="list-style-type: none"> • Frequency range $f = 150-1920$ MHz; • BTS antenna height $h_t = 30-1000$ m; • MS antenna height $h_r = 1-10$ m; • Distance between BTS and MS $r = 1-100$ km. 	
		$L_{rt} = L_{rt\ LOS} - A_{mu}(f, r) + H(h_t) + H(h_r) + G_{area}, \text{ dB} \quad (10)$	
		$P_r = P_r\ LOS - A_{mu}(f, r) + H(h_t) + H(h_r) + G_{area}, \text{ dBm} \quad (11)$	
		$H(h_t) = 20 \lg \left(\frac{h_t}{h_{m0}} \right), \text{ dB} \quad (12)$	
		$H(h_r) = \begin{cases} 10 \lg \left(\frac{h_r}{h_{rn0}} \right), & \text{if } h_r \leq 3 \text{ m} \\ 20 \lg \left(\frac{h_r}{h_{rn0}} \right), & \text{if } 3 < h_r < 10 \text{ m} \end{cases}, \text{ dB} \quad (13)$	
		where L_{rt} – path attenuation, dB; $A_{mu}(f, r)$ – median attenuation relative to free space (determined from the Fig. 4), dB; $H(h_t)$ – BTS antenna gain factor, dB; $H(h_r)$ – MS antenna gain factor, dB; h_{m0} – nominal BTS antenna height, $h_{m0} = 200$ m; h_{rn0} – nominal MS antenna height, $h_{rn0} = 3$ m; G_{area} – gain correction factor depending on the environment (Fig. 5), dB.	
		<p>Figure 4. Attenuation Factor for Urban Areas</p>	

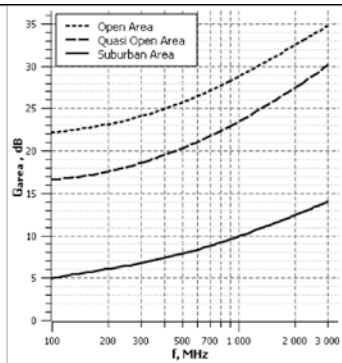


Figure 5. Gain Correction Factor

Limitations:

- Frequency range $f = 150 - 1500$ MHz;
- BTS antenna height $h_t = 30 - 200$ m;
- MS antenna height $h_r = 1 - 10$ m;
- Distance between BTS and MS $r = 1 - 20$ km.

$$P_{r\ Hata} = P_t - 69.55 - 26.16 \lg\left(\frac{f}{f_0}\right) - \left(44.9 - 6.55 \lg\left(\frac{h_t}{h_0}\right)\right) \lg\left(\frac{r}{r_0}\right) + 13.82 \lg\left(\frac{h_t}{h_0}\right) + a(f, h_r) + K(f) \quad (14)$$

where $a(f, h_r)$ – correction factor for MS antenna height (Table 6), dB;
 K – correction factor that depends upon the environment (Table 7), dB.

Okumura-Hata model
 $LOS + \sum R_i / EMW_i$

6

Table 6. Parameters for the Okumura-Hata Model

Environment	$a(f, h_r)$, dB
Open, suburban, urban area	$\left(1.1 \lg\left(\frac{f}{f_0}\right) - 0.7\right) \cdot h_r - \left(1.56 \lg\left(\frac{f}{f_0}\right) - 0.8\right)$
Metropolitan area, $f < 300$ MHz	$8.29 \left[\lg\left(\frac{1.54 h_r}{h_{r0}}\right) \right]^2 - 1.1$
Metropolitan area, $f \geq 300$ MHz	$3.2 \left[\lg\left(\frac{11.75 h_r}{h_{r0}}\right) \right]^2 - 4.97$

Table 7. Parameters for the Okumura-Hata Model

Environment	$K(f)$, dB
Open area	$4.78 \left[\lg\left(\frac{f}{f_0}\right) \right]^2 - 18.33 \lg\left(\frac{f}{f_0}\right) + 40.94$

		<table border="1"> <tr> <td>Suburban area</td> <td>$2 \left[\lg \left(\frac{f}{28} \right) \right]^2 + 5.4$</td> </tr> <tr> <td>Urban, metropolitan area</td> <td>0</td> </tr> </table>	Suburban area	$2 \left[\lg \left(\frac{f}{28} \right) \right]^2 + 5.4$	Urban, metropolitan area	0											
Suburban area	$2 \left[\lg \left(\frac{f}{28} \right) \right]^2 + 5.4$																
Urban, metropolitan area	0																
7	COST 231 -Hata model $LOS + \Sigma R_i /$ EMW_i	<p>Limitations:</p> <ul style="list-style-type: none"> • Frequency range $f = 150-2000$ MHz; • BTS antenna height $h_t = 30-200$ m; • MS antenna height $h_r = 1-10$ m; • Distance between BTS and MS $r = 1-20$ km. $P_r = P_t - A - B \lg \left(\frac{f}{f_0} \right) + 13.82 \lg \left(\frac{h_t}{h_0} \right) + a(f, h_r) -$ $- \left[44.9 - 6.55 \lg \left(\frac{h_t}{h_0} \right) \right] \lg \left(\frac{r}{r_0} \right) - C, \text{ dBm} \quad (15)$ $a(f, h_r) = \left(1.1 \lg \left(\frac{f}{f_0} \right) - 0.7 \right) \cdot h_r - \left(1.56 \lg \left(\frac{f}{f_0} \right) - 0.8 \right), \text{ dB} \quad (16)$ <p>where A – attenuation component (Table 8), dB; B – frequency correction factor (Table 8); $a(f, h_r)$ – correction factor for MS antenna height, dB; C – attenuation correction factor, depends on the environment (Table 9), dB.</p> <p>Table 8. Parameters for the COST 231-Hata Model</p> <table border="1"> <thead> <tr> <th>Frequency Range f, MHz</th> <th>A, dB</th> <th>B</th> </tr> </thead> <tbody> <tr> <td>$150 \leq f < 1500$</td> <td>69.55</td> <td>26.16</td> </tr> <tr> <td>$1500 \leq f \leq 2000$</td> <td>46.33</td> <td>33.9</td> </tr> </tbody> </table> <p>Table 9. Parameters for the COST 231-Hata Model</p> <table border="1"> <thead> <tr> <th>Environment</th> <th>C, dB</th> </tr> </thead> <tbody> <tr> <td>Open, suburban, urban area</td> <td>0</td> </tr> <tr> <td>Metropolitan area</td> <td>3</td> </tr> </tbody> </table>	Frequency Range f , MHz	A, dB	B	$150 \leq f < 1500$	69.55	26.16	$1500 \leq f \leq 2000$	46.33	33.9	Environment	C, dB	Open, suburban, urban area	0	Metropolitan area	3
Frequency Range f , MHz	A, dB	B															
$150 \leq f < 1500$	69.55	26.16															
$1500 \leq f \leq 2000$	46.33	33.9															
Environment	C, dB																
Open, suburban, urban area	0																
Metropolitan area	3																
8	Walfisch- Ikegami model LOS/NLO $S+$ $\Sigma D_i/EMW_i$	<p>Limitations:</p> <ul style="list-style-type: none"> • Frequency range $f = 800-2000$ MHz; • BTS antenna height $h_t = 4-50$ m; • MS antenna height $h_r = 1-3$ m; • Distance between BTS and MS $r = 0.02-5$ km. 															

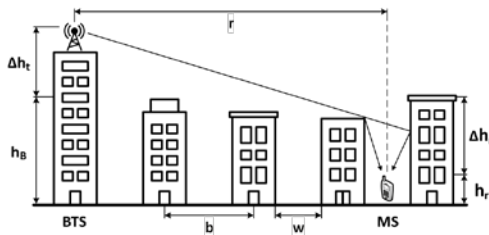


Figure 6. Definition of the Urban Parameters

$$L_{rt} = \begin{cases} L_0 + L_{rts} + L_{msd} & \text{if } L_{rts} + L_{msd} \leq 0 \text{ dB} \\ L_0 & \text{if } L_{rts} + L_{msd} > 0 \text{ dB} \end{cases} < 0, \text{ dB} \quad (17)$$

$$L_{LOS} = -42.6 - 261 \lg \left(\frac{r}{r_0} \right) - 201 \lg \left(\frac{f}{f_0} \right), \text{ dB} \quad (18)$$

$$L_0 = -32.45 - 201 \lg \left(\frac{f}{f_0} \right) - 201 \lg \left(\frac{r}{r_0} \right), \text{ dB} \quad (19)$$

$$L_{rts} = 16.9 + 10 \lg \left(\frac{w}{w_0} \right) - 10 \lg \left(\frac{f}{f_0} \right) - 201 \lg \left(\frac{\Delta h_r}{h_0} \right) - L_{ori}, \text{ dB} \quad (20)$$

$$L_{msd} = -L_{bsh} - k_a - k_d \lg \left(\frac{r}{r_0} \right) - k_f \lg \left(\frac{f}{f_0} \right) + 9 \lg \left(\frac{b}{b_0} \right), \text{ dB} \quad (21)$$

$$P_{rLOS} = P_r + L_{LOS} = P_t - 42.6 - 261 \lg \left(\frac{r}{r_0} \right) - 201 \lg \left(\frac{f}{f_0} \right), \text{ dBm} \quad (22)$$

$$P_{rNLOS} = P_t + L_{rt}, \text{ dBm} \quad (23)$$

where P_{rLOS} – received signal strength in line-of-sight (LOS) environment, dBm;

P_{rNLOS} – received signal strength in non-line-of-sight environment, dBm;

L_{rt} – path loss, dB;

L_{LOS} – path loss for free space (LOS) environment, dB;

L_{fs} – free space loss, dB;

L_{rts} – diffraction and scatter loss, dB;

L_{msd} – multiscreen diffraction loss, dB;

L_{ori} – orientation loss (Table 10), dB;

L_{bsh} – shadowing gain due to BTS antenna location (Table 11), dB;

b – average building separation, m;

w – average street width, usually $w = 0.5b$ m;

h_B – average building height, m;

Δh_t – BTS antenna height above the rooftops, $\Delta h_t = h_t - h_B$, m;

		<p> Δh_r – distance between MS antenna and rooftops, $\Delta h_r = h_B - h_r$, m; φ – road orientation with respect to the direct radio path, usually $\varphi = 90^\circ$ (worst case); k_a – correction factor (Table 11), dB; k_d – distance factor (Table 11); k_f – frequency factor (Table 12). </p> <p>Table 10. Parameters for the Walfisch-Ikegami Model</p> <table border="1" data-bbox="330 347 1002 470"> <thead> <tr> <th>Street Orientation Angle φ, °</th> <th>L_{ori}, dB</th> </tr> </thead> <tbody> <tr> <td>$0 \leq \varphi < 35$</td> <td>$-10 + 0.354\varphi$</td> </tr> <tr> <td>$35 \leq \varphi < 55$</td> <td>$2.5 + 0.075(\varphi - 35)$</td> </tr> <tr> <td>$55 \leq \varphi < 90$</td> <td>$4 - 0.114(\varphi - 55)$</td> </tr> </tbody> </table> <p>Table 11. Parameters for the Walfisch-Ikegami Model</p> <table border="1" data-bbox="330 534 1002 689"> <thead> <tr> <th>Height Δh_t, m; Distance r, km</th> <th>L_{bsh}, dB</th> <th>k_n</th> <th>k_d</th> </tr> </thead> <tbody> <tr> <td>$\Delta h_t > 0$</td> <td>$-18 \lg(1 + \Delta h_t)$</td> <td>54</td> <td>18</td> </tr> <tr> <td>$\Delta h_t \leq 0, r \geq 0.5$</td> <td rowspan="2">0</td> <td>$54 - 0.8 \cdot \Delta h_t$</td> <td rowspan="2">$18 - 15 \frac{\Delta h_t}{h_B}$</td> </tr> <tr> <td>$\Delta h_t \leq 0, r < 0.5$</td> <td>$54 - 0.8 \cdot \Delta h_t \cdot 2r$</td> </tr> </tbody> </table> <p>Table 12. Parameters for the Walfisch-Ikegami Model</p> <table border="1" data-bbox="330 753 1002 906"> <thead> <tr> <th>Environment</th> <th>k_f</th> </tr> </thead> <tbody> <tr> <td>Suburban, urban area</td> <td>$-4 + 0.7 \left(\frac{f}{925} - 1 \right)$</td> </tr> <tr> <td>Metropolitan area</td> <td>$-4 + 1.5 \left(\frac{f}{925} - 1 \right)$</td> </tr> </tbody> </table>	Street Orientation Angle φ , °	L_{ori} , dB	$0 \leq \varphi < 35$	$-10 + 0.354\varphi$	$35 \leq \varphi < 55$	$2.5 + 0.075(\varphi - 35)$	$55 \leq \varphi < 90$	$4 - 0.114(\varphi - 55)$	Height Δh_t , m; Distance r , km	L_{bsh} , dB	k_n	k_d	$\Delta h_t > 0$	$-18 \lg(1 + \Delta h_t)$	54	18	$\Delta h_t \leq 0, r \geq 0.5$	0	$54 - 0.8 \cdot \Delta h_t$	$18 - 15 \frac{\Delta h_t}{h_B}$	$\Delta h_t \leq 0, r < 0.5$	$54 - 0.8 \cdot \Delta h_t \cdot 2r$	Environment	k_f	Suburban, urban area	$-4 + 0.7 \left(\frac{f}{925} - 1 \right)$	Metropolitan area	$-4 + 1.5 \left(\frac{f}{925} - 1 \right)$
Street Orientation Angle φ , °	L_{ori} , dB																													
$0 \leq \varphi < 35$	$-10 + 0.354\varphi$																													
$35 \leq \varphi < 55$	$2.5 + 0.075(\varphi - 35)$																													
$55 \leq \varphi < 90$	$4 - 0.114(\varphi - 55)$																													
Height Δh_t , m; Distance r , km	L_{bsh} , dB	k_n	k_d																											
$\Delta h_t > 0$	$-18 \lg(1 + \Delta h_t)$	54	18																											
$\Delta h_t \leq 0, r \geq 0.5$	0	$54 - 0.8 \cdot \Delta h_t$	$18 - 15 \frac{\Delta h_t}{h_B}$																											
$\Delta h_t \leq 0, r < 0.5$		$54 - 0.8 \cdot \Delta h_t \cdot 2r$																												
Environment	k_f																													
Suburban, urban area	$-4 + 0.7 \left(\frac{f}{925} - 1 \right)$																													
Metropolitan area	$-4 + 1.5 \left(\frac{f}{925} - 1 \right)$																													
9	Hata-Davidson model $LOS + \Sigma R_i / EMW_i$	<p>Limitations:</p> <ul style="list-style-type: none"> • Frequency range $f = 150 - 1500$ MHz; • BTS antenna height $h_t = 30 - 2500$ m; • MS antenna height $h_r = 1 - 10$ m; • Distance between BTS and MS $r = 1 - 300$ km. <p> $P_r = P_{r \text{ Hata}} - A(h_t, r) + S_1(r) + S_2(h_t, r) + S_3(f) + S_4(f, r)$, dBm (24) </p> <p> $P_{r \text{ Hata}} = P_t - 69.55 - 26.16 \lg \left(\frac{f}{f_0} \right) - \left(44.9 - 6.55 \lg \left(\frac{h_t}{h_0} \right) \right) \lg \left(\frac{r}{r_0} \right) +$ $+ 13.82 \lg \left(\frac{h_t}{h_0} \right) + a(f, h_r) + K(f)$ (25) </p> <p> $S_2(h_t, r) = \begin{cases} 0.00784 \cdot \lg \left(\frac{9.98}{r} \right) \cdot (h_t - 300) & \text{if } h_t \geq 300 \text{ m} \\ 0 & \text{if } h_t < 300 \text{ m} \end{cases}$ (26) </p>																												

$$S_3(f) = \frac{f}{250 \lg\left(\frac{1500}{f}\right)}, \quad (27)$$

where $P_{r\ HATA}$ – received signal strength for Okumura-Hata model, dBm;
 $A(h_t, r)$, $S_1(r)$ – distance correction factor (Table 13), dB;
 $S_2(h_t, r)$ – BTS antenna height correction factor, dB;
 $S_3(f)$ – frequency correction factor, dB;
 $S_4(f, r)$ – frequency and distance correction factor (Table 14), dB.

Table 13. Parameters for the Hata-Davidson Model

Distance r, km	A(h _t , r), dB
r < 20	0
20 ≤ r < 300	$0.62137 \cdot (r - 20) \cdot \left[0.5 + 0.15 \lg\left(\frac{h_t}{121.92}\right) \right]$

Table 14. Parameters for the Hata-Davidson Model

Distance r, km	S ₁ (r), dB	S ₄ (f, r), dB
r < 64.38	0	0
64.38 ≤ r < 300	$0.174 \cdot (r - 64.38)$	$0.112 \lg\left(\frac{1500}{f}\right) \cdot (r - 64.38)$

10

RWP through the forest [1, 17]
 $LOS_1 + A/L$
 $OS_F + Late$
 $ral/EMW +$
 LOS_2

$$P_r \approx P_{I\ LOS} - \alpha_{mF} \cdot r + 10 \lg[\zeta(r)], \quad (28)$$

where α_{mF} – path attenuation factor (Table 15), dB/km;
 $\zeta(r)$ – correction factor (Table 15).

Table 15. Parameters Describing RWP through the Forest

Operational Frequency f, MHz	α_{mF} , dB/m	$\zeta(r)$
900	0.185	100.1
1800	0.36	

11

RWP in free space with the reflection from the forest [17]
 $NLOS + \Sigma R$
 $/EMW_i$

$$P_r = \frac{P_t \cdot G_t \cdot G_r \cdot \lambda^2}{(4\pi)^2} \int_{V_F} \frac{F_t(\theta)^2 \cdot \sigma(\varphi)}{r_1^2 \cdot r_2^2 \cdot V_0} dV, \quad W \quad (29)$$

where $F_t(\theta)$ – BTS antenna radiation pattern, $F_t(\theta) = 1$ for omnidirectional antennas;
 V_F – volume of the forest, m³;
 V_0 – norming volume, $V_0 = 1$ m³;
 φ – angle between the direction of the main wave and its reflection from the forest, °;
 $\sigma(\varphi)$ – effective scattering area of a volume unit, m²;
 r_i – distance between BTS and forest, km;

3. RWP Calculation Algorithm

Algorithm for numerical calculation of Radio wave propagation parameters is based on the examined mathematical models and shown in Fig. 7. The values of received signal strength for every model are found by using formulas from Table 2 with appropriate standard BTS and MS parameters.

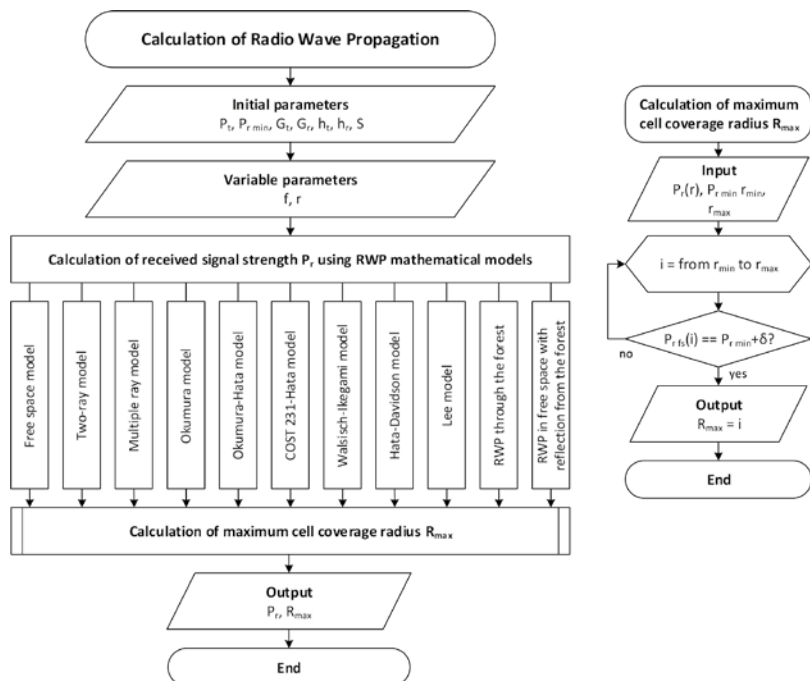


Figure 7. Algorithm for RWP Calculation Using Mathematical Models and Standard BTS and MS Parameters [18]

As it follows from the right part of the flowchart (Fig. 7), maximum radius of BTS radio coverage area R_{max} depends on the values of MS receiver sensitivity $P_{r_{min}}$ and received signal strength $P_r(r)$ for corresponding RWP model.

3.1. Numerical Calculation of the BTS Coverage Radius

To estimate the influence of environment on the received signal strength (for downlink transmission), calculations of proposed algorithm were made using Mathcad engineering software. Comparison of the results, visible in Fig. 8, are true for the following parameters:

- BTS transmitter power: $P_t = 50 \text{ W}$ (47 dBm);
- Operational frequency: $f = 900 \text{ MHz}$ ($\lambda = 0.33 \text{ m}$);
- BTS, MS antenna gain: $G_t = 10, G_r = 1$.

Environmental conditions, specific for every mathematical model (Table 2), were taken into account too.

Received signal strength is heavily dependent on wave propagation con-

ditions, and it changes drastically with an increase in distance between BTS and MS (as it is shown in Figure 8). For example, when distance $r = 3$ km, signal strength $P_r = -44$ dBm for the LOS model, but in dense urban area conditions (Walfisch-Ikegami model) – $P_r = -95$ dBm. The stability of the radio communication link depends on the MS receiver sensitivity level ($P_{r\min}$). If $P_{r\min} = -93$ dBm and the real path loss is equal to the calculated one (using Walfisch-Ikegami model), then signal from BTS couldn't be received by MS at 3 km distance already.

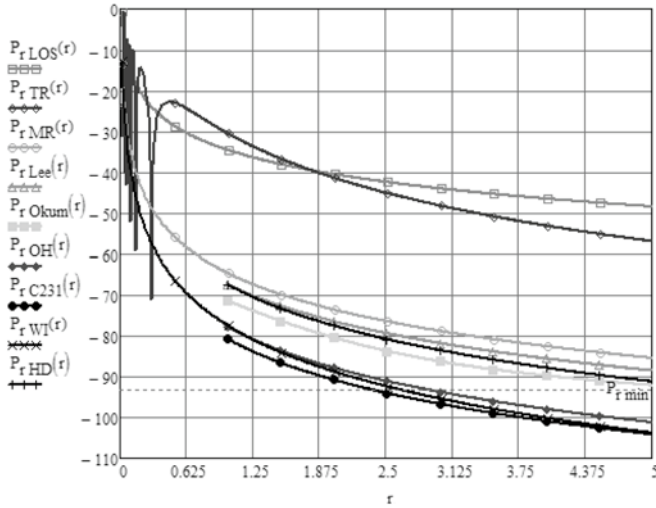


Figure 8. Comparison of the Received Signal Strength for 9 RWP Models

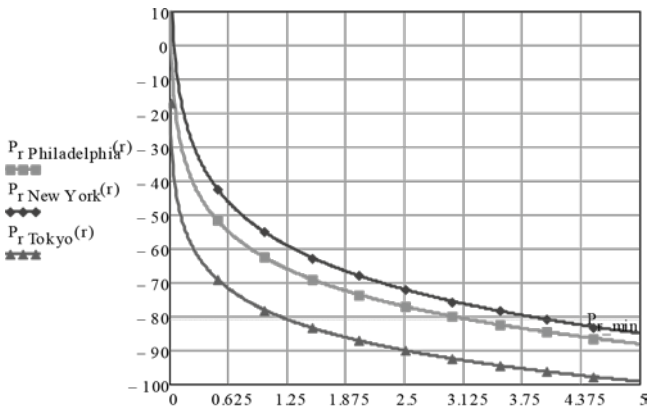


Figure 9. Received Signal Strength Relation to the Distance between BTS and MS (r , km) for Selected Urban Areas [11]

Experimentally gathered RWP data for Philadelphia, New York and Tokyo are shown in Fig. 9 to compare the real signal attenuation slopes with the calculated results (Fig. 8).

As it can be seen in Fig. 9, Lee model is most suitable for environmental conditions in Philadelphia, while COST 231-Hata model gives the best results for Tokyo.

Maximum cell coverage radiuses of calculated path loss models for two different MS receiver sensitivity settings ($P_{r \min} = [-81; -100]$ dBm) are visible in the Table 16.

Table 16. Maximum Cell Coverage Radius R_{\max} for Given MS Receiver Sensitivity

RWP Model	$P_{r \min}$	
	R_{\max} , km $P_{r \min} = -81$ dBm	R_{\max} , km $P_{r \min} = -100$ dBm
Free space model	210.253	1873.883
Two-ray model	19.911	59.452
Multiple ray model	3.536	15.202
Lee model	2.818	11.827
Okumura model	2.077	9.141
Okumura-Hata model	1.255	4.584
COST 231-Hata model	1.026	3.748
Walfisch-Ikegami model	1.253	3.962
Hata-Davidson model	3.794	13.858

4. Conclusions

1. This work discusses relatively simple RWP models for CNMC, which, using provided algorithm along with appropriate software (such as Mathcad), allows determining maximum coverage area of BTS cells, based on environment characteristics.

2. Every mathematical model is presented as a function of the received signal strength P_r ($P_r, G_r, G_p, f, h_p, h_r, r$). Unlike in most of the models recommended by ITU-R и CEPT, formulas in Table 2 comply with the strict mathematical laws. Expressions with common logarithms are reduced to dimensionless forms by introduction of the norming parameters. For instance, received path loss formula for the COST 231-Hata model in the book [7] looks like this:

$$L_{COST\ 231} = 46.3 + 33.9 \cdot \lg(f) - 13.82 \cdot \lg(h_t) - a(h_r) + \left[44.9 - 6.55 \cdot \lg(h_t) \right] \cdot \lg(r) + C \quad (30)$$

That path loss $L_r = 10 \cdot \lg(P_t / P_r) > 0$ doesn't match the actual reduction in the signal level when distance between BTS and MS increases, and expression (30), in overall, is mathematically incorrect, because quantities with dimensions are used as a logarithm arguments. This work proposes the following formula (31):

$$P_r \text{ COST } 231 = P_t - 46.3 - 33.9 \cdot \lg\left(\frac{f}{f_0}\right) + 13.82 \cdot \lg\left(\frac{h_t}{h_0}\right) + a(f, h_r) - \left[44.9 - 6.55 \cdot \lg\left(\frac{h_t}{h_0}\right)\right] \cdot \lg\left(\frac{r}{r_0}\right) - C \quad (31)$$

which corresponds to the path loss $L_{tr} = 10 \cdot \lg(P_r / P_t) < 0$ and shows the real attenuation of the radio signal.

3. Path loss in Philadelphia, as shown in the Fig. 10 where experimental (Fig. 9) and calculated (Fig. 8) data are compared, can be expressed using Lee model, whereas COST 231-Hata model is a match to the measured path attenuation in Tokyo. Visible errors are possible due to the choice of parameters (BTS and MS antenna height, antenna gain, urban parameters, etc.).

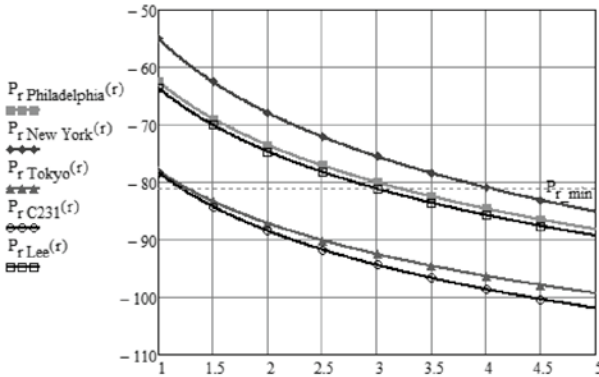


Figure 10. Comparison of the Experimental and Calculated Values of Received Signal Strength in Selected Cities

4. As a first approximation, calculations of the median received power $P_r(r)$ (for downlink regime) can be made using estimated engineering formula (32):

$$P_r = P_r \text{ LOS} - 10 \cdot n \cdot \lg\left(\frac{r}{r_0}\right), \quad (32)$$

where r_0 – norming distance, $r_0 = 1$ km;
 n – attenuation exponent which depends on the environment (Table 17).

Table 17. Attenuation Exponent Values in Given Environment

Environment	Attenuation Exponent n
Free space (LOS)	2
Suburban area (LOS/NLOS)	2.7–3.5
Urban area (NLOS)	3–5

For example, if the received power in LOS environment $P_{r,LOS} = -60$ dBm at the distance $r = 1$ km from the BTS, and $P_r = -75$ dBm when the distance is $r = 3$ km, then the slope of experimentally acquired data for RWP in New York (Fig. 9) can be roughly expressed using formula (32) with the coefficient $n = (-75 + 60) / (-10 \cdot 0.48) = 3.125$.

5. The developed algorithm (Fig. 7), which was implemented using Mathcad software, allows not only calculating maximum coverage area of BTS cell towers, but also performing network clustering and designing the coverage map afterwards [1, 2, 13].

References

1. Popovs V., GSM standarta šūnu mobilo sakaru sistēmas: Projektēšanas problēmas. Rīga: RTU Izdevniecība, 2003, 362 lpp.
2. Попов В. И., Основы сотовой связи стандарта GSM. Москва: Эко-Трендз, 2005, 296 с.
3. Blaunstein N., Radio Propagation in Cellular Networks. London, United Kingdom: Artech House, 2000, 384 p.
4. COST Action 231, “Digital Mobile Radio Towards Future Generations Systems,” European Commission, Technical Report EUR 18957, 1999, 474 p.
5. gvSIG Association, Portal gvSIG, 2006-2016 [cited—February 20, 2016]. Available at: <http://www.gvsig.com/>
6. Lee W. C. Y., Mobile Communications Design Fundamentals, 2nd ed. New York, USA: Wiley, 1993, 398 p.
7. Parsons J. D., The Mobile Radio Propagation Channel, 2nd ed. New York, USA: Wiley, 2000, 436 p.
8. Popovs V., Tehniskā elektrodinamika, 4. daļa: Radioviļņu izplatīšanas, Lektijas. Rīga: RTU DzTI, 2003–2005, 74 lpp.
9. Popovs V., Skudnovs V., Vasiljevs A., “Antenna Systems of Base Transceiver Stations in Cellular Mobile Networks. Modern State and Perspective of Development,” in Proceedings of the 56. International Scientific Conference of the Riga Technical University, 2015.
10. Попов В. И., Скуднов В. А., Васильев А. С., “Антенны базовых станций в сотовых сетях мобильной связи. Современное состояние и перспективы развития,” Евразийский Союз Ученых (ЕСУ), №11 (20), часть 3, 2015, с. 138–150.
11. Singh Y., “Comparison of Okumura, Hata and COST-231 Models on the Basis of Path Loss and Signal Strength,” International Journal of Computer Applications, vol.59 (no.11), Dec. 2012.
12. Siwiak K., Bahreini Y., Radio wave Propagation and Antennas for Personal Communications, 3rd ed. London, United Kingdom: Artech House, 2007, 494 p.
13. Весоловский К., Системы подвижной радиосвязи. Москва: Горячая линия–Телеком, 2006, 536 с.
14. Шабунин С. Н., Лесная Л. Л., Распространение радиоволн в мобильной связи. Методические указания по курсу “Распространение радиоволн и ан-

- тенно-фидерные устройства в системах мобильной связи”. Екатеринбург: УГТУ, 2000, 38 с.
15. Alim M. A., Rahman M. M., Hossain M. M., Al-Nahid A., “Analysis of Large-Scale Propagation Models for Mobile Communications in Urban Area,” *International Journal of Computer Science and Information Security (IJCSIS)*, vol.7 (no.1), 2010.
 16. Утц В. А., “Исследование потерь при распространении радиосигнала сотовой связи,” *Вестник Балтийского государственного университета им. И. Канта*, №5, 2011, с. 44 – 49.
 17. Попов В. И., *Распространение радиоволн в лесах*. Москва: Горячая линия–Телеком, 2015, 392 с.
 18. Попов В. И., Скуднов В. А., *Основы проектирования сотовых сетей мобильной связи*. Москва: Горячая линия–Телеком, 2017, 400 с.

СЕКЦИЯ 6. ИНФОРМАЦИОННЫЕ ТЕХНОЛОГИИ В ЕСТЕСТВЕННЫХ НАУКАХ

УДК 514.18

Бурлов В. В.,

к.т.н., профессор, ФГБОУ «Пензенский государственный технологический университет», г. Пенза, Россия, vladimir-burlov@yandex.ru, www.penzgtu.ru

Косолапов В. В.,

к.т.н., ГБОУ ВО «Нижегородский государственный инженерно-экономический университет», г. Княгинино, Россия, Vladimir.kosolapov@mail.ru, www.ngiei.ru

Косолапова Е. В.,

к.с.-х.н., ГБОУ ВО «Нижегородский государственный инженерно-экономический университет», г. Княгинино, Россия, K-art-inka@yandex.ru, www.ngiei.ru

Ремонтова Л. В.,

к.т.н., доцент, ФГБОУ «Пензенский государственный технологический университет», г. Пенза, Россия, remontova@mail.ru, www.penzgtu.ru

МОДЕЛИРОВАНИЕ ПОВЕРХНОСТЕЙ НА ОСНОВЕ КРИВЫХ ВТОРОГО ПОРЯДКА

Аннотация: В работе подробно изложены алгоритмы построения кривых второго порядка – параболы и гиперболы на основе применения их канонических и параметрических уравнений, а также моделирования производных 3D-поверхностей в системе трёх мерного моделирования КОМПАС-3D. Целями работы являются более глубокое изучение цикла геометрографических дисциплин и продвижение отечественного IT-продукта в сфере образования.

Ключевые слова: математическое моделирование, САПР, кривые второго порядка, алгоритмы, программное обеспечение.

V. V. Burlov,

Candidate of Technical Sciences, Professor, Penza State Technological University, Penza, Russia, vladimir-burlov@yandex.ru, www.penzgtu.ru

V. V. Kosolapov,

Candidate of Technical Sciences, Nizhny Novgorod State Engineering and Economic University, Knyaginino, Russia, Vladimir.kosolapov@mail.ru, www.ngiei.ru

E.V. Kosolapova,

Ph.D., Ph.D., Nizhny Novgorod State Engineering and Economic University, Knyaginino, Russia, K-art-inka@yandex.ru, www.ngiei.ru

L.V. Remontova,

Candidate of Technical Sciences, Associate Professor, FSBEI "Penza State Technological University", Penza, Russia, remontova@mail.ru, www.penzgtu.ru

MODELING SURFACES ON THE BASIS SECONDARY ORDER CURVES

Abstract: *The paper describes in detail the algorithms for constructing second-order curves – parabolas and hyperbolas based on the application of their canonical and parametric equations, as well as the modeling of derivatives of 3D surfaces in the three-dimensional modeling system КОМПАС-3D. The objectives of the work are a deeper study of the cycle of geometric-graphic disciplines and the promotion of the domestic IT-product in the field of education.*

Keywords: *mathematical modeling, CAD, second order curves, algorithms, software.*

Введение

Развитие современных вычислительных систем математического моделирования и САПР позволяют упростить процесс проектирования технически сложных конструкций и обеспечить высокую точность расчетов и моделей. Однако для корректного расчета виртуальных моделей и их отображения необходима серьезная теоретическая база и приемы построения геометрических моделей, которые имеют в своей основе дисциплинарные науки инженерная графика и начертательная геометрия.

Общеизвестно, что поверхности второго порядка образуются с помощью соответствующих кривых, меняющих свое положение по установленным правилам. Основы построения кривых и поверхностей второго порядка с помощью различных инструментов затрагиваются в многочисленных работах [1, 2, 3, 4, 5].

В среде САПР КОМПАС-3D на панели **Геометрия** из кривых второго порядка предусмотрены только **Круг** и **Эллипс**. Однако при моделировании поверхностей парабола и гипербола не менее востребованы.

Существует несколько алгоритмов создания кривых второго порядка – с помощью конических сечений, с применением **канонических и параметрических уравнений**.

Начиная с v. 15.1 программы Компас-График введен инструмент **Каноническая кривая**, позволяющая построить параболические кривые. Однако применение данного инструмента в частных случаях затрудняется тем, что необходимо знать: положение точки пересечения касательных (т.3), определяемой углами наклона касательных к конической кривой, формой кривой определяемой положением точки т.4 или параметрами высота/коэффициент. Кроме того, малейшая ошибка симметрии параболы, вызванная случайным сдвигом центра пересечения касательных, делает невозможным построение заданного геометрически верного параболоида, что исключено при использовании выше перечисленных способов создания параболы.

В данной работе предлагаем рассмотреть алгоритмы создания однополостного и двуполостного гиперболоидов, эллиптического и гиперболического параболоидов с помощью кривых второго порядка – гиперболы и параболы соответственно, заданных с помощью **канонических и параметрических уравнений**.

Алгоритмы создания кривых и поверхностей второго порядка

Использование канонических уравнений параболы

В КОМПАС-3D среди множества библиотек имеется библиотека **FTDraw**, входящая в меню библиотек **Прочие**. С помощью этой библиотеки по каноническому уравнению кривой можно построить график функции, т.е. изображение кривой параболы по заданным условиям. Затем полученное изображение можно использовать для 3D-моделирования поверхностей второго порядка.

Геометрия и каноническое уравнение параболы

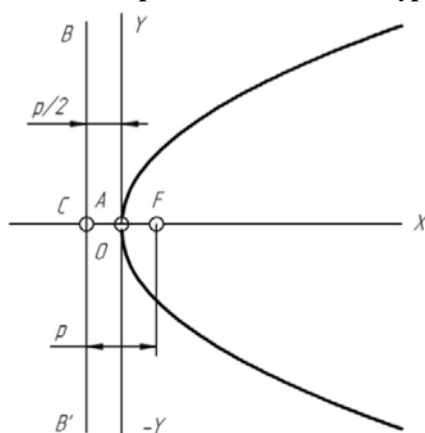


Рис. 1. Парабола

Элементы параболы (рис. 1): кривая параболы; *A* – вершина параболы, она совпадает с началом 0 координат; *BB'* директриса параболы; *F* – фокус параболы.

Параметры параболы: $FC=p$ – фокальный параметр, т.е. расстояние от фокуса *F* до директрисы *BB'*.

Каноническое уравнение параболы:

$$y^2=2px \quad (1)$$

Чем больше величина *p*, т.е. чем дальше фокус от вершины, тем шире раствор ветви параболы.

Сначала перепишем канонические уравнения (1) в другой форме, т.е. выразим для параболы зависимость координаты *y* от *x*:


$$y = \sqrt{2px} \quad (2)$$

Теперь, задав значение фокального параметра *p* кривой и подставив его в формулу (2), можно получить её изображение на экране монитора с помощью команд из библиотеки **FTDraw**.

Алгоритм создания параболы

Открываем документ **Фрагмент** и присваиваем ему имя **Парабола**. Назначим для параболы параметр *p* (в примере его параметр *p* равен 10 мм). В этом случае вершина параболы будет совпадать с началом координат, а по оси *X* на расстоянии 5 мм вправо будет расположен фокус параболы (рис. 1). Тогда по формуле (2) $y = \sqrt{2 \cdot 10 \cdot x}$.

Приступаем к работе с библиотекой **FTDraw** (рис. 2). Обращаемся к командам: **Менеджер библиотек** (📁) → **Прочие** (📁 Прочие). В списке библиотек **Прочие** щёлкаем курсором на **Библиотека FTDraw**.

В диалоговом окне (ДО) **Библиотека FTDraw1.1** (рис. 2, а) выполняем двойной щелчок мышью на кнопке  (**Построение графиков функциональных зависимостей по уравнению в декартовых координатах**). Появляется новое ДО **Построение графика функциональных зависимостей** (рис. 2, б).

Используя форму записи формул, принятую в системе КОМПАС-3D, записываем в этом ДО уравнения параболы $\sqrt{2*10*x}$ (рис. 2, б).

Заказываем пределы изменения x (в примере установлены пределы от 10 до 40 мм) и количество точек кривой (для обеспечения более высокой точности графика в примере выбрано 30 точек). В ДО (рис. 2, в) **Построение графиков функциональных зависимостей** щёлкаем на кнопке **Просмотр расчёта** (📄). Если расчёт выполнен по заданной формуле и при заданных границах, то просто закрываем ДО **Просмотр результатов расчёта**.

В ДО **Построение графиков функциональных зависимостей** (рис. 2, в) щёлкаем на кнопке **Указать положение базовой точки графика** (📍).

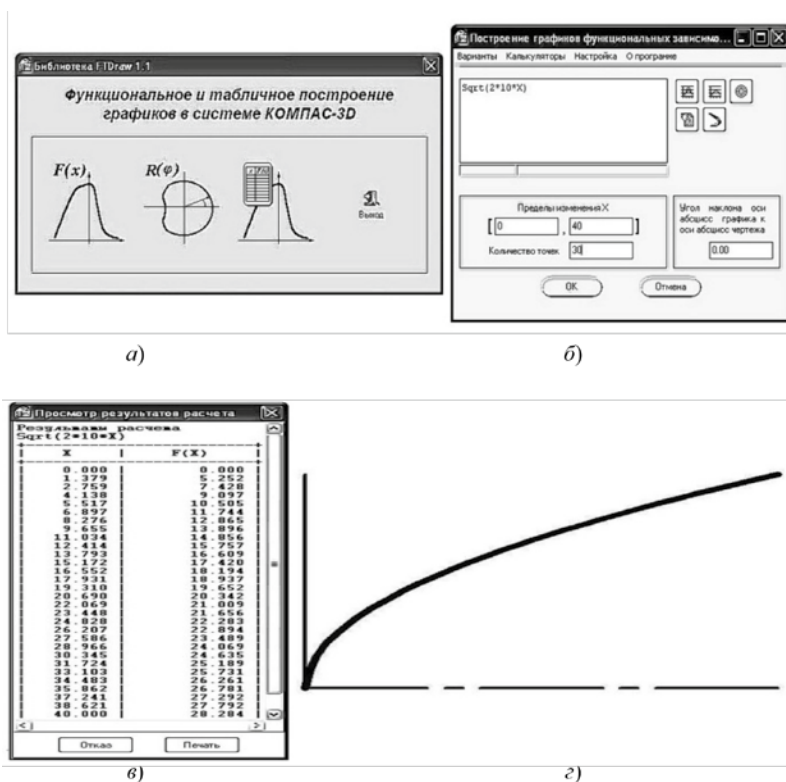



Рис. 2. Построение параболы в программе КОМПАС-3D

На экране появляется контекстное меню  Ввод базовой точки графика. Курсором привязываемся к началу координат документа **Парабола**. Затем в ДО (рис. 2, в) щёлкаем по кнопке **Построить график** (📈) и по кнопке **ОК**. В результате получаем изображение верхней половины ветви параболы (рис. 2, г). Закрываем диалоговые окна.

Это изображение части параболы создано системой в виде **Макроэлемента**. Чтобы его использовать для 3D моделирования, кривую выделяем и командой **Разрушить** из меню **Редактор** преобразуем в **кривую Безье**, а через вершину кривой проводим ось симметрии кривой командой **Отрезок** с типом линии **Осевая**. Выделяем фигуру командой **Выделить** и копируем её в буфер обмена с помощью команды **Копировать всё** с привязкой к вершине кривой.

Алгоритм создания эллиптического параболоида


Создаём новый документ **Деталь**, в **Дереве модели** присваиваем детали имя **Параболоид** и под этим именем сохраняем файл.

1. За плоскость эскиза выбираем фронтальную плоскость проекций **XУ**.

2. Выбираем команду **Эскиз** (для создания эскиза параболы) и командой **Вставить** вставляем эскиз параболы с привязкой к началу координат.

3. Выходим из режима **Эскиз**.

4. Обращаемся к команде **Операция вращения**. На рис. 3 показана **Панель свойств** команды с назначенными параметрами. Опцией команды **Создать объект (8)** создаем 3D-модель параболоида (рис. 3.2, а).

Можно, например, создать из параболоида оболочку с толщиной стенки 1мм (рис. 3, б). Для этого на **Панели свойств** выбираем команду **Оболочка** .

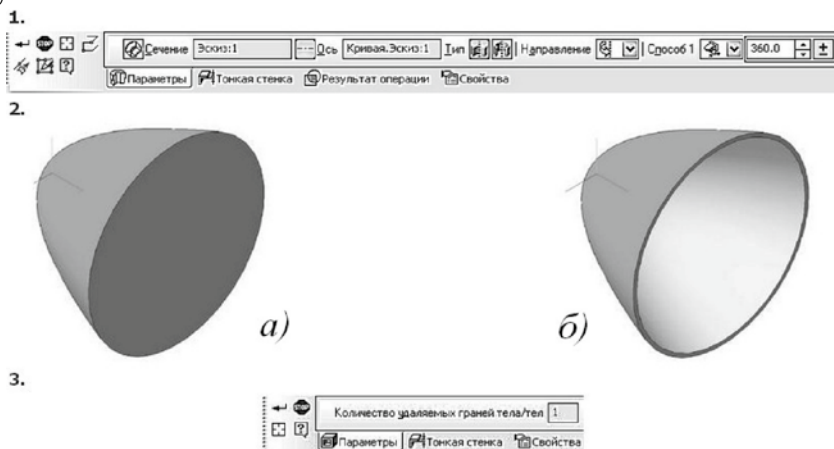


Рис. 3. Построение эллиптического параболоида

Использование параметрических уравнений параболы

В графическом редакторе КОМПАС-3D при работе с документом **Деталь** в меню **Пространственные кривые (2)** имеется команда **Кривая по закону (3)**. Эта команда позволяет получить изображение кривой второго порядка, заданной в параметрической форме в

пространственной прямоугольной системе $(\sqrt[2]{\frac{x}{y}})$ координат. Так как имеем дело с плоской кривой, то при задании координат X и Y в параметрической форме координате Z необходимо задать значение равное нулю.

Чтобы использовать эту кривую для моделирования поверхности второго порядка необходимо спроецировать её на плоскость XU эскиз. Созданные в эскизе проекции **Кривой по закону** ассоциативно связаны со своими исходными объектами. Например, при редактировании параметров кривой будет изменяться и эскиз. Это проекционная связь не даёт возможность редактировать эскиз на месте.

Удаление этой проекционной связи возможно двумя путями: *вырезать созданный эскиз в буфер и обратно вернуть плоскость эскиза, либо после создания эскиза удалить Кривую по закону.*

Параметрическое уравнение параболы

Используя формулу (2) канонического уравнения параболы введём следующие параметрические зависимости:

$$x = t^2; \quad y = t \cdot \sqrt{2 \cdot p}; \quad z = 0. \quad (3)$$

Это и есть параметрическое уравнение параболы, расположенной в плоскости XU , при этом параметр t определяет фактически длину параболы по оси X , параметр p – ширину раствора параболы.


Алгоритм создания гиперболического параболоида


Для создания гиперболического параболоида нужны две параболы – одна в роли направляющей, а вторая в роли образующей. Их плоскости должны располагаться взаимно перпендикулярно.

Открываем документ *Деталь. В Дереве модели* с помощью команды **Свойства модели** присваиваем файлу имя **Гиперболический параболоид** и сохраняем файл под этим именем.

Параметрическое уравнение зададим в прямоугольной $(\sqrt[2]{\frac{x}{y}})$ системе координат. Для создания параболы введём, например:

- интервал изменения параметра – $t = [0; 7]$,
- расстояние от фокуса параболы до директрисы – $p = 7$.

Для этого активизируем инструментальную панель **Пространственные кривые** (2). В этой панели выбираем команду **Кривая по закону** (3) и в **Панели свойств** команды, представленной в виде двух строк (рис. 4) записываем $x=t^2$; $y=t*\text{sgrt}(2*p)$; $z=0$. Завершаем работу команды опцией **Создать объект** .

1. За плоскость эскиза выбираем фронтальную плоскость проекции XU .
2. Выбираем команду **Эскиз** (для создания эскиза параболы) и:
 - командой  **Спроецировать объект** (из меню **Геометрия**) создаём проекцию параболы с **Кривой по закону**;

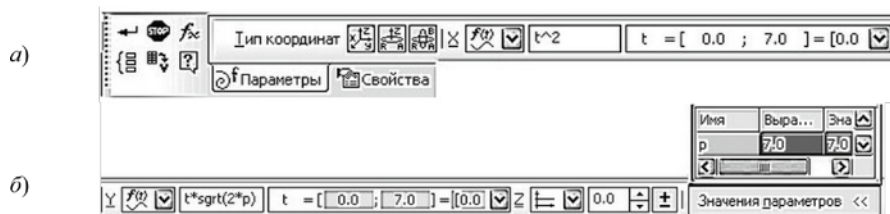


Рис. 4. Фиксация параметрического уравнения параболы в программе КОМПАС-3D
(а – верхняя панель свойств команды Кривая по закону; б – нижняя панель свойств команды Кривая по закону)

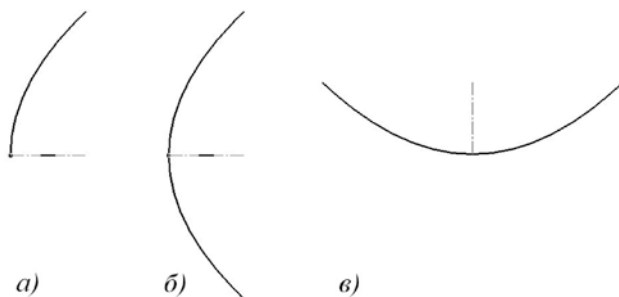


Рис. 5. Построение эскиза первой параболы

- командой **Выделить** выделяем эскиз параболы;
- командой **Вырезать** убираем эскиз параболы в буфер обмена;
- командой **Вставить** вставляем эскиз параболы с привязкой к началу координат (рис. 5, а) В результате связь эскиза с кривой по закону удалена;
- проводим ось симметрии (рис. 5, б) и выделяем половину параболы, командой **Симметрия** создаём её вторую половину;
- командой **Поворот** поворачиваем параболу на угол 90° , получаем эскиз первой параболы (рис. 5, в).

1. Выходим из режима Эскиз.

Выполняем создание эскиза второй параболы по аналогичному алгоритму, только с другими данными (рис. 6, а-в):

- интервал изменения параметра – $t = [0; 7]$,
- расстояние от фокуса параболы до директрисы – $p = 3$.

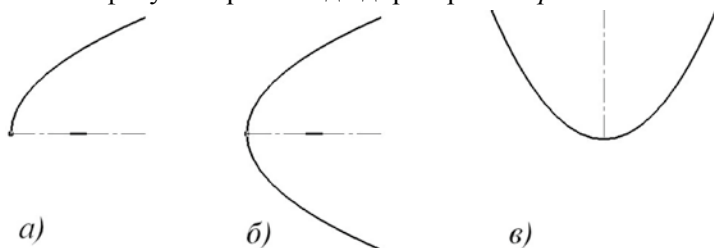


Рис. 6. Построение эскиза второй параболы

Если при заполнении **Строки свойств** команды **Кривая по закону** невозможно изменить величину p , то можно изменить имя p , например, на имя k и после этого в **Строке свойств** задать значение k (в примере значение $k = 3$) (рис. 7).

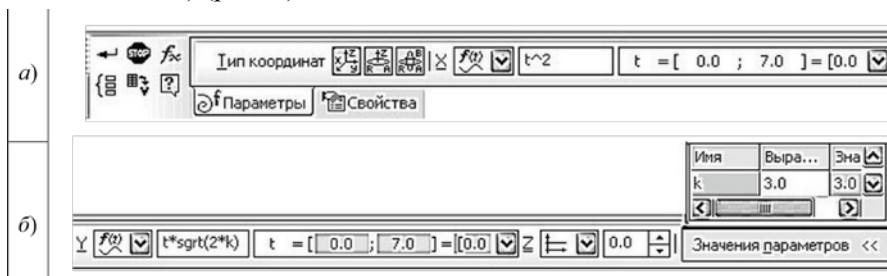






Рис. 7. Введение данных для построения эскиза второй параболы



Эскиз первой параболы будет служить направляющей гиперboloида в команде **Кинематическая операция**. Для эскиза второй параболы необходима плоскость, перпендикулярная к плоскости XY и проходящая через конечную точку первой параболы, созданную в эскизе 1.

Командой () **Плоскость через вершину параллельно другой плоскости** из меню () **Вспомогательная геометрия** создаём плоскость (**Параллельная плоскость 1**), проходящую через конец параболы и параллельную профилевой плоскости проекций ZY.

Командой **Редактировать** эскиз 2 открываем изображение второй параболы, командой **Выделить** выделяем изображение, командой **Вырезать** переносим эскиз с привязкой к началу координат в буфер обмена. Закрываем пустой эскиз 2. Выделяем созданную параллельную плоскость 1 и создаём в ней эскиз 3, т.е. из буфера обмена вставляем эскиз 2 с привязкой к началу координат. Закрываем эскиз 3.

Выбираем команду **Кинематическая операция** () (рис. 8, а). В качестве эскиза-сечения выделяем эскиз 3, а в качестве траектории выделяем эскиз 1, завершаем работу опцией  **Создать объект** (рис. 8, б).

В результате получим изображение поверхности гиперболического параболоида, не имеющей толщины (рис. 8, б).

На созданном изображении скрыты изображения Кривой по закону:1 и Кривой по закону:2, а также конструктивные плоскости. Толщина гиперболического параболоида для примера назначена в размере 0,01 мм. Для получения изображения не тела, а поверхности можно воспользоваться командой **Кинематическая поверхность** () (рис. 8, в) из инструментальной панели **Поверхности** (.

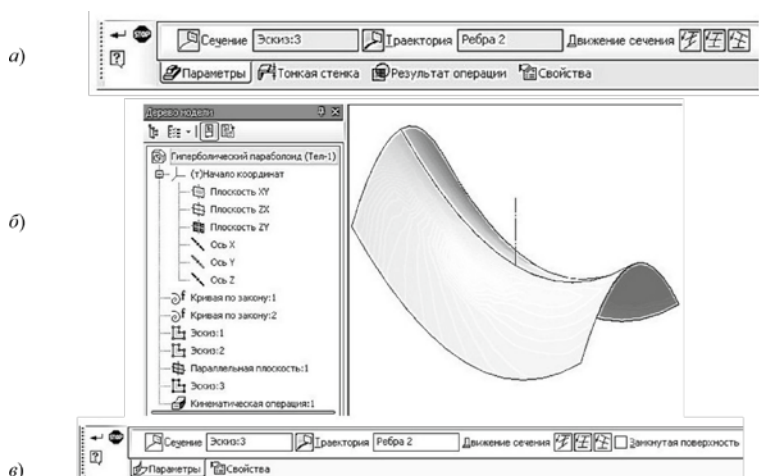


Рис. 8. Создание гиперболического параболоида

Использование канонических уравнений гиперболы

В КОМПАС-3D среди множества библиотек имеется библиотека **FTDraw**, входящая в меню библиотек **Прочие**. С помощью этой библиотеки по каноническому уравнению кривой можно построить график функции, т.е. изображение кривой гиперболы или параболы по заданным условиям. Затем полученное изображение можно использовать для 3D – моделирования поверхностей 2-го порядка или 4-го порядка.

Геометрия и каноническое уравнение гиперболы

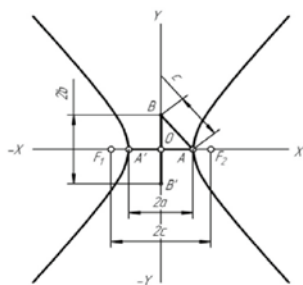


Рис. 9. Гипербола

Элементы гиперболы (рис. 9): ветви – левая и правая; центр гиперболы 0; F_1 и F_2 – фокусы; A и A' – вершины гиперболы; отрезок AA' – действительная ось гиперболы; B и B' – концы мнимой оси гиперболы.

Параметры гиперболы: $2a$ – размер действительной оси; $2b$ – размер мнимой оси; $2c$ – расстояние между фокусами.

Каноническое уравнение гиперболы:

$$\frac{x^2}{a^2} - \frac{y^2}{b^2} = 1. \quad (4)$$

При этом известна зависимость: $c^2 = a^2 + b^2$, т.е. зная величины полуосей гиперболы, можно узнать величину межфокусного расстояния c .

Мнимая ось $2b$ может быть меньше, равна или больше действительной оси $2a$. Если оси равны ($a=b$), то гипербола называется равносто-

ронней или равнобокой. Чем больше действительная полуось a по отношению к мнимой полуоси b , тем уже раствор ветви гиперболы.

Выразим для гиперболы зависимость координаты y от x из формулы (4):

$$\frac{y^2}{b^2} = \frac{x^2}{a^2} - 1 \rightarrow y^2 = b^2 * \left(\frac{x^2}{a^2} - 1\right) \rightarrow y = b\sqrt{\frac{x^2}{a^2} - 1}; \quad (5)$$

Теперь, используя формулу (5), задав нужные параметры кривой, можно получить её изображение на экране монитора с помощью команд из библиотеки **FTDraw**.

Алгоритм создания гиперболы

Открываем документ **Фрагмент** и присваиваем ему имя **Гипербола**. Назначим для гиперболы, например, следующие параметры: размер действительной полуоси $a = 10$, мнимой полуоси $b = 12$, тогда по формуле (5):

$$y = 12\sqrt{\frac{x^2}{100} - 1}.$$

Приступаем к работе с библиотекой **FTDraw** (рис. 10). Обращаемся к командам: **Менеджер библиотек** (📁) → **Прочие** (📁 Прочие). В списке библиотек **Прочие** щёлкаем курсором на **Библиотека FTDraw** (рис. 10, а).

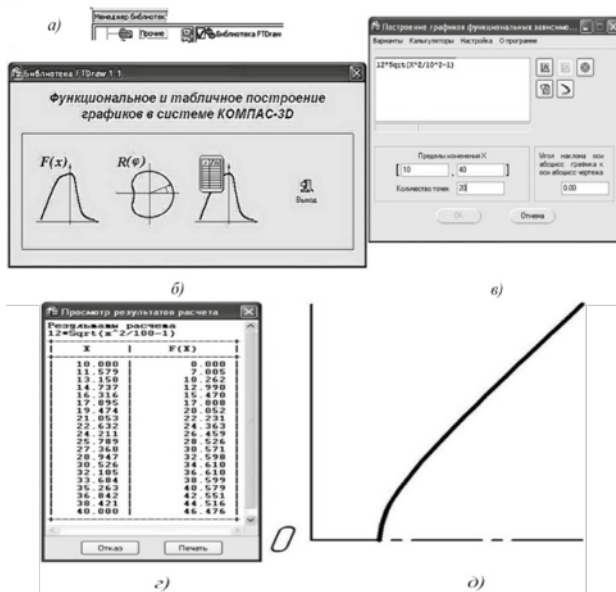





Рис. 10. Построение гиперболы в программе КОМПАС-3D



В диалоговом окне (ДО) *Библиотека FTDraw1.1* (рис. 10, б) выполняем двойной щелчок мышью на кнопке  (*Построение графиков функциональных зависимостей по уравнению в декартовых координатах*). Появляется новое ДО – *Построение графика функциональных зависимостей* (рис. 10, в).

Используя форму записи формул, принятую в системе КОМПАС-3D, записываем в этом ДО уравнения гиперболы – $12*\text{Sqrt}(x^2/10^2-1)$ (рис. 10, в).

Судя по параметрам уравнения, вершина гиперболы удалена от оси Y на 10 мм. Поэтому заказываем пределы изменения x от 10 до, например, 40 мм и количество точек кривой 20 (для обеспечения более высокой точности графика).

В ДО (рис. 10, в) имеется пять команд для управления процессом создания графиков. В ДО *Построение графиков функциональных зависимостей* щёлкаем на кнопке *Просмотр расчёта* (). Если расчёт выполнен по заданной формуле и при заданных границах (рис. 10, з), то просто закрываем ДО *Просмотр результатов расчёта*.

В ДО *Построение графиков функциональных зависимостей* (рис. 10, в) щёлкаем на кнопке *Указать положение базовой точки графика* ().

На экране появляется контекстное меню  Ввод базовой точки графика. Курсором привязываемся к началу координат документа *Фрагмент*. Затем в ДО (рис. 10, в) щёлкаем по кнопке *Построить график* () и по кнопке *ОК*. В результате получаем изображение верхней половины правой ветви гиперболы (рис. 10, д). Закрываем диалоговые окна.

Это изображение части гиперболы создано системой в виде *Макроэлемента*. Чтобы его использовать для 3D моделирования, кривую выделяем и командой *Разрушить* из меню *Редактор* преобразуем в кривую Безье. Проводим оси гиперболы через начало координат: действительную ось линией *Осевая*, а мнимую ось – *Тонкой линией*. **Выделяем** всё и **Копируем** в буфер обмена с привязкой к центру гиперболы.

Алгоритм создания гиперболического цилиндра

Верхняя половина правой ветви гиперболы создана (рис. 11, а). Однако для моделирования гиперболического цилиндра необходимо иметь полное изображение гиперболы.

Командой *Симметрия относительно оси X* из меню *Редактор* создаём нижнюю часть правой ветви гиперболы (рис. 11, б), затем командой *Симметрия относительно оси Y* из меню *Редактор* создаём левую ветвь гиперболы (рис. 11, в). **Выделяем** полное изображение гиперболы. Командой *Копировать* из меню *Редактор* копируем гиперболу в буфер обмена с привязкой к центру O гиперболы.

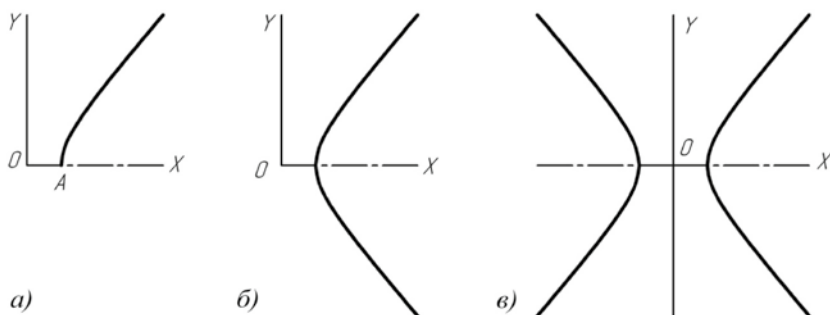


Рис. 11. Построение гиперболы



Рис. 12. Поверхность гиперболического цилиндра

Создаём новый документ *Деталь*, в дереве модели присваиваем детали имя *Гиперболический цилиндр* и сохраняем файл под этим именем.




1. За плоскость эскиза выбираем горизонтальную плоскость проекций *ZX*.

2. Выбираем команду *Эскиз* (для создания эскиза гиперболы) → командой **Вставить** вставляем эскиз гиперболы с привязкой к началу координат.

3. Выходим из режима *Эскиз*.

Обращаемся к команде **Операция выдавливания**. Назначаем **необходимые параметры**, например, с тонкой стенкой толщиной **0,01 мм** и опцией **Создать объект** создаем 3D модель гиперболического цилиндра (рис. 12).

Использование параметрических уравнений гиперболы

В графическом редакторе КОМПАС-3D при работе с документом *Деталь* в меню *Пространственные кривые* () имеется команда **Кривая по закону** (). Эта команда позволяет получить изображение кривой 2-го порядка, заданной в параметрической форме в пространственной прямоугольной системе () координат. Так как имеем дело с плоской кривой, то при задании координат *X* и *Y* в параметрической форме координате *Z* необходимо задать значение равное нулю.

Чтобы использовать эту кривую для моделирования поверхности 2-го порядка необходимо спроецировать её на плоскость *XY* Эскиза. Созданные в эскизе проекции **Кривой по закону** ассоциативно связаны со своими исходными объектами. Например, при редактировании параметров кривой будет изменяться и эскиз. Это проекционная

связь не даёт возможности редактировать эскиз на месте. Её удаление возможно двумя путями: вырезать созданный эскиз в буфер и обратно вернуть плоскость эскиза или после создания эскиза удалить Кривую по закону.

Параметрическое уравнение гиперболы

Используем форму записи **параметрического** уравнения гиперболы с помощью гиперболических функций:

$$x = a * sh(t), \quad y = b * ch(t), \quad (6)$$

При форме записи (6) действительная ось гиперболы совпадает с осью X.

Здесь: t – параметр уравнений, a – размер действительной полуоси гиперболы, b – размер мнимой полуоси гиперболы, а гиперболические функции определены следующими зависимостями:

$$sh(t) = \frac{e^t + e^{-t}}{2}, \quad ch(t) = \frac{e^t - e^{-t}}{2}, \quad \text{где } e=2,718281828 \text{ – число Эйлера.}$$

Существует также форма записи параметрического уравнения гиперболы и в тригонометрических функциях.

Алгоритм создания однополостного гиперболоида

Открываем документ *Деталь*. В *Дереве модели* с помощью команды **Свойства модели** присваиваем файлу имя **Однополостный гиперболоид** и сохраняем файл под этим именем.

Создадим **Эскиз** гиперболы в плоскости XU по параметрическому уравнению (6), например, при следующих данных: $a=15$, $b=25$, $t=1,5$.

Для этого активизируем инструментальную панель **Пространственные кривые**. В этой панели выбираем команду **Кривая по закону**.

В панели **Свойств** команды записываем исходные данные гиперболы в прямоугольной системе координат (рис. 13, а). Завершаем работу опцией **Создать объект**.

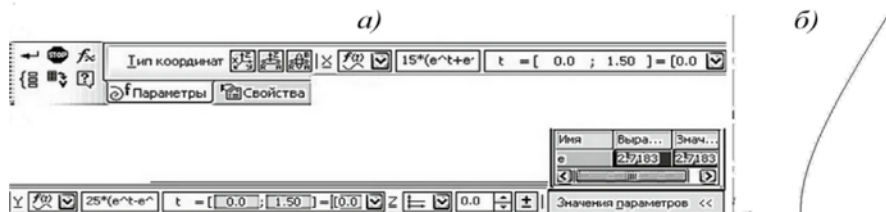




Рис. 13. Внесение исходных данных гиперболы

В результате создана часть гиперболы в плоскости XU системы (рис. 13, б).

Создаём эскиз гиперболы:

► выделяем плоскость XU в дереве модели и выбираем команду **Эскиз**;

► в инструментальной панели **Геометрия** выбираем команду  **Специализировать объект**. Выделяем созданную кривую. Завершаем работу кнопкой **Stop** ;

► **выделяем** этот эскиз и командой **Вырезать** из меню **Редактор** удаляем эскиз в буфер обмена;

► командой **Вставить** из меню **Редактор** возвращаем изображение в эскиз плоскости XY. В результате связь с **Кривой по закону** разорвана;

► через начало координат проводим горизонтальную прямую стилем *Вспомогательная* и вертикальную прямую стилем *Осевая*;

► **выделяем** кривую и командой **Симметрия** создаём полное изображение правой ветви гиперболы;

► командой **Эскиз** завершаем создание изображения правой ветви гиперболы.

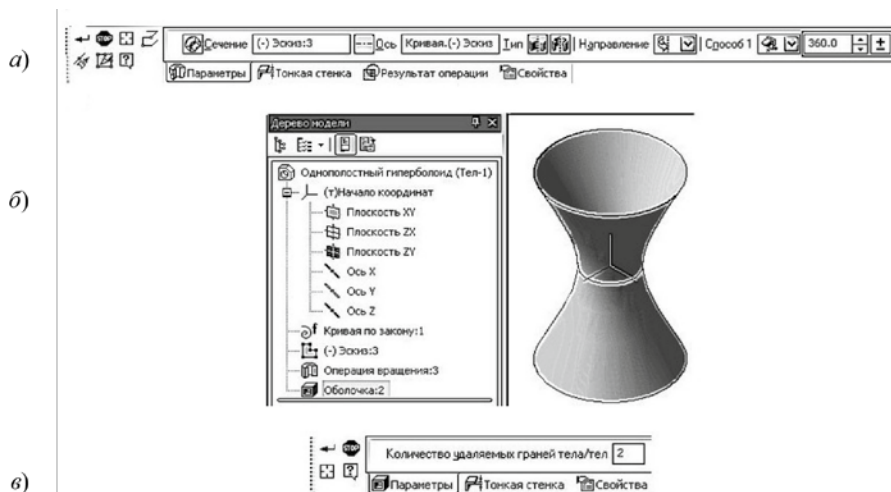


Рис. 14. Создание однополостного гиперболоида

Для моделирования однополостного гиперболоида вращения используем команду **Операция вращения** (рис. 14, а) – вращение выполнено вокруг мнимой оси гиперболы. В результате получаем тело ограниченное поверхностью гиперболоида. Для получения тонкостенного изображения гиперболоида использована команда **Оболочка** (рис. 14, в). В результате получено 3D изображение тонкостенного однополостного гиперболоида вращения с толщиной стенки 0,01 мм (рис. 14, б).

Алгоритм создания двуполостного гиперболоида вращения

Открываем документ **Деталь**. В **Дереве модели** с помощью команды **Свойства модели** присваиваем файлу имя **Двуполостный гиперболоид** и сохраняем файл под этим именем.

Кривая по закону создана с параметрами $a = 25$, $b = 15$ и $t = 2$ (рис. 15).



Рис. 15. Параметры кривой

С помощью этой кривой создан эскиз гиперболы с поворотом на 90° . Затем командой **Операция вращения** (рис. 16, а) создана 3D – модель гиперboloида и операцией **Оболочка** (рис. 16, в) получено изображение гиперboloида с толщиной стенки равной 0,01 мм (рис. 16, б). Для наглядности изображена его действительная ось.

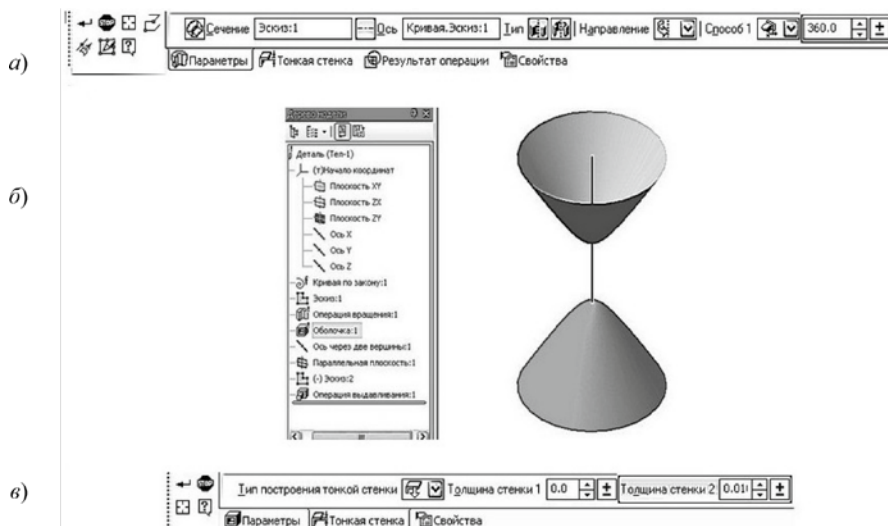


Рис. 16. Создание двухподошного гиперboloида

Заключение

В работе подробно изложены алгоритмы построения кривых второго порядка – параболы и гиперболы на основе применения их **канонических и параметрических уравнений**, а также моделирования производных 3D-поверхностей в системе трёх мерного моделирования КОМПАС-3D.

Данная работа направлена на более глубокое изучение цикла геометро-графических дисциплин, и продвижение отечественного IT-продукта в сфере образования. Она будет интересна студентам, преподавателям и специалистам в области геометрического моделирования.

Список литературы

1. Дорохов А. С. Компьютерное проектирование в системе AutoCAD. М.: Российский государственный аграрный университет – МСХА им. К.А. Тимирязева (Москва), 2016. 80 с.
2. Дорохов А. С. Выполнение чертежей с использованием системы «Компас-3D». М.: Российский государственный аграрный университет – МСХА им. К.А. Тимирязева (Москва), 2016. 76 с.
3. КОМПАС-3D V13: руководство пользователя в 3-х т. М.: Изд-во ИТАР-ТАСС, 2012. Т.3. 424 с.
4. Иванов Г. С. Теоретические основы начертательной геометрии: учеб. пособие. М.: Машиностроение, 1998. 160 с.
5. Фролов С. А. Начертательная геометрия: учеб. для вузов. 3-е изд., перераб. и доп. М.: Машиностроение, 2008. 240 с.

УДК 531.5

В.О. Гроппен

Северо-Кавказский горно-металлургический институт (государственный технологический университет)

ИСПОЛЬЗОВАНИЕ ИНФОРМАЦИОННЫХ ТЕХНОЛОГИЙ ДЛЯ АНАЛИЗА ДИАГРАММЫ ХАББЛА

Аннотация: Целью является демонстрация возможности использования современных информационных технологий для создания модели Вселенной, альтернативной Стандартной Модели и свободной от ряда противоречий последней. Предлагаемый подход позволяет трактовать закон Хаббла, как отображение сочетания сравнительно стабильной Вселенной с вариабельностью эталонов, используемых для измерения расстояний.

Ключевые слова: Вселенная, эталон, закон Хаббла, красное смещение, «усталость» света, ускоренное разбегание галактик, вариабельность.

V.O. Groppen

North Caucasus Mining and Metallurgical Institute (State Technological University)
groppen@mail.ru www.skgmi-gtu.ru

USING INFORMATION TECHNOLOGY TO ANALYZE THE HUBBLE CHART

Abstract. The goal is to demonstrate the possibility of using modern information technologies to create a model of the Universe, which on the one hand is an alternative to the Standard Model and on the other hand is free from a number of contradictions of the latter. The proposed approach allows us to interpret the Hubble law as a mapping of a combination of a relatively stable Universe with the variability of the standards used to measure distances.

Keywords: Universe, standard, Hubble's law, redshift, light "fatigue", accelerated scattering of galaxies, variability.

Введение

Принятая сегодня Стандартная Модель Вселенной базируется на

опубликованных в двадцатых годах прошлого века идеях Александра Фридмана о расширяющейся и пульсирующей Вселенной [1], на предложенной Джоржем Леметром в 1927 году идее Большого Взрыва и, как следствие, расширяющейся Вселенной [2], на работах Георгия Гамова, предсказавшего реликтовое излучение [3]. Эти идеи получили подтверждение в опубликованном в 1929 году экспериментальном законе Хаббла [4], зафиксировавшем линейную зависимость радиальных скоростей галактик от их расстояний до земного наблюдателя (Рис. 1), и заставившем отбросить теорию стационарной Вселенной, предложенную Германом Бонди, Томасом Голдом и развитую Фредом Хойлом [5, 6]. Основу закона Хаббла составляют два принципа:

а) красное смещение однозначно определяет скорость разбегания галактик; б) используемые для измерения времени и расстояний эталоны неизменны во времени.

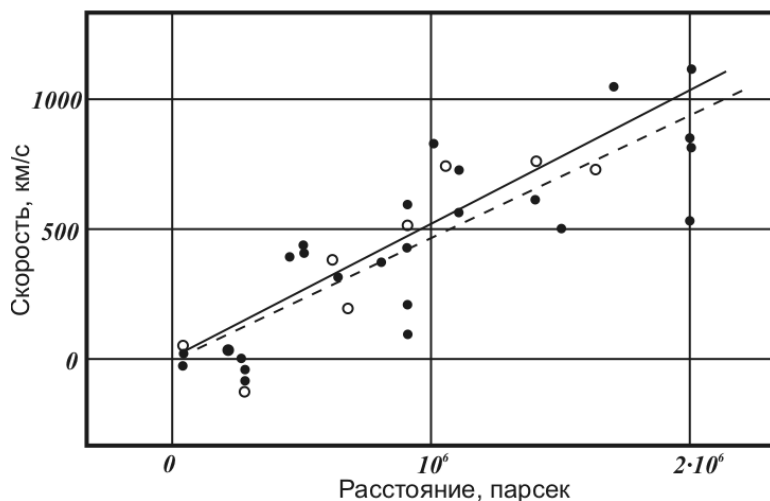


Рис.1. Диаграмма, иллюстрирующая закон Хаббла.

Поводами, позволяющими усомниться в адекватности Стандартной Модели, являются шесть присущих ей и приводимых ниже противоречий [7-10]:

- Равная плотность распределения материи во Вселенной, если последнюю представить разделенной на кубы с длиной ребра не менее 10^9 световых лет.

- Во второй половине прошлого века астроном Алан Сандаж выяснил, что закон Хаббла справедлив и для областей Вселенной, в которых материя распределена неравномерно, что противоречит результатам Фридмана.

- На расстояниях от центра Местного скопления галактик, не превышающих 2 Мпк, закон Хаббла не выполняется.

- Красное смещение оказалось независимым от длины волны.
- Отсутствует объяснение причин потери со временем световым лучом энергии, причем фокусировка луча сохраняется.
- Анизотропия реликтового излучения противоречит идее Большого Взрыва.

Основой предлагаемого ниже альтернативного подхода является возможность представить любой измерительный процесс алгоритмом, состоящим из трех шагов:

Шаг 1. Величина «А» измеряемого параметра фиксируется в условных единицах.

Шаг 2. Аналогично величина «В» используемого для измерения этого параметра эталона фиксируется в условных единицах.

Шаг 3. Результат, фиксируемый наблюдателем, равен «С»: $C = A/V$.

Примером может служить использование для измерения расстояний лазерного дальномера. Величина «А» в этом случае равна числу импульсов, сгенерированных генератором за время, прошедшее от начала измерения, до возвращения отраженного сигнала, величина «В» равна числу импульсов, сгенерированных генератором за время, прошедшее от начала измерения, до возвращения отраженного сигнала при условии, что предмет, от которого отражается свет, находится от его источника на расстоянии, равном одному метру. Очевидно, что «С» в этом случае – расстояние между генератором световых импульсов и отражающим их объектом, измеренное в метрах. Далее используются следующие обозначения, определения и допущения.

Обозначения, определения и допущения

t – текущее время, фиксируемое с помощью часов наблюдателя;

r – эталон длины, измеренный в условных единицах;

R_i – расстояние между наблюдателем и i -м объектом в условных единицах;

L_i – расстояние между наблюдателем и i -м объектом, определяемое отношением R_i и r :

$L_i = R_i / r$; (1) H – постоянная Хаббла [4]: $H \approx 2.282 \cdot 10^{-18}$ (с.⁻¹); (2)

V_i – радиальная скорость i -го космического объекта: $V_i = dL_i/dt$; (3)

I – множество индексов, принадлежащих объектам, для которых справедлив закон Хаббла, т.е. таких, для которых справедливо: $\forall i \in I: L_i > D$;

h – постоянная Планка ($h \approx 6.626 \cdot 10^{-34}$ джоулей в секунду);

\mathcal{R} – радиус видимой Вселенной ($\mathcal{R} \approx 13.75 \cdot 10^9$ световых лет);

λ – длина электромагнитной волны, зафиксированная наблюдателем;

λ_0 – длина волны генератора электромагнитных колебаний;

c – скорость света в вакууме ($c \approx 299\,792\,458$ м / сек.);

ν – частота электромагнитных колебаний: $\nu \cdot \lambda = \tilde{\nu}$; (4)

Ниже игнорируется наличие у космических объектов пекулярных скоростей.

Стационарная модель Вселенной с варьируемым эталоном расстояний

В основе приводимых ниже выкладок лежат два допущения, противоположные тем, которые были использованы при построении модели расширяющейся Вселенной:

- допущение о стационарности Вселенной;
- допущение о вариабельности эталона, используемого для измерения расстояний.

Эти допущения, в соответствии с принятыми выше обозначениями, приводят к системе:

$$\begin{cases} \forall i \in I : \frac{dL_i}{dt} = HL_i; \\ \forall i : L_i = \frac{R_i}{r}; \\ \forall i \in I, \frac{dR_i}{dt} = 0. \end{cases} \quad (6)$$

Решением системы (5) является: $r = r_0 \cdot \exp(-Ht)$, (6)

где r_0 – размер эталона, используемого для измерения расстояний в условных единицах, в момент в начале измерительного процесса, т.е., когда $t = 0$.

Этот результат приводит к следующим выводам:

• закон Хаббла можно объяснить не только расширением Вселенной, но и экспоненциальным сокращением эталона длины в стационарной Вселенной;

• так как кратчайшее расстояние между любыми двумя точками твердого тела можно использовать в качестве эталона расстояний, можно утверждать, что линейные размеры любого твердого тела со временем уменьшаются в соответствии с (6);

• вследствие того, что линейные размеры любого твердого тела со временем изменяются так же, как и используемые для их измерения эталоны, зафиксировать эти изменения не представляется возможным:

$$\forall i : L_i = \frac{R_i}{r_i} = \frac{R_0 \exp(-Ht)}{r_0 \exp(-Ht)} = \frac{R_0}{r_0}. \quad (7)$$

Уравнение (6) в стационарной Вселенной порождает у наблюдателя ряд приведенных ниже иллюзий.

Расстояние, скорость, ускорение

Ниже приводится описание ряда иллюзий в стационарной Вселенной, порождаемых экспоненциальным сокращением величины этало-

на длины. Для этого далее полагаем, что расстояние R_0 между наблюдателем и космическим объектом, измеренное в условных единицах, не меняется со временем. Тем не менее, благодаря (6), наблюдатель зафиксирует экспоненциальный рост расстояния L между ним и этим объектом:

$$L = \frac{R_0}{r} = \frac{R_0}{r_0} \exp\{Ht\} = L_0 \exp\{Ht\}. \quad (8)$$

Используя (8), легко показать, что скорость этого роста расстояния отвечает закону Хаббла:

$$\frac{dL}{dt} = \frac{R_0}{r_0} \frac{d}{dt}(\exp\{Ht\}) = L_0 H \exp\{Ht\} = LH. \quad (9)$$

Это означает, что закон Хаббла нельзя рассматривать, как окончательное доказательство расширения Вселенной: альтернативной его трактовкой является сокращение длины эталонов, используемых для измерения расстояний в стационарной Вселенной. Продолжением этого подхода является иллюзия ускоренного разбегания галактик:

$$\frac{d^2L}{dt^2} = \frac{HR_0}{r_0} \frac{d}{dt}(\exp\{Ht\}) = L_0 H^2 \exp\{Ht\} = LH^2. \quad (10)$$

Следует заметить, что в [11] приводится иное объяснение (10), использующее замедление времени, которое, однако, тоже не опирается на Стандартную Модель.

«Усталость» света и красное смещение

Предложенный выше подход позволяет по-новому объяснить потерю энергии фотонами – объединив (4) и (10), получим систему:

$$\begin{cases} \lambda = \lambda_0 \exp\{Ht\}; \\ E = h\nu; \\ \lambda = \frac{c}{\nu}, \end{cases} \quad (11)$$

из которой следует экспоненциальная зависимость энергии фотона от времени его существования: $E = E_0 \cdot \exp\{-Ht\}$. (12)

Иными словами, потеря энергии фотоном вызвана не его взаимодействиями, а постепенным увеличением длины волны, вызванным сокращением длины эталона, используемого для измерения расстояний, отвечающего (6). Близкий, но полученный с помощью иного подхода, результат, приведен в [12].

Предложенная выше модель, сочетающая стационарность Вселен-

ной с вариабельностью эталонов длины, позволяет также объяснить независимость красного смещения от длины волны. Величина красного смещения z определяется выражением:

$$z = (\lambda - \lambda_0) / \lambda_0. \quad (13)$$

Подставляя (1) в (13), получим:

$$z = \frac{\frac{R_0}{r} - \frac{R_0}{r_0}}{\frac{R_0}{r_0}} = \frac{r}{r_0} - 1. \quad (14)$$

Подставляя (6) в правую часть (14) и учитывая, что $t \cdot H \ll 1$, получим выражение, отображающее зависимость величины красного смещения от пройденного пути и независимость этой величины от длины волны:

$$\forall i : z = \frac{HL_i}{c}. \quad (15)$$

Аналогичный результат, но полученный на основании иных предпосылок, приведен в [12].

Заключение

Модель, сочетающая относительную стационарность Вселенной с вариабельностью эталонов длины, позволяет:

1. Объяснить закон Хаббла не прибегая к теории Большого взрыва. Другими словами, закон Хаббла нельзя рассматривать, как окончательное доказательство расширения Вселенной: его альтернативной трактовкой является экспоненциальное сокращение длины эталонов, используемых для измерения расстояний в стационарной Вселенной.

2. Объяснить ряд явлений, игнорируемых Стандартной моделью, таких, как «усталость» света, анизотропию реликтового излучения и независимость красного смещения от длины волны.

3. Трактовать закон Хаббла и ускоренное разбегание галактик, как иллюзии, являющиеся следствием экспоненциального сжатия со временем эталонов длины.

Можно утверждать, что предложенная модель вводит относительность понятия «расстояние» – последнее имеет смысл только в соотношении со временем измерения.

Литература

1. Friedman, A. On the Curvature of Space. Z. Phys. 10 (1), 1922, pp. 377–386.
2. Lemaître, G. Un Univers homogène de masse constante et de rayon croissant rendant compte de la vitesse radiale des nébuleuses extra-galactiques. Annales de la Société Scientifique de Bruxelles, 1927 47: 49. Bibcode:1927ASSB...47...49L.
3. Gamov G. Expanding Universe and the Origin of Elements // Physical Review. 1946. N 70. p. 572–573.

4. Hubble, E.P. A relation between distance and radial velocity among extra-galactic nebulae, Proceedings of the National Academy of Sciences of the United States of America, vol. 15, no. 3, 1929, pp. 167–173.
5. Joe Burchfield. Lord Kelvin and the Age of the Earth. London, MacMillan, 1975.
6. Kenneth Lang. Essential Astrophysics. Heigelberg, Springer, 2013.
7. Martinez-Vaquero, L.A., Yepes G., Hoffman Y. and Gottlober G. Near Field Cosmological Simulations: Is Dark Energy Playing a Role in Our Local Neighborhood? Proceedings of the International Conference “Mathematics and Astronomy: a Joint Long Journey”, Madrid, Spain 23-27 November 2009, pp. 166 – 174.
8. LaViolette, P. A. Is the universe really expanding? Astrophysical Journal, Part 1 (ISSN 0004-637X), vol. 301, Feb. 15, 1986, pp. 544-553.
9. Wetterich, C. Universe without expansion. arXiv:1303.6878v3 [astro-ph.CO] 30 July, 2013.
10. V.O. Groppen “Universe Modeling by the Variable Measurement Standards”. International Journal of Applied Physics, volume 1, 2016, pp. 85 – 89.
11. Marc Mars, Jos´e M. M. Senovilla and Raul Vera. Accelerating expansion and change of Signature. <http://arxiv.org/abs/0712.1462>.
12. Sergey V. Alimanov. Hubble’s Quantum law $v_n = nH_0$ (Quantum law of the cosmological redshift). The General Science Journal <http://gsjournal.net/Science-Journals/Research%20Papers-Astrophysics/Download/5048>

УДК 004.942:51.72

Леонтьев А.Л.

Южный федеральный университет, Таганрог, Россия

Таганрогский институт имени А.П. Чехова (филиал) «Ростовского государственного экономического университета (РИНХ)», Таганрог, Россия

Email:Leontyev_Anton@mail.ru

МATHCAD КАК СРЕДСТВО ВИЗУАЛИЗАЦИИ УРАВНЕНИЙ МАТЕМАТИЧЕСКОЙ ФИЗИКИ

Аннотация: Статья посвящена рассмотрению способов визуализации задач курса «Уравнения математической физики». Для визуализации разработаны алгоритмы построения графических моделей гиперболических уравнений разной сложности, с использованием метода разделения переменных и сеточного метода. Средством визуализации и средой программирования выбрана программа MathCAD.

Ключевые слова: Уравнения математической физики, гиперболические уравнения, визуализация, MathCAD, графики функций.

Leontyev A.L.

Southern Federal University, Taganrog, Russia

Taganrog Institute named after A.P. Chekhov (branch) “Rostov State University of Economics (RINH)”, Taganrog, Russia

Email: Leontyev_Anton@mail.ru

MATHCAD AS A MEANS OF VISUALIZATION OF THE EQUATIONS OF MATHEMATICAL PHYSICS

Abstract: *The article is devoted to the consideration of ways to visualize the tasks of the course "Equations of Mathematical Physics". For visualization, algorithms for constructing graphical models of hyperbolic equations of different complexity have been developed, using the method of separation of variables and the grid method. The program MathCAD was chosen as a visualizer and programming environment.*

Keywords: *Equations of mathematical physics, hyperbolic equations, visualization, MathCAD, graphs of functions.*

Введение

Последние 100 лет мировой истории обозначены временем бурного роста технического и научного прогресса, что находит отражение в педагогическом процессе. В высших учебных заведениях вводятся новые предметы, расширяется содержание уже преподаваемых курсов. Для того чтобы получаемая в процессе обучения информация могла быть успешно применена она должна быть доступна, наглядна и понятна. Таким образом учебный курс должен быть построен так, чтобы соблюдались ФГОС. Однако в математической и физической теориях существуют разделы, которые при аналитическом рассмотрении не могут быть наглядно представлены графически, то есть нарушается принцип наглядности информации, что в свою очередь ведет к непониманию у учащихся. Одним из таких разделов является математическая физика.

Математическая физика – теория математических моделей физических явлений [2].

Физика трактует данную теорию как аппарат описания физических моделей, математика как самостоятельный раздел. Понятийный аппарат математической физики формируется из математических методов и физических понятий, используемых при построении и изучении математических моделей, описывающих множество физических явлений. Данная теория широко применяется в теоретической механике, гидродинамики и теории упругости. Начиная с 20-го века, математическая физика широко применяется в квантовой механике, квантовой теории поля, квантовой статической физике, теории относительности, гравитации и синергетики.

Постановка задачи

Основной задаче математической физики является математическое моделирование. Конечная модель должна описывать физическое явление и связанные с ним закономерности. Реализация данной задачи заключается в выводе уравнений, описывающих физический процесс. При математическом моделировании используются только основные законы отражающие фундаментальные характеристики или закономерности, изучаемые непосредственно в рамках рассматриваемой задачи. Такое рассмотрение показало, что при наличии схожих характеристик

у разных физических процессов и явлений, к ним, с определенной долей погрешности, применимы одинаковые математические модели.

Уравнения математической физики описывают множество однотипных явлений и процессов, вследствие того, что дифференциальные выражения на которых строится математическая физика, имеют множество решений. Частное решение рассматриваемого физического явления отбирается из множества решений при помощи системы начальных и граничных условий [1].

Подавляющее число физических явлений и процессов реализуется в форме дифференциалов второго порядка с частными производными. Фундаментальные законы физики записываются как вторые производные. Методика исследования уравнений в математической физике зависит от принадлежности решаемого уравнения к тому или иному типу. Исторически выделено три основополагающих типа дифференциалов в частных производных второго порядка, решение которых качественно отличаются:

- параболического;
- гиперболического;
- эллиптического.

Уравнение с частными производными второго порядка, с независимыми переменными, имеет вид:

$$a(x, y) \frac{\partial^2 \omega}{\partial x^2} + 2b(x, y) \frac{\partial^2 \omega}{\partial x \partial y} + c \frac{\partial^2 \omega}{\partial y^2} = F \left(x, u, w, \frac{\partial \omega}{\partial x}, \frac{\partial \omega}{\partial y} \right). \quad (1)$$

где a, b, c – функции от x, y , имеющие непрерывные производные первого и второго порядка включительно.

Уравнение (1) в точке (x, y) , является уравнением:

• параболического типа, при $b^2 - ac = 0$; В каноническом виде это уравнение выглядит следующим образом:

$$\frac{\partial^2 \omega}{\partial \eta^2} = F_1 \left(\xi, \eta, \omega, \frac{\partial \omega}{\partial \xi}, \frac{\partial \omega}{\partial \eta} \right), \quad (2)$$

где $\xi = \varphi(x, y)$, $\eta = \psi(x, y)$ – независимые переменные. Помимо того, $\eta(x, y)$ – рассматриваемая в данной области функция, дифференцируемая дважды. Это уравнение имеет один член высшей производной, как и уравнение теплопроводности.

• гиперболического типа, при $b^2 - ac > 0$; В каноническом виде это уравнение выглядит следующим образом:

о первая каноническая форма:

$$\frac{\partial^2 \omega}{\partial \xi \partial \eta} = F_2 \left(\xi, \eta, \omega, \frac{\partial \omega}{\partial \xi}, \frac{\partial \omega}{\partial \eta} \right). \quad (3)$$

где $\xi = \varphi(x, y), \eta = \psi(x, y)$ – независимые переменные.

о вторая каноническая форма:

$$\frac{\partial^2 \omega}{\partial t^2} - \frac{\partial^2 \omega}{\partial z^2} = F_3 \left(t, z, \omega, \frac{\partial \omega}{\partial t}, \frac{\partial \omega}{\partial z} \right), \quad (4)$$

где $\xi = t + z, \eta = t - z, F_3 = 4F_2$. Левая часть данного уравнения представляет собой часть волнового уравнения.

• эллиптического типа, при $b^2 - ac < 0$; В каноническом виде это уравнение выглядит следующим образом:

$$\frac{\partial^2 \omega}{\partial \xi^2} - \frac{\partial^2 \omega}{\partial \eta^2} = F_4 \left(\xi, \eta, \omega, \frac{\partial \omega}{\partial \xi}, \frac{\partial \omega}{\partial \eta} \right), \quad (5)$$

где $\xi = \varphi(x, y), \eta = \psi(x, y)$ – независимые переменные. Левая часть данного уравнения представляет собой левую часть уравнения Лапласа.

Способы решения перечисленных типов уравнений математической физики делятся на две группы:

1. аналитическое решение, основанное на преобразовании уравнений в частных производных к системе уравнений или к одному простому уравнению;

2. численное решение с использованием вычислительных машин.

Аналитические методы решения включают в себя:

- метод характеристик;
- метод разделения переменных;
- метод Фурье;
- метод Даламбера;
- метод интегральных преобразований;
- преобразования Лапласа;
- использование функции Грина.

Методы реализации уравнений математической физики с применением вычислительной техники включают в себя:

- метод сеток;
- метод конечных разностей;
- метод переменных направлений;
- метод дробных шагов;
- метод Эйлера;
- методы Рунге-Кутты;
- методы Адамса;
- символьно-численные методы.

Для того чтобы доступность знаний была реализована необходимо иметь возможность наглядного изложения информации по курсу «Уравнения математической физики».

Одним из средств визуализации является программа MathCAD. Это универсальный программный комплекс, позволяющий реализовывать математические операции любой сложности, данное программное обеспечение используется для расчетов в различных областях, в том

числе при математическом моделировании. Важным отличием программы является пошаговая визуализация процесса внесения данных в программу. Решение задач и уравнений в MathCAD является элементарным процессом даже для незнакомого с программой пользователя. Визуализация процессов максимальна, в виду чего процесс работы в MathCAD максимально приближен к работе в обычном редакторе формул или же к аналитическим записям. Программный комплекс MathCAD обладает широкими возможностями и включает свыше двух сотен операторов и функций, используемых для решения задач в физико-математической области высокой сложности как в численном, так и в символьном видах [3].

MathCAD – это математически ориентированная система. Кроме операции вычисления, система позволяет решать задачи, нерешаемые, или же решаемые нерационально в текстовых редакторах, электронных таблицах. Высокое удобство и работоспособность данной системы подтверждается ее использованием при подготовке множества научных статей, монографий, квалификационных работ, разного рода проектов, книг и даже диссертационных работ, в добавок к этому сложнейшие математические записи реализуются максимально просто и при желании пользователя представляются графически.

Программа MathCAD является оптимизированным полноценным визуализатором, используемым для описания и создания алгоритмов и методов решения задач физики и математики. Благодаря дружелюбному русифицированному интерфейсу MathCAD может быть применен как в процессе обучения школьников, в качестве расчётного комплекса, так и в процессе работы научных сотрудников, в качестве среды разработки новых физико-математических моделей.

Создатели MathCAD реализовали систему так, что она может быть расширена пользователем в зависимости от его потребностей. Расширения программы происходит по средствам приобретения специализированных библиотек и расширений. Каждая библиотека или пакет расширений имеют специальные возможности, встраиваемые в систему после установки. Также в программный комплекс встроено множество электронных книг с описаниями различных методов решения специализированных задач, с примерами работающих алгоритмов, которые могут быть использованы в расчетах.

В системе также реализован текстовый редактор. Текстовые поля создаются как комментарии, и математические выражения в них не выполняются. Тексты комментариев могут содержать словесные формулировки, численные значения и математические символы.

Процессор обработки формул имеет вертикальное строение, что схоже с привычной формой записи математических выражений. Последние версии программы поддерживают русскоязычные коммента-

рии и подписи графиков.

Встроенные в программу блоки вычислителей обеспечивают расчеты по заданным математическим формулам, имеют обширный набор встроенных математических выражений и функций, позволяют рассчитывать значения рядов, сумм, произведений, интегралов, производных, работать с комплексными данными, реализовывать решения линейных и нелинейных выражений, а также дифференциальных выражений и систем, проводить минимизирование и максимизирование функций, реализовывать операции над векторами и матрицами и многое другое.

Рассматриваемый программный комплекс является средством автоматизации математических расчетов. Программ схожих с MathCAD по принципам построения и направленности существует довольно много. Наиболее похожими системами компьютерного моделирования являются Maple, Mathematica и MATLAB. По факту эти программы создавались для разных целей и в них закладывалась отличная друг от друга идеология, что делает невозможным объективное сравнение.

При рассмотрении отличий программы от аналогов необходимо указать на графический режим ввода выражений, что не поддерживается в прочих программных комплексах. Ввод функций, команд и формул может быть осуществлен как с клавиатуры, так и с помощью кнопок на специальных панелях инструментов, расположенных на многочисленных вкладках. В обои вариантах расчетные формулы будут отображаться, аналогично книжному виду. Это свидетельствует о том, что программой может пользоваться абсолютно неподготовленный человек. Вычислительный процесс по вводимым формулам осуществляется по вызову пользователя, одновременно с вводом информации или по команде. Вычисления по формулам происходят вертикально вниз и горизонтально направо. Переменные, параметры, формулы изменяются моментально, что позволяет наглядно отображать происходящие изменения. Это позволяет организовать интерактивные вычислительные документы.

Визуализация по средствам моделирования

Рассмотрим ряд наиболее затруднительных с точки зрения визуализации и понимания студентами задач рассматриваемого курса, а именно задачи, основанные на волновом уравнении.

Одним из самых простых методов решения уравнений в системе MathCAD является решение при помощи **блока Given-Find**. Примером использования данного блока послужит задача об одномерном волновом уравнении. Однородная струна колеблется по следующему закону:

$$\frac{\partial v}{\partial t} = a^2 \frac{\partial^2 u}{\partial t^2} \quad (6)$$

При заданном размере струны, времени колебаний, начальных и граничных условиях задача описывается при помощи команд, представленных в листинге на рис. 1.

```

a := 2

L := 10  T := 10

Given

v_t(x,t) = a^2 * u_xx(x,t)

u_t(x,t) = v(x,t)

f(x) := (10x) / L

g(x) := if(x <= 5, f(x), -x + 10)

u(x,0) = g(x)
v(x,0) = 0

u(0,t) = 0
u(L,t) = 0

(u
 v) := Pdesolve([u
 v], x, [L], t, [T])
    
```

Рис. 1. Листинг решения уравнения (6)

Изменение положения струны в моменты времени $t=0\dots 5$ представлены на рисунке 2.

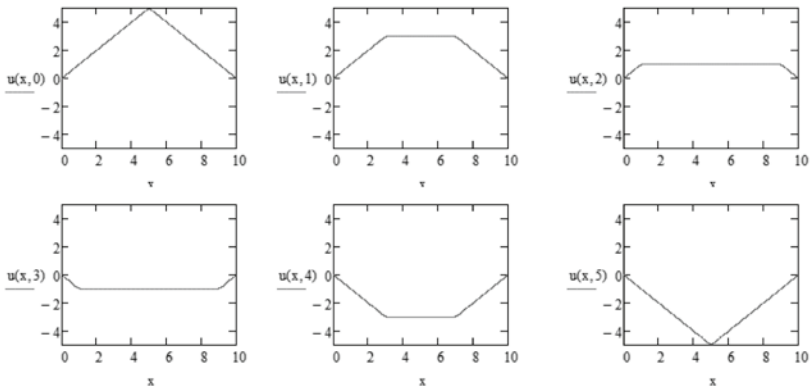


Рис. 2. Изменение положения струны

Для создания анимации движения струны необходимо использовать встроенную функцию FRAME (рис. 3).

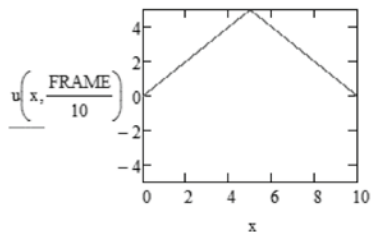


Рис. 3. Листинг применения функции FRAME, для создания анимации

Вторым рассматриваемым типом задач будут задачи гиперболического типа. В данном примере используем **метод сеток** для решения задачи. В конечно-разностном виде гиперболическое уравнение принимает вид:

$$\frac{u_{i+1,j} - 2u_{i,j} + u_{i-1,j}}{h^2} = \frac{1}{v^2} \frac{u_{i,j+1} - 2u_{i,j} + u_{i,j-1}}{\tau^2} \quad (7)$$

Уравнение (6) дает возможность определить значение функции u в $j+1$ момент времени:

$$u_{i,j+1} = v^2 \left(\frac{\tau}{h}\right)^2 (u_{i+1,j} - 2u_{i,j} + u_{i-1,j}) + 2u_{i,j} - u_{i,j-1}. \quad (8)$$

Полученная разностная схема является явной и устойчивой при $\tau \leq h/v$. Начальный и последний моменты времени описываются синусоидальной функцией. Начальные условия, граничные условия, а также количество измерений в MathCAD задаются согласно листингом на рис. 4.

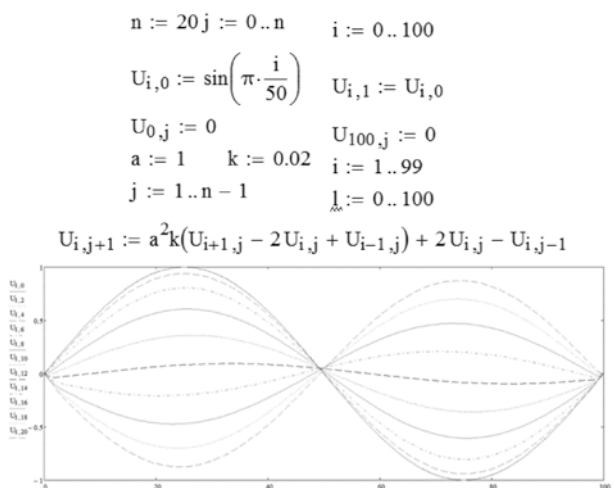


Рис. 4. Листинг и графическая интерпретация решения уравнения (8)

Представленный график может быть реализован как в качестве множества функций, накладываемых на один график, так и в виде анимации. Визуализация по средствам MathCAD позволяет посмотреть процесс изменения амплитуды колебания струны с течением времени, наглядно определять узловые точки струны, основные параметры (амплитуда, частота, период) колебательного процесса.

Следующим примером послужит задача о волновом уравнении на прямоугольнике. Плоская однородная мембрана колеблется по следующему закону:

$$u_{tt} = a^2(u_{xx} + u_{yy}). \quad (9)$$

Для *аналитического* решения данного уравнения необходимо задать начальные условия:

$$\begin{cases} u(x, y, 0) = \varphi(x, y) \\ u_t(x, y, 0) = \psi(x, y) \end{cases} \quad (10)$$

и граничные условия: $u(0, y, t) = 0, u(b_1, y, t) = 0, u(x, 0, t) = 0, u(x, b_2, t) = 0$.

Данная задача решается методом разделения переменных:

$$v(x, y) = X(x)Y(y), \quad (11)$$

где $X'' + \nu X = 0, X(0) = 0, X(b_1) = 0, Y'' + \mu Y = 0, Y(0) = 0, Y(b_2) = 0$.

Последние уравнения имеют следующие решения:

$$v_n = \left(\frac{n\pi}{b_1}\right) X_n = \sin \frac{n\pi}{b_1} x, \quad \mu_m = \left(\frac{m\pi}{b_2}\right) Y_m = \sin \frac{m\pi}{b_2} y. \quad (12)$$

Собственным значениям $\lambda_{n,m} = (n\pi/b_1)^2 + (m\pi/b_2)^2$ соответствуют собственные функции

$$v_{n,m}(x, y) = \sqrt{\frac{4}{b_1 b_2}} \sin \frac{n\pi}{b_1} x \cdot \sin \frac{m\pi}{b_2} y. \quad (13)$$

Уравнение (13) образует ортонормированную систему собственных функций прямоугольной мембраны. Как известно, любая непрерывная функция, дифференцируемая дважды $F(x, y)$, которая удовлетворяет граничному условию $F(0, y) = F(b_1, y) = F(x, 0) = F(x, b_2) = 0$, разлагается в равномерно и абсолютно сходящийся ряд по $v_{n,m}(x, y)$.

Решение данной задачи записывается рядом по собственным функциям $v_{n,m}(x, y)$:

$$u(x, y, t) = \sum_{m=1}^{\infty} \sum_{n=1}^{\infty} (A_{n,m} \cos(\sqrt{\lambda_{n,m}} at) + B_{n,m} \sin(\sqrt{\lambda_{n,m}} at)) \cdot v_{n,m}(x, y), \quad (14)$$

где

$$A_{n,m} = \sqrt{\frac{4}{b_1 b_2}} \int_0^{b_1} \int_0^{b_2} \varphi(x, y) \sin \frac{n\pi}{b_1} x \cdot \sin \frac{m\pi}{b_2} y dx dy, \quad (15)$$

$$B_{n,m} = \frac{1}{\sqrt{a^2 \lambda_{n,m}}} \sqrt{\frac{4}{b_1 b_2}} \int_0^{b_1} \int_0^{b_2} \psi(x, y) \sin \frac{n\pi}{b_1} x \cdot \sin \frac{m\pi}{b_2} y dx dy. \quad (16)$$

В частном случае уравнение колебаний запишем как: $u_{tt} = u_{xx}$, при $0 < x < 2$, $0 < y < 3$ и $\varphi(x, y) = xy(2-x)(3-y)/64$, $\psi(x, y) = 0$, $t > 0$. Рассмотрим реализацию данного случая в MathCAD (рис. 5).

$$\begin{aligned} & \text{ORIGIN} := 1 \\ & \text{kk} := 2 \quad \text{ly} := 3 \quad a := 1 \\ & \phi(x, y) := \frac{x \cdot y}{64} \cdot (2 - x) \cdot (3 - y) \quad \psi(x, y) := 0 \\ & n := 1..10 \quad m := 1..15 \\ & \lambda_{n,m} := \left(\frac{n \cdot \pi}{\text{kk}} \right)^2 + \left(\frac{m \cdot \pi}{\text{ly}} \right)^2 \\ & A_{n,m} := \sqrt{\frac{4}{\text{kk} \cdot \text{ly}}} \int_0^{\text{ly}} \int_0^{\text{kk}} \phi(x, y) \cdot \sin \left(\frac{n \cdot \pi}{\text{kk}} \cdot x \right) \cdot \sin \left(\frac{m \cdot \pi}{\text{ly}} \cdot y \right) dx dy \\ & B_{n,m} := \frac{1}{\sqrt{a^2 \lambda_{n,m}}} \cdot \sqrt{\frac{4}{\text{kk} \cdot \text{ly}}} \int_0^{\text{ly}} \int_0^{\text{kk}} \psi(x, y) \cdot \sin \left(\frac{n \cdot \pi}{\text{kk}} \cdot x \right) \cdot \sin \left(\frac{m \cdot \pi}{\text{ly}} \cdot y \right) dx dy \\ & u(x, y, t) := \sum_{k=1}^5 \sum_{l=1}^5 \left[A_{k,l} \cdot \cos(\sqrt{\lambda_{k,l}} \cdot a \cdot t) + B_{k,l} \cdot \sin(\sqrt{\lambda_{k,l}} \cdot a \cdot t) \right] \cdot \sin \left(\frac{k \cdot \pi}{\text{kk}} \cdot x \right) \cdot \sin \left(\frac{l \cdot \pi}{\text{ly}} \cdot y \right) \\ & i := 1..61 \quad j := 1..61 \quad X_i := (i - 1) \cdot 0.1 \quad Y_j := (j - 1) \cdot 0.1 \\ & Z_{i,j} := u(X_i, Y_j, 0) \quad Z1_{i,j} := u(X_i, Y_j, 0.5) \quad Z2_{i,j} := u(X_i, Y_j, 1) \quad Z3_{i,j} := u(X_i, Y_j, 1.5) \\ & Z4_{i,j} := u(X_i, Y_j, 2) \quad Z5_{i,j} := u(X_i, Y_j, 2.5) \quad Z6_{i,j} := u(X_i, Y_j, 3) \quad Z7_{i,j} := u(X_i, Y_j, 3.5) \\ & Z8_{i,j} := u(X_i, Y_j, 4) \quad Z9_{i,j} := u(X_i, Y_j, 4.5) \quad Z10_{i,j} := u(X_i, Y_j, 5) \quad Z11_{i,j} := u(X_i, Y_j, 5.5) \end{aligned}$$

Рис. 5. Листинг визуализации волнового уравнения на прямоугольнике

Изменение положения мембраны для моментов времени $t = 0, 0.5, \dots, 5.5$ представлена на рис.6.

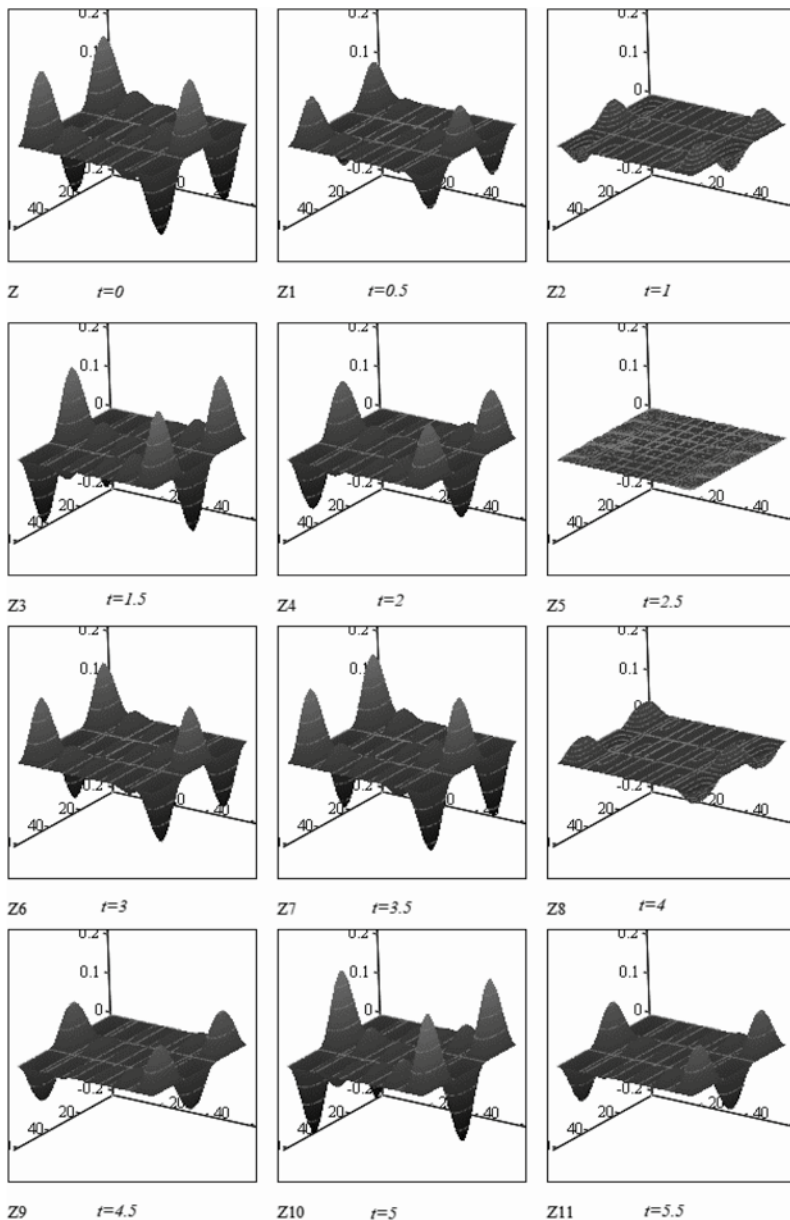
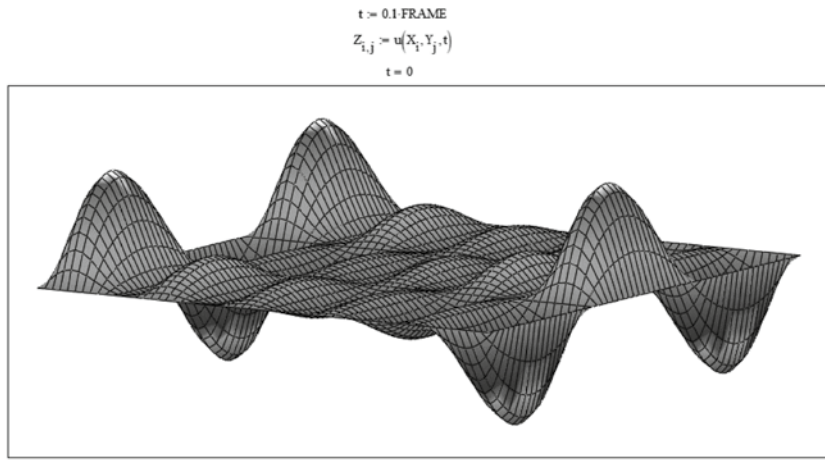


Рис. 6. Визуализация волнового уравнения на прямоугольнике в моменты времени $t = 0, 0.5, \dots, 5.5$

Также, как и для уравнения на круге реализация динамического представления колебаний мембраны происходит по средствам встроенной функцией FRAME. Применение функции показано на рисунке 7.

Стоит отметить что в то же время аналитическое решение данной, и предыдущих, задач потребовала огромного количество времени так как для каждого момента времени необходимо рассчитать 61 значение функции, а для отслеживания динамики необходим массив данных(таблица значений функции) размером 61 на 61, то есть необходимо произвести 3721 вычисление. Пакет MathCAD позволяет визуализировать этот массив данных (рис. 8) и при этом затратить на это считанные секунды.



З
Рис. 7. Листинг создания анимации волнового уравнения на прямоугольнике в MathCAD

	40	40	47	48	49	50	51	52	53	54	55	56	57	58	59	60	61
22	9.999E-7	8.305E-7	5.811E-7	2.739E-7	0	2.309E-7	0	1.969E-7	0.016	0.048	0.1025	0.103	0.033	0.033	0.033	0.033	0
23	0.018	-0.016	0.011	3.309E-7	0	2.322E-7	0	2.029E-7	0.012	0.015	0.049	0.105	-0.063	0.068	0.068	0.068	0
24	0.025	-0.023	0.016	7.522E-7	0	3.342E-7	0	0.041	0.028	0.15	0.17	0.085	0.09	0.083	0.084	0.085	0
25	0.032	-0.029	0.02	9.507E-7	0	4.106E-7	0	0.094	0.056	0.162	0.168	0.112	0.112	0.104	0.108	0.109	0
26	0.037	-0.032	0.024	0.011	0	4.919E-7	0	0.165	0.102	0.173	0.173	0.123	0.122	0.127	0.129	0.131	0
27	0.041	-0.037	0.026	0.014	0	5.672E-7	0	0.253	0.147	0.181	0.171	0.129	0.128	0.135	0.137	0.139	0
28	0.045	-0.04	0.029	0.015	0	6.084E-7	0	0.359	0.195	0.188	0.174	0.131	0.129	0.136	0.138	0.14	0
29	0.047	-0.043	0.03	0.014	0	6.299E-7	0	0.481	0.253	0.193	0.177	0.137	0.135	0.142	0.144	0.146	0
30	0.049	-0.045	0.031	0.013	0	6.486E-7	0	0.611	0.315	0.196	0.178	0.14	0.137	0.144	0.146	0.148	0
31	0.049	-0.046	0.032	0.013	0	6.659E-7	0	0.747	0.382	0.197	0.177	0.14	0.137	0.144	0.146	0.148	0
32	0.049	-0.046	0.031	0.012	0	6.809E-7	0	0.881	0.45	0.196	0.176	0.14	0.137	0.144	0.146	0.148	0
33	0.047	-0.045	0.03	0.014	0	6.939E-7	0	1.013	0.52	0.193	0.173	0.139	0.137	0.144	0.146	0.148	0
34	0.043	-0.04	0.029	0.013	0	7.039E-7	0	1.141	0.59	0.188	0.169	0.134	0.131	0.138	0.14	0.142	0
35	0.041	-0.037	0.026	0.012	0	7.129E-7	0	1.265	0.67	0.181	0.161	0.13	0.128	0.135	0.137	0.139	0
36	0.037	-0.033	0.024	0.011	0	7.199E-7	0	1.385	0.74	0.172	0.152	0.127	0.125	0.132	0.134	0.136	0
37	0.033	-0.029	0.02	0.012	0	7.249E-7	0	1.501	0.81	0.162	0.143	0.12	0.118	0.125	0.127	0.129	0
38	0.025	-0.023	0.016	7.522E-7	0	7.289E-7	0	1.611	0.88	0.15	0.137	0.115	0.114	0.121	0.123	0.125	0
39	0.023	-0.02	0.014	3.209E-7	0	7.349E-7	0	1.715	0.92	0.145	0.129	0.11	0.109	0.116	0.118	0.12	0
40	9.999E-7	8.305E-7	5.811E-7	2.739E-7	0	7.399E-7	0	1.811	0.97	0.138	0.121	0.104	0.103	0.11	0.112	0.114	0
41	0	0	0	0	0	7.449E-7	0	1.901	1.01	0.13	0.115	0.097	0.095	0.103	0.105	0.107	0
42	-0.009E-7	-8.305E-7	-5.811E-7	-2.739E-7	0	7.499E-7	0	1.991	1.07	0.125	0.108	0.09	0.089	0.097	0.099	0.101	0
43	-0.018	-0.016	-0.011	-3.309E-7	0	7.539E-7	0	2.071	1.12	0.118	0.101	0.083	0.082	0.09	0.091	0.093	0
44	-0.025	-0.023	0.016	7.522E-7	0	7.579E-7	0	2.141	1.17	0.111	0.094	0.075	0.074	0.081	0.082	0.084	0
45	-0.032	-0.029	0.02	9.507E-7	0	7.619E-7	0	2.201	1.22	0.104	0.087	0.068	0.067	0.074	0.075	0.077	0
46	-0.037	-0.032	0.024	0.011	0	7.659E-7	0	2.261	1.27	0.097	0.08	0.06	0.059	0.066	0.067	0.069	0
47	-0.041	-0.037	0.026	0.014	0	7.699E-7	0	2.311	1.32	0.09	0.073	0.053	0.052	0.059	0.06	0.062	0
48	-0.045	-0.04	0.029	0.013	0	7.739E-7	0	2.361	1.37	0.083	0.066	0.046	0.045	0.052	0.053	0.055	0
49	-0.047	-0.043	0.03	0.014	0	7.779E-7	0	2.411	1.42	0.076	0.059	0.04	0.039	0.046	0.047	0.049	0
50	-0.049	-0.045	0.031	0.013	0	7.819E-7	0	2.461	1.47	0.069	0.052	0.034	0.033	0.04	0.041	0.043	0
51	-0.049	-0.046	0.032	0.013	0	7.859E-7	0	2.511	1.52	0.062	0.045	0.028	0.027	0.034	0.035	0.037	0
52	-0.049	-0.046	0.031	0.012	0	7.899E-7	0	2.561	1.57	0.055	0.038	0.022	0.021	0.028	0.029	0.031	0
53	-0.047	-0.045	0.03	0.014	0	7.939E-7	0	2.611	1.62	0.048	0.031	0.016	0.015	0.022	0.023	0.025	0
54	-0.047	-0.046	0.031	0.013	0	7.979E-7	0	2.661	1.67	0.041	0.024	0.01	0.009	0.016	0.017	0.019	0
55	-0.045	-0.04	0.029	0.013	0	8.019E-7	0	2.711	1.72	0.034	0.017	0.004	0.003	0.01	0.011	0.013	0
56	-0.043	-0.037	0.026	0.012	0	8.059E-7	0	2.761	1.77	0.027	0.01	0	0	0.007	0.008	0.01	0
57	-0.042	-0.039	0.024	0.011	0	8.099E-7	0	2.811	1.82	0.02	0.003	0	0	0.004	0.005	0.007	0
58	-0.042	-0.039	0.023	0.01	0	8.139E-7	0	2.861	1.87	0.013	0.002	0	0	0.001	0.002	0.004	0
59	-0.042	-0.039	0.022	0.01	0	8.179E-7	0	2.911	1.92	0.006	0.001	0	0	0	0.001	0.003	0
60	-0.04	-0.037	0.02	0.012	0	8.219E-7	0	2.961	1.97	0.001	0	0	0	0	0	0.001	0
61	-0.04	-0.038	0.019	0.011	0	8.259E-7	0	3.011	2.02	0	0	0	0	0	0	0	0
62	0	0	0	0	0	8.299E-7	0	3.061	2.07	0	0	0	0	0	0	0	0

Рис. 8. Массив данных используемый для визуализации задачи о колебании прямоугольной мембраны

Заключение

В данной статье показан и реализован один из важнейших принципов обучения и подачи информации – наглядность. Рассмотрена малая часть всевозможных задач о колебаниях. Таким образом, построены графические интерпретации задач, основанные на уравнениях гиперболического типа.

В работе рассмотрены задачи о колебании струны, о свободных колебаниях на круге и о свободных колебаниях на прямоугольнике. Предложены и реализованы алгоритмы визуализации вышеупомянутых задач в системе компьютерного моделирования MathCAD. Помимо покадровой визуализации предложен алгоритм создания анимации, который может быть применен в процессе обучения бакалавров в рамках курса «Уравнения математической физики».

Исследование волновых уравнений в пакете MathCAD ускоряет процесс обучения, повышая скорость вычислений. Также увеличивает заинтересованность в процессе обучения, так как визуализация преподаваемого материала, делает его значительно понятнее и доступнее.

Необходимо также указать на существование межпредметной связи при использовании визуализатора MathCAD. Так как MathCAD это язык программирования, то обучаясь на педагогических направлениях, далеких от программирования, у студентов появляется возможность изучить азы ввода, обработки и хранения данных, что в большей степени присуще инженерно-техническим специальностям. Таким образом применение данного пакета расширяет жизненный опыт студентов, при этом не усложняя процесс обучения, в виду того, что интерфейс программы интуитивно понятен и приспособлен для людей незнакомых с программированием вообще.

Библиографический список

1. Владимиров, В.С. Уравнения математической физики: учебник для вузов / В.С. Владимиров, В.В. Жаринов. – М.:ФИЗМАТЛИТ, 2004. – 400 с.
2. Математическая физика. –URL: https://ru.wikipedia.org/wiki/Математическая_физика (дата обращения: 12.02.2018).
3. Очков, В.Ф. Mathcad 14 для студентов, инженеров и конструкторов / В.Ф. Очков. – СПб.: БХВ-Петербург, 2007. – 368 с.

УДК 654.026

Н. А. Пылаев

Военная академия связи, Санкт-Петербург, Россия, kursant631@yandex.ru

МЕТОДИКА ОЦЕНКИ ВЕРОЯТНОСТНО-ВРЕМЕННЫХ ХАРАКТЕРИСТИК ПЕРЕДАЧИ СООБЩЕНИЙ В СОСТАВНЫХ РАДИОЛИНИЯХ С ИСПОЛЬЗОВАНИЕМ РЕТРАНСЛЯТОРОВ СВЯЗИ НА БЕСПИЛОТНЫХ ЛЕТАТЕЛЬНЫХ АППАРАТАХ

Аннотация: В работе представлена методика позволяющая оценить вероятностно-временные характеристики передачи сообщений в составных радиолиниях с использованием ретрансляторов связи на БПЛА. Математическое моделирование зоны покрытия в рамках предложенного подхода подтвердило его жизнеспособность.

Ключевые слова: зона покрытия, мобильные ретрансляторы связи, беспилотные летательные аппараты, математическое моделирование.

N.A. Pylaev

Military Academy of Communications, St. Petersburg, Russia,
kursant631@yandex.ru

METHODOLOGY FOR ASSESSMENT OF PROBABILITY-TEMPORAL CHARACTERISTICS OF MESSAGE TRANSMISSION IN COMPOUND RADIOLINIA USING COMMUNICATION REPEATERS FOR UNMANNED AIRCRAFT

Abstract: The paper presents a technique permitting to evaluate the probability-time characteristics of the transmission of messages in composite radio links using communication UAV repeaters. Mathematical modeling of the coverage area within the framework of the proposed approach confirmed its viability.

Keywords: coverage area, mobile communication repeaters, unmanned aerial vehicles, mathematical modeling.

Введение

В настоящее время беспилотные летательные аппараты (БПЛА) находят свое применение во все более различных сферах. Одним из перспективных направлений применения БПЛА является их использование в качестве носителей ретрансляторов связи (РС). Использование РС на БПЛА в сравнении с наземными ретрансляторами позволяет значительно увеличить дальность связи в УКВ диапазоне и в кратчайшие сроки развертывать сети радиосвязи на необорудованной территории. Особую актуальность применение РС на БПЛА приобретает при обеспечении связи между удаленными друг от друга подвижными объектами (ПО).

Проведенные в Военной академии связи испытания показали возможность организации связи с помощью РС на БПЛА с использованием комплекса средств связи в диапазоне частот 1,5...1,75 ГГц. Однако было выявлено, что в данном диапазоне значительное влияние на состояние

радиоканала между ПО и РС на БПЛА оказывает рельеф местности, что требует необходимости учета данного фактора.

Модель составной радиолинии подвижный объект – ретранслятор связи на беспилотном летательном аппарате – подвижный объект

Модель составной радиолинии, обеспечивающей передачу сообщений между двумя подвижными объектами с использованием РС на БПЛА, можно представить в виде двух независимых радиоканалов – между первым ПО (ПО1) и РС на БПЛА, и между РС на БПЛА и вторым ПО (ПО2) [1].

Влияние рельефа местности на состояние радиоканала при перемещении ПО представляет собой сложный процесс и зависит от многих факторов. При перемещении ПО изменяется степень затенения трассы распространения сигнала рельефом местности и растительностью, вследствие чего уровень мощности сигнала в точке приема изменяется случайным образом.

В настоящее время общепринятым является утверждение, что изменение медианного значения уровня мощности принимаемого сигнала описывается логарифмически-нормальным законом распределения со стандартным отклонением, зависящим от рельефа местности и характера городской застройки [2]:

$$\omega(y) = \frac{1}{2\sigma_y \pi} \exp\left(-\frac{(y - \bar{y})^2}{2\sigma_y^2}\right)$$

где \bar{y} – среднее значение уровня мощности сигнала, дБ; σ_y – среднеквадратическое отклонение уровня мощности сигнала от своего среднего значения, дБ.

В результате, в отдельные промежутки времени, когда подвижный объект будет находиться в зоне затенения, значение уровня мощности сигнала в точке приема будет опускаться ниже порогового значения, в результате чего состояние радиоканала будет непригодным для передачи сообщений. Таким образом, радиоканал между ПО и РС на БПЛА при перемещении подвижного объекта представляет собой канал прерывистой связи.

Динамику функционирования канала прерывистой связи удобно описывать последовательной сменой состояний: пригодного для передачи сообщений с достоверностью не хуже заданной $S_{\text{при}}$ и непригодного состояния радиоканала $S_{\text{нпри}}$ с длительностями $\tau_{\text{при}}$ и $\tau_{\text{нпри}}$, соответственно [3]. Модель составной радиолинии, в таком случае, состоящей из двух независимых радиоканалов, можно представить в виде потока совпадений $X_{\Sigma}(t)$ двух независимых случайных импульсных потоков $X_1(t)$ и $X_2(t)$.

Случайный импульсных поток $X_1(t)$ характеризует состояние радиоканала ПО1–РС на БПЛА, где длительность импульса $X_2(t)$ соответствует длительности пригодного состояния радиоканала, а длительность паузы

$\tau_{\text{нпр}1}$ соответствует длительности непригодного состояния радиоканала. Случайный импульсных поток $X_2(t)$ характеризует состояние радиоканала РС на БПЛА–ПО2, где длительность импульса $\tau_{\text{пр}2}$ соответствует длительности пригодного состояния радиоканала, а длительность паузы $\tau_{\text{нпр}2}$ соответствует длительности непригодного состояния радиоканала.

Решение задачи о совпадении импульсов ряда стационарных (в широком смысле) и независимых потоков может быть найдено на базе математической модели, предложенной в [4]. Модель составной радиолинии ПО1 – РС на БПЛА – ПО2 в виде потока совпадений случайных импульсных потоков представлена в [5]. На рис. 1 представлена временная реализация предложенной модели.

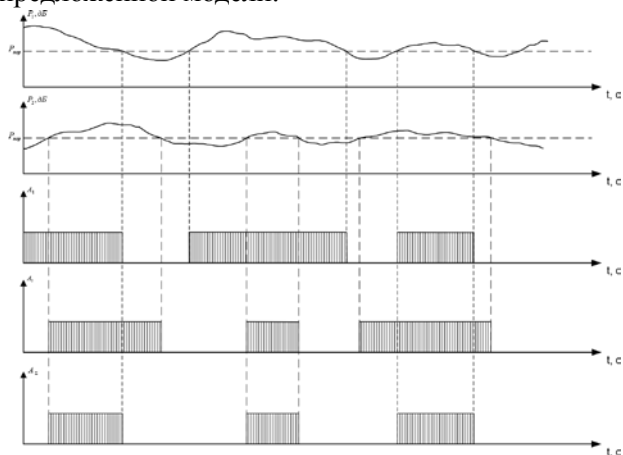


Рис. 1. Модель составной радиолинии ПО1–РС на БПЛА–ПО2 в виде потока совпадений двух независимых случайных импульсных потоков

На основе данной модели была разработана методика оценки вероятностно-временных характеристик передачи сообщений в составных радиолиниях с использованием ретрансляторов связи на беспилотных летательных аппаратах.

Методика оценки вероятностно-временных характеристик передачи сообщений в составных радиолиниях с использованием ретрансляторов связи на беспилотных летательных аппаратах

В настоящее время отсутствуют какие-либо экспериментальные данные подтверждающие теоретические предположения о распределении длительностей пригодных и непригодных состояний радиоканалов с БПЛА. Реализация натурного эксперимента для сбора статистических данных о длительностях состояний является дорогостоящей. Поэтому для определения средних значений длительностей $\tau_{\text{пр}i}$ и $\tau_{\text{нпр}i}$ целесообразно использовать учебные цифровые картографические модели местности, имеющие матрицу высот.

Используя ЦКМ и вводя ряд допущений возможно получить оценки

параметров распределения длительностей пригодного и непригодного для передачи сообщений состояний радиоканалов между ПО и БПЛА. Полученные данные используются для моделирования потока совпадений случайных импульсных потоков состояний радиоканалов между ПО и РС на БПЛА составной радиолинии ПО – РС на БПЛА – ПО, которая лежит в основе методики расчета ВВХ радиолинии.

Методика расчета ВВХ содержит следующие этапы.

1. Расчет радиуса зоны покрытия РС на БПЛА.

Для расчета зоны покрытия существует множество различных методик, обзор которых представлен в [6]. В данной методике для расчета зоны покрытия использовался численный метод, представленный в [7].

2. Ввод исходных данных о местоположении БПЛА

Ввод исходных данных осуществляется с помощью встроенных опций в ГИС «Панорама». Исходные данные включают в себя координаты и высоту местоположения БПЛА.

3. Построение радиуса зоны покрытия РС на БПЛА

Осуществляется с помощью встроенных опций ГИС «Панорама».

4. Построение зон затенения внутри зоны покрытия

Осуществляется с помощью встроенных опций ГИС «Панорама».

5. Прокладка случайных маршрутов перемещения ПО

6. Сбор статистических данных о времени нахождения ПО в зонах видимости и зонах затенения

Для упрощения проведения эксперимента были введены следующие допущения:

– БПЛА барражирует в достаточно малом относительно рассматриваемого участка местности районе, что позволяет считать координаты его местоположения постоянными;

– при распространении электромагнитных волн на интервале ПО–РС на БПЛА отсутствует дифракция на естественных препятствиях местности, поэтому состояние радиоканала ПО–РС на БПЛА определяется наличием или отсутствием оптической видимости между корреспондентами;

– подвижные объекты перемещаются равномерно со скоростью 10 м/с.

7. Обработка статистических данных и оценка параметров распределения длительностей пригодного и непригодного состояний радиоканала.

Результаты обработки представлены на рис. 2–5.



Рис. 4 – Распределение длительностей пригодного состояния радиоканала ПОИ–РС на БПЛА



Рис. 5. Распределение длительностей непригодного состояния радиоканала ПОИ–РС на БПЛА



Рис. 6. Распределение длительностей пригодного состояния радиоканала РС на БПЛА–ПО2



Рис. 7. Распределение длительностей непригодного состояния радиоканала РС на БПЛА–ПО2

8. Ввод полученных данных о длительностях импульсов и пауз в модель совпадения случайных импульсных потоков.

Исходными данными модели являются плотности распределения длительностей следующих случайных интервалов времени, полученных с помощью ЦКМ:

$\tau_{1пр}$ – длительности пригодного для передачи сообщений состояния радиоканала ПО1–РС на БПЛА;

$\tau_{1нпр}$ – длительности непригодного состояния радиоканала ПО1–РС на БПЛА;

$\tau_{2пр}$ – пригодного для передачи сообщений состояния радиоканала РС на БПЛА–ПО2;

$\tau_{2нпр}$ – длительности непригодного состояния радиоканала РС на БПЛА–ПО2.

9. С помощью модели осуществляется расчет вероятностно-временных характеристик передачи сообщений в составной радиолинии с использованием ретранслятора связи на беспилотном летательном аппарате.

Заключение

Таким образом, была разработана методика позволяющая оценить вероятностно-временные характеристики передачи сообщений в составных радиолиниях с использованием ретрансляторов связи на БПЛА.

Список литературы

1. Дмитриев В. И, Пылаев Н. А. Модель радиолинии через ретранслятор связи на беспилотном летательном аппарате в виде потока совпадений случайных импульсных последовательностей // Сборник трудов 73-й Всероссийской научно-технической конференции, посвященной Дню радио. – СПб.: Изд-во СПбГЭТУ «ЛЭТИ», 2018. – С. 106–107.
2. Дмитриев, В. И. Линии и сети связи через средневысотные ретрансляторы [Текст] / В. И. Дмитриев. – СПб.: ВАС, 1993. – 324 с.
3. Семисошенко, М. А. Управление автоматизированными сетями декаметрового диапазона связи в условиях сложной помеховой обстановки [текст] : монография / М. А. Семисошенко. – СПб.: ВАС, 1997. – 364 с.

4. Седякин, М. Н. Элементы теории случайных импульсных потоков [текст] : монография / М. Н. Седякин. – М.: Издательство «Советское радио», 1965. – 263 с.
5. Пылаев, Н. А. Модель, обеспечивающая аппроксимацию состояния радиоканала составной радиолинии потоком совпадений бинарных случайных импульсных последовательностей / Н. А. Пылаев, В. И. Дмитриев // XXIV Международная научно-техническая конференция «Радиолокация, навигация, связь». Сборник трудов. – Том 1. – Воронеж, 2017. – С. 65-72.
6. Весоловский, К.. Системы подвижной радиосвязи [Текст] / К. Весоловский; пер. с польск. И. Д. Рудинского; под ред. А. И. Ледовского. – М.: Горячая линия – Телеком, 2006. – 368 с.
7. Долуханов, М. П. Распространение радиоволн / М. П. Долуханов. – М.: Связь, 1972. – 336 с.

УДК 517.9

Черняев А.П.

Доктор физ.-мат. наук, профессор, Московский физико-технический институт (государственный университет), Москва, Россия. E-mail: chernyaev40@yandex.ru

Черняева С.А.

Студентка, Московский Государственный Технический Университет им. Н.Э. Баумана, Москва, Россия. E-mail: sofee2@rambler.ru

ПОСТАНОВКА И ЧИСЛЕННЫЙ АНАЛИЗ НЕКОТОРЫХ ЗАДАЧ ДЛЯ КНОИДАЛЬНОЙ ВОЛНЫ И СОЛИТОНА, КАК РЕШЕНИЙ УРАВНЕНИЯ КОРТЕВЕГА ДЕ ФРИЗА

Аннотация: В работе с помощью системы компьютерной алгебры «Wolfram Mathematica» исследуются периодические решения уравнения Кортевега де Фриза конкретного вида. Численные решения задач Коши для ОДУ первого, второго и третьего порядка сравниваются с соответствующим точным решением.

Ключевые слова: уравнение Кортевега де Фриза, эллиптическая функция Якоби, кноидальная волна, бегущая волна.

Chernyaev A.P.

Doctor Phys.-Mat. Sci., Professor, Moscow Institute of Physics and Technology (State University), Moscow, Russia. E-mail: chernyaev40@yandex.ru

Chernyaeva S.A.

Student, Moscow State Technical University. N.E. Bauman, Moscow, Russia.
E-mail: sofee2@rambler.ru

FORMULATION AND NUMERICAL ANALYSIS OF SOME PROBLEMS FOR THE KNOIDAL WAVE AND SOLITON, AS SOLUTIONS OF CORTEVEG DE VIEZE EQUATION

Abstract: In the paper, using the Wolfram Mathematica computer algebra system, periodic solutions of the Korteweg de Vries equation of a particular type are investigated. Numerical

solutions of Cauchy problems for ODEs of the first, second and third order are compared with the corresponding exact solution.

Keywords: Korteweg de Vries equation, Jacobi elliptic function, cnoidal wave, traveling wave.

Введение

Уравнение в частных производных Кортевега де Фриза (КдФ) было получено для описания волн на воде [1,2]. В данной работе рассмотрены периодические решения КдФ – кноидальные волны, а также солитон – уединенная волна. Для кноидальных волн характерны острые гребни и плоские впадины. Решение уравнения КдФ для описания волн справедливо только на определенном расстоянии от места образования волны и на определенном промежутке времени. Поэтому уравнение КдФ можно рассматривать как некоторую приближенную математическую модель, соответствующую с определенной степенью точности реальному процессу распространения волн на воде. Однако, это уравнение используется и во многих разделах нелинейной механики и теоретической физики для описания одномерных нелинейных волн с дисперсией без диссипации. В частности, на уравнении КдФ основано математическое моделирование волн умеренной амплитуды на поверхности неглубокой жидкости.

Поскольку, уравнение КдФ имеет третий порядок, постановки задач для него необычны, и могут быть индуцированы задачами для обыкновенных дифференциальных уравнений, к которым оно может быть сведено.

Основные обыкновенные дифференциальные уравнения, к которым может быть сведено уравнение Кортевега де Фриза

Рассматривается нелинейное уравнение КдФ в частных производных в виде [3, с. 282–284]

$$u_t - 6uu_x + u_{xxx} = 0. \quad (1)$$

Решение уравнения (1) ищется в виде

$$u(x, t) = u(\xi), \quad \xi = x - ct, \quad (2)$$

где $u(x, t)$ – общее решение типа бегущей волны. Т. о., уравнение (1) сводится к обыкновенному дифференциальному уравнению третьего порядка

$$u'''(\xi) = 6u(\xi)u'(\xi) + cu'(\xi), \quad (3)$$

Умножив обе части уравнения (3) на $d\xi$, интегрируя и преобразовывая полученное, имеем

$$u''(\xi) = 3u^2(\xi) + cu(\xi) + A, \quad A \in \mathfrak{R}, \quad (4)$$

где A – множество действительных чисел и A – константа интегрирования. Умножая обе части уравнения (4) на $u'(\xi)d\xi$, интегрируя и упрощая полученное, имеем уравнение первого порядка

$$u'(\xi) = \pm \sqrt{2P(u(\xi))}, \quad (5)$$

где

$$P(u) = u^3 + \frac{c}{2}u^2 + Au + B, \quad A \in \mathfrak{R}, B \in \mathfrak{R}. \quad (6)$$

Пусть корни u_1, u_2, u_3 – корни многочлена $P(u)$. Тогда

$$P(u) = (u - u_1)(u - u_2)(u - u_3) \quad (7)$$

Уравнение (4) в этом случае в силу (6) примет вид

$$u'' = 3u^2 + cu + A = P'(u) = (u - u_2)(u - u_3) + (u - u_1)(u - u_3) + (u - u_1)(u - u_2) = 3u^2 - 2(u_1 + u_2 + u_3)u + u_1u_2 + u_2u_3 + u_3u_1 \quad (8)$$

Исходя из (8), получим

$$c = -2(u_1 + u_2 + u_3), \quad A = u_1u_2 + u_2u_3 + u_3u_1. \quad (9)$$

Полагая в (6) и (7) $u = 0$, получаем

$$B = -u_1u_2u_3 \quad (10)$$

Точные решения уравнения КдФ типа кноидальной волны, сравнение с ними численных решений ОДУ первого, второго и третьего порядков

Сразу заметим, что $u = u_1, u = u_2, u = u_3$ (11)

стационарные решения уравнения (5). Если $u \neq u_1, u \neq u_2, u \neq u_3$, то из (5) и (7) следует, что

$$\frac{du}{\sqrt{(u - u_1)(u - u_2)(u - u_3)}} = \pm \sqrt{2} d\xi. \quad (12)$$

Сделаем в (12) замену $u - u_1 = \eta^2, u = u_1 + \eta^2$, (13)

которая приводит к выражениям

$$du = 2\eta d\eta, u - u_2 = u_1 - u_2 + \eta^2, u - u_3 = u_1 - u_3 + \eta^2.$$

Тогда (12) приводится к виду

$$\pm d\xi = \frac{\sqrt{2} d\eta}{\sqrt{(u_2 - u_1 - \eta^2)(u_3 - u_1 - \eta^2)}} \quad (14)$$

Если в (14) сделать замену

$$\eta = \sqrt{u_2 - u_1} s, s = \frac{\eta}{\sqrt{u_2 - u_1}} = \sqrt{\frac{u - u_1}{u_2 - u_1}}, \quad (15)$$

то

$$u_2 - u_1 - \eta^2 = (u_2 - u_1)(1 - s^2), u_3 - u_1 - \eta^2 = u_3 - u_1 - (u_2 - u_1)s^2 = \\ = (u_3 - u_1)(1 - k^2 s^2), \quad k = \sqrt{\frac{u_2 - u_1}{u_3 - u_1}} \quad (16)$$

С учетом (15) и (16) уравнение (14) приводится к виду

$$\pm \sqrt{\frac{u_3 - u_1}{2}} d\xi = \frac{ds}{\sqrt{(1 - s^2)(1 - k^2 s^2)}}. \quad (17)$$

Из (15) следует, что если $u = u_1$, то $s = 0$, а если $u = u_2$, то $s = 1$. Пусть u_1, u_2, u_3 действительны, различны и $u_1 < u_2 < u_3$. Возьмем интеграл от обеих частей (17), выбирая при этом c_1 так, чтобы $\xi + c_1 = 0$ при $s=0$

$$\pm \sqrt{\frac{u_3 - u_1}{2}} (\xi + c_1) = \int_0^s \frac{d\tau}{\sqrt{(1 - \tau^2)(1 - k^2 \tau^2)}}. \quad (18)$$

Пусть
$$F(s, k) = \int_0^s \frac{d\tau}{\sqrt{(1 - \tau^2)(1 - k^2 \tau^2)}} \quad (19)$$

неполный эллиптический интеграл правой части (18) с модулем k , тогда $f(s, k)$ – функция обратная к $sn(F, k)$ – эллиптической функции Якоби [2, с. 44]. Из (18) следует

$$sn\left(\pm \sqrt{\frac{u_3 - u_1}{2}} (\xi + c_1), k\right) = s = \sqrt{\frac{u - u_1}{u_2 - u_1}},$$

или

$$sn^2\left(\pm \sqrt{\frac{u_3 - u_1}{2}} (\xi + c_1), k\right) = \frac{u - u_1}{u_2 - u_1},$$

и тогда

$$u = u_1 + (u_2 - u_1) sn^2\left(\pm \sqrt{\frac{u_3 - u_1}{2}} (\xi + c_1), k\right). \quad (20)$$

или

$$u(\xi) = u_1 + (u_2 - u_1) \left(1 - cn^2\left(\pm \sqrt{\frac{u_3 - u_1}{2}} (\xi + c_1), k\right) \right) = \\ = u_2 - (u_2 - u_1) cn^2\left(\pm \sqrt{\frac{u_3 - u_1}{2}} (\xi + c_1), k\right).$$

Здесь мы воспользовались равенством $sn^2(F, k) + cn^2(F, k) = 1$, где $sn(F, k)$ – эллиптическая функция Якоби, а k определяется правой формулой (16). Из (18) получим период T функции $u(\xi)$ по координате ξ [2, с. 44]:

$$T = \frac{1}{\sqrt{(u_3 - u_1)/2}} \int_0^1 \frac{d\tau}{\sqrt{(1 - \tau^2)(1 - k^2\tau^2)}} = \frac{K(k)}{\sqrt{(u_3 - u_1)/2}}. \quad (21)$$

где $K(k)$ – полный эллиптический интеграл первого рода. В этом случае ограниченное решение уравнения (1) описывает периодическую волну с периодом (21) и амплитудой $u_3 - u_1$. Из-за наличия в (20) эллиптической функции Якоби $sn(F, k)$ соответствующая волна называется кноидальной.

Положим. Тогда $k = 1/\sqrt{2}$. Уравнение (5) с учетом (7) будет иметь вид

$$u' = \pm \sqrt{2(u-1)(u-2)(u-3)}, \quad (22)$$

а его решение согласно (20)

$$u = 2 - cn^2\left(\pm(\xi + c_1), 1/\sqrt{2}\right). \quad (23)$$

Константу $c_1 \approx 0,826018$ находим численно из трансцендентного уравнения $cn^2\left(\pm c_1, 1/\sqrt{2}\right) = 0,5$. Итак, строим график в системе компьютерной алгебры «Wolfram Mathematica» функции $u(\xi) = 2 - cn^2\left(\pm(\xi + 0,826018), 1/\sqrt{2}\right)$, см. рис.1.

Решим дифференциальное уравнение первого порядка (22) численно, выбрав перед правой частью плюс. Для этого поставим задачу Коши, добавив к уравнению (22) условие

$$u(\xi_0) = u_0. \quad (24)$$

В качестве ξ_0 в (24) нельзя брать значения, в которых кноидальная волна достигает минимума или максимума, поскольку в этих точках нарушается единственность решения задачи Коши. Это происходит потому, что нарушается условие Липшица правой части (22) по переменной u [4, 5], так как производная правой части (24) по u в окрестностях корней неограничена [6]. Т. о., при выборе начального условия исключаются точки минимума и максимума функции $u(\xi)$, так как через них проходят также решения $u = 1$ и $u = 2$ соответственно. В (24) зададим $\xi_0 = 0$, тогда $u(0) = 1,5$. Решаем задачу Коши с указанным начальным условием с помощью команды `NDSolve`, в качестве метода зададим метод Рунге-Кутты 4-го порядка [3,4]. Далее выбираем отрезок, на котором будем получать численное решение. Ближайшая точка минимума от начала координат – $(-0.846594; 1)$, точка максимума – $(1.23938; 2)$ (рис 1).

На рис. 2 приведены графики аналитического $u(\xi)$ и численного $u^*(\xi)$ решений уравнения (22) с начальным условием $u(0) = 1.5$ на отрезке $[-0.825018, 1.02706]$. В формуле (22) перед корнем стоит знак плюс.

Уравнение второго порядка (4) с учетом (8) и (9) при $u_1=1, u_2=2, u_3=3$ будет иметь вид

$$u'' = 3u^2 - 12u + 11. \quad (25)$$

Задача Коши для уравнения (25) будет иметь два начальных условия

$$u(\xi_0) = u_0, u'(\xi_0) = u'_0. \quad (26)$$

Правая часть (25) удовлетворяет условию Липшица по u и не зависит от u' . Единственность решения задачи (25), (26) не нарушается, и условия (26) можно задавать везде. Отметим, что если взять отрезок $[-10, 10]$, то видно, что численное решение является периодическим рис. 3.

Уравнение третьего порядка (3) с учетом (9) при $u_1=1, u_2=2, u_3=3$ будет иметь вид

$$u''' = 6uu' - 12u'. \quad (27)$$

Задача Коши для уравнения (27) будет иметь три начальных условия

$$u(\xi_0) = u_0, u'(\xi_0) = u'_0, u''(\xi_0) = u''_0. \quad (28)$$

Правая часть (27) удовлетворяет условию Липшица по u и u' не зависит от u'' . Единственность решения задачи (27), (28) не нарушается, и условия (27) можно задавать везде. Отметим, что если взять отрезок $[-10, 10]$, то видно, что численное решение является периодическим рис. 4.

Точные решения уравнения КдФ типа солитон-волны, сравнение с ними численных решений ОДУ первого, второго и третьего порядков

Пусть

$$u_1, u_2, u_3 \in \mathfrak{R} : u_1 < u_2 = u_3.$$

Тогда, (7) будет иметь вид $P(u) = (u - u_1)(u - u_2)^2$. Заметим, что $u = u_1, u = u_2$, по-прежнему решения уравнения (5). Поскольку, решение должно быть вещественным, то будем считать, что u заключено между u_1 и u_2 . Из (14) следует, что

$$\pm d\xi = \frac{\sqrt{2}}{\eta^2 - (u_2 - u_1)}. \quad (29)$$

После интегрирования (29), получим

$$\pm \sqrt{2}(\xi + c_2) = \frac{1}{\sqrt{u_2 - u_1}} \ln \left| \frac{\eta - \sqrt{u_2 - u_1}}{\eta + \sqrt{u_2 - u_1}} \right|, \quad c_2 \in \mathfrak{R} \quad (30)$$

Мы здесь воспользовались тем, что если $u = u_2$, то $\eta = 0$, поэтому выбрали $\xi + \tilde{\eta}_2 = 0$ при $\eta = 0$. Из (30) будем иметь

$$\pm \sqrt{2(u_2 - u_1)}(\xi + c_2) = \ln \left| \frac{\eta - \sqrt{u_2 - u_1}}{\eta + \sqrt{u_2 - u_1}} \right| = \ln \frac{\sqrt{u_2 - u_1} - \eta}{\sqrt{u_2 - u_1} + \eta} = \ln \frac{\sqrt{u_2 - u_1} - \sqrt{u - u_1}}{\sqrt{u_2 - u_1} + \sqrt{u - u_1}}, \quad c_2 \in \mathfrak{R}.$$

Выбирая в левой части последнего знак минус, после преобразования получим

$$e^{-\sqrt{2(u_2 - u_1)}(\xi + c_2)} = \frac{\sqrt{u_2 - u_1} - \sqrt{u - u_1}}{\sqrt{u_2 - u_1} + \sqrt{u - u_1}}, \quad c_2 \in \mathfrak{R}.$$

Разрешая последнее относительно u , имеем

$$u(\xi) = u_1 + (u_2 - u_1) \left(\frac{1 - e^{-\sqrt{2(u_2 - u_1)}(\xi + c_2)}}{1 + e^{-\sqrt{2(u_2 - u_1)}(\xi + c_2)}} \right)^2 = u_1 + (u_2 - u_1) \left(\frac{e^{\sqrt{2(u_2 - u_1)}(\xi + c_2)} - 1}{e^{\sqrt{2(u_2 - u_1)}(\xi + c_2)} + 1} \right)^2 \quad (31)$$

По аналогии с [2, с. 46] мы будем процесс, определяемый (31), называть уединенной волной, или солитоном. Задавая в (31) $u_1=1$, $u_2=u_3=2$ построим график соответствующего аналитического решения, рис. 5.

Если мы поставим задачу Коши для ОДУ первого порядка, то в качестве ξ_0 в (24) нельзя брать значение, в котором солитон достигает минимума, поскольку в этой точке нарушается единственность решения задачи Коши. Т. о., при выборе начального условия исключается точка минимума функции (31), т. к. через нее также проходит решение $u=u_1$. Зададим $\xi_0=0$, тогда из (31) при $u_1=1$, $u_2=u_3=2$ следует, что $u(0)=1,5$. На рис. 6 приведен график аналитического и численного решений уравнения

$$u' = \pm(u-2)\sqrt{2(u-1)} \quad (32)$$

при соответствующих начальных условиях на отрезке $[-1,24545; 3,75355]$ ($(-1,24645; 1)$ – точка минимума (31) при $u_1=1$, $u_2=2$).

Если мы поставим задачу Коши для ОДУ второго порядка (4), то правая часть этого уравнения удовлетворяет условию Липшица по u и не зависит от u' , поэтому единственность решения задачи Коши с условиями (26) не нарушается. Снова зададим $\xi_0=0$, $u(0)=u_0=1,5$, тогда $u'(0)=u'_0=0,5$. Ищем решение на симметричном отрезке относительно точки минимума для решения типа солитон $[-6,24645; 3,75355]$. Графики численного и аналитического решения представлены на рис. 7.

Если мы поставим задачу Коши для ОДУ третьего порядка (3), то правая часть этого уравнения удовлетворяет условию Липшица по u и по u' , а также не зависит от u'' , поэтому единственность решения задачи Коши с условиями (28) не нарушается. Зададим $\xi_0=0$, $u(0)=u_0=1,5$, $u'(0)=u'_0=0,5$. Снова ищем решение на отрезке $[-6,24645; 3,75355]$. Графики численного и аналитического решения представлены на рис. 8.

На рисунках 2, 3, 4, 6, 7, 8 численное решение изображено точками, а аналитическое сплошной линией. Во многих местах этих рисунков графики этих решений практически сливаются. Практическая неразличимость графиков численного и аналитического решений происходит из-за малой погрешности численных решений на отмеченных участках.

Заключение

В работе с помощью системы компьютерной алгебры «Wolfram Mathematica» исследуются периодические решения уравнения КдФ конкретного вида. Численные решения задач Коши для ОДУ первого, второго и третьего порядка сравниваются с соответствующим точным решением.

Для ОДУ первого порядка, полученного из уравнения КдФ в случае кноидальной волны задачу Коши нельзя ставить в точках минимума и

максимума, так как в этих точках правая часть уравнения не удовлетворяет условию Липшица. Поэтому начальное условие ставим в точке роста, а отрезок выбираем таким образом, чтобы он лежал между точками экстремума.

Затем изучаем ОДУ второго порядка, которое удовлетворяет условию Липшица, и показываем, что задачу Коши для этого уравнения в случае кноидальной волны можно ставить в любой точке. Кроме этого, отрезок для численного решения значительно расширяется, так что наблюдается периодичность.

Для ОДУ третьего порядка условие Липшица выполняется и, поэтому, задача Коши также может ставиться в любой точке.

Для решения типа солитон условие Липшица выполнено даже для уравнения первого порядка в окрестности кратного корня.

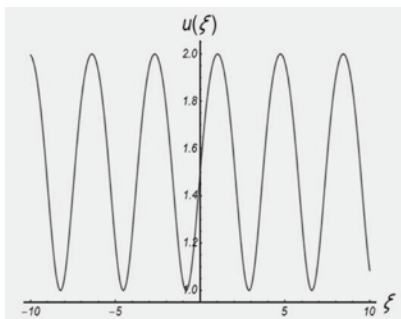


Рис. 1. Кноидальная волна при $u=u_1$, $u=u_2$, $u=u_3$.

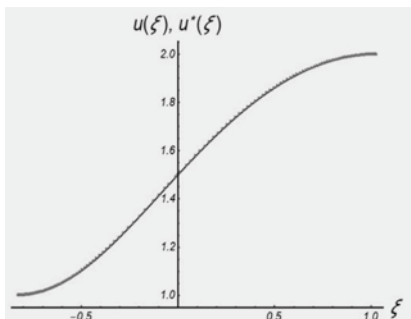


Рис. 2. Численное и аналитическое решение ОДУ первого порядка для кноидальной волны

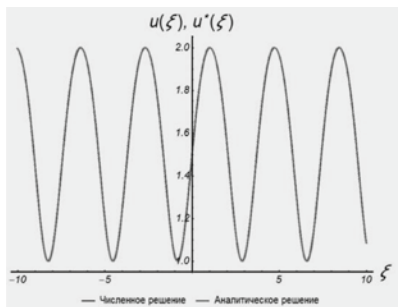


Рис. 3. Численное и аналитическое решение ОДУ второго порядка для кноидальной волны.

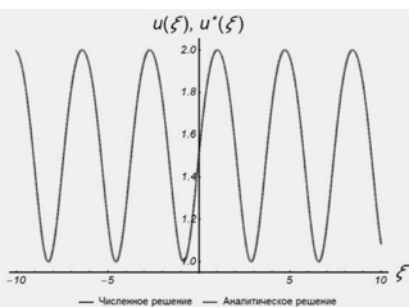


Рис. 4. Численное и аналитическое решение ОДУ третьего порядка для кноидальной волны.

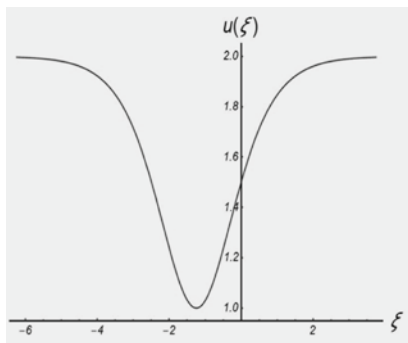


Рис. 5. График аналитического решения (солитон) при $u_1 = 1$, $u_2 = 2$, $u_3 = 2$, $c_2 \approx 1,24645$.

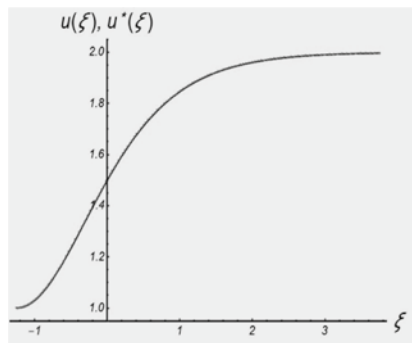


Рис. 6. Численное и аналитическое решения ОДУ первого порядка для солитона (в (32) в правой части взят знак плюс).

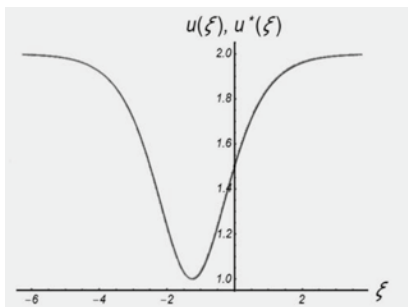


Рис. 7. Численное и аналитическое решения ОДУ второго порядка для солитона.

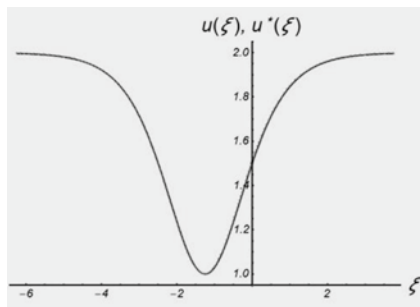


Рис. 8. Численное и аналитическое решение ОДУ третьего порядка для солитона.

Список литературы

1. Уизем Дж. Линейные и нелинейные волны.–М.: Мир, 1977.–624с.
2. Бхатнагар П. Нелинейные волны в одномерных диспергирующих системах. М.: Мир, 1983.–136с.
3. Полянин А.Д., Зайцев В.Ф. Справочник по нелинейным уравнениям математической физики: точные решения. М.: Физматлит, 2002. – 432 с.
4. Амосов А.А., Дубинский Ю.А., Копченова Н.В. Вычислительные методы для инженеров: Учеб. пособие – М.: Высш. шк., 1994.–544с.
5. Галанин М.П. Методы численного анализа математических моделей/М.П. Галанин, Е.Б. Савенков.–М.: Изд-во МГТУ им. Н.Э. Баумана, 2010.–591, [1] с.: ил. (Математическое моделирование в технике и технологии).
6. Полянин А.Д., Зайцев В.Ф. Справочник по нелинейным уравнениям математической физики: точные решения. М.: Физматлит, 2002. – 432 с.

СЕКЦИЯ 7. ИНФОРМАЦИОННЫЕ ТЕХНОЛОГИИ В ГУМАНИТАРНОЙ СФЕРЕ

УДК 61

**Valentīns Popovs¹, Vladimirs Skudnovs², Aleksejs Vasiljevs³,
Aleksejs Ševčenko⁴**

Riga Technical University, Institute of Transport
Azenes street 12a, Riga, LV-1048, Latvia

¹popovs@latnet.lv; ²vladimir.skudnov@sotus.net; ³aleksejs.vasiljevs@ldz.lv, ⁴
alexey@cryptolab.net

PROBLEMS OF TELEMEDICINE IN LATVIA

Abstract: *This work was written based on the current Latvian heterogeneous telecommunication network features and experimental results with medical data, which were obtained from the RTU Railway transport and communications laboratory. Achieved results can be used Latvian eHealth (electronic health) the new range of services for the implementation, which have not existed.*

Keywords: *health systems, telecommunication network, medical diagnosis data, wireless transmission technologies.*

1. Introduction

The emerging development of mobile health systems in the last decade was made possible due to the recent advances in wireless and network technologies, linked with recent advances in nanotechnologies, compact biosensors, wearable devices and clothing, and pervasive and ubiquitous computing systems. These advances will have a powerful impact on some of the existing healthcare services and will reshape the work flow and practices in the delivery of these services [1].

In result this work will provide an overview of recent heterogeneous telecommunication network in Latvia and use of this network to transmit and receive high resolution format (Blu-ray) medical diagnosis data.

E. Kyriacou, M.S. Pattichis, C.S. Pattichis, A. Panayides, and A. Pitsilides, in their work [2] suggests that “*Wireless telemedicine systems and services are expected to enhance traditional emergency care provision not only within the emergency department but also in a variety of prehospital emergency care situations where geographically remote consultation and monitoring can be implemented*”. A timely and effective way of handling emergency cases can prove essential for a patient’s recovery or even for a patient’s survival. Especially in cases of serious injuries of the head, the spinal cord, and internal organs, the way of transporting and generally the way of providing care are crucial for the future of the patient.

This is especially actual problem for small countries like Latvia, where due

bad economic situation people are migrating from villages and small towns to towns with more population and opportunities or capital in search for a job, in result of this migration local hospitals and medical clinics are receiving less finance. In result less medical services are available at local place, especially, professional specialists' consultation, forcing patients in need to search for a health services in huge clinical hospital centers or specialized clinics, in Latvia those clinical hospitals generally are centered around the capital of Latvia – Riga.

1.1. Methods

In order to write this article, the more recent research (published 2013 Aug. Cisco projection of global mobile data growth) in growth of mobile data and heterogeneous telecommunication network and potential use in health services were overviewed.

In preparation of this work about eHealth (electronic health) and mobile health services situation topic in Latvia, East Clinical University Hospital Medical Technology Director Uldis Jaspers (uldis.jaspers@aslimnica.lv) tel. 67303113 was interviewed on topic of medical data exchange and manage between Latvian hospitals, as well, what mobile health services right now are available for patients of the East Clinical University hospital in Latvia.

All the information regarding heterogeneous telecommunication and mobile network technologies, applications were obtained from legal sources – books, publications, conference abstracts from the internet websites. Used sources in preparation of this work are mentioned in the „References” section.

2. Wireless transmission technologies

The digital era in mobile cellular networks was introduced by the Global System for Mobile Communications (GSM), signifying the transition from analog first-generation (1G) to digital second-generation (2G) mobile telecommunications. The general packet radio service (GPRS) and enhanced data rates for GSM evolution (EDGE) 2.5G technologies facilitated the data-transfer rates that allowed mobile Health systems to integrate medical image and video transmission, in addition to biomedical signals. Toward this direction, third-generation mobile telecommunications (3G) systems (Universal Mobile Telecommunications System [UMTS]) set the foundations of establishing telemedicine systems in standard clinical practice, exploiting extended coverage for a virtually always-on service provision.[3] Today's 3.5G and near-4G (fourth generation mobile telecommunications) wireless networks, namely, high-speed downlink packet access (HSDPA), high-speed uplink packet access (HSUPA), high speed packet access (HSPA), HSPA+, mobile WiMAX, and Long-Term Evolution (LTE), allow the deployment of responsive mobile Health and eEmergency systems, minimizing delays and extending upload data-transfer rates that can accommodate high-diagnostic quality medical video that can rival the quality of in-hospital examinations. Ongoing deployment of 4G wireless networks and emerging fifth generation mobile telecom-

munications (5G) systems theoretical upload data-transfer rates comparable to wired infrastructure will expedite the adoption of such systems and services in standard clinical practice.[3]

The described wireless networks facilitate incremental data-transfer rates while minimizing end-to-end delay. Evolving wireless communications networks' theoretical upload data rates range from 50 kbps to 86 Mbps. In practice, typical upload data rates are significantly lower. More specifically, typical upload data rate ranges are (I) for GPRS, 30–50 kbps; (II) for EDGE, 80–160 kbps; (III) for evolved EDGE, 150–300 kbps; (IV) for UMTS, 200–300 kbps; (V) for HSPA, 500 kbps–2 Mbps; (VI) for HSPA+, 1–4 Mbps; and (VII) for LTE, 6–13 Mbps [3].

Figure 1 shows a Cisco projection of global mobile data growth through 2017, measured in exabytes (billion gigabytes) per month, demonstrating traffic growing at a compound annual rate of 66% – resulting in thirteenfold growth over that period.

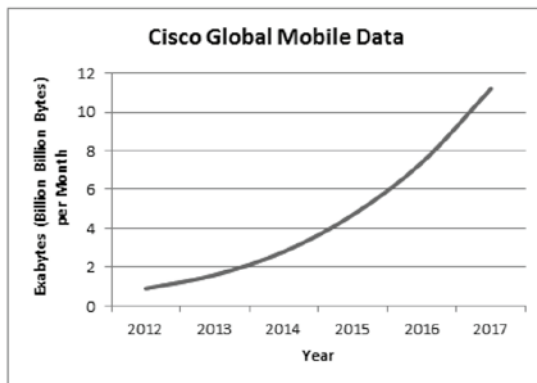


Figure 1. Global Mobile Data Growth[3]

Figure 2 shows another data projection, predicting 50% annual growth in data for the 2012 to 2018 period, resulting in twelvefold growth.

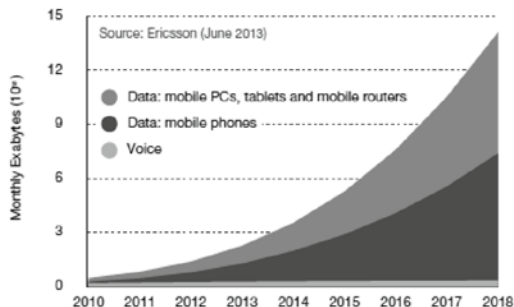


Figure 2. Global Mobile Traffic for Voice and Data 2010 to 2018 [3]

2.1. Mobile computing platforms

In a recent research, where mobile and fixed computer use by doctors and nurses on hospital wards was investigated, it was found that the choice of device was related to clinical role, nature of the clinical task, and degree of mobility required. Nurses work and clinical tasks performed by doctors during ward rounds require highly mobile computer devices, and they showed a strong preference for generic computers mobile on wheels (including laptops) over all other devices. Tablet personal computers (PCs), some devices are shown at Figure 3, were selected by doctors for only a small proportion of clinical tasks. Note that even when using mobile devices, clinicians completed a very low proportion of observed tasks at the bedside [4].



Figure 3. Tablet personal computers [4]

A systematic review of personal digital assistant (PDA) usage surveys by healthcare providers was carried out by Garritty and El Emam in their work “Who’s using PDAs? Estimates of PDA use by health care providers: A systematic review of surveys,” [5]. It was documented that younger physicians and residents and those working in large and hospital-based practices are more likely to use a PDA. Moreover, it appeared that professional PDA use in healthcare settings involved more administrative and organizational tasks than those related to patient care. Garritty and El Emam concluded that physicians are likely accustomed to using a PDA; however, there is still a need to evaluate the effectiveness and efficiency PDA-based applications [5].

2.2. Transmission of digital images

The use of digital images in medicine has benefited from the formation of the DICOM committee [8]. The committee was formed in 1983 by the American College of Radiology (ACR) and the National Electrical Manufacturers

Association (NEMA). For still images, DICOM has adopted various JPEG variants such as lossless JPEG (JPEG-LS) [9] and JPEG 2000 [10].

We first consider lossless image-compression methods that provide for exact reconstruction of the input images. Lossless image-compression eliminates the need for diagnostic validation of compression artifacts. Unfortunately, lossless methods provide limited compression ratios, usually ranging between 2 and 3.7. Lossy image-compression methods can provide much better compression ratios. However, the use of lossy image compression requires a careful evaluation of the effect of compression artifacts on diagnostic performance. While not directly relevant to diagnostic performance, lossy image compression of general images attempts to be perceptually lossless.

AUTHOR	YEAR	RESOLUTION, FRAME RATE, BITRATE	ENCODING STANDARD	WIRELESS NETWORK	MEDICAL VIDEO MODALITY
CHU ET AL. [17] ^F	04	(320 × 240 AND 160 × 120) < 5 FRAMES/S 30-80 Kbps	M-JPEG	3G-CDMA	TRAUMA VIDEO
GAIWAT ET AL. [18] ^{1,3}	06	176 × 144 @ 5 FRAMES/S 18.5-60 Kbps	H.263	3G-UMTS	CARDIAC ULTRASOUND
ALREJAD ET AL. [16] ^F	12	(176 × 144, 352 × 288) @ 10/20 FRAMES/S (220, 430) Kbps, 1.3 Mb/s	WINDOWS MEDIA VIDEO (WMV)	MOBILE WIMAX, HSDPA	CARDIAC ULTRASOUND
ISTERNIAN ET AL. [19] ^{1,3}	09	176 × 144 @ 8-10 FRAMES/S 50-130 Kbps	H.264/AVC	3G	ABDOMEN ULTRASOUND
PANAYIDES ET AL. [20] ^{1,3}	13	(176 × 144, 352 × 288, 560 × 416) @ 15 FRAMES/S, 64-768 Kbps	H.264/AVC	HSPA	CAROTID ARTERY ULTRASOUND
PANAYIDES ET AL. [6] ^{1,3}	13	560 × 416 @ 40 FRAMES/S, UP TO 2 Mb/s	HEVC	3.5G AND BEYOND	CAROTID ARTERY ULTRASOUND
RAO ET AL. [8] ^{1,2,4}	09	360 × 740 @ 30 FRAMES/S 500 Kbps	MPEG-2	3G AND BEYOND	PEDIATRIC RESPIRATOR DISTRESS RELATED VIDEO
MARTINI ET AL. [9] ^{1,4}	10	480 × 256 @ 15 FRAMES/S 300 Kbps	H.264/AVC	MOBILE WIMAX	CARDIAC ULTRASOUND
PANAYIDES ET AL. [10] ^{1,2,4}	11	352 × 288 @ 15 FPS 197-421 Kbps	H.264/AVC	3G AND BEYOND	CAROTID ARTERY ULTRASOUND
KHIRE ET AL. [12] ^{1,2,4}	12	720 × 480 @ 30 FRAMES/S, 125-200 Kbps	H.264/AVC	3G AND BEYOND	MAXILLOFACIAL SURGERY CLIPS
DEBONO ET AL. [14] ^{1,4}	12	640 × 480 @ 25 FRAMES/S	H.264/AVC	MOBILE WIMAX	CARDIAC ULTRASOUND
PANAYIDES ET AL. [11] ^{1,2,4}	13	704 × 576 @ 15 FRAMES/S 768 Kbps-1.5 Mb/s	H.264/AVC	MOBILE WIMAX	CAROTID ARTERY ULTRASOUND
CAVERO ET AL. [13] ^{1,3}	13	720 × 576 @ 25 FRAMES/S 40 Kbps (M-MODE), 200 Kbps (D-MODE)	SPHIT	3G AND BEYOND	CARDIAC ULTRASOUND
CAVERO ET AL. [15] ^{1,3}	12	720 × 576 @ 25 FRAMES/S 200 Kbps	SPHIT	HSPA, MOBILE WIMAX	CARDIAC ULTRASOUND

1-User, 2-Real Time, 3-Clinical Evaluation, 4-ROI.

Table 1. Selected mobil

Here, a compressed image is termed perceptually lossless if an (average) human observer cannot differentiate it from its uncompressed version. In general, optimal performance requires the study of the impact of lossy compression in different clinical scenarios. A general lossy compression approach that can be directly applied to medical images is to use diagnostic regions of interest (ROIs). Here, the parts of an image that are of diagnostic interest will see little or no compression. On the other hand, the parts that are not of diagnostic interest can be compressed significantly. For example, if the region of interest covers about 20% of the entire image, average compression ratios of about 15:1 have been reported

using JPEG-LS, while an average compression ratio of only 2.58 was achieved when using the entire image as the region of interest [11]. Similar to perceptually lossless compression for general images, another approach is to use lossy compression that does not allow clinicians to differentiate between the compressed image from the uncompressed one. Clearly, if the uncompressed image cannot be identified, then the (clinical) visual inspection of the compressed images should not impact the diagnosis [11]. This technique leads to near-lossless techniques where the uncompressed image differs from the original in only a small number of levels (± 1 and ± 2 out of possible 4096 levels). For comparison, JPEG-LS in lossless mode provides for an average compression ratio of 2.58 that improves to 3.83 in the near-lossless mode (± 1 levels) [11]. In addition, for a region of interest that covers 20% of the image region, the average compression ratio improves from 15.1 to 22.0 [11].

2.3. Transmission of digital video

The successful and efficient deployment of medical video communication systems relies on timely integration of video coding standards technologies and adaptation to the underlying wireless channel's capacity. Given the fact that wireless networks impose a limit on the available upload data-transfer rates, exploiting the trade-off between compression efficiency and clinical video quality is of primary importance. In contrast to wireless communication of conventional video and multimedia applications, clinical video quality cannot be compromised, as the latter would translate to misdiagnosis. Consequently, mobile Health medical video communication systems need to be diagnostically driven [11]. The objective of diagnostically driven systems is to maximize the clinical capacity of the communicated video by adapting the employed algorithms for video encoding, wireless communication, and both objective and clinical video quality assessments on the underlying medical video modality. While sharing common concepts and principles, diagnostically driven approaches are often medical video modality specific.

In terms of video resolutions facilitated by video coding standards, the H.261 standard (first video coding standard—early 1990s) supported the quarter common intermediate format (QCIF) (176×144) and the common intermediate format (CIF) (352×288) video resolutions. Subsequently, the H.262 standard (released in 1995) extended video resolutions support up to 4CIF (720×576) and 16CIF (1408×1152). Then, the H.263 standard (introduced in 1996) provided for improved quality at lower bit rates while also allowing lower, sub-QCIF (128×96) video resolution encoding. The aforementioned video coding standards were employed in early medical video communications systems of limited clinical capacity. This was partly attributed to the inability to encode medical video at video resolutions and frame rates (to match the available bit rate of wireless channels) that would not compromise diagnostic quality and clinical motion, respectively. The H.264/advanced video coding (AVC) standard [12] dominated mobile Health systems and services of

the past decade (released in 2003). The H.264/AVC standard, linked with 3G and 3.5G wireless networks, enabled efficient and timely encoding to match the available data rates and conform to real-time transmission requirements. In addition to coding efficiency and error-resilience coding tools, H.264/AVC introduced the network abstraction layer (NAL), a novel concept enabling network-friendly adaptation of the encoded content to candidate heterogeneous networks and/or storage devices (and cloud infrastructure), a significant feature toward H.264/AVC success. High-resolution and high-frame rate medical video communication is feasible using H.264/AVC over 3.5G wireless channels, leading to a plethora of mobile Health systems and services of high diagnostic value, being a milestone for mobile Health systems and services. However, in contradiction to initial projections and enthusiasm, the adoption of wireless telemedicine systems in standard clinical practice remains limited. The new high efficiency video coding (HEVC) standard specifically designed for beyond high-definition video communications that provides 50% bitrate gains for comparable visual quality compared to the H.264/AVC standard. Together with 4G and beyond (5G) wireless networks deployment, the HEVC standard is expected to play a decisive role toward wider adoption of such systems and services. To achieve this, mobile Health medical video communication systems that can transmit medical video at the clinically acquired resolution and frame rate that can be robustly transmitted in low delay without compromising clinical quality are envisioned.

3. Mobile health services in Latvia

From the personal interview of East Clinical University Hospital Medical Technology Director Uldis Jaspers (uldis.jaspers@aslimnica.lv), in Latvia around 90% of hospitals are using so called Arstabiroids [13] <http://www.arstabiroids.lv> provided by Meditec.lv company. By his words, they are monopolists in Latvia eHealth service market.

The only mobile health applications what they are providing according to their official site <http://www.meditec.lv/HP> is HealthPoint [14], description from meditec.lv site: “HealthPoint is a unique application, developed for smartphones, which allows dealing with a variety of health-related issues. The application allows you to make calls to the emergency number 113, while transferring the caller’s location data (GPS coordinates) and basic information about the caller to the emergency medical service dispatchers. The app allows you to call your family doctor or make a call to the consultative phone of NHS general practitioners.”

Here I would like to say, this is not true, that their application is unique. All what this application does is:

Making preset call to 113 emergency service and sending GPS coordinates of the caller, but there is a trick. Your phone should have an active GPS module and application should have access to gather that information from smartphone. If GPS is turned off in your smartphone or application won’t have

access rights to it, then application won't send any coordinates and instead of it after 5 seconds will just make a phone call to 113 emergency service as you could do it just from ordinary phone or mobile phone, not necessary from smartphone or tablet.

Instead of calling 113, second option of mentioned application is to call your family doctor, but before you can do that you need to preset his phone number.

Here comes the question, why I need the application if I can do it from Contacts?

Needless to say, as for latest version of HealthPoint have tested on 17.04.2016, its broken on Android 5.1, if HealthPoint can't access GPS data from smartphone, instead of calling after 5 seconds, application just hangs with black screen and no call to emergency service 113 is made, so it's not even safe to use for health emergency call, it's very dependable application for proper configuration of GPS module, which can be problem for non IT specialist users to use, especially for aged people.

4. Results and Discussion

The heterogeneous telecommunication network experiment for the research was placed at the RTU railway transport telecommunication laboratory in Riga, Ierikustreet 8A, Figure 4-5. [17]



Figure 4. General view of the heterogeneous network, Riga, Ieriku street 8A.

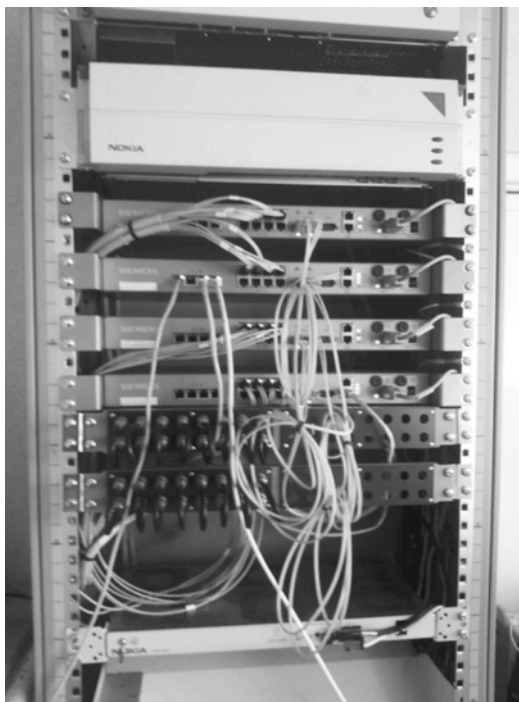


Figure 5. SIEMENS SURPASS hiT7020 – 4 multiplexers with subscribers system connections are reconnected to the ring via optical STM-1 interfaces.

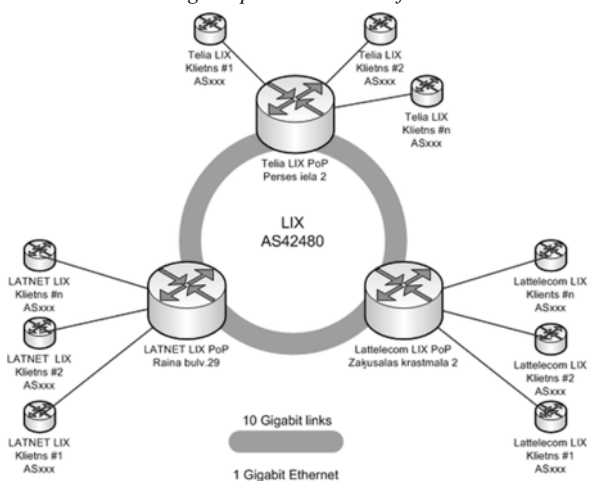


Figure 6 Internet exchange point of Latvia (Latnet, Telia, Lattelecom)

The experiment object was medical high-resolution video (Blu-ray) stream in DICOM H.264/AVC format over optic network and over mobile network using two largest mobile operators in Latvia, Tele2 and LMT – 2G/3G/4G

connection standard in their network.

Latvian internet exchange point to global internet, Figure 6.

The results of streaming high resolution medical data video format – DI-COM H.264/AVC in heterogeneous network:

Through(server-terminal) Optical Connection, Figure 7.

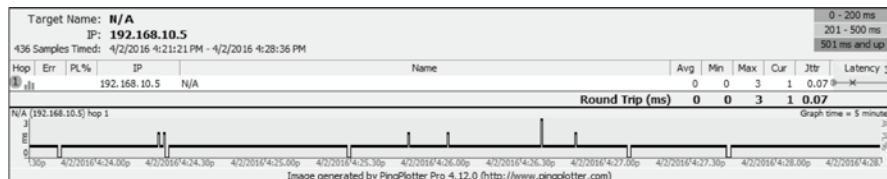


Figure 7. Optical connection



Fig. 8. LMT 4Gconnection

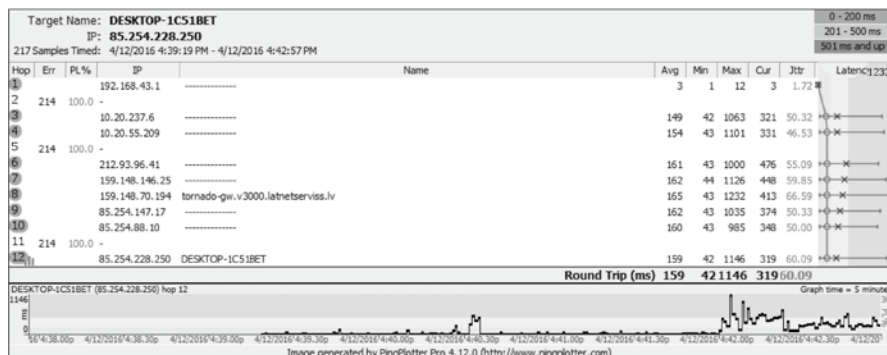


Figure 9. LMT 3G (HSPA)

LMT 2G connection (ED

GE)Figure 10.



Figure 10. LMT 2G (EDGE)

The results from TELE2 network:
 TELE2 3G connection (WCDMA) Figure 11.



Figure 11. TELE 3G (WCDMA)

TELE2 2G connection Figure 12.

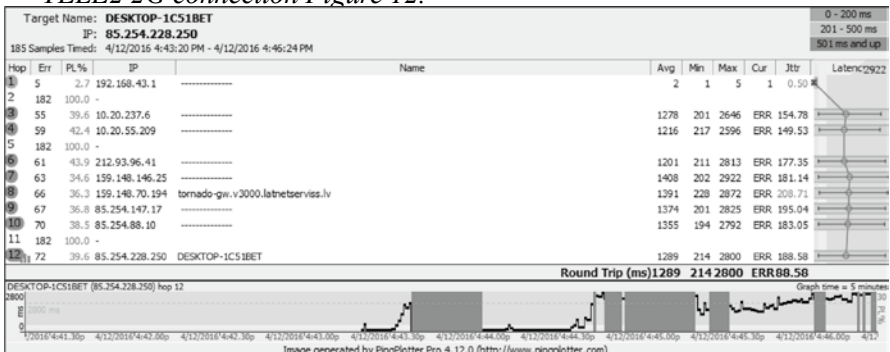


Figure 12. TELE2 2G

3. Conclusions

1. The experiment showed, that even in worst case scenario as 2G connection, it's like somewhere in the middle of unpopulated area in Latvia and some kilometers away from transmit/receive tower, it's possible to transmit and receive high resolution video format (Blu-ray, H.264/AVC DICOM) although it need some time for buffering, but at the end, have received video and audio stream without quality or data lost.

LMT and TELE2 representatives are declaring on their sites – www.lmt.lv and www.tele2.lv that their 4G coverage is 98% of the area of Latvia.[15-16]

2. From the experiment gained data, it's clear that Latvian heterogeneous telecommunication network infrastructure is ready for mobile medicine-based platform applications, that could improve the quality of service in the Latvian Health system and more people from unpopulated areas could receive the part of health services same as they would live in capital, technology allows that. The problem is in management, all stored data over clinical hospitals that are using “Arstabirojs” in Latvia and that is 80%-90% [13] of total healthcare system institutions is stored and managed by private company Meditec [13]. So independent developers can't access and integrate their mobile application due to lack of access to databases and as private company Meditec is not interested in sharing this access, because this is main source of income for this company [13]. Latvian government should step between and regulate this access by law, in this way mobile health application market in Latvia could gain healthy concurrence which is good for patients and overall for industry. Right now, Latvia has no any integrated mobile health applications between government owned clinical hospitals and patients with authorizing access for patients to medical databases to overview their medical examination/diagnostic data or gather data from wearable biosignal devices and upload it to their medical data profile, as an example later on medical file could be examined by doctor. The only mobile health application for patients that Meditec is providing (HealthPoint) isn't integrated with the database data from clinical hospitals, its standalone application with 2 poor designed functions.

Gratitude

The authors are grateful to Professor Dr.habil.phys. V.Popov for posing the problem, constant assistance and advice at all stages of this work.

Благодарность.

Авторы выражают искреннюю признательность профессору В.Попову за постановку задачи, постоянную помощь и консультирование на всех этапах проведения данной работы.

References

- [1] R.H. Istepanian, S. Laxminarayan, and C.S. Pattichis, Eds., M-Health: Emerging Mobile Health Systems, Springer, NewYork, 2006.
- [2] E. Kyriacou, M.S. Pattichis, C.S. Pattichis, A. Panayides, and A. Pitsillides,

- “m-Health e-Emergency systems: Current status and future directions,” *IEEE Antenn. Propag. Mag.*, vol. 49, no. 1, pp. 216–231, 2007.
- [3] Rysavy Research, LLC, “Mobile Broadband Explosion: The 3GPP Wireless Evolution,” Aug. 2013, <http://www.4gamerica.org/>. Access: 15.04.2016
- [4] P. Andersen, A.M. Lindgaard, M. Prgomet, N. Creswick, and J.L. Westbrook, “Mobile and fixed computer use by doctors and nurses on hospital wards: Multi-method study on the relationships between clinician role, clinical task, and device choice,” *J. Med. Internet Res.*, vol. 11, no. 3, p. e32, Aug. 4, 2009.
- [5] C. Garrity and K. El Emam, “Who’s using PDAs? Estimates of PDA use by healthcare providers: A systematic review of surveys,” *J. Med. Internet Res.*, vol. 8, no. 2, p. e7, Apr.–Jun. 2006. Published online May 12, 2006. doi:10.2196/jmir.8.2.e7.
- [8] National Electrical Manufacturers Association (NEMA), *Digital Imaging and Communications in Medicine (DICOM) Publication PS 3*, Rosslyn, Virginia, U.S., 1996.
- [9] ISO/IEC 14995-1 Final Draft International Standard (FDIS), “Information Technology—Lossless and Near-Lossless Coding of Continuous-Tone Still Images: Baseline,” *JPEG-LS Standard, Part 1*, Mar. 1999.
- [10] M. Boliek, C. Christopoulos, and E. Majani, Eds., “JPEG2000 Part I Final Draft International Standard” (ISO/IEC FDIS I5444-1), ISO/IEC JTC1 / SC29 / WG11 N855.
- [11] A.S. Panayides, M.S. Pattichis, and C.S. Pattichis, “Mobile-health systems used in diagnostically driven medical video technologies,” *IEEE Signal Proc. Mag.*, vol. 30, no. 6, pp. 163–172, Nov. 2013. doi:10.1109/MSP.2013.2276512.
- [12] ITU-T, “Advanced Video Coding for Generic Audiovisual Services,” ITU-T and ISO/IEC 14496-10 Recommendation H.264 (MPEG4 AVC), May 2003.
- [13] Arsta Birojs by Meditec <http://www.arstabirojs.lv/index.php/about-us>
- [14] Health Point application. Access: 15.04.2016 <http://www.meditec.lv/HP>
- [15] LMT, access: <http://www.lmt.lv>
- [16] Tele2, access: <http://www.tele2.lv>
- [17] Popovs V., Skudnovs V., Vasiljevs A. Modeling of modern heterogeneous data networks. RTU starptautiskā zinātniskā konference, Transportasekcija, Oktobris, 2014

СЕКЦИЯ 8. ТЕХНОЛОГИИ ИСКУССТВЕННОГО ИНТЕЛЛЕКТА И ИХ ПРИМЕНЕНИЕ

УДК 658.562.012.7

Жуков Д.А., Клячкин В.Н.,

д-р техн. наук, профессор

Ульяновский государственный технический университет, Ульяновск, Россия

v_kl@mail.ru, www.ulstu.ru

ОТБОР ЗНАЧИМЫХ ПОКАЗАТЕЛЕЙ ПРИ ДИАГНОСТИКЕ ТЕХНИЧЕСКОГО ОБЪЕКТА С ПРИМЕНЕНИЕМ МАШИННОГО ОБУЧЕНИЯ

Аннотация: Диагностика технического объекта часто сводится к оценке исправности объекта, то есть способности выполнять заданные функции. Естественно стремление сократить время диагностики, уменьшив число учитываемых факторов. Результаты проведенного исследования, однако, показывают, что удаление предположительно незначимых факторов влияет на качество бинарной классификации неоднозначно, и для каждого метода подбор значимых факторов индивидуален.

Ключевые слова: бинарная классификация, незначимые факторы, эффективность диагностики, коррелированность показателей.

Zhukov DA, Klyachkin VN,

Dr. Tech. sciences, professor

Ulyanovsk State Technical University, Ulyanovsk, Russia

v_kl@mail.ru, www.ulstu.ru

SELECTION OF SIGNIFICANT INDICATORS IN DIAGNOSTICS OF TECHNICAL OBJECT USING MACHINE TRAINING

Abstract: *Diagnostics of a technical object often comes down to assessing the health of the object, that is, the ability to perform specified functions. Naturally, the desire to reduce the time of diagnosis, reducing the number of factors to be taken into account. The results of the study, however, show that the removal of supposedly insignificant factors affects the quality of the binary classification is ambiguous, and for each method the selection of significant factors is individual.*

Key words: *binary classification, insignificant factors, diagnostic efficiency, correlated indicators.*

Введение

С целью повышения надежности технического объекта проводится диагностика его состояния. Диагностика часто сводится к оценке исправности объекта, то есть способности выполнять заданные функции. С точки зрения машинного обучения – это задача бинарной классификации (объект исправен или неисправен). При этом оценка состояния объекта проводится в условиях эксплуатации, и для принятия решения

используются различные методы распознавания.

В качестве исходных данных рассматриваются известные результаты оценки состояния системы: при заданных значениях контролируемых показателей функционирования объект признается исправным или неисправным. Таким образом, имеется множество ситуаций с заданными показателями функционирования и множество возможных состояний системы. Известна конечная совокупность пар «набор показателей, состояние» – исходная выборка данных. На основе этих данных требуется построить алгоритм, способный для любого объекта выдать достаточно точный ответ о его состоянии. При этом всегда имеется риск ложной тревоги или пропуска цели. Это задача машинного обучения по прецедентам (с учителем) [1-5].

Подобный подход используется и при оценке клиентов банка для решения вопроса о возможности выдачи кредита: клиенты классифицируются на кредитоспособных и некредитоспособных [6-7].

Для измерения точности классификации вводится функционал качества, в роли которого может использоваться средняя ошибка прогнозирования. При этом исходная выборка разбивается на собственно обучающую, по которой строится алгоритм поиска искомой зависимости, и контрольную (тестовую), по которой оценивается средняя ошибка. Оценивается количество неправильных прогнозов в контрольной выборке.

Часто используется кросс-валидация, когда исходная выборка разбивается на несколько частей, одна из которых рассматривается как контрольная, остальные – как обучающая выборка; последовательно рассматриваются все варианты, и ошибка прогнозирования усредняется.

Для решения задач технической диагностики используются методы обучения с учителем: применяются как классические статистические методы [8-9]: дискриминантный анализ, наивный байесовский классификатор, логистическая регрессия, так и специальные методы машинного обучения: нейронные сети, метод опорных векторов, случайные деревья решений и другие.

Задача классификации технического объекта решается по следующему принципу: объект считается исправным $Y = 1$, если модель спрогнозировала вероятность $P\{Y = 1 | X\} > 0,5$, и неисправным $Y = 0$ – в противном случае. При этом в качестве граничного значения может быть использовано число, отличное от 0,5.

Исходная выборка для рассматриваемого объекта (системы водоочистки) включала 348 наблюдений за восьмью показателями функционирования [10] – характеристиками водоисточника. Система считалась исправной, если показатели качества питьевой воды соответствовали требуемым. Исправность системы Y оценивалась по показателям качества питьевой воды в зависимости от физико-химических показателей водоисточника: X_1 – температуры, X_2 – цветности, X_3 – мутности, X_4 –

значения pH, X_5 – щёлочности, X_6 – окисляемости, и доз добавляемых реагентов: X_7 – коагулянта и X_8 – флокулянта. Расчет проводился в системе Matlab.

Отбор показателей с учетом их коррелированности

Показатели функционирования объекта могут существенно отличаться по их влиянию на его исправность. Для отбора наиболее значимых факторов могут использоваться различные подходы. Один из них – анализ корреляционной матрицы.

В табл. 1 показана корреляционная матрица, характеризующая степень тесноты линейной связи как между самими показателями функционирования $X_1 - X_8$, так между ними и состоянием системы Y .

Таблица 1. Корреляционная матрица

	X_1	X_2	X_3	X_4	X_5	X_6	X_7	X_8	Y
X_1	1,000	-0,545	0,226	0,220	-0,328	0,270	-0,561	-0,964	-0,566
X_2	-0,545	1,000	-0,057	-0,024	0,318	0,069	0,999	0,578	0,224
X_3	0,226	-0,057	1,000	0,190	-0,128	0,078	-0,061	-0,045	-0,071
X_4	0,220	-0,024	0,190	1,000	-0,097	0,273	-0,035	-0,189	-0,009
X_5	-0,328	0,318	-0,128	-0,097	1,000	-0,022	0,320	0,287	0,222
X_6	0,270	0,069	0,078	0,273	-0,022	1,000	0,062	-0,245	-0,144
X_7	-0,561	0,999	-0,061	-0,035	0,320	0,062	1,000	0,594	0,236
X_8	-0,964	0,578	-0,045	-0,189	0,287	-0,245	0,594	1,000	0,563
Y	-0,566	0,224	-0,071	-0,009	0,222	-0,144	0,236	0,563	1,000

Анализ этой матрицы показывает, что показатели X_3 (мутность) и X_4 (значение pH) практически не оказывают влияние отклик Y : соответствующие значения коэффициентов корреляции $r(X_3, Y) = -0,071$ и $r(X_4, Y) = -0,009$ (выделены жирным шрифтом). По-видимому, эти показатели не должны оказывать существенного влияния на исправность системы водоочистки.

Результаты расчета для некоторых методов (логистической регрессии и двух вариантов бустинга [11]) представлены в табл. 2. Приведены проценты ошибок, полученные для контрольной выборки. Видно, что удаление факторов не приводит к улучшению качества прогнозирования.

Таблица 2. Результаты исследования (процент несовпадений) с исключением факторов X_3 и X_4

Метод	Все факторы	Без X_3	Без X_4	Без X_3 и X_4
Логистическая регрессия	16,17	16,47	16,58	16,45
AdaBoost	11,53	11,9	13,09	11,90
LogitBoost	8,39	9,23	8,95	9,16

Еще одна особенность корреляционной матрицы – практически линейная связь показателя X_1 (температуры) с X_8 (дозой флокулянта): коэффициент корреляции $r(X_1, X_8) = -0,964$, а также показателя X_2 (цветности) с X_7 (дозой коагулянта): коэффициент корреляции $r(X_2, X_7) = 0,999$. Действительно, начальная доза коагулянта определяется по цветности, а начальная доза флокулянта подбирается с учетом температуры воды. Таким образом, из каждой пары таких показателей имеет смысл использовать лишь один.

Соответствующие результаты расчета представлены в табл. 3. Видно, что удаление факторов привело к некоторому улучшению качества прогнозирования для метода Логистическая регрессия (без учета показателя X_8 процент ошибок снизился с 16,17 до 13,90) и GentleBoost (без учета показателя X_7 процент ошибок снизился с 8,00 до 7,26).

Таблица 3. Результаты исследования (процент несовпадений) с исключением факторов X_7 и X_8

Метод	Все факторы	Без X_7	Без X_8	Без X_7 и X_8
Логистическая регрессия	16,17	20,00	13,90	19,81
LogitBoost	8,39	8,58	9,52	8,81
GentleBoost	8,00	7,26	8,49	7,54

Отбор показателей по регрессионной модели

Оценка значимости показателей функционирования может быть проведена с использованием критерия Стьюдента для оценки значимости регрессоров в регрессионном анализе [8]. Рассматривая показатели $X_1 - X_8$ как регрессоры, а Y – как отклик, можно построить регрессионную модель и оценить значимость регрессоров. На рис.1 представлены результаты расчета, полученные в системе Statistica. Видно, что незначимы в модели показатели X_3 (мутность), X_5 (щёлочность), X_6 (окисляемость) и X_8 (доза флокулянта): с позиций регрессионного анализа эти факторы не оказывают значимого влияния на отклик Y – исправность системы (p -значения для этих факторов больше, чем 0,05).

Результаты расчета представлены в табл. 4. Видно, что удаление факторов привело к некоторому улучшению качества прогнозирования для метода Логистическая регрессия (процент ошибок снизился с 16,17 до 14,50). Для всех трех методов бустинга результаты ухудшились: процент ошибок в контрольной выборке увеличился.

Таблица 4. Результаты исследования (процент несовпадений) с исключением факторов, незначимых по критерию Стьюдента

Метод	Все факторы	Без X_3, X_5, X_6 и X_8
Логистическая регрессия	16,17	14,50
AdaBoost	11,53	13,87
LogitBoost	8,39	12,91
GentleBoost	8,00	12,38

Regression Summary for Dependent Variable: Y (исх in						
R= ,62395309 R ² = ,38931746 Adjusted R ² = ,38460631						
F(8, 1037)=82,637 p<0,0000 Std. Error of estimate: ,32756						
	b*	Std.Err. of b*	b	Std.Err. of b	t(1037)	p-value
Intercept			-15,6943	2,533189	-6,19548	0,000000
X1	-0,35435	0,125831	-0,0244	0,008674	-2,81610	0,004953
X2	-4,31049	0,684271	-0,3766	0,059777	-6,29938	0,000000
X3	-0,01380	0,036471	-0,0025	0,006488	-0,37827	0,705304
X4	0,07642	0,026576	0,3097	0,107710	2,87541	0,004118
X5	0,00100	0,026054	0,0117	0,304472	0,03829	0,969463
X6	-0,01883	0,025816	-0,0012	0,001665	-0,72952	0,465845
X7	4,20968	0,692540	3,8174	0,628001	6,07861	0,000000
X8	0,21167	0,125169	3,2610	1,928365	1,69109	0,091120

Рис.1. Результаты регрессионного анализа

Заключение

Результаты проведенного исследования показывают, что удаление предположительно незначимых факторов влияет на качество бинарной классификации неоднозначно, и для каждого метода подбор значимых факторов индивидуален. Необходимо также отметить, что при большом количестве показателей, оказывающих влияние на функционирование технического объекта, может быть эффективен метод главных компонент [8].

Список литературы

1. Witten I.H., Frank E. Data mining: practical machine learning tools and techniques. 2nd ed. – San Francisco: Morgan Kaufmann Publishers, 2005. – 525 p.
2. Воронина В.В. Теория и практика машинного обучения : учебное пособие /
3. В. В. Воронина, А. В. Михеев, Н. Г. Ярушкина, К. В. Святков. – Ульяновск: УлГТУ, 2017. – 290 с.
4. Клячкин В.Н., Кувайскова Ю.Е., Жуков Д.А. Использование агрегированных классификаторов при технической диагностике на базе машинного обучения // Информационные технологии и нанотехнологии: сборник трудов III международной конференции и молодежной школы. – Самарский национальный исследовательский университет имени академика С.П. Королева. – 2017. – С. 1770-1773
5. Жуков Д.А., Клячкин В.Н., Кувайскова Ю.Е. Сравнительный анализ методов машинного обучения при прогнозировании состояния технического объекта // Радиоэлектронная техника. – 2017. – № 1 (10). – С. 189-192.
6. Жуков Д.А., Хорева А.С., Кувайскова Ю.Е., Клячкин В.Н. Формирование контрольных выборок при технической диагностике объекта с применением машинного обучения // Математические методы и модели: теория, приложения и роль в образовании международная научно-техническая конференция : сборник научных трудов. – Ульяновск, 2016. – С. 44-48.

7. Шунина Ю.С., Клячкин В.Н. Прогнозирование платежеспособности клиентов банка на основе методов машинного обучения и марковских цепей // Программные продукты и системы. 2016. № 2. С. 105-112.
8. Клячкин В.Н., Шунина Ю.С. Система оценки кредитоспособности заемщиков и прогнозирования возврата кредитов // Вестник компьютерных и информационных технологий. 2015. № 11 (137). С. 45-51.
9. Клячкин В.Н., Кувайскова Ю.Е., Алексеева В.А. Статистические методы анализа данных. – М.: Финансы и статистика, 2016. – 240 с.
10. Клячкин В.Н. Статистические методы в управлении качеством: компьютерные технологии. – М.: Финансы и статистика, ИНФРА-М, 2009. – 304 с.
11. Бубырь Д.С., Клячкин В.Н. Разработка математических моделей и программ для обеспечения надежности функционирования системы управления водочисткой // Информатика, моделирование, автоматизация проектирования: Сборник научных трудов VII Всероссийской школы-семинара аспирантов, студентов и молодых ученых. – Ульяновск, 2015. – С. 85-89.
12. Жуков Д.А., Клячкин В.Н. Алгоритмы бустинга в задачах технической диагностики // Перспективные информационные технологии: труды Международной научно-технической конференции. – Самара: Издательство Самарского научного центра РАН. – 2017. – С. 787-790.

УДК 004. 81

Проскурин А. Е., Токарева И. В.
ИСПОЛЬЗОВАНИЕ МОДЕЛИ «НЕОКОГНИТРОНА» В
МЕТОДАХ РАСПОЗНАВАНИЯ ОБЪЕКТОВ

Данная статья посвящена исследованию возможностей нейросетевого алгоритма «Неокогнитрон» в области распознавания контуров объектов. В соответствии зрительной нервной системы был предложен новый алгоритм. Этот алгоритм может быть реализован с помощью многослойной сети, состоящая из нейронов-клеток. Сеть, далее называемая «Неокогнитрон», является самоорганизующейся самообучаемой сетью, и обладает способностью распознавать образы.

Ключевые слова: Нейрон; Неокогнитрон; Многослойная нейросеть; Модель нейронной сети; Зрительная нервная система.

Proskurin A. E., Tokareva I. V.
USING OF THE “NEOCOGNITRON” MODEL IN THE OBJECT
RECOGNITION METHODS

Abstract: This article is devoted to the study of the capabilities of the Neocognitron neural network algorithm in the field of object contour recognition. In accordance with the visual nervous system, a new algorithm has been proposed. This algorithm can be implemented using a multilayer network consisting of neuron cells. The network, hereinafter referred to as “Neocognitron,” is a self-organizing, self-learning network, and has the ability to recognize patterns.

Keywords: neuron; Neocognitron; Multilayered neural network; Neural network model; The visual nervous system.

Проблема эффективного распознавания образов имеет важное значение в сферах автоматизации определенных процессов человеческой деятельности, связанных с идентификацией различных объектов окружающего мира, например, авторизация рабочего персонала по отпечаткам пальцев или сетчатке глаза, системы охраны, верификация кредитных карточек, идентификация продукта и расчет цены в магазине по штрих-коду, данная технология крайне важна и в робототехнике и военном деле, поскольку позволит разделять людей на «свой» – «чужой» или разделять на классы по уровню доступа и так далее.

В ходе развития кибернетических наук было разработано множество методов по распознаванию образов. Однако они обладают существенными недостатками, такими как узкая специализированность применения, а также необходимость использования дополнительных алгоритмов обработки изображений и т.д. В результате последних медицинских исследований появилась возможность смоделировать нейросеть, структура которой основывается на строении зрительной коры головного мозга. Модель этой сети – неокогнитрон – показала высокую эффективность ее применения для распознавания образов[6].

С технической точки зрения неокогнитрон является иерархической нейронной сетью состоящей из достаточно большого числа слоев идущих друг за другом и имеющий неполные (можно даже сказать достаточно редкие) связи между слоями [2]. Входной образ подается на входной слой и далее передается каждому последующему. Слоевая структура неокогнитрона показана на рис. 1.

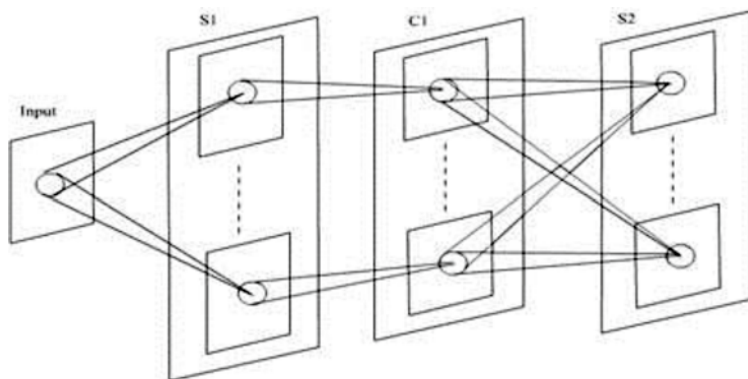


Рисунок 1 – Структура неокогнитрона

Неокогнитрон состоит из двух типов слоев: простые и комплексные. Нейроны входящие в простые слои настраиваются на распознавание каких-либо определенных образов. Комплексные же матрицы созданы для того, чтобы объединить однотипные результаты распознавания простых матриц. Рассмотрим пример объединения результатов распознавания на

базе двух шаблонов для первого слоя, каждый из которых представляет собой отрезок, наклоненный на 22 градуса от горизонтали (см. рис. 2, 3).



Рисунок 2 – Шаблоны распознавания

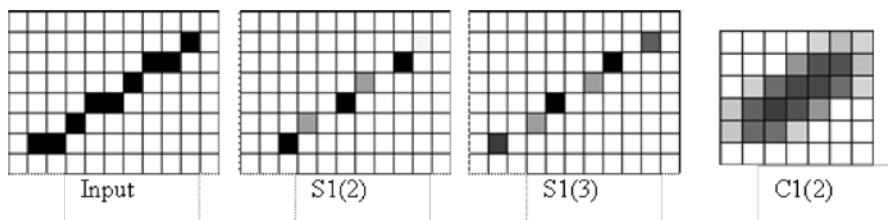


Рисунок 3 – Пример объединение результатов распознавания

Еще одной особенностью неокогнитрона является то, что он обучается слой за слоем, это позволяет ввести достаточно большое количество слоев и обучить их за непродолжительный отрезок времени. Довольно часто в неокогнитроне пользуются 8 слоями (4 простых и 4 сложных слоя [1, 3, 4]) с приблизительным количеством в 15000 нейронов. Использование такого количества нейронов и слоев, например, в сети обратного распространения ошибки привело бы к очень долгому (почти невозможному) обучению нейронной сети.

Каждый нейрон как простого, так и комплексного слоя получает сигналы не от всех нейронов, а лишь от некоторых, что также влияет на увеличение скорости распознавания образов неокогнитроном. Такие нейроны называются «областью видимости» [1]. Размерность областей видимости может быть различна, однако в результате исследований было выявлено, что оптимальные размеры областей видимости это 3x3 и 5x5. Для центрального нейрона каждой матрицы, центральный элемент области видимости располагается в центре матрицы предыдущего слоя (см. рис. 4).

Различают слои со сжатием и без него. Для слоев со сжатием характерно то, что размерность его матриц меньше, чем размерность матриц предыдущего слоя. Для сжатия могут использоваться различные методы. В создаваемой модели для слоев со сжатием используется следующий метод: центр области видимости нейрона, смещенного на 1 позицию от центрального нейрона, смещается на 2 единицы относительно центра матрицы предыдущего слоя (см. рис. 5). Для примера приведенного на рисунке 4 получается, что в область видимости угловых элементов входят только 4 нейрона.

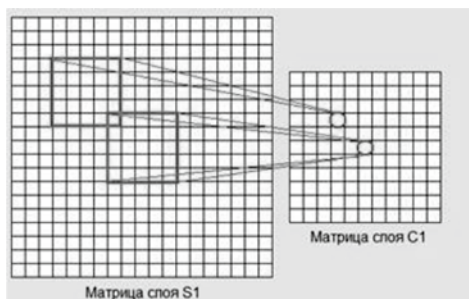


Рисунок 4 – Области видимости

У нейронов, расположенных в матрицах слоев без сжатия, индекс центрального элемента области видимости соответствует его индексу в матрице.

В неокогнитроне не все коэффициенты являются настраиваемыми, также имеются фиксированные коэффициенты, расчет которых производится в самом начале моделирования. Настраиваемыми коэффициентами в неокогнитроне являются коэффициенты от нейронов комплексного слоя к нейронам простого слоя и коэффициент от подавляющего нейрона к нейрону простого слоя. Коэффициенты от комплексных нейронов к подавляющему нейрону и от простых нейронов к комплексным являются фиксированными. В связи с этим возникает проблема выбора фиксированных коэффициентов. Изначально, предполагается лишь, что значения коэффициентов должно уменьшаться при удалении от центра области видимости. При настройке весовых коэффициентов используется простая формула расчета Δ , которая зависит от величины входных сигналов (c):

$$\Delta w_{i,j} = a * t_{i,j} * c_{i,j}$$

$$w_{i,j}^{k+1} = w_{i,j}^k + \Delta w_{i,j}$$

[3], здесь t – фиксированные коэффициенты от комплексных нейронов к тормозящему нейрону, и a – скорость обучения. Величина Δ , для весового коэффициента подавляющего нейрона выбирается так, чтобы процесс обучения мог завершиться. В проделанной работе выбрана следующая формула:

$$\Delta w_{\phi} = \alpha \frac{v}{2}.$$

Таким образом, неокогнитрон, являясь универсальным средством для распознавания графических образов, имеет такие важнейшие качества как быстрая скорость обучения и распознавания образов, расширяемость модели, для получения более точных результатов распознавания. Эти качества позволяют использовать эту нейронную сеть для распознавания контуров объектов.

Литература

1. Lauren Fausett. Fundamentals of Neural Networks: Architectures, Algorithms and Applications. New Jersey: Prentice Hall 398-417 pp. 2000.

2. Ф. Уоссермен. Нейрокомпьютерная техника. Теория и практика. М. : Мир, 1992 – 260 с.
3. S. Satoh, J. Kuroiwa, H. Aso and S. Miyake. Recognition of Hand-written Patterns by Rotation-invariant Neocognitron Proc. of ICONIP'98, 1, pp. 295-299, 1998.
4. S. Satoh, J. Kuroiwa, H. Aso and S. Miyake, Pattern Recognition System with Top-Down Process of Mental Rotation, Proc. of IWANN'99, 1, pp. 816-825, 1999.
5. Ф. Уоссермен, Нейрокомпьютерная техника. Теория и практика. – М. : Мир, 1992. – 260.

СЕКЦИЯ 9. РАЗВИТИЕ ИНФОРМАЦИОННЫХ ТЕХНОЛОГИЙ

УДК: 004.421.5.

Ацамаз Аланович Датиев¹

¹Федеральное государственное бюджетное образовательное учреждение высшего образования «Северо-Кавказский горно-металлургический институт (государственный технологический университет)», 362031, РСО-Алания, г. Владикавказ, Николаева 44

ЭКСПЕРИМЕНТАЛЬНЫЙ АНАЛИЗ ЭФФЕКТИВНОСТИ КОМПОЗИТНОЙ ВЕРСИИ МЕТОДА НЕЯВНОГО ПЕРЕБОРА ПРИМЕНИТЕЛЬНО К ЗАДАЧЕ О РАНЦЕ

Аннотация: Экспериментально исследуется эффективность композитной модификации метода динамического программирования применительно к задаче о ранце. Этот алгоритм совмещает технологии вычисления оценок, присущие методам типа ветвей и границ и принципы отсеечения «бесперспективных» планов динамического программирования. Экспериментально продемонстрирована высокая эффективность предложенного подхода, причем показано, что превосходство композитной версии алгоритма неявного перебора над традиционной реализацией этого метода возрастает с ростом размерности задачи о ранце.

Ключевые слова: глобально оптимальное решение, булевы переменные, композитные алгоритмы, динамическое программирование, экспериментальные исследования, задача о ранце, методы типа ветвей и границ.

Atsamaz Alanovich Datiev¹

¹Federal State Budgetary Educational Institution of Higher Education "North Caucasus Mining and Metallurgical Institute (State Technological University)", 362031, RNO-Alania, Vladikavkaz, Nikolaev 44

EXPERIMENTAL ANALYSIS OF THE EFFICIENCY OF THE COMPOSITE VERSION OF THE METHOD OF INVISIBLE BREAKFAST APPLICABLE TO THE PROBLEM OF THE RANCH

Abstract: The effectiveness of composite modification of a method of a dynamic programming in relation to a task about a satchel is experimentally investigated. This algorithm combines the technologies of calculation of estimates inherent in methods of type of branches and borders and the principles of a dissecting away of "unpromising" plans of a dynamic programming. High performance of the offered approach is experimentally shown, and it is shown that superiority of the composite version of an algorithm of implicit search over traditional realization of this method increases with body height of dimension of a task about a satchel.

Keywords: the optimal solution, Boolean variables, composite algorithms, a dynamic programming, the pilot studies, a task on a satchel, methods of type of branches and borders is global.

1. Введение

Методы неявного перебора являются одним из основных инструментов решения прикладных задач, сводимых к математическим моделям с дискретно меняющимися переменными [1-3]. Ниже приведены экспериментальные результаты сравнительного анализа эффективности предложенных в [4] композитной версии динамического программирования применительно к задаче о ранце [5] по сравнению с ее «классической» реализацией. Далее к процедурам такого рода отнесено динамическое программирование [6-10], а под ее композитной версией, в соответствии с [4], понимается процедура, в которой одновременно присутствуют компоненты двух технологий: технология отсечения «бесперспективных» планов, используемая в динамическом программировании [9], дополнена учетом оценок, генерируемых методами типа ветвей и границ [6 – 8]. При этом ниже не анализировалась эффективность попыток сократить время счета методами неявного перебора за счет распараллеливания вычислений [11-14]. Сокращение времени поиска оптимальных решений композитными версиями методов неявного перебора достигается за счет сокращения числа анализируемых и хранимых в памяти планов, что позволяет надеяться на сокращение не только времени счета, но и объема используемой оперативной памяти компьютера. Схема проведения экспериментов по проверке сравнительной эффективности вышеописанного алгоритма была следующей:

1. Размерность решавшихся задач “ n ” менялась в диапазоне $3 \div 32$.

2. При каждой фиксированной размерности решалось десять различных задач вида (1), коэффициенты которых генерировались с помощью генератора случайных чисел. В памяти фиксировались среднеарифметические величины времени счета T применительно к каждой величине “ n ”.

3. Эффективность предложенного подхода отображалась отношением времен поиска решения «классической» и композитной версии.

Далее представлены результаты экспериментальной проверки сравнительной эффективности классического и композитного процедур поиска глобально оптимальных решений задачи о ранце, в рамках которых применяются следующие обозначения, допущения и определения.

2. Обозначения, допущения и определения

Ниже рассматривается задача о ранце с вектором Z булевых переменных вида:

$$\begin{cases} F = \sum_i C_i z_i \rightarrow \max; \\ \sum_i b_i z_i \leq a; \\ \forall i : z_i = 1, 0, \end{cases} \quad (1)$$

где: $Z = \{z_1, z_2, \dots, z_n\}$; $Z' \in Z$; C_i, b_i, a – константы, причем далее полагаем, что все константы неотрицательны, причем верхняя оценка величины F при условии, что в базис введено подмножество $Z' \in Z$ переменных, значениям которых соответствует вершина x_k дерева поиска, ниже определяется выражением:

$$\Delta(x_k) = \Delta(Z') = \begin{cases} \sum_{i \in I_1(x_k)} C_i z_i + \sum_{j \in I(Z) \setminus I_1(x_k)} C_j, & \text{если } \sum_{i \in I_1(x_k)} b_i z_i \leq a; \\ -\infty, & \text{в противном случае,} \end{cases} \quad (2)$$

где $I_1(x_k)$ обозначает множество индексов введенных в базис булевых переменных, значения которых определяются положением $x_k \in X$ на дереве поиска; $I(Z)$ – множество индексов всех булевых переменных решаемой задачи.

Аналогично определяется нижняя оценка $\delta(I_j)$ величины F задачи (1), отвечающая вершине x_k дерева поиска:

$$\delta(x_k) = \begin{cases} \sum_{i \in I_1(x_k)} C_i z_i, & \text{если } \sum_{i \in I_1(x_k)} b_i z_i \leq a; \\ -\infty, & \text{если } \sum_{i \in I_1(x_k)} b_i z_i > a. \end{cases} \quad (3)$$

Еще одной характеристикой вершины x_k дерева поиска является ресурс $\mu(x_k)$:

$$\mu(x_k) = a - \sum_{i \in I_1(x_k)} b_i z_i. \quad (4)$$

Теперь если на дереве поиска существуют две вершины x_k и x_p , такие, что справедливо хотя бы одно из приводимых ниже условий а), б):

$$\text{а) } \Delta(x_k) < \delta(x_p); \quad (5)$$

или

$$\text{б) } \begin{cases} I(x_k) = I(x_p); \\ \mu(x_k) \leq \mu(x_p); \\ \delta(x_k) \leq \mu(x_p), \end{cases} \quad (6)$$

то подмножество векторов переменных, отвечающих вершине x_k , можно исключить из дальнейшего рассмотрения.

3. Динамическое программирование

Содержательно решение экстремальной задачи с булевыми переменными с помощью динамического программирования может быть описано последовательным, ярус за ярусом, построением сети $G(X, U)$ применительно к задаче (1). Ниже приводится агрегированное описание «классической» процедуры такого рода.

Алгоритм 1

1. На множестве всяких и не помеченных вершин $X \subseteq X_4$ построенной части дерева ветвлений $G_4(X_4, U_4)$ выбирается вершина x_j . Если

таковой вершины нет, то перейти к шагу 4. Если выбор осуществляется на первой итерации, то такой вершиной *a priori* считается корневая вершина дерева.

2. Осуществляется ветвление из выбранной на шаге 1 последней итерации вершины x_j и с использованием (3) и (4) вычисляются компоненты $\delta(x_k)$ и $\mu(x_k)$ каждого вектора $R(x_k)$, отвечающего каждой висячей вершине x_k построенного «куста» (величина $h=1$).

3. Все вершины куста, построенного на шаге 2 последней итерации, помечаются, перейти к шагу 1

4. Убираются все пометки подмножества висячих вершин дерева ветвлений.

5. Если на множестве висячих вершин существует вершина x_k , для которой на том же множестве вершин существует вершина x_q такая, что справедлива система (6), то вершина x_k вычеркивается.

6. Если число введенных в базис переменных висячих вершин дерева ветвлений равно числу переменных решаемой задачи, то перейти к следующему шагу, в противном случае перейти к шагу 1.

7. На множестве не вычеркнутых висячих вершин выбирается вершина x_k с наибольшей $\delta(x_k)$ компонентой соответствующего ей вектора $R(x_k)$.

8. Алгоритм закончен. Вектор переменных Z , соответствующий выбранной на шаге 7 вершине x_k , является оптимальным.

Ниже, на рис. 1, представлена экспериментальная зависимость среднего времени T поиска решения задачи (1) алгоритмом 1 от числа переменных n этой задачи.

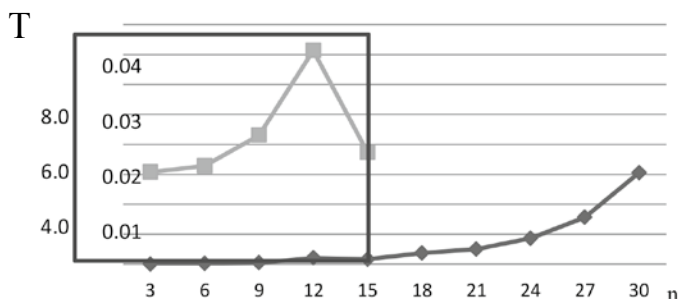


Рис. 1. Экспериментальная зависимость $T(n)$

Композитная версия алгоритма 1 отличается реализацией только шагов 2 и 5, которые далее приводятся с штриховой индексацией:

2'. Осуществляется ветвление из выбранной на шаге 1 последней итерации вершины x_j и с использованием формул (2), (3) и (4) вычисляются три компоненты $\Delta(x_k)$, $\delta(x_k)$ и $\mu(x_k)$ каждого вектора $R(x_k)$, отвечающего каждой висячей вершине x_k построенного «куста» (величина $h=1$).

5°. Если на множестве всеяких вершин существует вершина x_k , для которой на том же множестве вершин существует вершина x_q такая, что справедливо хотя бы одно из условий (5) и (6), то вершина x_k вычеркивается.

Ниже, на рис. 2, представлена экспериментальная зависимость среднего времени T' (в секундах) поиска решения задачи (1) композитной версией алгоритма 1 от числа переменных n этой задачи.

T

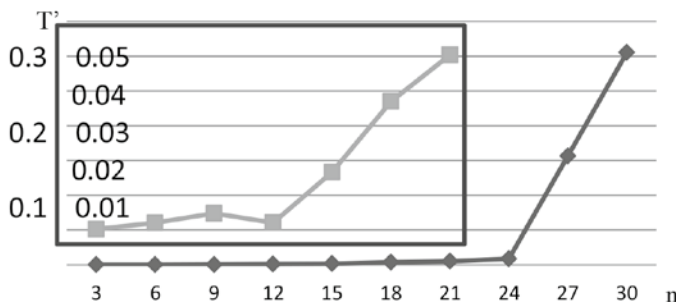


Рис. 2. Экспериментальная зависимость среднего времени T' поиска решения задачи о ранце композитной версией алгоритма 1 от числа переменных n в секундах

Величина выигрыша во времени поиска решения композитной версией динамического программирования по сравнению его традиционной реализацией определяется отношением $\eta = T(n)/T'(n)$, зависимость которого от размерности задачи о ранце представлена ниже на рис. 3.

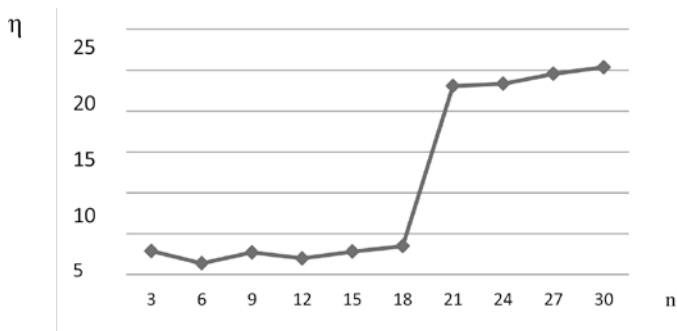


Рис. 3. Выигрыш η во времени поиска решения композитной версией динамического программирования по сравнению его традиционной реализацией применительно к задаче о ранце, как функция числа переменных этой задачи

Заключение

Экспериментальный анализ эффективности предложенной в [1] композитной модификации динамического программирования применительно к задаче о ранце подтвердил его превосходство над традиционной процедурой такого рода [6-8,10]. К экспериментально подтвержденным достоинствам композитного алгоритма можно также отнести

рост его эффективности с ростом размерности решаемых задач (рис. 3), что позволило более чем на 25% расширить рекомендованный в [10] диапазон применения методов типа ветвей и границ.

Библиографический список

1. Nowozin, Sebastian; Lampert, Christoph H. "Structured Learning and Prediction in Computer Vision". *Foundations and Trends in Computer Graphics and Vision*. 2011, **6** (3–4), pp. 185–365.
2. Achterberg, T.; T. Koch; A. Martin. "Branching rules revisited". *Operations Research*, 2005, **33** (1), pp. 42–54.
3. Mehlhorn, Kurt; Sanders, Peter (2008). *Algorithms and Data Structures: The Basic Toolbox*. 2008, Springer. p.249.
4. Гроппен В.О. Композитные алгоритмы поиска глобально оптимальных решений экстремальных задач с булевыми переменными. Материалы семинара "IT- технологии: теория и практика", Изд. Терек, Владикавказ, 2017г., с. 21-38.
5. Caccetta, L. & Kulanoot, A. "Computational Aspects of Hard Knapsack Problems". *Nonlinear Analysis*. 2001, **47**, pp. 5547–5558.
6. Land, A.H. & Doig, A.G. An automatic method of solving discrete programming problems, *Econometrica*, Vol. 28, No. 3. 1960, pp.497-520.
7. Little, John D. C.; Murty, Katta G.; Sweeney, Dura W.; Karel, Caroline (1963). "An algorithm for the traveling salesman problem". *Operations Research*, 1963, **11** (6), pp. 972–989.
8. Balas, Egon; Toth, Paolo (1983). *Branch and bound methods for the traveling salesman problem. Report in the Carnegie Mellon University, Graduate School of Industrial Administration.*, 1983.
9. Bellman, R. The Theory of Dynamic Programming. The RAND Corporation, 1954.
10. Корбут А.А., Финкельштейн Ю.Ю. Дискретное программирование. М.: Наука, 1969. – 368 с.
11. Casti J., Richardson M. & Larson R. Dynamic programming and parallel computers. – JOTA, Vol. 12, No. 4, 1973, pp. 423–438.
12. Brochard, L. Efficiency of some parallel numerical algorithms on distributed systems. *Parallel Computing*, 1989, Vol. 12, No.1, pp. 21–44.
13. Groppen V. O. Efficiency of parallel computations in solving extreme combinatorial problems. First World Conference on Parallel Computing in Engineering and Engineering Education. UNESCO, Paris, France, October 8-12, 1990, p. 127-131.
14. Bader, David A.; Hart, William E.; Phillips, Cynthia A. "Parallel Algorithm Design for Branch and Bound". In Greenberg, H. J. *Tutorials on Emerging Methodologies and Applications in Operations Research*. Kluwer Academic Press, 2004.

УДК 651:002.6

Жуков Р.В., Жуков И.В.,
НИУ «БелГУ», Белгород, Россия

КОНЦЕПЦИЯ АВТОМАТИЗАЦИИ ПРОЦЕССА ПОДГОТОВКИ ДОКУМЕНТАЦИИ НА ИНФОРМАЦИОННЫЕ СИСТЕМЫ

Аннотация: На основании анализа процессов генерации различных пакетов документов и вспомогательных файлов, авторы формулируют требования к программным средствам, используемым для автоматизации этих процессов. Наиболее перспективным в этом плане авторам представляется пакет Help&Manual: в статье приводятся достоинства этого пакета.

Ключевые слова: система автоматизации подготовки документации, трудозатраты, структурные элементы документов, Help&Manual.

R. Zhukov, I. Zhukov,
Belgorod State National Research University "Bel GU", Belgorod, Russia

THE CONCEPT OF AUTOMATING THE PROCESS OF PREPARING DOCUMENTATION ON INFORMATION SYSTEMS

Abstract: Based on the analysis of the generation processes of various packages of documents and auxiliary files, the authors formulate the requirements for the software used to automate these processes. The most promising in this regard, the authors of the package Help & Manual: the article presents the advantages of this package.

Key words: documentation automation system, labor costs, structural elements of documents, Help & Manual.

По мере роста числа создаваемых и вводимых в действие информационных систем все более остро встает вопрос сокращения сроков и снижения трудоемкости разработки и поддержания в актуальном состоянии, в том числе и документации.

Разработка документации на информационную системы предполагает подготовку целого ряда документов: концептуальных, отчетных, проектных, рабочих, эксплуатационных и организационно-распорядительных. Итоговый объем документации системы, включающей в себя текстовую и графическую информацию, как правило, составляет десятки, а иногда и сотни тысяч страниц.

Разработка и поддержание документации в актуальном состоянии на всех стадиях и этапах жизненного цикла информационной системы вызывает немало проблем. В большей степени они связаны, с необходимостью многократного внесения в документацию, состоящую из множества отдельных файлов, всевозможных поправок, дополнений и изменений.

Корректировка документов – трудоемкая операция, требующая от технического писателя аккуратности, высокой сосредоточенности, отнимающая массу времени, сил и нервов.

Ситуация отягощается еще и тем, что:

– «государственный» заказчик, финансирующий львиную долю разрабатываемых информационных систем, все чаще и чаще требует от исполнителя строжайшего и неформального соблюдения требований ГОСТ при разработке технической документации;

– требования отечественных ГОСТов несоизмеримо более жесткие по сравнению с требованиями зарубежных стандартов;

– специалистов, способных удовлетворить требованиям заказчика (соблюдение требований ГОСТ), в ВУЗах практически не готовят, так как их подготовка все больше сводится к так называемому «процессному подходу».

Автоматизация процесса подготовки технической документации – единственный разумный путь, гарантирующий стабильное качество документации, существенное сокращение сроков ее разработки, снижение трудоемкости поддержания комплекта документов в актуальном состоянии.

Предпосылками к автоматизации процесса подготовки документации, в соответствии с формулировками ГОСТов 34-й серии, являются:

– Приложение 1 ГОСТ 34.201–89: «Документация на автоматизированную систему – комплекс взаимосвязанных документов, в котором полностью описаны все решения по созданию и функционированию системы, а также документов, подтверждающих соответствие системы требованиям технического задания и готовность ее к эксплуатации (функционированию)» [1].

– п. 5.1 ГОСТ 34.003–90: «Комплект взаимосвязанных документов, полностью определяющих технические требования к АС, проектные и организационные решения по созданию и функционированию АС» [2].

В чем же заключается «взаимоувязанность» документов комплекса (комплекта)? Практически каждый документ включает в себя фрагменты, идентичные фрагментам, содержащимся в других документах комплекса. Документация на систему оказывается «насквозь пронизанной» жесткими связями на уровне структурных единиц – разделов, подразделов, пунктов и подпунктов отдельных взятых документов.

На рисунке показана связь между подразделами всего лишь трех документов системы: технического задания (ТЗ), пояснительной записки (ПЗ) и общего описания системы (ПО) [1]. Изменения, вносимые в ТЗ, неизбежно повлекут изменения в ПЗ и ПО. Иначе и быть не может, поскольку документация на систему должна быть согласованной – сведения, содержащиеся в ПО, не должны противоречить решениям, обоснованным в ПЗ, и требованиям, предъявляемым ТЗ. Решения, обоснованные в ПЗ, в свою очередь, не должны противоречить требованиям ТЗ. Согласованность, выражающаяся в непротиворечивости, является наиважнейшим показателем качества технической документации.

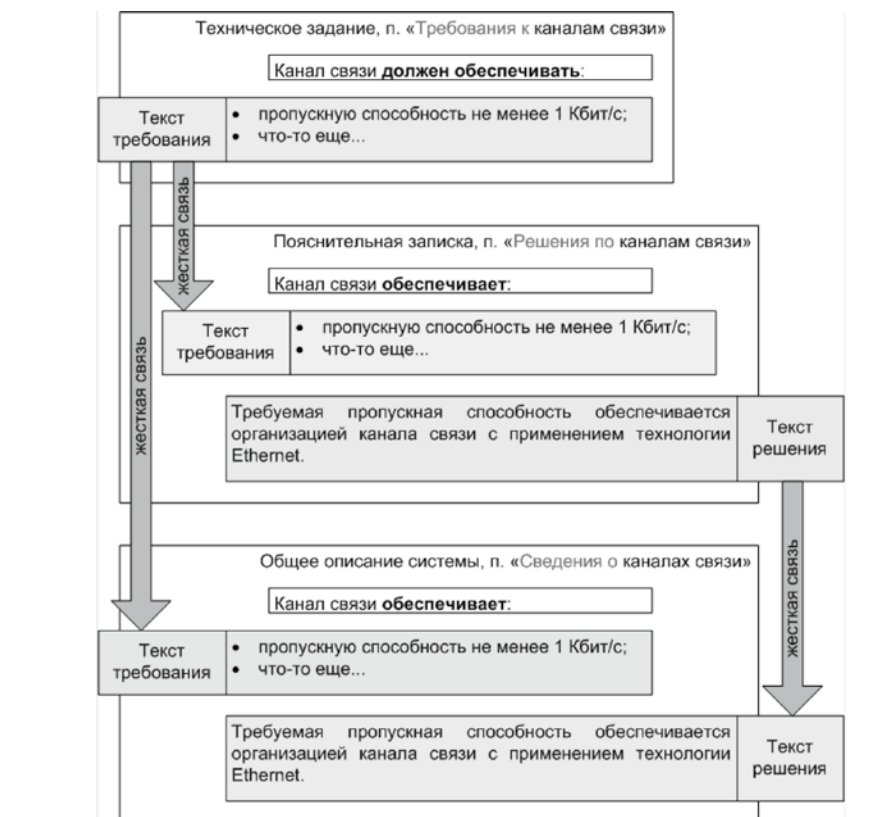


Рисунок 1 – Связь между документами

Каждый документ, создаваемый в отдельном файле, при даже незначительном изменении в требованиях ТЗ потребует ручной корректировки двух других документов. Возникает вопрос, во что выльются трудозатраты, если изменения в ТЗ придется вносить многократно? А если указанный текст содержится не в трех, а в шести (или более) документах, как, например, подраздел «Перечень объектов автоматизации» ТЗ [3], имеющий также место в пяти других документах П2, П4, ПД, ИЗ и Б1.

В практике технических писателей распространено понятие “принцип единого источника” (single source publishing). Это концепция публикации документов, согласно которой один и тот же фрагмент может быть использован в разных выпускаемых документах или в разных форматах публикации (рисунок 2).

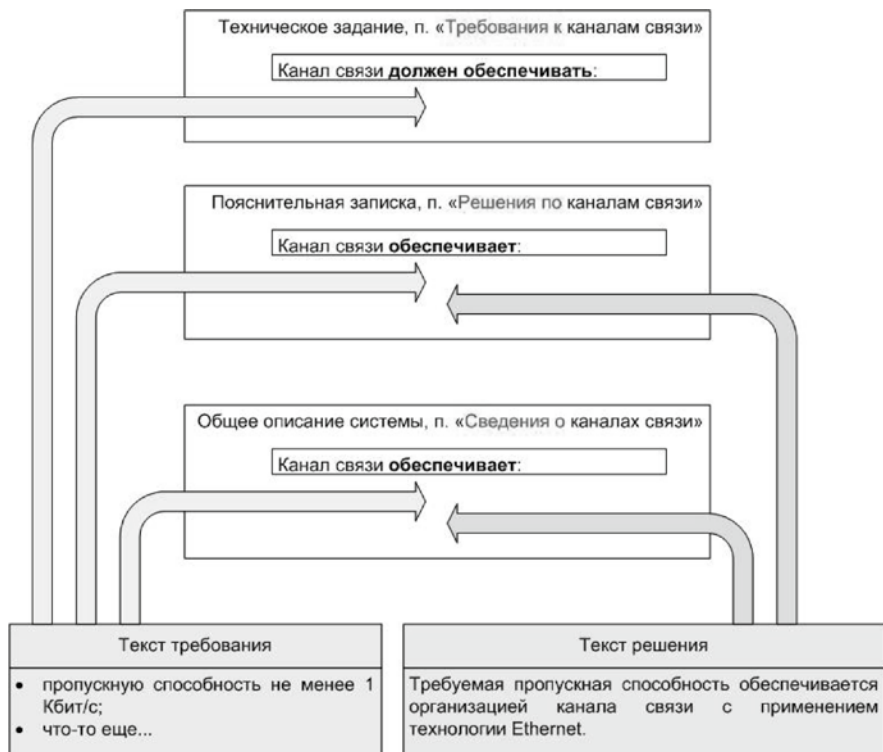


Рисунок 2 – Использование принципов единого источника

Все фрагменты, из которых состоят выходные документы, находятся в общей базе – едином источнике. Каждый выходной документ представляет собой упорядоченную выборку из единого источника, имеющего удобную для составителей структуру, при этом напрямую в структуру какого-либо документа не трансформирующуюся. Иначе говоря, соотношение между единым источником и документом примерно такое же, как между базой данных и отчетом. Создать документ в такой системе – значит описать его структуру и правила формирования из фрагментов единого источника. Фрагментами могут являться: страницы или абзацы текста, графический материал, отдельные предложения или числа.

Разработка технической документации сводится к однократной реализации связей между структурными элементами документов комплекта, т.е. к созданию библиотеки взаимоувязанных документов. Связи между структурными элементами реализуются единожды, внедрением одного или множества общих фрагментов в структуры множества документов. Корректировка текста требования и (или) текста решения приводит к синхронному изменению содержимого всех взаимоувязанных документов.

В итоге, трудозатратная работа по редактированию текста проводится однажды в одном месте (внутри единого источника). Трудозатраты на поддержание документации в актуальном состоянии сводятся к минимуму, снижаются требования к внимательности и аккуратности исполнителя, сокращается повторение ошибок. Исключаются трудозатраты на оформление документов – их публикация может осуществляться с применением единых шаблонов разметки, однажды созданных, к примеру, согласно требованиям ГОСТ 2.104 и ГОСТ 2.105.

Но самое главное состоит в том, что библиотека единого источника с течением времени превращается в настоящую универсальную базу знаний – знания накапливаются в библиотеке от проекта к проекту, документы оптимизируются, качество документации становится выше.

Автоматизация разработки технической документации путем создания каскада взаимоувязанных документов (библиотеки), удовлетворяющих ряду условий создания, жизни, актуализации и утилизации (в том числе требованиям ГОСТов) возможна с использованием одного из современных инструментариев подготовки документации.

Возможности современных программ очень обширны и далеко не исчерпываются мощным инструментом создания документации, генерации help-файлов, а также удобной системой управления контентом web-сайтов.

Современные системы автоматизации подготовки документации могут обеспечивать следующие возможности:

- регулировать процесс создания, редактирования, отображения и сохранения текстов электронной документации;
- обеспечить возможность сборки документов из модулей данных в файлы различных форматов;
- печати электронных версий;
- обеспечить возможность управления процессом – организация и назначение задач, управление рабочими потоками, управление продуктом в целом;
- обеспечение возможности управления базой данных, управления правами пользователей, контроля версий, создания отчета.

Основываясь на возможностях СУБД, принципах концепции единого источника и других факторах автоматизации возможно проведение анализа современных технологий подготовки документации по следующим параметрам:

- Внутренний формат.
- Многопользовательская работа.
- Пользовательский интерфейс.
- Требования к квалификации технического писателя.
- Возможность импортирования во внутренний формат файлов из формата MS Word.

- Поддержка необходимых выходных форматов.
- Трудоемкость перехода на автоматизированную систему.
- Стоимость автоматизации.

Наиболее удобным средством достижения указанных целей автоматизации является Help&Manual. Выбор данного инструментария обусловлен рядом факторов:

– Интерфейс Help&Manual очень похож на интерфейс Microsoft Office 2010. При этом Help&Manual более интуитивно понятен, что позволит повысить производительность труда. Интерфейс имеет большие значки, всплывающие подсказки для быстрого доступа и доступные кнопки команд (которые не спрятаны где-то в меню, где их невозможно найти).

– Исходным форматом Help&Manual является XML, это позволяет при необходимости использовать исходные файлы в формате XML в другом инструментарии.

– Help&Manual является WYSIWYG редактором XML, который не требует знания языка XML или навыков XML-кодирования. Работа осуществляется в знакомом текстовом редакторе типа MS Word, но также в случае необходимости есть доступ к файлу в формате XML.

– Несколько пользователей могут работать над одним проектом одновременно. При этом редактирование определенного топика может осуществляться только одним автором, другой автор этот топик может только просматривать.

– Help&Manual может редактировать несколько проектов одновременно, что позволит легко настроить использование повторяющегося текста в разных документах.

– Разработчики Help&Manual усовершенствовали систему для обработки больших проектов (с многотысячными страницами в разделах). Усовершенствования затронули возможность динамической загрузки в систему проектов готовых справок, поиск и замену, настройку статуса раздела (топика), многократное использование стилей и настроек.

– Help&Manual в качестве источников для документа может использовать разделы и главы из других проектов, различные фрагменты из импортируемых файлов и фрагменты из центрального хранилища данных. При изменении источников подтянутые фрагменты обновляются автоматически.

– Для вывода готовых документов в Help&Manual предусмотрено создание шаблонов HTML и PDF документов, при этом также предусмотрены настройки вывода различных видов справок.

– Система поддерживает вывод в наиболее распространенные форматы (HTML Help, Cross-platform Help, Adobe PDF, Visual Studio Help, Winhelp, Word Documents, e-Books, Print manuals) нажатием на 1 кнопку.

Список литературы

1. ГОСТ 34.201–89. Виды, комплектность и обозначение документов при создании автоматизированных систем. – Введ. 01.01.1990. – М.: Межгосударственный совет по стандартизации, метрологии и сертификации, 2001. – 9 с.
2. ГОСТ 34.003–90 Автоматизированные системы. Комплекс стандартов на автоматизированные системы. Термины и определения. – М.: ИПК Изд-во стандартов, 1992.
3. ГОСТ 34.602–89. Техническое задание на создание автоматизированной системы. – Введ. 01.01.1990. – М.: Межгосударственный совет по стандартизации, метрологии и сертификации, 2001. – 15 с.

УДК 004.7

Ибрагимов Байрам,

д.т.н., профессор, Азербайджанский Технический
Университет, Баку, E-mail:i.bayram@mail.ru

Керимов Вусал,

Институт Систем Управления НАН Азербайджана, Баку

ОБ ОДНОМ ПОДХОДЕ К ОЦЕНКЕ РЕСУРСОВ В МУЛЬТИСЕРВИСНЫХ СЕТЯХ СВЯЗИ НА БАЗЕ ЕДИНОГО ИНФОКОММУНИКАЦИОННОГО ПРОСТРАНСТВА

Аннотация: В работе предложена математическая модель в виде системы массового обслуживания общего типа, описывающая мультисервисные сети связи на базе единого инфокоммуникационного пространства, обеспечивающего наращивание единой многооператорской среды для оказания мультимедийных услуг. Установлено, что чем больше отличались по своей пропускной способности минимальный разрез и остальные разрезы мультисервисной сети, тем больший выигрыш по среднему значению времени пребывания потоков пакетов трафика обеспечивало использование критерия оптимальности.

Ключевые слова: ресурсы мультисервисной сети, математическая модель, критерий эффективности, трафик, многооператорская инфраструктура, экономические затраты.

Ibrahimov Bayram,

Doctor of Technical Sciences, Professor, Azerbaijan Technical
University, Baku, E-mail: i.bayram@mail.ru

Kerimov Vusal,

Institute of Control Systems of Azerbaijan National Academy of Sciences, Baku

ABOUT ONE APPROACH TO THE ASSESSMENT OF RESOURCES IN MULTISERVICE COMMUNICATION NETWORKS ON THE BASIS OF A SINGLE INFOCOMMUNICATION SPACE

Abstract: The paper proposes a mathematical model in the form of a general type queuing system that describes multi-service communication networks on the basis of a single information and communication space, providing for building up a single multi-operator

environment for the provision of multimedia services. It has been established that the more the minimum section and the remaining sections of the multiservice network differed in their carrying capacity, the greater the gain in the average value of the residence time of the traffic packet flows was provided by the use of the optimality criterion.

Keywords: *multi-service network resources, mathematical model, efficiency criterion, traffic, multi-operator infrastructure, economic costs.*

Введение

В настоящее время развитие мультисервисных сетей связи на базе архитектурной концепции FN (Future Network) с использованием конвергенции телекоммуникационных технологий требует создания единого инфокоммуникационного пространства, обеспечивающего наращивание глобальной единой многооператорской среды для оказания мультимедийных услуг [1-3].

При построении единой многооператорской инфраструктуры и обслуживания работы телекоммуникационных сетей следует учитывать тот факт, что основную долю нагрузки генерируют мультимедийные приложения, которые состоят из полезных и служебных трафиков. Мультимедийный трафик занимает все большую долю и в трафике операторов связи телекоммуникационных компаний, в том числе ведомственных сетей, при этом основное направление развития которых направлено на создание масштабируемых, мобильных, надежных и безопасных сетей с контекстно-зависимым предоставлением услуги «Triple Play services».

Известно [2, 3, 4], что мультисервисных сетей связи (МСС) представляет собой сложную распределенную систему, управление информационными и сетевыми ресурсами которой требует решения целого спектра разнообразных задач FCAPS (Fault, Configuration, Accounting, Performance and Security). Эти задачи в рекомендациях ITU-T, M.3400 (International Telecommunications Union-Telecommunications) разделены на пять функциональных направлений: управление конфигурацией, неисправностями, расчетами за услуги связи, производительностью и безопасностью.

Из перечисленных выше функциональных задач видно, что особое место занимает управление информационными и сетевыми ресурсами в мультисервисных сетях связи на базе концепции FN с использованием современных ИКТ. Совместное использование современных ИКТ (Информационно-коммуникационных технологий) операторами связи ресурсов в МСС на базе FN позволяет организовать единое инфокоммуникационное пространство и оптимизировать затраты на их создание [5, 6].

На основе исследования установлено [3, 6, 7, 8], что создание единой многооператорской инфраструктуры и создание единого инфокоммуникационного пространства в МСС минимизирует экономические затраты на его эксплуатацию.

В рамках доклада рассматривается один из подходов решения задачи исследования и оценки ресурсов в мультисервисных сетях связи буду-

щего поколения FN на базе современных ИКТ при создании единого инфокоммуникационного пространства (ЕИП).

Постановка задачи и описание единой многооператорской среды

В настоящее время развитие мультисервисных сетей связи на базе концепции FN, увеличение числа мультимедийных приложений и характера нагрузки в современных инфокоммуникационных сетях требует обеспечение качества обслуживания (QoS – Quality of Service) трафиков услуги «Triple Play services», которое является одной из наиболее важных задач для построения единой многооператорской среды.

Показано в [1, 2-4, 6], что перспективной телекоммуникационной технологией для создания единого инфокоммуникационного пространства предлагается использование современных ИКТ, таких как подсистема IMS (IP multimedia subsystem), программно-конфигурируемая сеть (ПКС) SDN (Software Defined Network) и, так называемая NFV (Network Functions Virtualization), которые являются основой современных ИКТ для создания единой многооператорской инфраструктуры.

В мультисервисных сетях связи IMS-архитектура с использованием сервера домашних абонентов HSS (Home Subscriber Server) представляет собой набор функций, соединенных стандартными интерфейсами, и обеспечивает унифицированный доступ к приложениям независимо от технологии сети доступа, а также решает вопросы безопасности доступа со стороны терминальных устройств и эффективно управляет профилями абонентов [3, 5, 6]. В архитектуре IMS используется основной протокол сигнализации SIP. В частности, при наличии надежной транспортной сети достаточно одного логического ядра IMS на всю сеть связи.

Функциональные компоненты IMS могут быть территориально-распределенными в зависимости от точек концентрации пользовательского трафика и вычислительных ресурсов ЦОДов (Центр обработки данных). Такая возможность обеспечивает гибкий подход к созданию /модернизации инфраструктуры мультисервисных сетей связи [3, 4].

Второй важной технологией является программно-определяемая сеть SDN [4], которая обеспечивает администрирование МСС, позволяющее управлять услугами сети, когда функционал управления (Control Plane) отделен (абстрагирован) от нижележащего уровня пересылки пакетов (Data Plane). Планирование сети и управление служебным и полезным трафиком происходит программным путем. Для приложений верхнего уровня предоставляются интерфейсы прикладного программирования API (Application Programming Interface).

С целью создания единого инфокоммуникационного пространства в МСС целесообразно использовать третью важную технологию NFV – виртуализацию функции сети, представляющую собой фундаментальный сдвиг в сетевой модели [6]. Эта технология предполагает реализацию функций управления сетями и предоставление услуг в виде про-

грамм, а не специализированного оборудования.

Следовательно, как SDN, так и NFV используют облачные и интернет-технологии. Здесь SDN позволяет конфигурировать плоскость передачи данных программным способом, а NFV, также программным образом, позволяет задавать роли виртуальных сетевых устройств [3-8].

Учитывая вышеизложенные концепции для создания единой многооператорской среды мультисервисных сетей связи будущего поколения, последние представляют собой систему массового обслуживания (СМО) с некоторыми допущениями при критической загрузке, $i = \overline{1, K}$, где K – число источников нагрузки от различных операторов связи телекоммуникационных компании при установлении соединения и при оказании услуги «Triple Play services».

Данная СМО представляет собой единую многооператорскую среду, описывающую процессы передачи и обработки потоков пакетов трафика с использованием сетевых и канальных ресурсов мультисервисных сетей связи будущего поколения на базе архитектурной концепции подсистемы IMS, программно-конфигурируемых сетей SDN и NFV.

Системно-технический анализ показывает, что исследуемая СМО общего типа часто характеризуется двумя важными показателями [9-11]:

- с точки зрения пользователя, передающего потоки пакетов трафиков для обслуживания системы, важно минимизировать средние значения потери пакетов в связи с ожиданием начала обслуживания трафиков:

$$E[P_m(\rho_i)] \rightarrow \min_i P_m^{don.}(\rho_i), \quad i = \overline{1, K}, \quad (1)$$

где $E[P_m(\rho_i)]$ – среднее значение вероятности потерь пакетов при обслуживании i -го трафика в зависимости от критической загрузки ρ_i системы, $i = \overline{1, K}$; $P_m^{don.}(\rho_i)$ – допустимое значение потери пакетов в связи с ожиданием начала обслуживания трафика, $i = \overline{1, K}$.

- с точки зрения эксплуатации СМО желательно сделать число коммутаторов, каналов связи, аппаратно-программных средств, контроллеров и маршрутизаторов как можно меньшим от допустимого:

$$N_k(\rho_i) < N_k^{don.}(\rho_i) \rightarrow \min, \quad i = \overline{1, K} \quad (2)$$

где $N_k(\rho_i)$ – число абонентских и сетевых терминалов от начала до конца трактов систем передачи при обслуживании i -го трафика в зависимости от критической загрузки ρ_i системы, $i = \overline{1, K}$; $N_k^{don.}(\rho_i)$ – допустимое минимальное значение числа абонентских и сетевых терминалов от начала до конца трактов систем передачи при обслуживании i -го трафика.

Из выражений (1) и (2) видно, что эти два показателя в СМО являются противоположными действующим экономическим факторам, которые можно учесть для решения задач минимизации функционала следующим образом:

$$Q_m = W[\inf_i [A(T_{i.c.ож}, C_i^o, P_{i,m})]], \quad i = \overline{1, K}, \quad (3)$$

где $T_{i.c.ож}$ – среднее значение времени ожидания потоков пакетов i -го трафика; C_i^o – отношение стоимости эксплуатации число абонентских и сетевых терминалов от начало до конца трактов систем передачи при обслуживании i -го трафика и при критической загрузке ρ_i системы в единицу времени к убытку от ожидания потоков пакетов трафика за единицу времени; $P_{i.mn}$ – вероятность потерь пакетов при передаче i -го трафика.

На основе (3) в мультисервисных сетях связи будущего поколения FN, использующих современные ИКТ, существуют практически такие постановки задач, если $C_i^o \rightarrow \max C_i^o$, то оптимальное число $N_k(\rho_i)$ достаточно $N_k^{don}(\rho_i) \rightarrow \min$ мало, но все же таково, что загрузка меньше

$$\rho_i = (L_n \cdot \lambda_i / V_{i,k}) < 1, \quad i = \overline{1, K}, \quad (4)$$

где L_n – средняя длина передаваемого пакета трафика; $V_{i,k}$ – скорость работы системы передачи потоков пакетов i -го трафика с использованием пользовательских, канальных и сетевых терминалов, $i = \overline{1, K}$; λ_i – скорость поступления входящего потока пакетов i -го трафика.

Последнее выражение, определяющее коэффициент загрузки мультисервисных сетей связи при обслуживании потоков пакетов i -го трафика, играет важную роль в теории очередей [9].

Выражения (1), (2), (3) и (4) характеризуют эффективность создания единого инфокоммуникационного пространства на основе мультисервисных сетей связи будущего поколения FN, использующих современные ИКТ.

Поэтому СМО с малой загрузкой $\rho_i < 1$ весьма актуальна для исследования нового подхода к оценке ресурсов мультисервисных сетей связи будущего поколения FN на базе современных ИКТ [11].

Создание нового подхода к оценке ресурсов мультисервисных сетей связи на основе архитектурной концепции FN

Исследования показали [3, 5, 8], что создание ЕИП очень важно для операторов связи при оказании широкого набора мультимедийных услуг и способствует эффективности использования информационных и сетевых ресурсов единой многооператорской среды, которая удовлетворит потребностям стационарных и подвижных, корпоративных и ведомственных, государственных и частных пользователей услуг связи.

С учетом управления ресурсами в едином инфокоммуникационном пространстве сформулируем задачу распределения и эффективного использования пропускных способностей мультисервисных сетей связи будущего поколения в наиболее общем виде, следуя [3, 6, 11]. При этом необходимо минимизировать среднее пребывание трафиков для одного типа потока пакетов и определить набор оптимальных пропускных способностей сетей ПКС, коммутаторов ФСВ и ядра-IMS при условии, что суммарная пропускная способность ограничена. Здесь ограничения на суммарную пропускную способность означают, что заданы ограничения

на суммарную стоимость оборудования сетей ПКС, платформы IMS и системы NFV.

Для решения выше поставленной задачи и учитывая важность взаимодействия программно-аппаратных средств, оборудования терминальных, сетевых и канальных сетей ПКС, платформы IMS и системы NFV предлагаются математические модели, основанные на упрощениях в описании исследуемой концепции единого инфокоммуникационного пространства.

Математическая формулировка задачи предлагаемого нового подхода к оценке ресурсов программно-аппаратных средств мультисервисных сетей связи на основе архитектурной концепции FN описывается следующими целевыми функциями:

$$W_{эф} = W[\text{Arg min}_i(T_{i,yc})], i = \overline{1, K}, \quad (5)$$

при следующих ограничениях

$$N_k \leq N_{k,доп.}, T_{i,ср.з} \leq T_{i,ср.з,доп.}, C_{i,макс} \leq C_{i,макс,доп.}, \eta_i \leq \eta_{i,доп.}, C_{i,ап} \leq C_{i,ап,доп.}, i = \overline{1, K} \quad (6)$$

где $T_{i,yc}$ – среднее время установления соединения при оказании i -го мультимедийного услуга, $i = \overline{1, K}$; $C_{i,ап}$ – стоимость аппаратных и программных средств сетей ПКС, платформы IMS и системы NFV, $i = \overline{1, K}$; $T_{i,ср.з}$ – среднее время задержки при передаче i -го потока пакетов, $i = \overline{1, K}$; η_i – коэффициент эффективного использования информационных и сетевых ресурсов в единой многооператорской среде при передаче i -го потока пакетов, N_k – общее число используемых в системе абонентских, канальных и сетевых терминалов; $C_{i,ап,доп.}$, $T_{i,ср.з,доп.}$, $N_{k,доп.}$, $C_{i,макс,доп.}$ и $\eta_{i,доп.}$ – соответственно допустимое значение величины стоимости аппаратных и программных средств, среднее время задержки, число обслуживающих терминалов и коэффициент эффективного использования информационных и сетевых ресурсов единой многооператорской среды при передаче i -го потока пакетов полезного и служебного трафиков, $i = \overline{1, K}$.

Выражения (5) и (6) определяют сущность рассматриваемого нового подхода, на основе которого предлагается математическая модель (ММ) для оценки вероятностно-временных характеристик мультисервисных сетей связи, использующих концепции единого инфокоммуникационного пространства при оказании мультимедийных услуг.

Кроме того, (5) и (6) описывают особенности программно-аппаратных средств сетей ПКС, платформы IMS и технологий ВСФ, которые позволяют более точно учесть телекоммуникационные процессы, протекающие в исследуемой единой многооператорской инфраструктуре МСС будущего поколения FN.

Построения ММ с использованием архитектурной концепции сетей SDN, платформы IMS и технологий VNF

Проведенные исследования показали [1, 2, 3, 5, 8, 11], что использование вышеуказанных решений, как IMS и SDN, так и NFV в качестве инфраструктуры МСС будущего поколения, позволяет решить одну из важных телекоммуникационных задач FCAPS – создание единого инфокоммуникационного пространства, обеспечивающего эффективность использования доступных сетевых и информационных ресурсов при оказании услуг доставки информации пользователей.

В данном случае, сетей ПКС в основном состоит из коммутатора, контроллера и программного коммутатора типа Open vSwitches с использованием протоколов OpenFlow. В данной инфраструктуре технические средства, создающие единую многооператорскую среду, имеют дополнительный сервер с производительными вычислительными ресурсами.

В сетях ПКС для оказания услуги «Triple Play services» в буферную накопителя коммутатор-контроллера поступают потоки пакетов полезного и служебного трафиков.

В исследуемой системе с целью анализа процессов обработки и передачи трафиков построена математическая модель в виде СМО общего типа $M / M / N_k / R_{\text{он}} \leq \infty / K$, учитывающая характер потока пакета трафика, поступающего в ядро ПКС сетей, IMS и NFV. Здесь, ММ учитывает эффективность управления информационными и сетевыми ресурсами в МСС будущего поколения.

Предполагаем, что на буферный накопитель коммутатора сетей ПКС поступают стационарные гетерогенные пуассоновские потоки пакетов трафиков с параметрами λ_i , $i = \overline{1, K}$, создаваемые различными типами источников нагрузки (голос, факс, Internet, данные, видео и др. Однако, в частном случае делим нагрузки на две большие группы – полезные и служебные нагрузки).

На основе механизмов работы логических функций ядра IMS, сетей ПКС и ВСФ реализуется обслуживание гетерогенных полезных $\lambda_{i,n}$ и служебных трафиков $\lambda_{i,c}$, благодаря управлению ресурсами в мультисервисных сетях связи.

Таким образом, исследуемая система на основе алгоритмов работы функциональной архитектуры единого инфокоммуникационного пространства в общем случае представляет собой СМО общего типа $M / M / N_k / R_{\text{он}} \leq \infty / K$, которая базируется на архитектурной концепции сетей SDN, платформы IMS и технологий VNF.

На основе СМО общего типа $M / M / N_k / R_{\text{он}} \leq \infty / K$ коэффициент загрузки при обслуживании потоков пакетов i -го трафика выражается следующим образом [11]:

$$\rho_i = (\lambda_i / \mu_i \cdot N_k) < 1, \quad i = 1, 2, \dots, K. \quad (7)$$

Выражения (4) и (7) определяют, что стационарное распределение числа пакетов трафика в рассматриваемой СМО существует и гарантирующие условия выполняются.

Предполагаем, что на буферный накопитель коммутатора сети ПКС $R_{\text{он}}$ поступают ($i = 1, 2, \dots, K$) K – одинаково распределенных гетерогенных пуассоновских потоков пакетов трафика с интенсивностью λ , имеющих правила маршрутизации. Имеют прямой доступ к обслуживанию N_k (коммутатор-контроллер) с помощью коммутатор-контроллера сети ПКС, причем одинаковых обслуживающих программным коммутатором типа Open vSwitches с использованием протоколов OpenFlow с экспоненциальным законом распределением. Распределение длительности обслуживания потоков пакетов в этом случае имеет вид:

$$F(t, L_n) = (1/L_n) \cdot C_{\text{max}} \cdot \exp(-t \cdot C_{\text{max}} / L_n), \quad (8)$$

где L_n – средняя длина передаваемого пакета трафика в битах, $L_n = C_{\text{max}}/\mu$; μ – средняя длительность обслуживания пакетов трафика; C_{max} – максимальное значение пропускной способности коммутатор-контроллера сети ПКС.

Из (8) следует, что при воздействии пуассоновских потоков пакетов трафика и обслуживания на систему процесс, протекающий в ней, будет марковским ординарным с непрерывным временем, которое имеет K фазовых состояний.

В каждый момент времени K система имеет распределение вероятностей K фазовых состояний $P^T(K)$, причем $\sum P_{i,K} = 1$. Система начинает движение из начального состояния $K=0$ с распределением $P^T(0)$. Здесь данная последовательность задается не вероятностью $P^T(K)$, а условной вероятностью перехода $\pi_{ij}(K, K+1)$.

На основе ММ в стационарном состоянии вектор финальных вероятностей $p^T = [p_0, p_1, \dots, p_K]$ находится из векторно-матричного уравнения [9, 10]:

$$P^T(\pi - I) = 0, \quad (9)$$

где I – единичная диагональная матрица.

Здесь выражение (9) является алгебраический вид как результат перемножения вектор-строки финальных вероятностей на матрицу переходов и определяется следующим образом [10]:

$$p_j(\pi_{ij} - 1) = \sum_{i=0}^K p_i \cdot \pi_{ij} \cdot \delta_{ij}, \quad (10)$$

где δ_{ij} – символ Кронекера.

Полученные аналитические выражения (9) и (10) описывают поведение многолинейных СМО $M/M/N_k/R_{\text{он}} \leq \infty/K$ в переходном режиме.

Учитывая условную вероятность перехода и особенности многолинейных СМО переходим к оценке показателя системы.

Оценка показателя математической модели на основе многолинейных СМО

Известно [1-3, 10], что исследуемой математической моделью на основе СМО общего типа являются $M/M/N_k/R_{\text{он}} \leq \infty/K$ многолиней-

ные системы обслуживания с пуассоновскими входящими K потоками (K -входа) и потоками обслуживания N_k – коммутатор-контроллерами, использующими протоколы OpenFlow. В данной системе каждый коммутатор-контроллер сетей ПКС имеет в буферном накопителе $R_{\text{бу}}$ свою секцию памяти и $N_k \leq R_{\text{бу}} \leq K$. Здесь при конечном $R_{\text{бу}}$ система имеет $R_{\text{бу}} + 1$ состояние, включая состояние пустого буферного накопителя с вероятностью p_0 . Тогда системы можно отождествить с состоянием процесса i с длиной очереди q .

Для оценки показателей ВВХ сетей ПКС на основе предлагаемой ММ с использованием многолинейных СМО $M / M / N_k / R_{\text{бу}} \leq \infty / K$ необходимо определить компоненты вектора финальных вероятностей $P(q)$ с учетом длины очереди q в системе обслуживания.

Решая систему разностных уравнений для многолинейных СМО типа $M / M / N_k / R_{\text{бу}} \leq \infty / K$ получаем компоненты вектора финальных вероятностей $P(q)$ в следующем виде [9]:

• В первом случае в системе длина очереди начиная от 0 увеличивается [9, 10]:

$$P(q) = p_0 \cdot \rho^q \cdot \frac{K!}{q!(K-q)!} \quad \text{при } 0 \leq q \leq N_k \quad (11)$$

где p_0 –вероятность пустого буферного накопителя коммутатор-контроллера сетей ПКС.

Допустим, что $q=K$. Это означает, что многолинейные СМО, работающие в мультисервисных сетях связи с использованием единой многооператорской среды, имеют емкость буфера $R_{\text{бу}}$, достаточную для размещения K потоков пакетов трафика. Последние реализованы благодаря эффективному использованию буферных ресурсов сетей ПКС, которые представляют собой организованные на коммутатор-контроллерах и программном коммутаторе типа Open vSwitches очереди потоков пакетов трафика.

Тогда выражение (11) примет следующий компактный вид:

$$P(q) = p_0 \cdot \left[\frac{\lambda}{C_{\text{max}} \cdot N_k} \cdot L_n \right]^q, \quad (12)$$

Выражение (12) характеризует эффективность управления ресурсами пропускной способности сетей ПКС и определяет их информационную эффективность.

• Во втором случае в системе длина очереди изменяется до максимума емкости буферного накопителя $R_{\text{бу}}$:

$$P(q) = p_0 \cdot \rho^q \cdot \frac{K!}{(K-q)!} \cdot \frac{(N_k)^{N_k-q}}{N_k!} \quad \text{при } N_k \leq q \leq R_{\text{бу}} \quad (13)$$

На основе предлагаемой ММ, полученные аналитические выражения позволяют оценить вероятностно-временные характеристики мульти-

сервисных сетей связи на базе единого инфокоммуникационного пространства, обеспечивающего эффективность использования доступных сетевых и информационных ресурсов.

Оценка вероятностно-временных характеристик СМО

Учитывая, что обслуживаемые пакеты состоит из полезного L_{nt} и служебного L_{ct} трафиков, тогда среднее значение времени обслуживания потоков пакетов трафика определяется следующим образом [11]:

$$E[\mu, L_n] = (L_{nt} + L_{ct}) / C_{\max} \quad (14)$$

На основе (7) и (12) можно определить для СМО общего типа $M/M/N_k/R_{\text{бу}} \leq \infty/K$ загрузки по всей мультисервисной сети, характеризующие эффективность использования пропускной способности единой многооператорской среда и выражается как

$$\rho = \sum_{i=1}^K \lambda_i \cdot \frac{L_{i,n}}{N_k} \cdot C_{\max}^{-1} < 1, \quad i = \overline{1, K} \quad (15)$$

Выражения (14) и (15) определяют сетевые ресурсы сети ПКС и характеризуют информационную эффективность единого инфокоммуникационного пространства мультисервисных сетей связи будущих поколений FN.

Следует отметить, что сетевые и информационные ресурсы для создания единой многооператорской среды МСС будущего поколения и единого инфокоммуникационного пространства являются одним из важных показателей. Среди этих показателей особое место занимают средние значения времени пребывания потоков пакетов трафика в режиме «End to end»:

$$E[T_{cbn}(\lambda)] = \frac{\rho}{C_{\max}} + \frac{\lambda}{N_k \cdot (1 - p_0)} \cdot \frac{L_n}{\mu(1 - \rho)} \cdot C_{\max}^{-1}, \quad \rho < 1 \quad (16)$$

Выражение (16) характеризует эффективность использования ресурсов пропускной способности в мультисервисных сетях связи на базе единой многооператорской среды и определяет среднее значение времени установления соединения $E[T_{yc}]$ при оказании мультимедийных услуг «Triple Play services».

Таким образом, на основе ММ получены аналитические выражения (10), ..., (16) которые характеризуют вероятностно-временные характеристики и являются показателем QoS мультисервисных сетей связи на базе единой многооператорской среды при оказании мультимедийных услуг.

Заключение

1. В результате исследования предложена ММ в виде СМО общего типа $M/M/N_k/R_{\text{бу}} \leq \infty/K$, описывающая мультисервисные сети связи на

базе единого инфокоммуникационного пространства, обеспечивающего наращивание единой многооператорской среды для оказания мультимедийных услуг.

2. Установлено, что чем больше отличались по своей пропускной способности минимальный разрез и остальные разрезы мультисервисной сети, тем больший выигрыш по среднему значению времени пребывания потоков пакетов трафика обеспечивало использование критерия оптимальности (5) и (6).

3. На основе предложенной ММ получены аналитические выражения, позволяющие оценить показатели качества обслуживания QoS трафика и анализировать вероятностно-временные характеристики мультисервисных сетей связи будущего поколения при оказании мультимедийных услуг.

Литература

1. Росляков А.В., Ваняшин С.В. Будущие сети (Future Networks). Самара. ПГУ-ТИ. 2015. – 274 с.
2. Ефимушкин В.А., Ледовских Т.В., Щербакова Е.Н. Инфокоммуникационное технологическое пространство цифровой экономику // Т-COMM: Телекоммуникации и транспорт. – 2017. -Т.11. -№5. – с.15-20.
3. Ибрагимов Б.Г., Гумбатов Р.Т., Ибрагимов Р.Ф. Исследование и управление ресурсами в мультисервисных NGN/IMS сетях телекоммуникаций// Сборник трудов XII-Международная НТК «Технологии информационного общества». МТУСИ. 2018. – 44 – 46 с.
4. Логинов С.С. Об уровнях управления в программно конфигурируемой сети (SDN) // T-Comm: Телекоммуникации и транспорт. 2017. Том 11. №3. – с. 50-55.
5. Волкова С.В. Использование современных ИКТ для создания единого инфокоммуникационного пространства // Электросвязь, №12. Москва. 2017. – с.98 – 99.
6. Ибрагимов Б.Г., Керимов В.Р. Эффективность использования ресурсов в мультисервисных сетях связи на базе единого инфокоммуникационного пространства // Сборник трудов XII-Международная НТК «Технологии информационного общества». МТУСИ. 2018. – 47-49 с.
7. Олифер В., Олифер Н. Компьютерные сети. Принципы, технология, протоколы. Учебник для вузов. 5-е изд. – СПб.: Питер. 2016. – 992 с.
8. Chowdhury N. M. K. and Boutaba R. A Survey of Network Virtualization. Computer Networks, vol. 54, no. 5, 2010. – pp.862-876.
9. Карташевский В.Г. Основы теории массового обслуживания. – М.: Горячая линия – Телеком, 2013. – 130 с.
10. Крылов В.В., Самохвалова С.С. Теория телетрафика и ее приложения. – СПб.:БХВ – Петербург. – 2005. – 288 с.
11. Ibrahimov B.G., Ismaylova S.R. The Effectiveness NGN/IMS Networks in the Establishment of a Multimedia Session //American Journal of Networks and Communications. Vol. 7, No. 1. 2018. – pp.1-5.

УДК 004.853

В.Р. Кристалинский,

кандидат физ.-мат. наук, доцент, Смоленский государственный университет,
г. Смоленск, Российская Федерация, e-mail: kristvr@rambler.ru

А. Н. Колюсь,

магистрант, Смоленский государственный университет, г. Смоленск,
Российская Федерация, e-mail: kolyus@mail.ru

ОБ ИСПОЛЬЗОВАНИИ АССОЦИАТИВНЫХ ПРАВИЛ ДЛЯ НЕЧЕТКОГО МОДЕЛИРОВАНИЯ

Аннотация: В работе показано, что ассоциативные правила есть другое представление нечетких продукционных правил, которые могут использоваться для прогнозирования взаимосвязей между объектами. Нечеткий характер полученных правил связан с принципиальной непредсказуемостью поведения интересующих нас объектов, с тем, что мы не можем достоверно сказать, будет ли данный набор данных реализован в каждом конкретном случае. Для математического описания процесса выбора ассоциативных правил и принятия решений на их основе предлагается использовать аппарат нечеткой логики и теории нечетких множеств.

Ключевые слова: ассоциативные правила, теория нечетких множеств, нечеткая логика, нечеткое моделирование.

V.R. Kristalinsky

Candidate Phys.-Mat. Sci., Smolensk State University, Smolensk, Russian Federation,
e-mail: kristvr@rambler.ru

A. N. Kolaus,

Undergraduate, Smolensk State University, Smolensk, Russian Federation,
e-mail: kolyus@mail.ru

ON THE USE OF ASSOCIATIVE RULES FOR FUZZY MODELING

Abstract: The paper shows that association rules are another representation of fuzzy production rules that can be used to predict the relationships between objects. The fuzzy nature of the rules obtained is related to the fundamental unpredictability of the behavior of the objects of interest to us, to the fact that we cannot reliably say whether this data set will be implemented in each particular case. For a mathematical description of the process of choosing associative rules and making decisions based on them, it is proposed to use the apparatus of non-clear logic and the theory of fuzzy sets.

Keywords: association rules, theory of fuzzy sets, fuzzy logic, fuzzy modeling.

Введение

В настоящее время стремительно возрастает интерес соответствующих специалистов к нахождению закономерностей данных в больших базах данных. Особенно это наблюдается в прогнозировании поведения покупателей для повышения прибыли предприятий. В результате чего возникает потребность обработки больших массивов данных для извлечения закономерностей – ассоциативных правил. Для анализа данных широко используются интеллектуальные системы, в частности, нечеткие модели. Достигнутые в последнее время впечатляющие успехи в применении нечетких технологий для решения самых разнообразных

практических задач позволяют утверждать, что нечеткое моделирование различных сложных систем – эффективная альтернатива традиционным моделям и методам.

Целью данной статьи является показать возможность интерпретации утверждений, описывающих ассоциативные правила, с нечеткими высказываниями в том смысле, в котором принято рассматривать это понятие в нечеткой логике.

Для достижения поставленной цели необходимо выполнить задачи:

- изучить литературу по нечеткому моделированию и ассоциативным правилам;
- провести анализ полученных знаний;
- сделать выводы об использовании нечеткого моделирования для описания ассоциативных правил.

Результаты и обсуждение

Пусть $I = \{i_1, i_2, \dots, i_n\}$ – множество всех возможных значений объектов, входящих в транзакции; $D = \{d_1, d_2, \dots, d_N\}$, $d_j \subset I$ – множество, элементами которого являются транзакции, содержащие наборы объектов из множества I . Тогда ассоциативным правилом является следующая импликация: $X \rightarrow Y$, где $X \subset I$, $Y \subset I$ и $X \cap Y = \emptyset$. Другими словами (см., например, [1]), ассоциативные правила – это закономерности, описывающие связь между событиями, описанными в виде базы данных. Часто такими базами являются данные о покупке товаров в крупных супермаркетах. В таком случае суть ассоциативных правил заключается в анализе потребительской корзины. Например, при покупке товаров нередко складывается ситуация, что покупка одних товаров с большой долей уверенности влечет за собой покупку других товаров. Такого рода связи могут быть выявлены на основе анализа больших данных. Но на данный момент применения ассоциативных правил далеко не ограничено анализом потребительской корзины, их использование широко распространено в различных областях. Выявление ассоциативных правил может быть использовано для организации торговли, исследования покупательского поведения, организации каталогов, выбора стратегии рекламных кампаний и т.д.

Существуют различные алгоритмы поиска ассоциативных правил. Это алгоритмы Apriori [2], ДНР [3], Partition [4] и другие. Работа всех этих алгоритмов основана на статистическом анализе базы данных в поиске часто встречающихся предметных наборов. Главными характеристиками в ассоциативных правилах являются уровень поддержки (support) и достоверности (confidence).

Нетрудно понять, что утверждения, описывающие ассоциативные правила (например «если куплены фасоль и кукуруза, то будут куплены и помидоры») являются нечеткими высказываниями в том смысле, в каком это понятие рассматривается в нечеткой логике. А имен-

но, нечетким высказыванием называется высказывание, которому можно придать степень истинности, лежащую на отрезке $[0, 1]$ (см. [5]). Такие высказывания часто используются для описания утверждений, которые характеризуются неопределенностью. Но именно это и можно сказать об ассоциативных правилах. Действительно, указанное нами правило не является ни однозначно истинным, ни однозначно ложным, а выполняется с той степенью истинности, которая описывается степенями поддержки и достоверности правила. Таким образом, для описания ассоциативных правил можно, на наш взгляд, использовать хорошо известные методы нечеткой логики, теории нечетких множеств и нечеткого вывода. Например, если степень достоверности указанного в примере нечеткого правила равна 75%, то можно сказать, что мы имеем связанное с ним нечеткое высказывание, степень истинности которого равна 0,75. Более того, на наш взгляд можно приписать эту степень истинности данному нечеткому высказыванию «вообще», уже вне связи с конкретной базой данных. Построенная на этой основе система нечеткого вывода может использоваться для прогнозирования поведения объектов исследования в аналогичных ситуациях.

Системы нечеткого вывода и системы нечетких продукций (см. [6]) широко применяются для прогнозирования в различных сферах человеческой деятельности. Основой системы нечетких продукций является множество правил, описывающих взаимосвязь между входными и выходными переменными. С помощью этой системы правил на основе известных алгоритмов нечеткого вывода на основе степеней истинности входных переменных получают значения истинности выходных переменных, причем сами правила также имеют определенные степени истинности. При построении системы нечетких продукций ключевым моментом является именно построение множества правил. Часто это делается с привлечением квалифицированных экспертов в данной предметной области. На наш взгляд, для этой цели могут использоваться ассоциативные правила, описание которых формализовано на языке нечеткой логики.

Пусть у нас имеется ассоциативное правило {спаржа, фасоль} \rightarrow {кабачки} со степенью поддержки 67 %. На его основе можно построить правило для системы нечетких продукций: «Если покупатель покупает спаржу И покупатель покупает фасоль ТО покупатель покупает кабачки» со степенью истинности 0,67. Используя какой-либо алгоритм поиска ассоциативных правил, например, алгоритм Априори, мы можем на основе всех полученных правил построить систему нечетких продукций и определить степени истинности правил этой системы на основе степеней поддержки ассоциативных правил. Затем эта система может быть использована для прогнозирования значения истинности заключе-

ния для каких-то других наборов данных. При этом частота появления групп данных-условий будет не чем иным, как степенью истинности посылки правила нечетких продукций. Этот подход можно использовать и для прогнозирования поведения наборов данных в будущем. Например, предположив, что степень истинности утверждения о появлении набора-условия {спаржа, фасоль} составляет 0,6, на основе правила получения заключения для импликации Мамдани получаем, что степень истинности утверждения, что в этом случае появится объект {кабачки} составит $\min\{0,67; 0,6\}=0,6$. В случае, если одно и то же следствие появляется в нескольких правилах, для получения значения истинности результата можно использовать нечеткое ИЛИ. Как известно, одним из способов вычисления значения истинности для операции ИЛИ является определение максимума из значений истинности операндов. Например, если наряду с рассмотренным правилом имеется правило {огурцы, морковь} \rightarrow {кабачки} со степенью истинности 0,7 и степень истинности утверждения «появился набор {огурцы, морковь}» равна 0,62, то мы заключаем, что для данного правила заключение получается со степенью истинности $\min\{0,62; 0,7\}=0,62$. А на основании двух правил получаем, что степень истинности заключения «кабачки» составляет $\max\{0,6; 0,62\}=0,62$.

Ясно, что этот подход может широко использоваться в маркетинговых исследованиях. Построение системы правил для нечетких продукций на основе ассоциативных правил является хорошей альтернативой экспертному методу. В случае, если мы используем для построения правил экспертов, мы зачастую не можем сказать, на основании чего эксперты делают свои заключения. Экспертный подход работает как «черный ящик»: мы не можем сказать, откуда результат, который получен на выходе. В случае использования ассоциативных правил ответ на вопрос о происхождении правил ясен: они получены на основе известных и апробированных алгоритмов теории ассоциативных правил.

Заключение

Таким образом, ассоциативные правила есть не что иное, как другое представление нечетких продукционных правил, которые могут использоваться для прогнозирования взаимосвязей между объектами. Нечеткий характер полученных правил связан с принципиальной непредсказуемостью поведения интересующих нас объектов, с тем, что мы не можем достоверно сказать, будет ли данный набор данных реализован в каждом конкретном случае. Для математического описания процесса выбора ассоциативных правил и принятия решений на их основе можно использовать аппарат нечеткой логики и теории нечетких множеств. Это позволит сделать теорию ассоциативных правил более обоснованной и поставить ее на прочный математический фундамент.

Список литературы

1. R. Agrawal, T. Imielinski, A. Swami. 1993. Mining Associations between Sets of Items in Massive Databases. In Proc. of the 1993 ACM-SIGMOD Int'l Conf. on Management of Data, 207-216.
2. R. Agrawal, R. Srikant. "Fast Discovery of Association Rules", In Proc. of the 20th International Conference on VLDB, Santiago, Chile, September 1994.
3. A. Savasere, E. Omiecinski, and S. Navathe, "An Efficient Algorithm for Mining Association Rules in Large Databases", In Proc. 21st Int'l Conf. Very Large Data Bases, Morgan Kaufmann, San Francisco, 1995.
4. J.S. Park, M.-S. Chen, and S.Y. Philip, "An Effective HashBased Algorithm for Mining Association Rules", In Proc. ACM SIGMOD Int'l Conf. Management of Data, ACM Press, New York, 1995.
5. Пегат А. Нечеткое моделирование и управление / А. Пегат ; пер. с англ. – 2-е изд. – М.: БИНОМ. Лаборатория знаний, 2013. – 798 с.: ил. – (Адаптивные и интеллектуальные системы).
6. Леоненков А. В. Нечеткое моделирование в среде MatLab и fuzzyTECH. СПб:БХВ-Петербург, 2003.

УДК 004.519.7

Левин В.И.,

д.т.н., проф.

Пензенский государственный технологический университет,
Россия, Пенза, vilevin@mail.ru

ИНТЕРВАЛЬНАЯ МАТЕМАТИКА И ВЫЧИСЛЕНИЯ В УСЛОВИЯХ НЕОПРЕДЕЛЕННОСТИ

Аннотация: В работе предлагается подход к решению неопределенных систем, в рамках которого параметры системы предварительно не детерминируются, а решение нужной задачи проводится в ее «естественной» форме, т.е. с учетом полного множества всех возможных значений недетерминированных параметров системы. Его достоинство – ориентировка получаемого решения на все множество возможных значений параметров изучаемой системы внутри их областей неопределенности, а недостаток – несколько более сложная интерпретация решения.

Ключевые слова: неопределенные системы, интервальная функция, ширина интервала, интервальное число.

Levin V.I.,

Doctor of Technical Sciences, prof.

Penza State Technological University, Russia, Penza, vilevin@mail.ru

INTERVAL MATHEMATICS AND CALCULATIONS IN UNCERTAINTY CONDITIONS

Abstract: The paper proposes an approach to solving indefinite systems, within which the parameters of the system are not previously deterministic, and the solution of the desired problem is carried out in its "natural" form, i.e. taking into account the full set of all possible

values of non-deterministic system parameters. Its advantage is the orientation of the obtained solution to the whole set of possible values of the parameters of the system under study within their areas of uncertainty, and the disadvantage is a somewhat more complex interpretation of the solution.

Keywords: *indefinite systems, interval function, interval width, interval number.*

В настоящее время наука накопила богатый опыт исследования различных систем с полностью определенными (детерминированными) параметрами. Соответствующие задачи обычно формируются как задачи расчета, анализа и синтеза тех или иных функций с детерминированными параметрами – характеристик изучаемых систем.

Однако на практике все чаще встречаются системы с неполностью определенными (недетерминированными) параметрами. Причины этого таковы: 1) естественная неопределенность, свойственная многим реальным процессам и системам; 2) неточное задание параметров большинства систем из-за погрешностей вычислений или измерений; 3) возникающая необходимость совместного исследования семейства однотипных систем, имеющих одинаковые функции-характеристики и различающихся лишь значениями параметров этих функций, 4) изменение во времени параметров систем.

Исследование неопределенных систем формируется в виде задач расчета, анализа и синтеза характеристик этих систем, имеющих вид тех или иных функций с недетерминированными параметрами – случайными, нечеткими, интервальными и т.д. Все эти задачи сложнее их детерминированных аналогов, что связано со значительным усложнением алгебраических операций над числами при переходе в область недетерминированных чисел.

В настоящее время наиболее распространены три различных подхода к исследованию неопределенных систем – детерминированный, вероятностный и нечеткий. Первый состоит в решении поставленной задачи для определенных значений параметров системы, взятых внутри соответствующих заданных областей неопределенности. Так, можно выбрать наихудшее значение или сочетание значений параметров (пессимистический подход), их наилучшее значений или сочетание значений (оптимистический подход) и др. Достоинство этого подхода – простота интерпретации получаемого решения (так, при pessimistic подходе получается наихудшее возможное решение); недостаток – ориентировка на какое-то одно определенное (чаще всего экстремальное) значение или сочетание значений параметров системы, которое на практике реализуется очень редко, что может обернуться неоправданной сложностью решения. Второй подход состоит в решении задачи для усредненных (ожидаемых) значений параметров системы, что предполагает задание вероятностных распределений этих параметров внутри соответствующих областей неопределенности. Достоинство подхода – ориен-

тировка получаемого решения хотя и на одно, но зато наиболее часто реализующееся значение или сочетание значений параметров системы, недостаток подхода – необходимость знания вероятностных распределений параметров системы, что далеко не всегда возможно. Третий подход идейно близок второму, но отличается тем, что вместо вероятностных распределений параметров изучаемой системы, являющихся объективными характеристиками значений этих параметров, используются нечеткие распределения параметров системы, получаемые экспертным путем.

Все три изложенных подхода объединяет предварительная «детерминизация» параметров системы, выполняемая перед решением поставленной задачи. Однако возможен и принципиально иной, четвертый подход, когда параметры системы предварительно не детерминируются, а решение нужной задачи проводится в ее «естественной» форме, т.е. с учетом полного множества всех возможных значений недетерминированных параметров системы. Именно такой подход излагается в данной работе. Его достоинство – ориентировка получаемого решения на все множество возможных значений параметров изучаемой системы внутри их областей неопределенности; недостаток – несколько более сложная интерпретация решения. При изложении данного подхода будем считать параметры всех функций – характеристик изучаемых систем – неопределенностями интервального типа, потому что интервальные оценки неизвестных параметров систем наиболее просты и доступны для получения.

На протяжении большей части своей истории человечество занималось интеллектуальной деятельностью, существенно используя понятие числа. Это понятие, совместно с хорошо известными действиями над числами – сложением, вычитанием, умножением, делением и сравнением – позволяли решать разнообразные задачи по изучению различных систем, возникающие в естественных, технических и гуманитарных науках. Однако задачи, решаемые современной наукой и современными технологиями, значительно усложнились, и одной из главных причин этого является неопределенность изучаемых систем. Поэтому для успешного решения новых сложных задач базовое математическое понятие числа пришлось пересмотреть. В результате появилось понятие неопределенного числа. К настоящему времени известно три типа неопределенных чисел: случайные, нечеткие и интервальные. Случайные числа задаются некоторыми вероятностными распределениями их возможных значений; такие числа изучаются в теории вероятностей. Нечеткие числа задаются лингвистически сформулированными распределениями своих возможных значений; они изучаются в теории нечетких множеств. Наконец, интервальные числа задаются интервалами их возможных значений без указания какого-либо распределения возможных значений числа внутри

заданного интервала; они изучаются в интервальной математике. Очевидно, что интервальные числа содержат минимальную информацию о неопределенном числе, которую проще всего получить. Кроме того, операции над интервальными числами проще, чем операции над случайными и нечеткими числами. Отсюда – большой интерес, который представляют эти числа для различных приложений.

Любое интервальное число может быть записано в виде некоторого замкнутого вещественного интервала $\tilde{a} = [a_1, a_2]$. Произвольная интервальная функция может быть определена формально теоретико-множественной конструкцией

$$\tilde{v} = F(\tilde{x}, \tilde{y}, \dots, \tilde{z}) \equiv \{f(x, y, \dots, z) \mid x \in \tilde{x}, y \in \tilde{y}, \dots, z \in \tilde{z}\}. \quad (1)$$

здесь и ниже обозначено

$$\tilde{x} = [x_1, x_2] \quad \tilde{y} = [y_1, y_2] \quad \tilde{z} = [z_1, z_2] \quad \tilde{v} = [v_1, v_2]$$

Используя общее определение (1), можно ввести конкретные интервальные алгебраические функции: сложение и вычитание:

$$\tilde{v} = \tilde{x} \pm \tilde{y} = \{x \pm y \mid x \in \tilde{x}, y \in \tilde{y}\}. \quad (2)$$

умножение переменной на постоянную и умножение переменных

$$\tilde{v} = k\tilde{x} = \{kx \mid x \in \tilde{x}, k = \text{const}\} \quad \tilde{v} = \tilde{x} \cdot \tilde{y} = \{x \cdot y \mid x \in \tilde{x}, y \in \tilde{y}\}. \quad (3)$$

деление переменных

$$\tilde{a} \circ \tilde{b} \equiv [a_1, a_2] \circ [b_1, b_2] \equiv \{a \cdot b \mid a \in \tilde{a}, b \in \tilde{b}\}. \quad (4)$$

возведение в степень

$$\tilde{v} = \tilde{x}^n = \{x^n \mid x \in \tilde{x}\}. \quad (5)$$

и т.д. Основной задачей в интервальной алгебре является вычисление произвольной заданной интервальной функции, т.е. нахождение интервальных значений \tilde{v} интервальных функций вида (1) по данным интервальным значениям $\tilde{x}, \tilde{y}, \dots, \tilde{z}$ аргументов этих функций. Но произвольная интервальная функция может быть представлена в виде суперпозиции некоторых элементарных интервальных функций вида (2)–(5). Поэтому решение задачи вычисления произвольной интервальной функции всегда сводимо к вычислению элементарных интервальных функций, выполняемому по соответствующим простым правилам. Правила для элементарных алгебраических функций таковы.

$$\begin{aligned} [x_1, x_2] + [y_1, y_2] &= [x_1 + y_1, x_2 + y_2] \\ [x_1, x_2] - [y_1, y_2] &= [x_1 - y_2, x_2 - y_1] \end{aligned} \quad (6)$$

$$k[x_1, x_2] = \begin{cases} [kx_1, kx_2], & k \geq 0, \\ [kx_2, kx_1], & k < 0, \end{cases} \quad (7)$$

$$[x_1, x_2] \cdot [y_1, y_2] = [\min_{i,j}(x_i y_j) \quad \max_{i,j}(x_i y_j)] \quad (8)$$

$$[x_1, x_2] [y_1, y_2] = [x_1, x_2] \cdot [1/y_2, 1/y_1] \quad (9)$$

Множество всех интервальных чисел совместно с определенными на нем элементарными операциями (функциями) (2)–(5) образует алгебру интервальных чисел. Эта алгебра частично похожа на известную алгебру вещественных чисел A , поскольку в ней выполняются законы, аналогичные соответствующим в A :

$$\tilde{x} + \tilde{y} = \tilde{y} + \tilde{x}, \quad \tilde{x} \cdot \tilde{y} = \tilde{y} \cdot \tilde{x} \quad (\text{коммутативность}) \quad (10)$$

$$(\tilde{x} + \tilde{y}) + \tilde{z} = \tilde{x} + (\tilde{y} + \tilde{z}), \quad (\tilde{x} \cdot \tilde{y}) \cdot \tilde{z} = \tilde{x} \cdot (\tilde{y} \cdot \tilde{z}) \quad (\text{ассоциативность}) \quad (11)$$

$$\tilde{x} + \tilde{0} = \tilde{0} + \tilde{x} = \tilde{x} \quad (\tilde{0} = [0,0] - \text{единственный нейтральный элемент сложения}), \quad (12)$$

$$\tilde{x} \cdot \tilde{1} = \tilde{1} \cdot \tilde{x} = \tilde{x} \quad (\tilde{1} = [1,1] - \text{единственный нейтральный элемент умножения}), \quad (13)$$

$$\tilde{x} \cdot \tilde{y} = \tilde{0} \text{ влечет } \tilde{x} = \tilde{0} \text{ или } \tilde{y} = \tilde{0} \text{ и наоборот,} \quad (14)$$

$$k \cdot (\tilde{x} + \tilde{y}) = k \cdot \tilde{x} + k \cdot \tilde{y} \quad (\text{дистрибутивность умножения на вещественное число } k \text{ относительно сложения интервальных чисел}). \quad (15)$$

Заметим, что согласно (12) интервальное число $\tilde{0}$ (нуль-интервал) играет в алгебре интервальных чисел ту же роль, что и число 0 в алгебре вещественных чисел. Аналогично, согласно (13) интервальное число $\tilde{1}$ (единица-интервал) имеет в интервальной алгебре чисел такой же смысл, что и число 1 в алгебре вещественных чисел. Кроме того, в алгебре интервальных чисел действительны некоторые законы, отличающиеся от соответствующих в алгебре вещественных чисел.

$$\tilde{x} - \tilde{x} \ni 0; \text{ если } \tilde{x} - \tilde{x} = \tilde{0}, \text{ то } \tilde{x} = [x, x] \text{ и наоборот.} \quad (16)$$

Т.е. в общем случае разность равных интервальных чисел \tilde{x} не равна нулю, а лишь включает его. Эта разность равна нуль-интервалу \tilde{x} только если \tilde{x} – вырожденное интервальное число (интервал с совпадающим началом и концом).

$$\tilde{x} / \tilde{x} \ni 1; \text{ если } \tilde{x} / \tilde{x} = \tilde{1}, \text{ то } \tilde{x} = [x, x] \text{ и наоборот.} \quad (17)$$

Другими словами, в общем случае частное от деления равных интервальных чисел $\tilde{1}$ не равно единице, а лишь включает ее. Это частное равно единице-интервалу $\tilde{1}$ только если \tilde{x} – вырожденное интервальное число.

$$\text{Если } \tilde{x} + \tilde{y} = \tilde{0}, \text{ то } \tilde{x} = [x, x], \tilde{y} = [-x, -x] \text{ и наоборот.} \quad (18)$$

Т.е. в общем случае невырожденное интервальное число (интервал с

несовпадающими началом и концом) не имеет обратного интервального числа по сложению – его имеет только вырожденное интервальное число.

$$\text{Если } \tilde{x} \cdot \tilde{y} = \tilde{0}, \text{ то } \tilde{x} = [x, x], \tilde{y} = [1/x, 1/x] \text{ и обратно.} \quad (19)$$

Аналогично, невырожденное интервальное число не имеет обратного интервального числа по умножению – его имеет лишь вырожденное интервальное число.

$$\tilde{x} \cdot (\tilde{y} + \tilde{z}) \subseteq (\tilde{x} \cdot \tilde{y}) + (\tilde{x} \cdot \tilde{z})$$

(субдистрибутивность общего случая умножения на интервальное число относительно сложения таких чисел). (20)

$$\tilde{x} \cdot (\tilde{y} + \tilde{z}) = (\tilde{x} \cdot \tilde{y}) + (\tilde{x} \cdot \tilde{z}),$$

если $yz \geq 0$ для всех $y \in \tilde{y}$ и $z \in \tilde{z}$

(субдистрибутивность частного случая умножения на интервальное число относительно сложения таких чисел). (21)

Наконец, в алгебре интервальных чисел действительны некоторые законы, которые вообще не имеют аналогов в алгебре вещественных чисел. Наиболее характерен из них следующий закон монотонности включения.

$$(\tilde{x} \subseteq \tilde{x}', \tilde{y} \subseteq \tilde{y}', \dots, \tilde{z} \subseteq \tilde{z}') \Rightarrow [F(\tilde{x}, \tilde{y}, \dots, \tilde{z}) \subseteq F(\tilde{x}', \tilde{y}', \dots, \tilde{z}')] \quad (22)$$

В (22) F – любая одноместная интервальная функция, определяемая в виде:

$$F(\tilde{x}) = [\min_{x \in \tilde{x}} f(x), \max_{x \in \tilde{x}} f(x)] \quad (23)$$

с непрерывной функцией f , или 2-местная интервальная функция $-;$ или n -местная при $n \geq 2$ (в частности, 2-местная) интервальная функция $+$, \cdot .

Всякое интервальное число $\tilde{x} = [x_1, x_2]$ характеризуется следующими 2 параметрами – нижней x_1 и верхней x_2 границами или, что эквивалентно, средним значением $m(\tilde{x}) = (x_1 + x_2)/2$ и шириной $d(\tilde{x}) = x_2 - x_1$. Эта ситуация отлична от ситуации с вещественными числами, где каждое число характеризуется одним параметром – значением этого числа. Законы действий над интервальными числами во многом отличаются от законов действий над вещественными числами. Так, вычитание из интервального числа равного ему в общем случае не приводит к нуль-интервалу (16), как и деление интервального числа на равное ему не дает единицы-интервала (17). Интервальные числа не имеют в общем случае обратных чисел ни по сложению (18), ни по умножению (19). Наконец, интервальные числа в общем случае не обладают свойством дистрибутивности умножения относительно сложения (20).

Перечисленные особенности интервальных чисел приводят к значи-

тельной специфике их использования на практике, где типичной задачей является оценка множества решений некоторой проблемы при интервальных параметрах. Такова, например, задача оценки диапазона w_f изменения вещественной функции $f(x_1, \dots, x_n; a_1, \dots, a_m)$, аргументы x_i и коэффициенты a_i которой заданы интервалами \tilde{x}_i и \tilde{a}_i их возможных значений, причем \tilde{x}_i, \tilde{a}_i считаются независимыми. Оценочная функция в этом случае берется в виде $F(\tilde{x}_1, \dots, \tilde{x}_n; \tilde{a}_1, \dots, \tilde{a}_m)$, получаемой из исходной вещественной функции заменой точечных операндов x_p, a_i и операций над ними соответствующими интервальными операндами \tilde{x}_i, \tilde{a}_i и операциями. При этом на основании закона (22) имеет место следующее соотношение

$$w_f \subseteq F(\tilde{x}_1, \dots, \tilde{x}_n; \tilde{a}_1, \dots, \tilde{a}_m), \quad (24)$$

которое служит средством получения искомой оценки. Для этого вычисляется интервальная функция F от интервальных параметров в правой части (24), с использованием правил (6)–(19). Полученное интервальное число – значение функции F будет оценкой сверху диапазона w_f изменения вещественной функции f . В процессе вычисления интервальной оценочной функции F ширина интервала – величины F может в некоторых случаях достигать больших значений, многократно превышающих ширину величин ее аргументов. Этот эффект можно интерпретировать как возрастание неопределенности оценки вещественной функции по отношению к неопределенностям оценок ее аргументов. Такое возрастание может приводить к тому, что интервальные оценки некоторых вещественных функций становятся практически бесполезными. Другой эффект, связанный с получением интервальных оценок диапазона изменения вещественной функции f , заключается в зависимости ширины получаемой оценки от выбранного аналитического выражения для f . Это побуждает выбирать для f лишь те выражения, которые ведут к минимальной ширине оценки. Оба эффекта отсутствуют при работе с детерминированными вещественными (неинтервальными) функциями, где аналогом этой задачи является обычное вычисление функции.

Список литературы

1. Левин В.И. Методология оптимизации в условиях неопределенности методом детерминизации // Информационные технологии. 2014. № 5. С. 14–21.
2. Левин В.И. Математическая теория сравнения интервальных величин и ее применение в задачах измерения, контроля и управления // Измерительная техника. 1998. № 9. С. 45–58.
3. Левин В.И. Математические методы оптимизации систем в условиях неопределенности // IDO Science. 2010. № 2. С. 55–72.

УДК 004.94.519.7

Левин В.И.,д.т.н., проф., Пензенский государственный технологический университет,
Россия, Пенза, vilevin@mail.ru

ИНТЕРВАЛЬНАЯ МАТЕМАТИКА И ПРИБЛИЖЕННЫЕ ВЫЧИСЛЕНИЯ

Аннотация: В работе предлагается приближенный метод изучения поведения детерминированной функции, имеющей вид дроби и содержащей особые точки. Приближенность предлагаемого подхода объясняется тем, что при вырезании интервала, содержащего нуль, отбрасывается часть возможных значений подфункции – знаменателя изучаемой функции. Однако погрешность такого приближения может быть сделана как угодно малой путем уменьшения ширины вырезаемого интервала.

Ключевые слова: приближенный метод исследования, функция, вырезание интервала, погрешность приближения.

Levin V.I.,Doctor of Technical Sciences, prof.
Penza State Technological University, Russia, Penza, vilevin@mail.ru

INTERVAL MATHEMATICS AND APPROACH COMPUTATIONS

Abstract: *The paper proposes an approximate method for studying the behavior of a deterministic function that has the form of a fraction and contains singular points. The proximity of the proposed approach is explained by the fact that when cutting out the interval containing zero, some of the possible values of the subfunction, the denominator of the function being studied, are rejected. However, the error of such an approximation can be made arbitrarily small by reducing the width of the cut interval.*

Keywords: *approximate method of investigation, function, interval excision, approximation error.*

Введение

Появление в XX веке разнообразных сложных систем (системы управления экономикой, ракетно-космические системы, атомная энергетика и др.) выдвинуло новые сложные задачи по их изучению. Современная наука и практика обработки информации вполне успешно справляется с задачами исследования различных сложных систем с полностью определенными (детерминированными) параметрами. Эти задачи обычно формулируются как задачи расчета, анализа и синтеза тех или иных функций с детерминированными параметрами, служащих соответствующими характеристиками изучаемых систем. Но на практике достаточно часто встречаются другие системы – системы с неточно известными, т.е. неполностью определенными (недетерминированными) параметрами. Причины появления таких систем заключаются в естественной неопределенности, свойственной многим реальным процессам, происходящим в системах; в неточном задании параметров большинства систем из-за неизбежных погрешностей при их измерении или вычислении; в

изменении во времени параметров систем; в необходимости или целесообразности совместного исследования целых семейств однотипных систем, имеющих типовые функции-характеристики и различающихся лишь значениями параметров этих функций. Учет неопределенности систем очень важен при их проектировании, так как полная определенность в работе системы появляется лишь на последних этапах ее создания.

Исследование введенных неопределенных систем формулируется в виде задач расчета, анализа и синтеза различных функций с недетерминированными параметрами, служащих соответствующими характеристиками данных систем. Все эти задачи значительно сложнее их вышеупомянутых детерминированных аналогов, которые приходится решать при изучении систем с детерминированными параметрами. Это усложнение связано с тем, что алгебра недетерминированных чисел сложнее, чем алгебра детерминированных чисел. В связи с этим для решения указанных задач приходится применять тот или иной специальный математический аппарат: теорию вероятностей [1], теорию нечетких множеств [2], интервальную математику [3], многозначные функции [4]. Применение этого аппарата позволяет строить и исследовать более адекватные математические модели сложных систем с недетерминированными параметрами, учитывающие неопределенность поведения таких систем [5–13].

Однако на практике встречаются еще более трудные для изучения классы сложных систем, в которых даже математические модели с детерминированными параметрами приводят к задачам, не имеющим определенного решения. Таковы, например, сложные системы, изучение которых сводится к решению системы линейных уравнений, определитель которой в некоторых случаях может быть равен нулю. Именно для таких систем раздетерминизация, т.е. переход к соответствующей недетерминированной системе позволяет получить необходимое решение. Так, для сложных систем, изучение которых сводится к решению системы линейных уравнений (с возможно нулевым определителем), определитель после раздетерминизации становится численно равным интервалу, включающему, кроме нуля, также ненулевые значения, что открывает возможность получения решения.

В работе рассматриваются задачи изучения именно таких классов сложных систем. В качестве раздетерминизации используется процедура перехода от системы с детерминированными параметрами к системе с недетерминированными – интервальными параметрами. В качестве математического аппарата используется интервальная алгебра. Раздетерминизация является процедурой, обратной по отношению к детерминизации, широко используемой в работах автора по изучению поведения неопределенных систем [14–19].

1. Постановка задачи

Имеется некоторая практическая задача, сводящаяся с математической точки зрения к исследованию (расчету и анализу поведения) некоторой детерминированной функции одной независимой переменной – характеристики изучаемой системы

$$y = f(x), \quad (1)$$

однозначно отображающей заданное множество $X = \{x\}$ независимых переменных x в заданное множество $Y = \{y\}$ зависимых переменных y , в соответствии с законом f , который называется функцией. Хорошо известно, что задача расчета (вычисления значений) функции (1) принципиально всегда решается с помощью адекватного этой задаче математического аппарата алгебры вещественных чисел, с использованием подходящих методов вычислений, а задача анализа поведения функции (1) – с помощью адекватного ей аппарата классического дифференциального исчисления.

Рассмотрим далее распространенную ситуацию, когда изучаемая детерминированная функция (1) имеет вид дроби

$$f(x) = f_1(x) / f_2(x) \quad (2)$$

с числителем – функцией $f_1(x)$ и знаменателем – функцией $f_2(x)$. В этой ситуации расчет и анализ поведения функции (1) затрудняется, поскольку в каждой точке с нулевым знаменателем эта функция не существует (если в этой точке числитель не равен нулю) либо принимает бесконечное множество значений, т.е. не имеет определенного значения (если в этой точке числитель равен нулю). Очевидно, что для функций с указанными точками (их естественно называть особыми) должны быть разработаны специальные методы расчета и анализа поведения функций, позволяющие исключить эффект влияния особых точек. Задача состоит в построении двух систематических процедур, связанных с изучением поведения детерминированных функций вида (2). А именно: 1) процедура расчета (т.е. вычисления значений) детерминированной функции типа (2), содержащей особые точки; 2) процедура анализа поведения такой же детерминированной функции.

2. Обзор литературы

Ниже мы будем изучать поведение детерминированных функций типа (2), содержащих особые точки, имея в виду преодоление трудности, связанной с наличием таких точек, путем раздетерминизации, т.е. перехода от функции (2) к соответствующей недетерминированной функции (т.е. функции с недетерминированными, а точнее – с интервальными параметрами). В связи с этим представляет интерес обзор научной литературы посвященной изучению различных неопределенных объектов.

Проблема изучения объектов, характеризующихся той или иной неопределенностью, возникла впервые в начале Второй мировой войны, в связи с необходимостью управления огнем зенитной артиллерии, в ус-

ловиях случайного движения воздушных целей. Соответствующими задачами занимались выдающиеся математики-вероятности-ки Н. Винер [5] и А.Н. Колмогоров [6] и их многочисленные последователи. Однако широкое развитие исследований по изучению гражданских объектов, работающих в условиях неопределенности, началось только в конце 1950-х – начале 1960-х гг., в рамках математической статистики и ее новых направлений – обработка данных и планирование экспериментов [7, 8].

Исследования, выполненные в 1970-е–80-е гг., привели к более широкому пониманию неопределенности, включавшей теперь не только случайность, но и незнание возможных исходов, их неединственность, неопределенность целей, многокритериальность при решении оптимизационных задач. В связи с этим появились и новые подходы к описанию неопределенности, такие как теория нечетких множеств, принцип недоопределенной модели, принятие решений в многокритериальных задачах [2, 9, 10].

С 1980-х годов начал интенсивно применяться подход к описанию неопределенности, базирующийся на интервальной математике, позволяющей получать оценки характеристик неопределенных систем с гарантированной точностью [11–18]. При этом указанный подход применялся сначала в метрологии для определения интервального значения известной функции при интервальных значениях аргументов. Затем его развитие шло по двум направлениям. За рубежом интервальный подход развивался как средство автоматического учета ошибок округления при численном решении задач на компьютерах, в то время как в СССР и России ученые развивали его с целью нахождения области возможных значений результата вычислений с учетом структуры данных и функций, заданных в символьном виде.

Наконец, с 1990-х годов начала изучаться очень важная в практических приложениях задача исследования поведения произвольной недетерминированной функции с интервальными параметрами, являющейся аналогом известной задачи математического анализа – исследования поведения детерминированной функции методами дифференциального исчисления [20].

3. Используемые методы

Сначала изложим основную идею предлагаемого метода. Рассмотрим детерминированную функцию вида дроби (2), с возможными особыми точками, т.е. точками, в которых знаменатель функции (2) равен 0. В таких точках, как уже говорилось в п. 1, функция (2) либо не существует, либо не имеет определенного значения. Мы предлагаем метод, позволяющий придать функции (2) одно определенное значение во всех ее точках, включая и особые, тем самым исключается влияние особых точек на характер поведения функции. Предлагаемый метод состоит в переходе от детерминированной функции (2) к соответствующей неде-

терминированной – интервальной функции

$$\tilde{f}(x) = \tilde{f}_1(x) / \tilde{f}_2(x), \quad (3)$$

путем замены всех точно заданных параметров исходной функции соответствующими интервальными параметрами. Эту процедуру естественно назвать *раздетерминизацией*. В результате раздетерминизации точные значения числителя $y_1 = f_1(x)$ и знаменателя $y_2 = f_2(x)$ исходной функции $f(x)$ переходят в интервальные значения $\tilde{y}_1 = \tilde{f}_1(x)$, $\tilde{y}_2 = \tilde{f}_2(x)$, где \tilde{y}_1, \tilde{y}_2 – интервалы $Y = \tilde{f}(X)$, $y = f(x)$, а все точные значения самой исходной функции $y = f(x)$ – в соответствующие интервальные значения этой функции $\tilde{y} = \tilde{f}(x)$, где $\tilde{y} = [y_1, y_2]$. При этом все особые точки исходной функции после раздетерминизации можно исключить из рассмотрения. Действительно, в каждой такой точке знаменатель $f_2(x)$ исходной функции равен 0, но раздетерминизируемый знаменатель $\tilde{f}_2(x)$ в этой точке равен уже не нулю, а интервалу, содержащему нуль. Деление на такой интервал в интервальной математике не рассматривается и предполагается невозможным [3]. Однако это ошибочная точка зрения, поскольку, если вырезать из интервала малый подынтервал, содержащий нуль, то оставшаяся, большая часть интервала уже не будет содержать нуля и деление на такой интервал по методологии интервальной математики окажется вполне возможным. Таким образом, использование метода раздетерминизации исходной детерминированной функции (2) в сочетании с вырезанием нуля из интервала возможных значений раздетерминизируемого знаменателя этой функции позволяет ликвидировать все особые точки исходной функции (2) и применить к изучению поведения этой функции обычные методы изучения поведения интервальной функции [3, 11, 14, 19, 20]. Разумеется, предлагаемый нами метод изучения поведения детерминированной функции, содержащей особые точки, является приближенным, поскольку при вырезании интервала, содержащего нуль, отбрасывается часть возможных значений подфункции – знаменателя изучаемой функции. Однако погрешность такого приближения может быть сделана как угодно малой путем уменьшения ширины вырезаемого интервала.

Опишем теперь метод раздетерминизации подробнее. Как известно из интервальной математики [3], любая операция над интервалами вводится как теоретико-множественное обобщение соответствующей операции над точными вещественными числами. Другими словами, если a, b – точные вещественные числа, $\tilde{a} = [a_1, a_2]$, $\tilde{b} = [b_1, b_2]$ – интервалы, \circ – операция над точными вещественными числами, \circ – соответствующая операция над интервалами, по определению имеем

$$\tilde{a} \circ \tilde{b} \equiv [a_1, a_2] \circ [b_1, b_2] \equiv \{a \cdot b \mid a \in \tilde{a}, b \in \tilde{b}\}. \quad (4)$$

Также получаем формулу для операции деления двух интервалов

$$\tilde{a}/\tilde{b} \equiv [a_1, a_2]/[b_1, b_2] \equiv \{a/b \mid a \in \tilde{a}, b \in \tilde{b}\}, 0 \notin \tilde{b}. \quad (5)$$

Дополнительное требование относительно нуля в этой формуле связано с невозможностью деления вещественного числа на нуль. Так как деление вещественных чисел обратно умножению, формулу (5) можно записать в терминах операции умножения

$$\tilde{a}/\tilde{b} \equiv [a_1, a_2]/[b_1, b_2] \equiv [a_1, a_2] \cdot [1/b_1, 1/b_2], 0 \notin \tilde{b}. \quad (6)$$

Однако операция умножения интервалов, как известно из интервальной математики [3], выполняется по формуле следующего вида

$$\tilde{c} \cdot \tilde{d} \equiv [c_1, c_2] \cdot [d_1, d_2] = [\min_{i,j} (c_i \cdot d_j), \max_{i,j} (c_i \cdot d_j)]. \quad (7)$$

Соединяя вместе формулы (6), (7), мы получаем окончательную формулу для выполнения операции деления двух интервалов

$$\tilde{a}/\tilde{b} \equiv [a_1, a_2]/[b_1, b_2] = [\min_{i,j} (a_i / b_j), \max_{i,j} (a_i / b_j)], 0 \notin \tilde{b}. \quad (8)$$

Эта формула, однако, пригодна для выполнения операции деления интервалов лишь в тех случаях, когда интервал-делитель $\tilde{b} = [b_1, b_2]$ не содержит нуля. А что делать, если он содержит нуль? Интервальная математика не отвечает на этот вопрос [3].

Мы предлагаем такой ответ на него. Будем считать, что интервал-делитель содержит 0 внутри (а не на одном из концов). Вырежем из интервала-делителя $\tilde{b} = [b_1, b_2]$ в формуле (8) деления интервалов некоторый достаточно малый подынтервал $\tilde{b}^* = [b'_1, b'_2]$, содержащий нуль, тем самым заменив интервал \tilde{b} биинтервалом

$$\tilde{b}''' = \tilde{b} / \tilde{b}^* = \tilde{b}' \cup \tilde{b}'' \quad (9)$$

где $\tilde{b}' = [b_1, b'_1]$, $\tilde{b}'' = [b'_2, b_2]$ (см. рис. 1).

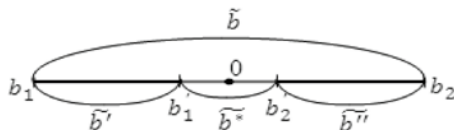


Рис. 1. Вырезание подынтервала из интервала

Произведенную операцию естественно назвать операцией вырезания подынтервала из указанного интервала. Объективный смысл этой операции – это приближение исходного интервала $\tilde{b} = [b_1, b_2]$ биинтервалом (объединением двух непересекающихся интервалов) вида $\tilde{b}' \cup \tilde{b}''$, выражающимся по формуле (9), таким образом, чтобы полученный биинтервал не содержал неприемлемого для нас множества точек. В данном случае это

множество – интервал $\tilde{a} = [a_1, a_2]$, содержащий нуль. Так, если мы хотим разделить интервал $\tilde{a} = [a_1, a_2]$ на интервал $b = [b_1, b_2]$, содержащий нуль, то надо заменить в общей формуле деления (5) интервал \tilde{b} биинтервалом \tilde{b}'' вида (9), не содержащим нуль. Явный вид этой формулы найдем, используя общий принцип теоретико-множественного обобщения операций над точными вещественными числами (4):

$$\tilde{a} / \tilde{b} \approx \tilde{a} / (\tilde{b} \setminus \tilde{b}^*) = \tilde{a} / (\tilde{b}' / \tilde{b}'') = \{a/b \mid a \in \tilde{a}, b \in \tilde{b}' \cup \tilde{b}''\} = \{(a/b') \cup (a/b'') \mid a \in \tilde{a}, b' \in \tilde{b}', b'' \in \tilde{b}''\} = \{a/b' \mid a \in \tilde{a}, b' \in \tilde{b}'\} \cup \{a/b'' \mid a \in \tilde{a}, b'' \in \tilde{b}''\} = (\tilde{a} / \tilde{b}') \cup (\tilde{a} / \tilde{b}'').$$

Итак, деление интервала \tilde{a} на интервал \tilde{b} , содержащий нуль, можно выполнить по приближенной формуле

$$\tilde{a} / \tilde{b} = (\tilde{a} / \tilde{b}') \cup (\tilde{a} / \tilde{b}''), \quad (10)$$

где \tilde{b}' и \tilde{b}'' – подынтервалы интервала \tilde{b} , не содержащие нуля, объединение которых приближенно равно \tilde{b}' (рис. 1). Поскольку \tilde{b}' и \tilde{b}'' не содержат нуля, обе скобки правой части (10) можно вычислять по формуле (8). В развернутом виде формула (10) переписывается как

$$[a_1, a_2] / [b_1, b_2] \equiv [a_1, a_2] / [b_1, b_1'] \cup [a_1, a_2] / [b_2', b_2]. \quad (11)$$

Список литературы

1. Гнеденко Б.В. Курс теории вероятностей. М.: Наука, 2004. 350 с.
2. Заде Л.А. Понятие лингвистической переменной и его применение к принятию приближенных решений. М.: Мир, 1976. 165 с.
3. Алефельд Г., Херцбергер Ю. Введение в интервальные вычисления. М.: Мир, 1987. 360 с.
4. Горбань И.И. Феномен статистической устойчивости. Киев: Наукова Думка, 2014. 370 с.
5. Wiener N. Extrapolation, Interpolation and Smoothing of Stationary Time Series. New-York: Technology Press and Wiley, 1949. 180 p.
6. Колмогоров А.Н. Интерполирование и экстраполирование стационарных случайных последовательностей // Известия АН СССР. Серия: математика. 1941. № 5. С. 3–14.
7. Канторович Л.В. О некоторых новых подходах к вычислительным методам и обработке наблюдений // Сибирский математический журнал. 1962. Т. 3. № 5. С. 3–14.
8. Налимов В.В., Чернова Н.А. Теория эксперимента. М.: Наука, 1971. 320 с.
9. Нариньяни А.С. Недоопределенность в системе представления и обработки знаний // Известия АН СССР. Техническая кибернетика. 1986. № 5. С. 3–28.
10. Hyyonen E. Constraint Reasoning Based on Interval Arithmetic: the Tolerance Propagation Approach // Artificial Intelligence. 1992. Vol. 58. P.19.
11. Вошинин А.П., Сотиров Г.Р. Оптимизация в условиях неопределенности. М.: МЭИ, София: Техника, 1989. 226 с.
12. Вошинин А.П., Бочков А.Ф., Сотиров Г.Р. Интервальный анализ данных // Заводская лаборатория. 1990. № 7. С. 76–81.
13. Куржанский А.Б. Задача идентификации – теория гарантированных оценок // Автоматика и телемеханика. 1991. № 4. С. 75–89.

14. Левин В.И. Дискретная оптимизация в условиях неопределенности // Автоматика и телемеханика. 1992. № 7. С. 97–106.
15. Левин В.И. Булево линейное программирование с интервальными коэффициентами // Автоматика и телемеханика. 1994. № 7. С. 111–122.
16. Левин В.И. Интервальное дискретное программирование // Кибернетика и системный анализ. 1994. № 6. С. 92–103.
17. Левин В.И. Нелинейная оптимизация в условиях интервальной неопределенности // Кибернетика и системный анализ. 1999. № 2. С. 138–146.
18. Левин В.И. Методы оптимизации систем в условиях интервальной неопределенности параметров // Информационные технологии. 2012. № 4. С. 52–59.
19. Левин В.И. Методология оптимизации в условиях неопределенности методом детерминизации // Информационные технологии. 2014. № 5. С. 14–21.
20. Левин В.И. Анализ поведения неточно заданных функций с помощью интервально-дифференциального исчисления // Информационные технологии. 2015. Т. 21. № 3. С. 163–170.

УДК 004.657

Носова Татьяна Николаевна,

МГТУ им. Г.И. Носова, старший преподаватель кафедры Информатики и информационных технологий

Федеральное государственное бюджетное образовательное учреждение высшего образования «Магнитогорский государственный технический университет им. Г.И. Носова» ФГБОУ ВО «МГТУ им Г.И. Носова», г. Магнитогорск, Челябинская обл., Россия. <http://magtu.ru>. e-mail: ntn.mgtu@mail.ru

СОЗДАНИЕ NET-ПРИЛОЖЕНИЯ ДЛЯ УВЕЛИЧЕНИЯ ЭФФЕКТИВНОСТИ ПОИСКОВЫХ ЗАПРОСОВ В БАЗАХ ДАННЫХ

Аннотация: В работе анализируются условия, повышающие эффективность запросов различных баз данных. На основании этого анализа разработано программное приложение, которое расширяет встроенные возможности СУБД MS SQL Server по индексированию данных и оптимизации поисковых запросов.

Ключевые слова: оптимизация запросов баз данных, индексирование данных, двоичные индексы, Net-приложение

Nosova Tatyana Nikolaevna,

MSTU. G.I. Nosova, Senior Lecturer, Department of Informatics and Information Technologies Federal State Budgetary Educational Institution of Higher Education "Magnitogorsk State Technical University. G.I. Nosova"
FSBEI of HE "MG TU them. I.I. Nosova, Magnitogorsk, Chelyabinsk Region, Russia.
<http://magtu.ru>. e-mail: ntn.mgtu@mail.ru

CREATING A NET APPLICATION TO INCREASE THE EFFICIENCY OF SEARCH REQUESTS IN DATABASES

Abstract: This paper analyzes conditions that increase the efficiency of queries of various

databases. Based on this analysis, a software application has been developed that extends the built-in capabilities of MS SQL Server DBMS for indexing data and optimizing search queries.

Keywords: *database query optimization, data indexing, binary indexes, Net-application*

Введение

При создании проектов, взаимодействующих с базами данных, следует уделять особое внимание оптимизации работы информационных систем. Практически во всех современных веб-приложениях эффективные способы доступа к данным являются критически важными задачами, влияющими на время отклика системы на запросы пользователей.

К самым распространенным средствам современных СУБД по оптимизированию поисковых запросов относятся: изучение плана выполнения запроса с оценкой его стоимости, использование индексных структур разного вида, анализ избирательности индексов [2-5].

План запроса создается в фазе оптимизации обработки данных компонентом ядра базы данных, называемым оптимизатором запросов. Последний, принимая во внимание множество различных факторов, пытается подобрать наиболее эффективный алгоритм выполнения запроса.

Одним из ключевых элементов оптимизации поиска и сортировки данных являются индексные структуры разного вида. Индекс формируется из значений одного или нескольких столбцов таблицы и указателей на соответствующие строки. Так как индексная таблица имеет упорядоченную структуру, то основным способ повышения быстродействия состоит в использовании алгоритма бинарного поиска в сбалансированном В-дереве вместо полного сканирования таблицы [6].

Кластеризованные и некластеризованные индексы помогают серверу баз данных находить результат значительно быстрее, используя для этого разные варианты сбалансированных В-деревьев и хеш-таблиц.

Многие источники рекомендуют правило формирования оптимальных систем, регламентирующее выставление индексов на внешних ключах табличных связей. Именно по этим полям система осуществляет поиск той или иной записи в разных таблицах [3].

Также следует проводить анализ запросов, используемых в подсистеме и индексировать именно те поля, которые используются для сортировки или поиска. Следует учитывать, что индексы замедляют выполнение команд DML, а их сопровождение требует времени и увеличивает стоимость обработки.

Многие СУБД блокируют использование индексов, если: индексное поле используется в вычисляемых выражениях, в качестве операнда сравнения со значениями неиндексированного поля, в операциях, использующих сравнение с неопределенным значением NULL, или является параметром встроенных или пользовательских функций [1].

Необходимо отметить, что в вопросе организации индексов, у разработчиков различных СУБД имеются свои подходы.

Известные на сегодняшний момент СУБД используют разные виды

индексов. СУБД Microsoft SQL Server работает со следующими видами индексов (табл.1) [7].

Таблица 1.

В следующей таблице приведен список типов индексов, доступных в SQL Server

Тип индекса	Описание
Хэш	При использовании хэш-индекса доступ к данным осуществляется через хэш-таблицу в памяти. Хэш-индексы используют фиксированный размер памяти, который зависит от числа контейнеров.
Некластеризованные индексы с оптимизацией для памяти	Для оптимизированных для памяти некластеризованных индексов потребление памяти является функцией от количества строк и размера ключевых столбцов индекса
Кластеризованный	Кластеризованный индекс сортирует и хранит строки данных таблицы или представления в порядке, определяемом ключом кластеризованного индекса. Кластеризованный индекс реализуется в виде сбалансированного дерева, которое поддерживает быстрое получение строк по значениям ключа кластеризованного индекса.
Некластеризованный	Некластеризованный индекс можно определить в таблице или представлении вместе с кластеризованным индексом или в куче. Каждая строка некластеризованного индекса содержит некластеризованное ключевое значение и указатель на строку. Этот указатель определяет строку данных кластеризованного индекса или кучи, содержащую ключевое значение. Строки в индексе хранятся в порядке, определяемом значениями ключа индекса, но до создания кластеризованного индекса в таблице нет никакой гарантии того, что строки данных будут расположены в каком-либо определенном порядке.
Уникальный	Уникальный индекс обеспечивает отсутствие повторяющихся значений ключа индекса, что, в свою очередь, приводит к тому, что каждая строка в таблице или представлении является в каком-то смысле уникальной.
columnstore	Индекс columnstore в памяти хранит данные и управляет данными с использованием основанного на столбцах хранилища данных и обработки запросов.
Индекс с включенными столбцами	Некластеризованный индекс, дополнительно содержащий кроме ключевых столбцов еще и неключевые.
Индекс на вычисляемых столбцах	Индекс на столбце, являющемся производным от одного или нескольких других столбцов или нескольких детерминированных источников.
Фильтруемый	Оптимизированный некластеризованный индекс, в особенности подходящий для покрытия запросов из хорошо определенного подмножества данных. Хорошо спроектированный отфильтрованный индекс позволяет повысить производительность запросов, снизить затраты на обслуживание и хранение индексов по сравнению с полнотабличными индексами.
Пространственный	Пространственный индекс обеспечивает возможность более эффективного использования определенных операций с пространственными объектами (пространственными данными) в столбце типа данных geometry . Пространственные индексы снижают количество объектов, к которым должны применяться пространственные операции, требующие больших затрат.

XML	Представление разобранное и сохраненное, большим двоичным объектам XML (BLOB) в xml столбец типа данных.
Полнотекстовый	Специальный тип функционального индекса, основанный на токене, построенный и поддерживаемый средством полнотекстового поиска (Майкрософт) для SQL Server. Он обеспечивает эффективную поддержку сложных операций поиска слов в символьных строковых данных.

Данные в таблице хранятся в отсортированном виде только в случае, если создан кластеризованный индекс для этой таблицы.

В большинстве случаев индексирование полей, используемых в условиях отбора, повышает производительность запросов. Однако в действительности это определяется значениями избирательности индекса и коэффициента повторяющихся значений.

Наилучшей избирательностью обладают поля первичных ключей и столбцы с требованием свойства уникальности. Наиболее полезными для оптимизатора являются критерии запроса по индексированным полям с высокой избирательностью (соответственно, низкой плотностью), т.к. позволяют надежно предсказать, какое количество операций ввода-вывода потребуется при выполнении запроса.

В общем случае, чем больше дубликатов в индексированном столбце, тем хуже работает индекс [4]. Поэтому, все вышеперечисленные рекомендации по увеличению производительности запросов не подходят для поиска данных по полям с низкой избирательностью.

Постановка задачи

Разработанный программный продукт тестировала на примере конкретной базы данных «Сотрудники МГТУ», логическая схема которой представлена на рис. 1.

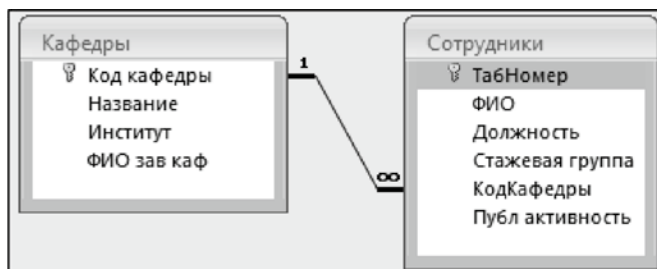


Рис. 1. Логическая схема базы данных «Сотрудники МГТУ»

К хранимым данным постоянно предъявляется перечень часто возникающих вопросов: «Сотрудники, каких кафедр или институтов имеют заданную должность, стажевую группу, публикационную активность и т.п.?». Т.е. требуются статистические показатели по подразделениям и лично по сотрудникам.

Для получения обобщенной информации о сотрудниках университе-

та составлен запрос SQL на слияние таблиц:

```
SELECT Сотрудники.*, Кафедры.[Название каф], Кафедры.Институт
FROM Кафедры INNER JOIN Сотрудники ON Кафедры.[Код кафедры]
= Сотрудники.[Код кафедры];
```

Фрагмент выполнения запроса представлен на рис. 2.

ТабНомер	ФИО	Должность	Стажевая гр	Кодf	Название	Институт	Публ акт
103	Грушин	ассист	I		2 Менеджмента	ЭиУ	
106	Соколов	ст преп	I		2 Менеджмента	ЭиУ	
101	Иванов	ассист	I		3 ИиИБ	ЭиАС	
102	Петров	ст преп	II		3 ИиИБ	ЭиАС	
107	Симонян	ассист	I		3 ИиИБ	ЭиАС	
111	Мохов	доктор	I		5 Механики	МиМ	
108	Якунин	ст преп	I		6 Управления	ЭиУ	
104	Слива	доктор	III		7 Физики	ГО	
105	Сирчук	ассист	I		7 Физики	ГО	
109	Игнатно	доцент	II		7 Физики	ГО	

Рис. 2. Фрагмент временной таблицы, используемой для реализации поисковых запросов с несколькими предикатами отбора

Приведенная выше таблица демонстрирует, что только столбец ТабНомер имеет уникальные значения и для него системой Microsoft SQL Server будет построен кластеризованный индекс.

Для увеличения скорости выполнения запросов по другим полям таблицы общей рекомендацией является создание некластеризованного индекса. В действительности такой подход к сокращению времени выполнения запроса не приводит, так как избирательность индекса по полям: «Должность», «Пол», «Стажевая группа», «Институт» – низкая. В тестовом примере поле «Институт» содержит всего восемь уникальных значений, «Должность» – четыре (ассистент, старший преподаватель, доцент, доктор), «Стажевая группа» – три (I, II, III), «Пол» – два.

Таблица 2.

Плотность распределения значений в полях таблицы «Сотрудники»

Название поля таблицы	Плотность распределения данных 1/число уникальных значений
Пол	0,5
Должность	0,25
Стажевая группа	0,333
Институт	0,128
Кафедра	0,016

Если плотность распределения значений в столбцах реляционных таблиц превышает 10%, то некластеризованные индексы (простые и составные) можно считать бесполезными [5].

Цель работы

Основной целью представленной работы является изучение и применение эффективного метода поиска данных в реляционных таблицах.

Для достижения этой цели было разработано программное средство, реализующее поисковые запросы в информационной системе «Сотрудники университета». Особенностью этой базы данных является наличие большого количества повторяющихся значений в столбцах таблиц.

Для решения проблемы поиска данных в таких таблицах по полям с низкой избирательностью, было разработано приложение для платформы Net на языке C#, реализующее эффективный способ поиска данных, и использующий для этого механизм битовых индексов (bitmap-index).

Обычно в B*-дереве имеется однозначное соответствие между записью индекса и строкой таблицы. Битовая карта используется для ссылки на большое количество строк таблицы одновременно. Такие индексы подходят для данных с небольшим количеством различных значений, подвергающиеся нечастому изменению.

Как следует из таблицы 1. в MS SQL Server не представлены индексы на основе битовых карт, следовательно, разработанное приложение расширяет встроенные средства СУБД по решению задач оптимизации запросов.

Объектом исследования являлись двоичные (bitmap)-индексы как средство повышения эффективности обработки поисковых запросов в СУБД MS SQL Server.

Метод битовых индексов заключается в создании отдельных битовых карт (последовательностей 0 и 1) для каждого возможного значения столбца, где каждому биту соответствует строка с индексируемым значением, а его значение равно 1 означает, что запись, соответствующая позиции бита содержит индексируемое значение для данного столбца или свойства.

Приложение содержит несколько методов, последовательно реализующих механизм битовых шкал (карт) для рассматриваемой таблицы.

```

namespace DB_BitMaps
{
    class Program
    {
        public static void Print_Empl(string[,] Mas)...

        /*Статический метод составления БИТОВОЙ ШКАЛЫ для заданного поля таблицы.
        Принимает в качестве параметров: массив сотрудников, массив
        ключей для словаря, название поля, по которому составляется битовая шкала*/

        static Dictionary<string, int[]> Ccreate_Bit_Map(string[,] arr_Emp,
            string[] arr_Keys, int number_fild)...

        //+++++
        /*МЕТОД ПОБИТОВОГО УМНОЖЕНИЯ
        Принимает в качестве параметров: ссылку на массив сотрудников,
        две битовых шкалы и ключи словарей*/
        public static void Devision_BitMaps(string[,] arr, Dictionary<string, int[]> bm1,
            string fild_bm1, Dictionary<string, int[]> bm2, string fild_bm2)...

        //+++++
        //Перегрузка метода полбитового умножения для трех ключей
        public static void Bit_Devision(string[,] arr, Dictionary<string, int[]> bm1,
            string fild_bm1, Dictionary<string, int[]> bm2, string fild_bm2,
            Dictionary<string, int[]> bm3, string fild_bm3)...

        //+++++
        //Метод поиска ПОЛНЫМ СКАНИРОВАНИЕМ ТАБЛИЦЫ
        public static void Finf_ScanTable(string[,] arr, int N, string fild1,
            string fild2)...

        //+++++
        static void Main(string[] args)...
    }
}

```

Рис. 3. Структура проекта

Первый статический метод класса Program, предназначен для создания битовой шкалы для заданного поля реляционной таблицы.

```
static Dictionary <string, int[] > Ccreate_Bit_Map (string[ , ], string[
], int) {...}
```

Принимает в качестве параметров: ссылку на реляционную таблицу, строковый массив, встречающихся в столбце уникальных значений и номер столбца, для которого создается битовая карта.

Метод возвращает готовый bitmap-индекс заданного поля таблицы, программно реализованный с использованием встроенного класса Dictionary – словаря данных.

Dictionary – аналог ассоциативных массивов в других языках программирования, работающих с парами {Key->Value} [8].

В нашем случае – это набор {string, int[] }, где ключом Key будет являться очередное уникальное значение поля, представленное в виде строки; в качестве Value используется целочисленный одномерный массив, содержащий 0 и 1. Единица на очередной позиции массива соответствует совпадению у очередной записи значения поля с заданным уникальным значением.

Для тестового примера для таблицы размером в 10 строк результат работы метода следующий:

```

files://D:/TANYA/Статьи/17/New/Empl_BitMaps/DB_BitMaps/bin/Debug/DB_BitMaps.EXE
id      ФИО      Должность  Стаж.группа
101     Иванов  ассист     I
102     Петров  ст преп    II
103     Грушин  ассист     I
104     Слива  доктор     III
105     Сирчук  ассист     I
106     Соколов ст преп    I
107     Симонян ассист     I
108     Якунин  ст преп    I
109     Игнатко доцент  II
111     Мохов   доктор     I

Битовая шкала для должностей: 'ассист','ст преп','доцент','доктор'
1010101000
0100010100
0000000010
0001000001

Битовая шкала для стажевых групп: I, II и III
1010111101
0100000010
0001000000

```

Рис. 4. Результат работы программы по формированию битовых шкал для двух полей реляционной таблицы

Следующий метод побитового умножения `public static void BitDevision()`, предназначен для реализации запроса с составными критериями отбора по полям с малым процентом уникальных значений.

Принимает в качестве параметров: ссылку на таблицу, ссылки на один или несколько битовых индексов и критерии поиска, представленных в виде текстовых строк.

```

public static void Devision_BitMaps(string[ ], Dictionary <string,
int[ ]> , string , Dictionary <string, int[]>, string)
{...}

```

Метод реализует алгоритм побитового умножения для нескольких битовых шкал и выводит на печать те строки таблицы, для которых выполняются заданный набор критериев отбора.

```

files://D:/TANYA/Статьи/17/New/Empl_BitMaps/DB_BitMaps/bin/Debug/DB_BitMaps.EXE

*** ПОИСК С ИСПОЛЬЗОВАНИЕМ БИТОВЫХ ШКАЛ ****
Сотрудники для которых выполняются критерии: ассист и I стажевая группа :
ID:      Фамилия:      Должность:      Стаж.группа
-----
101:     > Иванов           > ассист:         > I
103:     > Грушин            > ассист:         > I
105:     > Сирчук            > ассист:         > I
107:     > Симонян           > ассист:         > I

ВРЕМЯ ВЫПОЛНЕНИЯ ПОИСКА БИТОВЫМИ ШКАЛАМИ: 7 мс

**** ПОИСК МЕТОДОМ ПОЛНОГО СКАНИРОВАНИЯ ТАБЛИЦЫ ****
ID:      Фамилия:      Должность:      Стаж.группа
-----
101:     > Иванов           > ассист:         > I
103:     > Грушин            > ассист:         > I
105:     > Сирчук            > ассист:         > I
107:     > Симонян           > ассист:         > I

ВРЕМЯ ВЫПОЛНЕНИЯ ПОИСКА ПОЛНЫМ СКАНИРОВАНИЕМ: 11 мс

```

Рис. 5. Сравнение скорости выполнения запроса при использовании битовых шкал и полного сканирования таблицы

Для поддержки реализации запросов с составными условиями отбора, предусмотрено несколько перегрузок метода `Devision_BitMaps`: поиск в таблице по двум, трем и по любому количеству заданных критериев.

Тестирование созданного приложения на таблицах разного размера показало существенное сокращение времени выполнения запроса при использовании `bitmap`-индексов по сравнению с полным сканированием таблицы. Для подтверждения этого тезиса в программу добавлен метод, осуществляющий поисковый запрос полным сканированием таблицы (рис. 5).

Результаты работы программы для таблицы размером 100 и 1000 записей приведены в таблице 3.

Таблица 3.

Сравнение времени выполнения поискового запроса при использовании разных подходов

Использованный алгоритм поиска	Количество записей таблицы	
	100	1000
полное сканирование таблицы	5мс	23 мс
использование битовых индексов	2 мс	12 мс

Следует отметить, что некоторые ресурсы тратятся на составление и хранение самой битовой карты, поэтому двоичный индекс максимально эффективен на полях с нечастым изменением значений.

Результаты

1. Применение алгоритма битовых шкал является эффективным методом повышения производительности поисковых запросов в РТ по полям с малой долей уникальных значений.

2. Разработанное программное приложение расширяет встроенные возможности СУБД MS SQL Server по индексированию данных и оптимизации поисковых запросов.

Заключение

Большинство видов индексов достаточно эффективны применительно к полям РТ с уникальными значениями либо с низкой плотностью значений.

Применение бинарных индексов оправдано для полей с большой долей повторяющихся значений, в том случае, если имеется достаточное количество ресурсов дискового пространства и модификация таблиц не является интенсивной (статичных таблиц).

Список литературы

1. Оптимизация производительности выполнения запросов (SQL Server Compact) //Microsoft. Developer Network . [Электронный ресурс]. URL: <https://msdn.microsoft.com/ru-ru/library/ms172984.aspx> (дата обращения: 24.01.2017).
2. Носова Т.Н. Методологические аспекты изучения дисциплины «Информаци-

- онные технологии. Базы данных» // Педагогическая информатика. 2017. № 2. С. 99-109.
3. Туманов В. Основы проектирования реляционных баз данных // Интуит. Национальный открытый университет. [Электронный ресурс]. URL: <http://www.intuit.ru/studies/courses/1095/191/lecture/4997?page=4> (дата обращения: 24.01.2017).
 4. Мирошниченко Г. А. Реляционные базы данных: практические приемы оптимальных решений. СПб.: БХВ-Петербург, 2005. 400 с.
 5. Типичные причины неоптимальной работы запросов и методы оптимизации // 1С: ИТС. Информационно-технологическое сопровождение пользователей 1С: Предприятие. [Электронный ресурс]. URL: <https://its.1c.ru/db/metod8dev#content:5842:hdoc> (дата обращения: 24.01.2017).
 6. Защита баз данных Oracle /Носова Т.Н., Быкова Т.В., Булатов Р.Р., Михайлова У.В. //Актуальные проблемы современной науки, техники и образования. 2015. Т. 2. № 1. С. 188-191.
 7. Optimizing Database Structure [Электронный ресурс]. Режим доступа: URL: <https://dev.mysql.com/doc/refman/5.7/en/optimizing-database-structure.html> (дата обращения: 24.01.2017).
 8. Скит Д. C# для профессионалов: тонкости программирования. М.: Вильямс, 2014. 610 с.

УДК 004.89

Пименов В.И.,

д.т.н.,

Пименов И.В.,

к.т.н., Санкт-Петербургский государственный университет промышленных технологий и дизайна, Россия, v_pim@mail.ru, <http://sutd.ru>

Университет при Межпарламентской Ассамблее ЕврАзЭС, Санкт-Петербург, Россия, i-pim@mail.ru, <http://www.miep.edu.ru>

МЕТОДИКА АНАЛИЗА БОЛЬШИХ ДАННЫХ В СИСТЕМАХ ПОДДЕРЖКИ ПРИНЯТИЯ РЕШЕНИЙ

Аннотация: В статье рассмотрена методика, основанная на применении многомерного анализа данных, методов и алгоритмов для построения баз знаний, позволяющая повысить степень автоматизации создания интеллектуальных систем поддержки принятия решений.

Ключевые слова: многомерный анализ, большие данные, принятие решений, интеллектуальная система, база знаний.

Pimenov V.I.,

Doctor of Technical Sciences,

Pimenov I.V., Ph.D.

St. Petersburg State University of Industrial Technologies and Design, Russia, v_pim@mail.ru, <http://sutd.ru>

University of the EurAsEC Inter-Parliamentary Assembly, St. Petersburg, Russia, i-pim@mail.ru, <http://www.miep.edu.r>

METHOD OF ANALYSIS OF BIG DATA IN SUPPORT SYSTEMS FOR DECISION-MAKING

Abstract: *The article describes epy technique based on the use of multidimensional data analysis, methods and algorithms for building knowledge bases, which allows to increase the degree of automation in creating intelligent decision-making support systems.*

Keywords: *multidimensional analysis, big data, decision making, intellectual system, knowledge base.*

Введение

Быстрое принятие управленческих решений является важным инструментом конкурентной борьбы. Для повышения их качества нельзя полностью доверяться аналитикам. Оценки экспертов приходится соизмерять со статистическими данными, с оперативной обстановкой на рынке.

Анализ данных, приводящий к выработке конкретной модели восприятия и поведения, – ключевой подход к развитию бизнеса и достижению поставленных целей.

Огромные объемы данных генерируются информационными сайтами, интернетом вещей, различными сервисами обмена файлами и социальными сетями. Анализ больших данных является незаменимым при определении целевой аудитории, ее интересов и активности в процессе разработки маркетинговой стратегии. Однако это лишь малая часть от всего производимого объема данных.

Большие данные генерируются также из внутренней информации предприятий и организаций. При этом 85% всех данных, которыми оперируют сегодня предприятия, являются неструктурированными. Их переизбыток составляет главную проблему современного бизнеса.

Работа с большими данными подразумевает использование инструментов и методов обработки структурированных, неструктурированных, неопределенно структурированных данных огромных объемов и значительного многообразия с целью получения воспринимаемых человеком результатов, эффективных в условиях непрерывного прироста данных, распределенных по многочисленным узлам вычислительной сети.

Бизнес-решение, основанное на анализе больших данных, может быть реализовано в виде портала, каталога, электронного магазина – как система, обеспечивающая функционирование электронного бизнеса, основанного на интернет-технологиях.

Прогностическое моделирование дает возможность выработки рекомендаций на основе образцов покупок, сделанных пользователем интернет-магазина, или на основе сходства пользователей. Зная рейтинговые оценки одного товара, можно с высокой степенью вероятности сделать вывод о рейтинговых оценках другого товара.

При большом количестве клиентов нерационально анализировать их характеристики индивидуально. Создание профилей посетителей веб-

ресурса позволяет объединять их в группы с однородными признаками. После кластеризации можно узнать, какие группы наиболее активны, какие приносят наибольшую прибыль, а также выделить характерные для них признаки.

Идентификация групп пользователей в веб-приложении помогает нацеливать рекламу на конкретную аудиторию, улучшать взаимодействие с отдельными пользователями. Для каждого пользователя можно отображать сообщения от пользователей с похожими взглядами, облегчая создание на портале социальной сети.

Анализ базы данных клиентов сети сотовой связи может использоваться для разработки нового тарифного плана, который наилучшим образом устраивал бы абонентов малочисленного кластера.

Учет введенных ранее данных и прежнего поведения пользователя на протяжении некоторого периода времени требуется для оценки кредитоспособности клиента в приложении, обслуживающем ипотечное кредитование, при отнесении клиента по каким-то признакам к одной из групп риска.

Модель, обеспечивающая качество принимаемого решения, опирается на набор данных, в которых объекты можно определить в терминах значений атрибутов.

Моделирование процессов поиска типовых решений

Формальные качества объекта определяются множеством первичных характеристик. Например, при отнесении нового товара к той или иной товарной группе приходится описывать образ, назначение, конструкцию, эстетическую ценность, эргономичность и другие свойства.

Для выполнения запроса типа “Вывести список товаров-заменителей некоторой продукции” необходимо вывести значения характерных признаков объекта, а затем поиск по ним аналогичных объектов. Для запроса типа “Вывести список потенциальных покупателей некоторого товара” требуется сначала определить список посредников-продавцов, выполняющих продажу данного товара, а затем провести поиск связанных с ними покупателей.

Формулирование запроса осуществляется в диалоге с пользователем, позволяющем гибко определять значимые признаки анализа в зависимости от данных, накопленных от других пользователей.

Вербально формулируемой целью заказчика товара обычно является перечисление заданных предметных характеристик. Показатели качества Y имеют прагматический смысл, обеспечивая удовлетворение обусловленных заказчиком или предполагаемых потребностей: функциональности, стиля, технологичности и др. Качество многомерного объекта определяется также скрытым, неявным взаимодействием между его составными частями. Выявление законов, объясняющих, как значения первичных характеристик X и их сочетания влияют на показатели каче-

ства Y , необходимо при поиске прототипов и оценке их отличительных свойств.

Попытка формализации задачи поиска объектов-прототипов, товаров-заменителей, типовых решений сводится к многомерной математической модели многокритериальной оптимизации: требуется найти такие характеристики изделия X , которые удовлетворяют системе ограничений по используемой среде, условиям эксплуатации, восприятию объекта определенной группой потребителей, и обеспечивающие достижение максимального значения для некоторых функций, определяющих утилитарные потребности и эстетические запросы пользователей Y .

Значительные массивы данных “объекты–свойства”, накопленные в настоящее время в различных предметных областях, позволяют объективизировать оценку предлагаемых потребителю решений, опираясь на значения измеримых признаков найденного прототипа.

Инструментальной базой при извлечении из данных знаний для широкой совокупности объектов-прецедентов являются методы многомерного анализа данных и машинного обучения. Их комплексное использование является сложной системной задачей ввиду разнотипности признаков, наличия взаимосвязей между ними, необходимостью организации данных при передаче между методами. Результат в виде алгоритма распознавания определяет принадлежность объекта к одному из альтернативных классов посредством геометрического представления. Проблемой остается дальнейший переход к продукционным правилам, обеспечивающим семантическую интерпретацию решений в условиях многомерности описаний объектов, когда число атрибутов варьируется от ста до нескольких сотен, а геометрическое представление объекта становится равноудаленным от любого другого объекта (проклятие размерности) [1], и формирование базы знаний при построении интеллектуальной системы (ИС).

В статье рассмотрена методика, основанная на применении многомерного анализа данных, методов и алгоритмов для построения баз знаний, позволяющая повысить степень автоматизации создания интеллектуальных систем поддержки принятия решений.

Подходы к приобретению знаний

Обзор подходов и методов приобретения знаний можно свести к их краткому сравнительному анализу, представленному в таблице 1.

Направление Data Mining представляет целый спектр методов, предназначенных для извлечения из данных знаний различного типа. Он включает две основные группы методов – индуктивные методы машинного обучения и статистические методы. Каждый из методов предназначен для решения определенного круга задач – кластеризации, установления зависимостей между переменными, классификации, выявления закономерностей между связанными событиями. При этом задается мо-

дель заданной структуры (например, нейронная сеть, дискриминантные функции, линейная регрессия), а ее параметры подстраиваются под данные, описывающие наблюдаемые объекты.

Технология экспертных систем направлена на эксплицирование знаний специалистов с помощью ряда коммуникативных и текстологических методов. Ансамбли моделей способствуют поиску наилучшего решения, усложняя при этом интерпретацию результатов.

Вместе с тем, решение трудноформализуемых задач при создании интеллектуальных систем поддержки принятия решений связано с необходимостью выполнения ряда этапов по подготовке и первичной обработке разнородных данных, сжатию признакового пространства, определению структуры классов, построению правил принадлежности к классам, поиску локальных пространств, логического вывода на основе распознающей базы знаний, использованию знаний при проектировании и поиске изделий.

Повышение эффективности создания систем рассматриваемого класса при многоэтапном использовании методов многомерного анализа данных связано, в первую очередь, с необходимостью автоматизированного решения комплекса задач по организации и передаче данных между методами, выбору пространства описаний объектов, построению алгоритма для нахождения правил, обеспечивающих легкую семантическую интерпретацию и возможность автоматического формирования поля знаний.

Таблица 1

Сравнительный анализ подходов и методов приобретения знаний

Название подхода, метода	Автоматизация операций по приобретению знаний					
	Формирование последовательности действий для построения модели (извлечения знания)	Выбор пространства описаний объектов	Организация и передача данных между методами	Формирование поля знаний	Построение алгоритма для нахождения знаний	Семантическая интерпретация
Data Mining (добыча данных, интеллектуальный анализ данных). Knowledge discovery in databases, KDD (обнаружение знаний в базах данных)	+ Обнаружение в данных практически полезных и доступных интерпретации знаний	+	-	-	+ Ассоциативные правила, деревья решений, математические функции (аппроксимации), регрессионные модели и др.	+

Индуктивные методы искусственного интеллекта	+ Построение модели, пригодной для прогнозирования, корректировка свойств модели	+	-	-	+ Решающие правила, деревья решений, нейросетевые модели, эволюционные модели и др.	+
Статистические методы многомерного анализа данных	+ Обнаружение закономерностей, взаимосвязей между объектами	-	-	-	+ Сжатое описание, структура данных, модели зависимостей, модели временных рядов	-
Технология экспертных систем	+ Получение эксплицитных знаний	+	-	+	+ Тезаурусы, онтологические схемы. Структура фреймов, система продукционных правил	+
Ансамбли моделей	- Построение агрегированного классификатора для прогноза	-	-	-	- Комбинация экспертных оценок	-
Методика использования методов многомерного анализа данных	+ Построение распознающих баз знаний и интеллектуальных систем в предметной области	+	+	+	+ Модель знаний об объекте дизайна: описания классов объектов, обобщающие решающие правила, правила вывода на знаниях	+

Единым методологическим подходом для извлечения из данных знаний может служить теория распознавания образов. Исходными данными является массив “объекты–свойства” для достаточного числа образцов. Результат в виде решающего правила описывает принадлежность объекта к одному из классов.

Преобразование данных в процессе обучения интеллектуальной системы

В процессе извлечения знаний из базы данных и их использования в ИС должны выполняться следующие преобразования: данные → информация → решающие правила → знания, закономерности условной логики (conditional logic) → полученные решения (рисунок 1).

Исходные данные представляются посредством описания i -го объекта ω_i , $i = \overline{1, n}$, значениями признаков $(x_{i1}, \dots, x_{iJ}, \dots, x_{iN})$ априорного словаря $X_A = \{X_j | j = \overline{1, N_A}\}$.

Для обеспечения численного представления качественных свойств номинальные и порядковые переменные преобразуют к бинарному типу. В [2] показано, что такие признаки можно обрабатывать как количественные и вместе с ними. Вычисления можно проводить по формулам для количественных признаков с использованием евклидовой метрики.

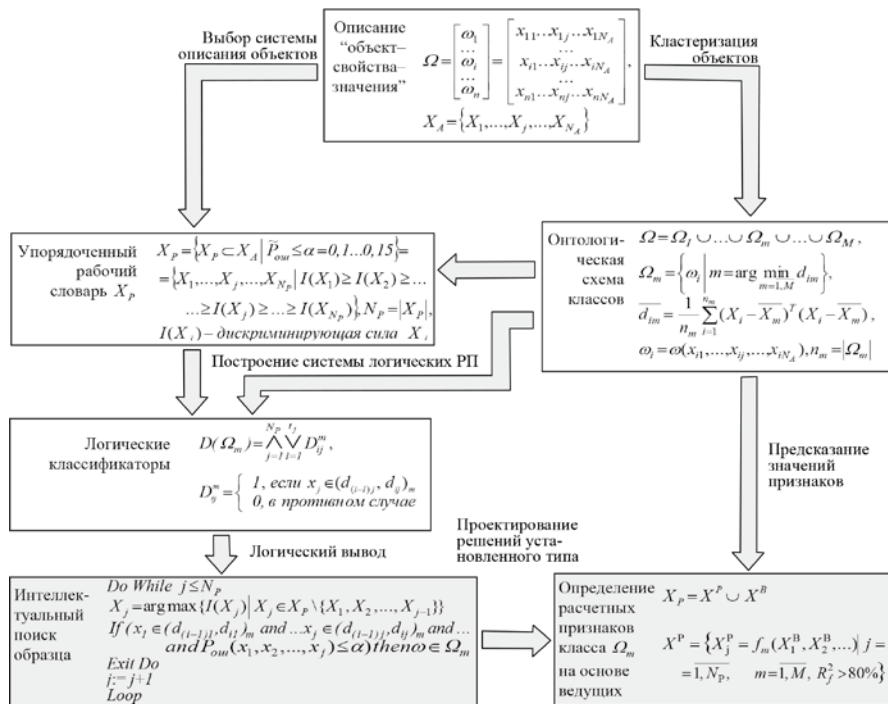


Рисунок 1 – Преобразование данных в процессе обучения и использования интеллектуальной системы

Онтологическая схема $\Omega = \Omega_1 \cup \dots \cup \Omega_m \cup \dots \cup \Omega_M$ определяет принадлежность объекта-прецедента ω_i к одному из классов Ω_m на основе среднего расстояния d_{im} между ω_i и всеми объектами из Ω_m . На следующем этапе выбирается рабочий словарь признаков X_p , упорядоченный по их дискриминирующей силе $I(X_j)$, $j = 1, N_p$, $N_p = |X_p|$, которая характеризует вклад признака в безошибочное разделение объектов обучающей выборки (ОВ).

Дальнейшая обработка заключается в автоматизированном построении решающих правил в виде логических классификаторов $D(\Omega_m)$ для каждого класса Ω_m . Такие классификаторы являются обобщенными описаниями кластеров в терминах отобранных признаков и устанавливают правила D_{ij}^m соответствия i -го интервала j -го признака классу Ω_m , где $(d_{(i-1)j}, d_{ij})_m$ – i -й интервал кодирования признака X_j для Ω_m .

Набор данных, представляющих онтологическую схему классов Ω ,

рабочий словарь X_D и логические классификаторы $D(\Omega_m)$, могут использоваться дальше для формирования ядра ИС.

В интеллектуальной поисковой системе в процессе логического вывода пользователю последовательно предъявляется показатель X_j , обладающий максимальной дискриминирующей силой среди еще не рассмотренной совокупности признаков:

$$X_j = \arg \max \{ I(X_j) \mid X_j \in X_p \setminus \{ X_1, X_2, \dots, X_{j-1} \} \}$$

Процесс останавливается, когда вероятность ошибки распознавания объектов ОВ $P_{ош}(x_1, x_2, \dots, x_j)$ не будет превышать заданного уровня α .

Принимается решение $\omega_i \in U_m$. Использование системы логических правил, раскрывающих закономерности между характеристиками объекта и его принадлежностью к определенной категории позволяет, в отличие от традиционной технологии фактографических баз данных, существенно сократить глубину поиска образца в многомерном признаковом пространстве.

Данное преобразование сводится к разработке алгоритма обработки массива параметров РП, поскольку в процедуре распознавания с помощью логических правил определяется отображение значений признаков в классы объектов, а иерархия классов задает структуру предметной области.

Выбор процедур и метрик

Конечным результатом применения методов многомерного анализа является построение РП в форме бинарной решающей матрицы (БРМ) [3] (рисунок 2). Для заполнения матрицы используются сведения о сочетании значений отобранных информативных признаков предъявляемого объекта и принадлежности его к заданному классу. Разбиение выборки на классы – группы схожих объектов – выполняется посредством кластерного анализа. После выявления стратификационной структуры рассматриваемого набора изделий, с помощью дискриминантного анализа определяются основные признаки, по которым различаются классы, – рабочий словарь X_p . Он позволяет в компактной форме представить центроиды классов, а также построить в каждой выделенной группе собственную модель взаимосвязей, учитывающую индивидуальные особенности группы.

Предварительный отбор признаков – расширенный рабочий словарь, осуществляется с применением метода кластеризации переменных и метода главных компонент.

Если после первичной обработки в распределении объектов обнаруживаются выбросы, для их сглаживания при определении основных кластеров в качестве меры расстояния между объектами может задаваться метрика, отличная от евклидовой.

Сжатие признакового пространства выполняется в два этапа. Сначала методами главных компонент, кластеризации переменных и дискрими-

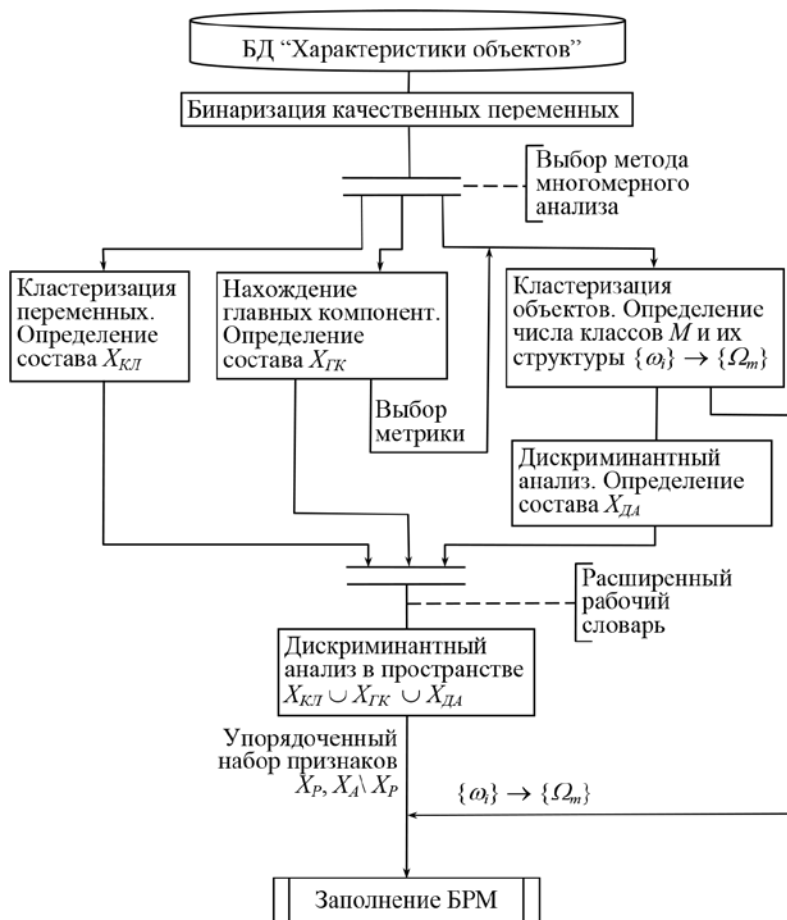


Рисунок 2 – Общая схема использования методов многомерного анализа для сжатия признакового пространства и структуризации данных

нантного анализа последовательно определяются подгруппы исходных признаков $X_{ГК}$, $X_{КЛ}$, $X_{ДА}$, проявляющих большую разделяющую силу.

После выделения кластеров, с помощью дискриминантного анализа определяется набор функций, который позволяет правильно разделить 90–95% наблюдений на заданное число классов. Выявляются наиболее значимые связи между номером кластера и значениями признаков – соответствующая подгруппа признаков $X_{ДА}$ определяется по наибольшим значениям стандартизированных коэффициентов отобранных дискриминантных функций.

На втором этапе выполняется дискриминантный анализ на объединенном множестве отобранных подгрупп признаков $X_{КЛ} \cup X_{ГК} \cup X_{ДА}$.

Ввиду того, что дискриминирующая сила набора признаков не сле-

дует принципу аддитивности, выполняется ранжирование признаков по росту накопленной части безошибочных разделений объектов.

Дискриминантные функции не могут в явном виде использоваться для построения баз знаний, поскольку не содержат сведений о порядке индуктивного логического вывода. Поэтому формируется БРМ, которая может автоматически обрабатываться в процессе логического вывода или транслироваться в поле знаний.

Процедура обучения

Поиск рабочего словаря признаков небольшой размерности, который позволяет безошибочно разделить большую часть объектов ОВ, является основным этапом при построении легко интерпретируемых логических решающих правил, осуществляющих разбиение пространства на области с помощью гиперплоскостей, ортогональных координатным осям.

Логическое правило в форме БРМ строится для априорного словаря, упорядоченного по убыванию информативности бинаризованных признаков (таблица 2).

Принадлежность рабочему словарю X_p	Исходная переменная	Бинарный признак	Ранг признака	Класс			
				Ω_1	Ω_2	...	Ω_M
1	X_N	X_T	1	0	1	...	1
1	X_I	X_I	2	1	0	...	1
...
0	X_j	X_k	r	1	0	...	1
...
0	X_j	X_2	T	0	1	...	0

Число строк матрицы $T = \sum_{j=1}^N t_j$, x_{ij} – i -й интервал кодирования X_j , $i = \overline{1, t_j}$, значение ячейки

$$x_{ij}(m) = \begin{cases} 1, \exists \omega \in \Omega_m, x_j \in (d_{(i-1)j}, d_{ij}), m = \overline{1, M}, i = \overline{1, t_j}, \\ 0, \text{ в противном случае.} \end{cases}$$

Алгоритм заполнения элементов $X_k(m)$ БРМ основан на анализе таблицы с результатами кластеризации данных. Значение $X_k(m)$, $k = \overline{1, T}$, $m = \overline{1, M}$, характеризует наличие или отсутствие бинарного признака X_k у объектов класса Ω_m либо принадлежность значений количественного признака X_j объектов класса Ω_m интервалу кодирования с номером t , тогда $k = t_1 + t_2 + \dots + t_{j-1} + t$, где t_j – число уровней (интервалов кодирования) признака X_j .

В отличие от решающего дерева, формируемого через описания классов по значениям признаков, БРМ заполняется по описаниям интервалов кодирования признаков теми номерами классов, значения признаков которых принадлежат данному диапазону. Распознавание осуществляется

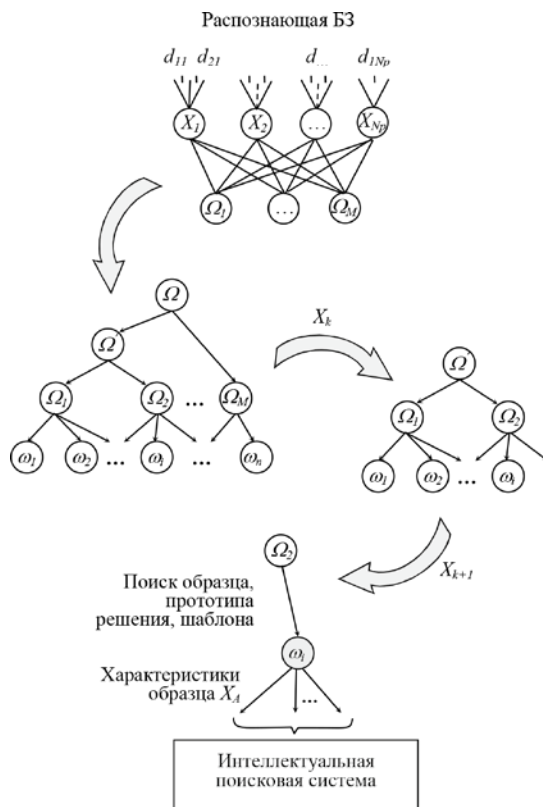


Рисунок 3 – Использование знаний в интеллектуальной поисковой системе

посредством поэлементной конъюнкции ячеек матрицы, на которые указывают значения признаков распознаваемого объекта, и выделения единичной ячейки, соответствующей коду класса. Предъявляемый объект распознается в пространстве “своих” градаций.

Простота процедуры обучения, которая сводится к установке интервальных порогов для количественных признаков и элементов матрицы, позволяет легко реализовать семантическую интерпретацию БРМ. При распознавании достаточно выполнить логические операции и операции сравнения.

Обработка и применение решающего правила

Алгоритм интеллектуального поиска основан на обработке решающего правила. Он реализует процедурную часть базы знаний интеллектуальной системы для выбора шаблона, прототипа решения.

В процессе интеллектуального поиска устанавливаются единичные значения для элементов строки с кодами классов, находящихся в рабочей памяти. Из БРМ выбирается строка с кодами классов и выполняется поэлементная конъюнкция с ячейками рабочей памяти. В них устанавливаются

ся единичные значения, указывающие на претендентов для дальнейшего рассмотрения. Через множества индексов классов, значения признаков которых принадлежат выбранной категории или интервалу, последовательно уточняется множество возможных решений (рисунок 3).

Для объекта характерного типа дается морфологическое описание – перечисляются названия номинальных (интервальных) признаков и их типичные значения.

Также для выбранного кластера указываются количество образцов и их локальные особенности, отражающие специфику кластера. С этой целью анализируется фрагмент матрицы для всех объектов одного кластера.

База знаний, сформированная в процессе обработки результатов многомерного анализа данных, включает знания как декларативного, так и процедурного типов. Декларативная часть содержит знания о структуре предметной области, морфологии объектов и разделительной силе их признаков. Процедурная же часть реализуется как универсальный алгоритм обработки БРМ, использующий вводимые пользователем интеллектуальной системы значения признаков.

Заключение

Методика, основанная на комплексном использовании методов многомерного анализа, обеспечивает при выработке решений извлечение зафиксированных в изготовленных объектах знаний из массива данных “объекты–свойства”. Помимо классификации данных в предметной области, структуризации понятий, представляющих объект, определения значимости признаков, выполнения их ранжирования и отбора, устанавливаются правила, раскрывающие причинно-следственные связи между свойствами изделия и его показателями качества, внешнего вида.

Функции, которые может выполнять обученная интеллектуальная система: нахождение решения (образца, прототипа) без перебора значений всех признаков; умение распознавать сходство ситуации с прошлым опытом (рассуждение с проведением аналогий); оптимизация запросов интеллектуальной базы данных (отбор наиболее эффективного пути следования запросов выполняется на основе знаний об информативности признаков); описание локальных особенностей выбранного морфологического типа, представленных набором входных признаков небольшой размерности для соответствующего кластерного решения.

Список литературы

1. Марманис Х., Бабенко Д. Алгоритмы интеллектуального Интернета. Переводные методики сбора, анализа и обработки данных. – Пер. с англ. – СПб.: Символ-Плюс, 2011. – 480 с.
2. В.М. Ефимов, В.Ю. Ковалева. Многомерный анализ биологических данных: учебное пособие. – СПб, 2008. – 102 с. – pixie.bionet.nsc.ru/ratdna/ckp/pособие_efimov_kovalev.doc.
3. Науман Э. Принять решение – но как?: Пер. с нем. – М.: Мир, 1987. – 198 с.

УДК 511.2

Полещук Ольга Митрофановна

доктор технических наук, профессор, poleshchuk@mgul.ac.ru

Тумор Сергей Владимирович,

аспирант, tumor.sergey@mail.ru

Московский государственный технический университет им. Н.Э.Баумана г. Москва,
Россия

ИСПОЛЬЗОВАНИЕ Z-ЧИСЕЛ (ЧИСЕЛ ЗАДЕ) ДЛЯ ПОДДЕРЖКИ ПРИНЯТИЯ РЕШЕНИЙ

Аннотация: В работе описывается применение чисел Заде для принятия решений в условиях неопределенности. Рассмотренный в статье пример показывает адекватность и эффективность подхода на основе Z-чисел для решения таких задач.

Ключевые слова: число Заде, неопределенность, принятие решений, оптимум по Парето, нечеткое множество.

Poleshchuk Olga Mitrofanovna

Doctor of Technical Sciences, Professor, poleshchuk@mgul.ac.ru

Tumor Sergey Vladimirovich,

graduate student, tumor.sergey@mail.ru

Moscow State Technical University. N.E. Baumana Moscow, Russia

USE OF Z-NUMBER (NUMBER OF GOES) TO SUPPORT DECISION-MAKING

Abstract: This paper describes the use of Zade numbers for decision making under uncertainty. The example considered in the article shows the adequacy and effectiveness of the approach based on Z-numbers for solving such problems.

Keywords: Zade number, uncertainty, decision making, Pareto optimum, fuzzy set.

Введение

В настоящее время все больше внимания уделяется проблеме выбора оптимального решения в условиях неопределенности высокого порядка. Для формализации таких неопределенностей в 2011 году профессор Лотфи Заде ввел понятие Z-числа, которое является упорядоченной парой нечетких чисел $Z=(A, B)$ [1]. В этой паре A – ограничение на значения некоторой нечеткой переменной X (X есть A), а B представляет собой оценку уверенности в том, что X есть A. После общего понятия Z-числа было введено понятие Z^+ числа, которое представляет собой сочетание нечеткого числа A и случайного числа R и записывается в виде упорядоченной пары $Z^+ = (A, R)$. В этой паре A играет ту же роль, что и в Z-числе, а R является распределением вероятностей случайной величины X. Использование введенного профессором Лотфи Заде понятия Z-числа открыло новые возможности для поддержки принятия решений в условиях неопределенности высокого порядка[1, 2].

Основная часть

Дадим необходимые определения.

Согласно [1] нечетким множеством \tilde{A} называется множество пар вида $\{(x, \mu_{\tilde{A}}(x)) : x \in X\}$, где $\mu_{\tilde{A}}(x) : X \rightarrow [0, 1]$ – функции принадлежности \tilde{A} .

Нечеткой переменной называется тройка

$$\{X, U, \tilde{A}\},$$

где X – название переменной; U – область ее определения (универсальное множество); \tilde{A} – нечеткое множество универсального множества, описывающее возможные значения нечеткой переменной [3].

На основе понятия нечеткой переменной вводится понятие лингвистической переменной.

Лингвистической переменной называется пятерка

$$\{X, T(X), U, V, S\}$$

где X – название переменной;

$T(X) = \{X_i, i = \overline{1, m}\}$ – терм-множество переменной X , то есть множество термов или названий лингвистических значений переменной X (каждое из этих значений – нечеткая переменная со значениями из универсального множества U);

V – синтаксическое правило, порождающее названия значений лингвистической переменной X ;

S – семантическое правило, которое ставит в соответствие каждой нечеткой переменной с названием из $T(X)$ нечеткое подмножество универсального множества U .

Термы $X_i, i = \overline{1, m}$ называют понятиями, которые образуют лингвистическую переменную [3]. Функции принадлежности нечетких множеств $\tilde{X}_i, i = \overline{1, m}$, описывающих возможные значения нечетких переменных с названиями $X_i, i = \overline{1, m}$, традиционно называют функциями принадлежности понятий $X_i, i = \overline{1, m}$ или функциями принадлежности термов $X_i, i = \overline{1, m}$. Согласно одному из принципов психолингвистики – принципу лингвистической дополнителности функции принадлежности одних и тех же понятий, но используемыми разными людьми, не обязательно совпадают.

Z – число состоит из упорядоченной пары нечетких чисел $Z = (A, B)$. Первое нечеткое число – это значение, которое принимает некоторая нечеткая переменная X . Второе нечеткое число является показателем надежности (достоверности) первого компонента. Другими словами, второе нечеткое число выражает значение уверенности (возможности, вероятности), что нечеткая переменная X принимает значение A . Как правило, числа A и B описываются на естественном языке. Пример: $Z =$ (около 45 минут, очень уверен).

Рассмотрим формализацию неопределенной информации и процесс принятия решений на основе Z -чисел. Для наглядности в качестве примера будем считать, что некоторая компания представила на рынок два новых программных продукта [4]. У компании есть три варианта, связанные с этими двумя продуктами: она вводит только продукт 1, только продукт 2 или оба эти продукта. Расходы на исследования и разработку для каждого из них составляют соответственно 190 000 и 160 000 рублей. Рост доходов и реакция потребителей повлияют на успех этих продуктов в следующем году. Если компания вводит продукт 1, то доход составит 600 000 рублей, 280 000 рублей и 170 000 рублей, что определяет высокий, средний и низкий уровни, соответственно. Аналогичным образом, когда будет введен продукт 2, доход составит 450 000 рублей, 260 000 рублей и 140 000 рублей, что определяет высокой, средний и низкий уровни, соответственно. Наконец, при введении обоих продуктов 1 и 2 доходы будут составлять 870 000 рублей, 460 000 рублей и 250 000 рублей для высокого, среднего и низкого уровней, соответственно. Эксперты компании вполне уверены, что вероятность высокого и среднего уровня доходов составляет около 0,30 и 0,50 соответственно. Проблема заключается в определении наилучшего решения.

Перейдем к формальному описанию рассматриваемой проблемы. Нечеткая информация в задаче будет формализована Z -числами. Набор вариантов:

$$A = \{f_1, f_2, f_3\},$$

где f_1 обозначает введение продукта 1, f_2 обозначает введение продукта 2, а f_3 обозначает введение обоих продуктов (1 и 2).

Множество возможных состояний это:

$$S = \{S_1, S_2, S_3\},$$

где S_1 обозначает высокий уровень доходов, S_2 обозначает средний уровень доходов, а S_3 обозначает низкий уровень доходов. Вероятности состояний соответствуют Z -числам $Z_{p(S_1)}$ (около 0.3, вполне уверен), $Z_{p(S_2)}$ (около 0.5, вполне уверен).

Набор возможных исходов:

$X = \{(\text{низкий, вероятно}), (\text{выше низкого, вероятно}), (\text{средний, вероятно}), (\text{ниже в сравнении с высоким, вероятно}), (\text{высокий, вероятно})\}$.

Нечеткая информация о значениях лингвистических переменных и вероятностях их появления, показана в таблице 1.

Таблица 1.

	S_1 (около 0.3, вполне уверен)	S_2 (около 0.5, вполне уверен)	S_3 (около 0.2, вполне уверен)
f_1	(высокий, вероятно)	(средний, вероятно)	(низкий, вероятно)
f_2	(ниже в сравнении с высоким, вероятно)	(средний, вероятно)	(низкий, вероятно)
f_3	(высокий, вероятно)	(выше низкого, вероятно)	(низкий, вероятно)

Соответствующая матрица решений на основе Z -чисел показана в таблице 2.

Таблица 2.

	S_1	S_2	S_3
	Z_{41}	Z_{42}	Z_{43}
f_1	Z_{11}	Z_{12}	Z_{13}
f_2	Z_{21}	Z_{22}	Z_{23}
f_3	Z_{31}	Z_{32}	Z_{33}

Построение функций принадлежности Z -чисел осуществляется на основе методов построения функций принадлежностей для обычных нечетких чисел и лингвистических переменных [5-7]. Существуют разные методы построения, связанные со способами получения информации. Исходная информация может быть статистической [6,7] или полученной на основе прямого опроса эксперта или группы экспертов [5]. Функции принадлежности Z -чисел из таблицы 2 строятся на основе методов, разработанных в [5].

Перейдем к решению задачи. Сначала необходимо определить неизвестное Z -число, исходя из вероятности $Z_{p(S3)} = Z_{43} = (A_{43}, B_{43})$ на основании того, что $Z_{p(S1)} = Z_{41}$ и $Z_{p(S2)} = Z_{42}$. Поскольку $Z_{p(S3)}$ полностью определяется через $Z_{p(S1)}$ и $Z_{p(S2)}$, то его степень надежности B_{43} будет такой же, как степень надежности B_{41} и B_{42} . Поэтому для окончательного определения $Z_{43} = (A_{43}, B_{43})$ необходимо вычислить A_{43} на основе A_{41} и A_{42} . Для вычисления A_{43} используем подход, предложенный в [8].

Основываясь на полученных ранее Z -числах, вычислим ожидаемую полезность для каждого из вариантов f_1, f_2, f_3 , следующим образом:

$$Z_U(f_1) = Z_{11} \times Z_{41} + Z_{12} \times Z_{42} + Z_{13} \times Z_{43},$$

$$Z_U(f_2) = Z_{21} \times Z_{41} + Z_{22} \times Z_{42} + Z_{23} \times Z_{43},$$

$$Z_U(f_3) = Z_{31} \times Z_{41} + Z_{32} \times Z_{42} + Z_{33} \times Z_{43}$$

Теперь необходимо выбрать наилучший вариант путем сравнения вычисленных Z -чисел. Для сравнения будем использовать подход, предложенный в [4]. В соответствии с этим принципом сначала получим степени оптимальности (degrees of optimality) трех наших вариантов:

$$do(f_1) = 1, \quad do(f_2) = 0, \quad do(f_3) = 0.92.$$

Как можно заметить, второй вариант не является оптимальным по Парето. Осталось сравнить первый и третий варианты. Предположим, что степень пессимизма при сравнении этих вариантов равна $\beta = 0,3$.

Тогда получим:

$$r(Z_{U(f_1)}, Z_{U(f_3)}) = 0.976 > \frac{1}{2}(do(Z_{U(f_1)}) + do(Z_{U(f_3)})) = 0.96.$$

Таким образом, лучший вариант – f_1 .

Заключение

Решения основаны на информации. Чтобы быть полезной, информация должна быть надежной. Понятие Z-числа непосредственно связано с вопросом обеспечения достоверности информации, поскольку позволяет оперировать с информацией и поддерживать принятие решений в условиях высокого уровня неопределенности. Вычисления с Z-числами можно рассматривать как обобщение вычислений с числами, интервалами, случайными и нечеткими числами и соответственно, как самый высокий уровень вычислений из всех выше перечисленных. Предложенная профессором Л.Заде концепция Z-чисел открывает новые возможности для задач принятия решений в различных сферах деятельности человека. Рассмотренный в статье пример показывает адекватность и эффективность подхода на основе Z-чисел для решения таких задач.

Список литературы

1. Zadeh, L.A. A Note on Z-numbers // Information Sciences. – 2011. – №181. – С. 2923-2932.
2. Полещук О.М., Чернова Т.М. Z-числа и их новые возможности для моделирования реального мира // В сборнике: Современные проблемы физико-математического образования. Сборник материалов VI Международной заочной научно-технической конференции. Государственный гуманитарно-технологический университет. – 2016. – С. 33-35.
3. Полещук О.М. О развитии систем обработки нечеткой информации на базе полных ортогональных семантических пространств // Вестник Московского государственного университета леса – Лесной вестник. 2003. № 1. С. 112-117.
4. Aliev, R.R., Mraiziq, D.A.T., Huseynov, O.H.: Expected utility based decision making under Z-information and its application. Comput. Intell. Neurosci. (2015).
5. Полещук О.М. Математическая модель обработки экспертных оценок // Вестник Московского государственного университета леса – Лесной вестник. 2005. № 6. С. 161-163.
6. Poleshuk O.M., Komarov E. G. Multiple hybrid regression for fuzzy observed data // В сборнике: Annual Conference of the North American Fuzzy Information Processing Society – NAFIPS'2008, – New York, New York. – 2008. С. 4531224.
7. Полещук О.М. Некоторые подходы к моделированию системы управления образовательным процессом // Телекоммуникации и информатизация образования. 2002. № 3. С. 4.
8. Zadeh L. A. Generalized theory of uncertainty (GTU)–principal concepts and ideas. Computational Statistics & Data Analysis. 2006;51(1):15–46. doi: 10.1016/j.csda.2006.04.029.

Научное издание

IT-ТЕХНОЛОГИИ: РАЗВИТИЕ И ПРИЛОЖЕНИЯ

**XV ЕЖЕГОДНАЯ МЕЖДУНАРОДНАЯ
НАУЧНО-ТЕХНИЧЕСКАЯ КОНФЕРЕНЦИЯ**

Сборник докладов

Компьютерная верстка: Провоторова Н.М.

Северо-Кавказский горно-металлургический институт
(государственный технологический университет),
362021, г. Владикавказ, ул. Николаева, 44.

Издательство «Веста»,
362021, г. Владикавказ, ул. Иристонская, 6.
vesta519@mail.ru

Сдано в печать 21.12.2018. Формат 60X 84/16.
Бумага «Снегурочка».
Печать офсетная. Усл. печ.л. 19,73.
Тираж 85 экз. Заказ №3.

Отпечатано в ИП Могилевский Е.С.
344064, г. Ростов-на-Дону, ул. 2-й Пятилетки 23У.

