

**РЕДАКЦИОННАЯ КОЛЛЕГИЯ СЕРИИ
«МАТЕМАТИКА. ИНФОРМАТИКА. ФИЗИКА»**

Рыбаков Юрий Петрович — Заслуженный деятель науки РФ, доктор физико-математических наук, профессор, заведующий кафедрой теоретической физики и механики РУДН — *главный редактор*

Севастьянов Леонид Антонович — доктор физико-математических наук, профессор, профессор кафедры прикладной информатики и теории вероятностей РУДН — *заместитель главного редактора*

Члены редколлегии

Ильгисонис Виктор Игоревич — доктор физико-математических наук, профессор, заведующий кафедрой прикладной физики РУДН, начальник отдела теории плазмы Российского научного центра «Курчатовский институт»

Самуйлов Константин Евгеньевич — доктор технических наук, профессор, заведующий кафедрой прикладной информатики и теории вероятностей РУДН

Скубачевский Александр Леонидович — доктор физико-математических наук, профессор, заведующий кафедрой прикладной математики РУДН

Датта Гупта Субхашиш — профессор Хайдерабадского университета (Индия)

Мартикайнен Олли Эрки — PhD., научный сотрудник Научно-исследовательского института экономики Финляндии (Финляндия)

Медведев М. В. — доктор физико-математических наук, профессор Канзасского университета (США)

Орландо Рафаэль Рамирез Инестрозо — профессор Барселонского университета (Испания)

Саха Биджан — доктор физико-математических наук, ведущий научный сотрудник лаборатории информационных технологий Объединённого института ядерных исследований, г. Дубна (Бангладеш)

Чулуунбаатар Очбадрах — доктор физико-математических наук, ведущий научный сотрудник Института математики, Монгольского Государственного Университета, начальник сектора лаборатории информационных технологий Объединённого института ядерных исследований, г. Дубна (Монголия)

EDITORIAL BOARD
Series “Mathematics. Information Sciences. Physics”

Rybakov Yu. P. — DSc in Physics and Mathematics, professor, Honored Scientist of Russia, Head to the Department of Theoretical Physics and Mechanics of PFUR — *Editor-in-Chief of the series*

Sevastianov L. A. — DSc in Physics and Mathematics, professor, professor of the Department of Applied Probability and Informatics of PFUR — *Vice Editor-in-Chief of the series*

Members of the editorial board

Il’gisonis V. I. — DSc in Physics and Mathematics, professor, Head to the Department of Applied Physics of PFUR, Head to the Department of Plasma Theory in Scientific Center of Russia “Kurchatov Institute”

Samouylov K. E. — DSc in Engineering, professor, Head to the Department of Applied Probability and Informatics of PFUR

Skubachevski A. L. — DSc in Physics and Mathematics, professor, Head to the Department of Applied Mathematics of PFUR

Datta Gupta Subhashish — PhD in Physics and Mathematics, professor of Hyderabad University (India)

Martikainen Olli Erkki — PhD in Engineering, member of the Research Institute of the Finnish Economy (ETLA, Helsinki)

Medvedev M. V. — DSc in Physics and Mathematics, professor of the Kansas University (USA)

Ramirez Inostrozo Orlando Raphael – DSc in Physics and Mathematics, professor of the Barcelona University (Spain)

Saha Bijan – DSc in Physics and Mathematics, leading researcher in Laboratory of Information Technologies of the Joint Institute for Nuclear Research (Dubna, Bangladesh)

Chuluunbaatar Ochbadrah — DSc in Physics and Mathematics, leading researcher in the Institute of Mathematics, State University of Mongolia, Head to the Department in Laboratory of Information Technologies of the Joint Institute for Nuclear Research (Dubna, Mongolia)

**Вестник
Российского
университета
дружбы народов**

НАУЧНЫЙ ЖУРНАЛ

Основан в 1993 году

Серия

Математика. Информатика. Физика

2016, № 1

Российский университет дружбы народов

СОДЕРЖАНИЕ

Прикладная математика

- Матюшенко С. И., Пяткина Д. А., Калениченко В. Н.** Анализ времени пребывания заявок в многоканальной экспоненциальной системе обслуживания с ограниченным накопителем и буфером перепорядочивания 5
- Горбачева А. В., Карамзин Д. Ю.** О некоторых классах задач управления с фазовыми ограничениями 11

Математическое моделирование

- Кулябов Д. С., Геворкян М. Н., Мачука Х. Р., Диаррасуба К., Дали Д. Т. Г.** Численное и имитационное моделирование дисциплин обслуживания очередей типа RED на маршрутизаторе 19
- Шарапова А. А., Кулябов Д. С.** Моделирование распространения электромагнитных волн методом конечных разностей с помощью openEMS 32
- Батгэрэл Б., Никонов Э. Г., Пузынин И. В.** Процедура построения симплектических численных схем для решения гамильтоновых систем уравнений 41

Информатика и вычислительная техника

- Самуйлов К. Е., Ботвинко А. Ю., Зарипова Э. Р.** Оценка времени установления сессии между пользователями при наличии межсетевого экрана 59
- Азофейфа Э. Х., Новикова Г. М.** Нечеткие концептуальные графы для представления знаний в процессно-ориентированной организации 67
- Нгуен Зуй Тхань** Инварианты в задачах распознавания графических образов 76

Физика

Молотков В. И. Лагранжиан лептонного и барионного секторов в нелинейной восьмиспинорной модели	86
Василеска И., Корнеева М. А., Стёпин В. П. Сравнительный анализ спектрально-фотометрических характеристик бытовых источников света	91
Правила оформления статей	101

Bulletin of Peoples' Friendship University of Russia

SCIENTIFIC JOURNAL

Founded on 1993

Series

Mathematics. Information Sciences. Physics
2016, No 1

Peoples' Friendship University of Russia

CONTENTS

Applied Mathematics

- Matyushenko S. I., Pyatkina D. A., Kalenichenko V. N.** Analysis of the Multichannel Exponential System of Service with the Limited Store and Resequencing of Requests 5
- Gorbacheva A. V., Karamzin D. Yu.** On Some Classes of Optimal Control Problem with State Constraints 11

Mathematical Modeling

- Kulyabov D. S., Gevorkyan M. N., Machuca J. R., Diarrasuba K., Dali D. T. G.** Numerical Modeling and Simulation of RED Active Queue Management Algorithms 19
- Sharapova A. A., Kulyabov D. S.** Simulation of Wave Propagation with openEMS 32
- Batgerel B., Nikonov E. G., Puzynin I. V.** A Procedure for Constructing Symplectic Numerical Schemes for Solving of Hamiltonian Systems of Equations 41

Computer Science and Engineering

- Samouylov K. E., Botvinko A. Yu., Zaripova E. R.** Session Setup Time Estimation in the Network with a Firewall 59
- Azofeifa E. J., Novikova G. M.** Fuzzy Conceptual Graphs for Knowledge Representation in Process-Oriented Organizations 67
- Nguyen Duy Thanh** Invariant in the Pattern Recognition 76

Physics

Molotkov V. I. Lagrangian Density of Lepton and Baryon Phases in Nonlinear 8-Spinor Model	86
Vasileska I., Korneeva M. A., Stepin V. P. Comparative Analysis of Spectral-Photometric Characteristics of Domestic Sources of Light	91
Guidelines for Authors	101

Прикладная математика

УДК 519.21

Анализ времени пребывания заявок в многоканальной экспоненциальной системе обслуживания с ограниченным накопителем и буфером переупорядочивания

С. И. Матюшенко, Д. А. Пяткина, В. Н. Калениченко

*Кафедра прикладной информатики и теории вероятностей
Российский университет дружбы народов
ул. Миклуто-Маклая, д. 6, Москва, Россия, 117198*

Рассматривается многоканальная система массового обслуживания с накопителем ограниченной ёмкости, на которую поступает пуассоновский поток заявок. Заявка, застоявшая все места в накопителе занятыми, теряется и в дальнейшем не оказывает влияния на функционирование системы. Длительности обслуживания заявок случайны, независимы между собой и имеют экспоненциальное распределение. При этом интенсивности обслуживания на приборах различны. Заявка, имеющая возможность выбора прибора, выбирает из всех свободных приборов тот, который имеет наибольшую интенсивность обслуживания. На выходе из системы располагается буфер, в котором происходит переупорядочивание заявок в соответствии с порядком их поступления.

Функционирование системы описывается однородным марковским процессом. В предположении, что интенсивности потока и обслуживания заявок конечны, финальные вероятности состояний марковского процесса существуют, строго положительны, не зависят от начального распределения и совпадают со стационарными вероятностями.

В предыдущей работе нами был разработан алгоритм для расчёта стационарных вероятностей состояний рассматриваемой системы. Основная задача данной работы состоит в том, чтобы, опираясь на результаты предыдущей работы, получить стационарные показатели производительности системы. В результате нами была получена функция распределения времени пребывания заявок в буфере переупорядочивания и проведён численный анализ зависимости средней задержки переупорядочивания от загрузки системы и от количества приборов.

Ключевые слова: система массового обслуживания, переупорядочивание заявок, стационарное распределение, функция распределения задержки переупорядочивания, начальные моменты задержки переупорядочивания.

1. Описание системы

Рассматривается многоканальная система массового обслуживания (СМО) с m обслуживающими приборами, $2 \leq m < \infty$, и общим накопителем ограниченной ёмкости. На систему поступает пуассоновский поток заявок с параметром λ . Времена обслуживания на приборе j независимы между собой, а также не зависят от времени обслуживания на других приборах и распределены по экспоненциальному закону с параметром μ_j , $j = \overline{1, m}$.

Накопитель системы имеет r мест для ожидания, $r < \infty$. Заявка, поступающая на систему, когда в ней находится $m+r$ заявок, теряется и в дальнейшем не влияет на её функционирование.

Далее, без ограничения общности, примем, что $\mu_1 \geq \dots \geq \mu_m$, т. е. интенсивность обслуживания заявок с возрастанием номера прибора не возрастает. Примем также, что заявка, имеющая возможность выбора прибора, выбирает из всех свободных приборов тот, который имеет наименьший порядковый номер. Заявки выбираются из очереди на обслуживание в порядке их прибытия в систему, т. е. согласно дисциплине FCFS.

Предполагается, что всем заявкам в момент поступления в систему присваивается порядковый номер. При этом если в момент окончания обслуживания заявки

с номером n (заявки n) продолжается обслуживание хотя бы одной заявки с номером, меньшим n , то заявка n помещается в буфер переупорядочивания (БП). В противном случае заявка n сразу покидает систему, и вслед за ней из БП уходят все заявки с номерами, отличающимися друг от друга на единицу, начиная с номера $n + 1$ (если таковые в этом буфере имеются). Последнее предположение позволяет моделировать механизм сохранения порядка заявок на выходе из системы, в соответствии с которым заявки поступают в неё.

В соответствии с обозначениями Кендалла рассматриваемую СМО будем кодировать как $M/M/m/r/res$, где буквы res являются сокращением от английского $reserquence$ — переупорядочивание.

Предположим, что в любой момент времени заявки, находящиеся в системе (в накопителе и на приборах), пронумерованы в порядке их поступления, начиная с единицы. Тогда стохастическое поведение рассматриваемой СМО можно описать однородным марковским процессом (МП) $X(t)$, $t \geq 0$, над пространством состояний

$$x^m = \bigcup_{k=0}^{m+r} x_k^m,$$

$$x_k^m = \left\{ (k, i_1, \dots, i_m), \quad i_j = \overline{0, k}, \quad \sum_{j=1}^m u(i_j) = k, \right.$$

$$\left. \text{при этом, если } i_j i_s > 0, \text{ то } i_j \neq i_s, \quad j, s = \overline{1, m} \right\}, \quad k = \overline{0, m-1},$$

$$x_k^m = \left\{ (k, i_1, \dots, i_m), \quad i_j = \overline{0, m}, \quad i_j \neq i_s, \quad j, s = \overline{1, m} \right\}, \quad k = \overline{m, m+r},$$

где $u(x)$ — функция Хевисайда.

Здесь для некоторого момента времени t : $X(t) = (k, i_1, \dots, i_m)$, если в системе находится k заявок, $k = \overline{0, m+r}$, $i_j = 0$ означает, что прибор j пуст, в противном случае i_j есть номер заявки, обслуживаемой прибором j , $j = \overline{1, m}$.

В работе [1] был разработан алгоритм расчёта стационарных вероятностей состояний МП $X(t)$. Теперь наша задача состоит в том, чтобы на основе этого распределения получить стационарные характеристики времени пребывания заявок в системе.

2. Стационарная функция распределения задержки переупорядочивания

Обозначим через $F_{res}(t)$ стационарную ФР времени пребывания заявок в БП. В дальнейшем время пребывания заявок в БП будем называть задержкой переупорядочивания.

Перейдём к нахождению $F_{res}(t)$. Для этого рассмотрим условную стационарную ФР задержки переупорядочивания заявки, обслуженной на приборе j , при условии, что в момент $\tau = 0$ окончания обслуживания этой заявки система находилась в состоянии (k, i_1, \dots, i_m) и обозначим её через $F_{(k, i_1, \dots, i_m), j}(t)$, $j = \overline{1, m}$, $(k, i_1, \dots, i_m) \in x^m$.

Ясно, что если непосредственно перед окончанием обслуживания заявки на приборе j система находилась в состоянии (k, i_1, \dots, i_m) , то задержка переупорядочивания этой заявки будет равна максимуму из времён дообслуживания заявок, пришедших в СМО раньше данной заявки, т. е. заявок, обслуживаемых на приборах с номерами из множества

$$y_{(k, i_1, \dots, i_m)} = \{s_n : 0 < i_{s_n} < i_j\}, \quad j = \overline{1, m}, \quad (k, i_1, \dots, i_m) \in x^m.$$

Так как обслуживание заявки на приборе j распределено по экспоненциальному закону с параметром μ_j , $j = \overline{1, m}$, то

$$F_{(k, i_1, \dots, i_m), j}(t) = \prod_{s_n \in Y_{(k, i_1, \dots, i_m), j}} (1 - e^{-\mu_{s_n} t}), \quad j = \overline{1, m}, \quad (k, i_1, \dots, i_m) \in x^m. \quad (1)$$

При этом $F_{(k, i_1, \dots, i_m), j}(t) = 1$, если $Y_{(k, i_1, \dots, i_m), j} = \emptyset$, так как заявка в этом случае не будет задержана.

Теперь обозначим через $\Pi_{D, j}^-(k, i_1, \dots, i_m)$ стационарную вероятность состояния (k, i_1, \dots, i_m) в момент $\tau = 0$ окончания обслуживания заявки на приборе j , а через $\lambda_{D, j}$ — интенсивность выхода заявок, обслуженных прибором j , $j = \overline{1, m}$. Нетрудно видеть, что

$$\lambda_{D, j} = \sum_{(k, i_1, \dots, i_m) \in x^m} p(k, i_1, \dots, i_m) u(i_j) \mu_j, \quad j = \overline{1, m}.$$

Тогда, на основании результатов [2], получаем

$$\Pi_{D, j}^-(k, i_1, \dots, i_m) = \frac{1}{\lambda_{D, j}} p(k, i_1, \dots, i_m) u(i_j) \mu_j.$$

Далее заметим, что вероятность выхода заявки с прибора j равна $\lambda_{D, j} / \lambda_D$, где $\lambda_D = \sum_{j=1}^m \lambda_{D, j}$.

Подытоживая теперь все предыдущие рассуждения и применяя формулу полной вероятности, приходим к искомому результату, который сформулируем в виде теоремы.

Теорема 1. *Функция распределения $F_{res}(t)$ задержки переупорядочивания в СМО $M/M/m/r/res$ в стационарном режиме её работы определяется выражением*

$$F_{res}(t) = \frac{1}{\lambda_D} \sum_{(k, i_1, \dots, i_m) \in x^m} p(k, i_1, \dots, i_m) \sum_{j=1}^m u(i_j) \mu_j, \quad \prod_{s_n \in Y_{(k, i_1, \dots, i_m), j}} (1 - e^{-\mu_{s_n} t}).$$

Обозначим через $\delta_{(k, i_1, \dots, i_m), j, \nu}$ условные начальные моменты порядка ν , $\nu = 1, 2, 3, \dots$, задержки переупорядочивания заявки, обслуженной на приборе j при условии, что в момент $\tau = 0$ окончания её обслуживания система находилась в состоянии (k, i_1, \dots, i_m) , т. е. положим

$$\delta_{(k, i_1, \dots, i_m), j, \nu} = \int_0^{\infty} t^{\nu} dF_{(k, i_1, \dots, i_m), j}(t), \quad j = \overline{1, m}, \quad (k, i_1, \dots, i_m) \in x^m. \quad (2)$$

Тогда из (2) с учётом (1) путём несложных вычислений приходим к следующему результату:

$$\delta_{(k, i_1, \dots, i_m), j, \nu} = \nu! \left[\sum_{n_1=1}^l \frac{1}{\mu_{s_{n_1}}^{\nu}} - \sum_{\substack{n_1, n_2=1 \\ n_1 \neq n_2}} \frac{1}{(\mu_{s_{n_1}} + \mu_{s_{n_2}})^{\nu}} + \dots + \right. \\ \left. + (-1)^{l+1} \frac{1}{(\mu_{s_{n_1}} + \dots + \mu_{s_{n_l}})^{\nu}} \right], \quad (3)$$

где $s_n \in y_{(k, i_1, \dots, i_m), j}$, $i = \overline{1, l}$, $l = |y_{(k, i_1, \dots, i_m), j}|$.

Получив соотношение для моментов $\delta_{(k, i_1, \dots, i_m), j, \nu}$, $(k, i_1, \dots, i_m) \in x^m$, $j = \overline{1, m}$, $\nu = 1, 2, 3, \dots$, можем сформулировать очевидное следствие из теоремы.

Следствие. Начальный момент δ_ν порядка ν задержки переупорядочивания в СМО $M/M/m/r/res$ в стационарном режиме её работы определяется выражением

$$\delta_\nu = \frac{1}{\lambda_D} \sum_{(k, i_1, \dots, i_m) \in x^m} p(k, i_1, \dots, i_m) \sum_{j=1}^m u(i_j) \mu_j \delta_{(k, i_1, \dots, i_m), j, \nu}, \quad \nu \geq 1.$$

3. Примеры численного исследования

Численное исследование показывает, что задержка переупорядочивания во многом зависит не только от суммарной интенсивности μ обслуживания заявок, но и от соотношения интенсивностей обслуживания μ_j , $j = \overline{1, m}$, между собой. Поясним это высказывание следующим примером.

Рассмотрим трёхканальную экспоненциальную систему обслуживания ёмкости $r = 10$. Будем считать, что на систему поступает пуассоновский поток заявок с интенсивностью λ . Далее рассмотрим три варианта параметров μ_j , $j = \overline{1, 3}$. В первом случае возьмём $\mu_1 = \mu_2 = \mu_3 = 1,0$, во втором — $\mu_1 = 1,715$, $\mu_2 = 0,855$, $\mu_3 = 0,430$, а в третьем — $\mu_1 = 2,08$, $\mu_2 = 0,69$, $\mu_3 = 0,23$. Заметим, что суммарная интенсивность обслуживания μ во всех трёх случаях равна трём. Однако в первом случае $\mu_1/\mu_2 = \mu_2/\mu_3 = 1$, во втором — $\mu_1/\mu_2 \approx \mu_2/\mu_3 \approx 2$, а в третьем — $\mu_1/\mu_2 \approx \mu_2/\mu_3 \approx 3$. Проанализируем во всех трёх случаях зависимость средней задержки переупорядочивания δ_1 от загрузки ρ , изменение которой будем достигать за счёт изменения λ .

Характер полученной зависимости отражён на рис. 1. Мы видим, что с ростом ρ наблюдается рост с последующей стабилизацией величины δ_1 . Однако в первом случае средняя задержка переупорядочивания значительно меньше, чем во втором, а тем более в третьем. Таким образом, рассмотренный пример позволяет сделать предположение о том, что чем меньше в системе различие между интенсивностями обслуживания μ_j , $j = \overline{1, m}$, тем меньше в такой системе средняя задержка переупорядочивания.

Рассмотрим ещё один пример, который показывает, как влияет на задержку переупорядочивания количество приборов в системе. Для этого предположим, что в m -канальной системе с накопителем ёмкости $r = 10$ все приборы одинаковы, а длительность обслуживания на j -ом приборе распределена по экспоненциальному закону с параметром $\mu_j = 1$, $j = \overline{1, m}$.

Проанализируем зависимость средней задержки переупорядочивания от загрузки ρ в системах с перечисленным выше фиксированным набором параметров, но с различным количеством приборов m , $m = 2, 3, 4, 5, 6$. Характер полученной зависимости отражён на рис. 2. Как видно из рисунка, при одинаковой загрузке имеется существенное различие между значениями средней задержки переупорядочивания. При этом чем больше приборов в системе, тем больше времени затрачивается на переупорядочивание заявок.

Таким образом, рассмотренные примеры позволяют сделать предположение о том, что наименьшая задержка переупорядочивания присутствует в системах с одинаковыми приборами. В двухканальных системах затраты времени на переупорядочивание заявок минимальны. Данное предположение пока не получило аналитического подтверждения и, по-видимому, может служить объектом дальнейших исследований.

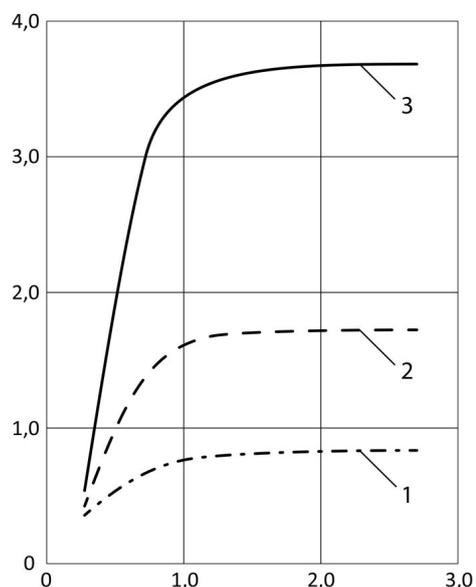


Рис. 1. Зависимость средней задержки переупорядочивания от загрузки системы при различных соотношениях интенсивностей обслуживания: кривая 1: $\mu_1 = \mu_2 = \mu_3 = 1$; кривая 2: $\mu_1 = 1,715$, $\mu_2 = 0,855$, $\mu_3 = 0,430$; кривая 3: $\mu_1 = 2,08$, $\mu_2 = 0,69$, $\mu_3 = 0,23$

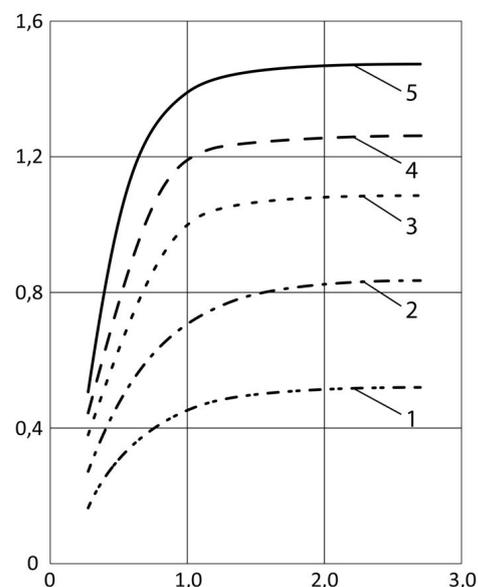


Рис. 2. Зависимость средней задержки переупорядочивания от загрузки в системах с различным количеством приборов m : кривая 1: $m = 2$; кривая 2: $m = 3$; кривая 3: $m = 4$; кривая 4: $m = 5$; кривая 5: $m = 6$

Литература

1. Матюшенко С. И. Анализ многоканальной системы обслуживания с ограниченным накопителем и переупорядочиванием заявок // Вестник Тверского государственного университета, серия «Прикладная математика». — 2010. — Т. 4, № 37. — С. 55–70.
2. Наумов В. А. О предельных вероятностях полумарковского процесса // Современные задачи в точных науках. — М.: УДН, 1975. — 1. — С. 35–39.

UDC 519.21

Analysis of the Multichannel Exponential System of Service with the Limited Store and Resequencing of Requests

S. I. Matyushenko, D. A. Pyatkina, V. N. Kalenichenko

*Department of Applied Probability and Informatics
Peoples' Friendship University of Russia
6, Miklukho-Maklaya str., Moscow, Russia, 117198*

The multi-channel finite-capacity queuing system with Poisson flow is considered. The customer caught all the places in the queue employed, is lost and will not affect the functioning of the system. The service times are random, independent and have exponential distribution. Intensities of service are different. The customer with the possibility of selecting a device, choosing from available devices one, that has the highest intensity of service. On leaving the system there is a buffer in which there is a resequencing of customers according to order of their receipt.

Functioning of the system is described by uniform Markov process. In the assumption that intensities of the flow and service of customers are finite the final probabilities of statuses of Markov process exist, are strictly positive, don't depend on initial distribution and match the stationary probabilities.

An algorithm to calculate the stationary probabilities of the state of the system we developed in previous work. The main objective of this work is to obtain the stationary performance of the system based on the results of previous work. The distribution function of the resequence time was obtained. The numerical analysis of dependence of the average resequence time from the system load and number of devices was considered.

Key words and phrases: queuing system, resequence of requests, stationary distributing, distribution function of the resequence time, initial moments of the resequence time.

References

1. S. I. Matyushenko, Analysis of Multi-Channel Queueing System with Limited Storage and Resequencing Customers, Vestnik of Tver State University, Series "Applied Mathematics" 4 (37) (2010) 55–70, in Russian.
2. V. A. Naumov, About the Limiting Probabilities of the Semi-Markov Process, in: Modern Problems in the Exact Sciences, no. 1, UDN, Moscow, 1975, pp. 35–39, in Russian.

О некоторых классах задач управления с фазовыми ограничениями

А. В. Горбачева*[†], Д. Ю. Карамзин[‡]

** Кафедра нелинейного анализа и оптимизации
Российский университет дружбы народов
ул. Миклухо-Маклая, д. 6, Москва, Россия, 117198*

*† Кафедра прикладной математики
Российский государственный социальный университет
ул. Вильгельма Пика, д. 4, стр. 6, Москва, Россия, 129226*

*‡ Вычислительный центр им. А.А. Дородницына ФИЦ ИУ РАН
ул. Вавилова, д. 40, Москва, Россия, 119333*

В принципе максимума для задач оптимального управления с фазовыми ограничениями возникает борелевская мера-множитель Лагранжа μ . В различных инженерных приложениях, в частности, в некоторых задачах кинематического управления одним из важных вопросов является вопрос о непрерывности или абсолютной непрерывности такой меры. Скорость в подобного рода задачах имеет смысл фазовой переменной. Если модуль скорости ограничен, например, сверху (что вполне естественно в задачах кинематического управления), то это приводит к фазовым ограничениями, и, следовательно, к упомянутой выше мере-множителю Лагранжа μ в необходимых условиях оптимальности. Методы, которые используются для решения таких задач, как правило, подразумевают непрерывность меры. В этой работе рассматриваются примеры задач управления с фазовыми ограничениями, для которых можно гарантировать а priori (то есть без вычисления экстремального процесса), что соответствующая мера непрерывна.

Ключевые слова: оптимальное управление, принцип максимума, фазовые ограничения, борелевская мера, условие Гельдера.

1. Постановка задачи и основные определения

Рассмотрим следующую задачу оптимального управления

$$\begin{cases} K_0(p) + \int_{t_1}^{t_2} f_0(x, u, t) dt \rightarrow \min, \\ \dot{x} = f(x, u, t), \quad t \in [t_1, t_2], \quad t_1 < t_2, \\ \varphi(x, t) \leq 0, \quad R(x, u, t) \leq 0, \\ K_1(p) = 0, \quad K_2(p) \leq 0, \\ p = (x_1, x_2, t_1, t_2). \end{cases} \quad (1)$$

Здесь отображения R , K_i принимают значения в пространствах размерности $d(R)$, $d(K_i)$ соответственно, функции K_0 , f_0 , φ скалярные, $\dot{x} = \frac{dx}{dt}$, $t \in [t_1, t_2]$ – время (концы времени t_1 и t_2 не предполагаются фиксированными), x есть фазовая переменная из R^n , и $u \in R^m$ – переменная управления. Вектор $p \in R^n \times R^n \times R^1 \times R^1$ называется *концевым*. Управляющая функция, или просто *управление*, есть измеримая существенно ограниченная функция $u(\cdot)$, т.е. элемент пространства $L_\infty([t_1, t_2])$.

Пусть $u(t)$, $t \in [t_1, t_2]$ – управление, $x(t)$, $t \in [t_1, t_2]$ – соответствующая траектория, и p – соответствующий концевой вектор. Тройка (p, x, u) называется *допустимым процессом*, если удовлетворены *концевые ограничения* $K_1(p) = 0$, $K_2(p) \leq 0$, *смешанные ограничения* $u(t) \in U(x(t), t)$ для п.в. $t \in [t_1, t_2]$, и *фазовое ограничение* $\varphi(x(t), t) \leq 0 \quad \forall t \in [t_1, t_2]$. Здесь, $U(x, t) = \{u : R(x, u, t) \leq 0\}$.

Статья поступила в редакцию 13 января 2016 г.

Работа выполнена при поддержке Российского фонда фундаментальных исследований (проекты 16-01-00283, 16-31-60005), гранта Президента РФ № МД-4639.2016.1.

Допустимый процесс называется *оптимальным*, если значение минимизируемого функционала является наименьшим возможным на множестве всех допустимых процессов.

Относительно функций, вовлечённых в постановку задачи, предположим следующее: функции $K_0, K_1, K_2, f_0, f, R, \varphi$ непрерывно дифференцируемы, функция φ дважды непрерывно дифференцируема, а функции f, f_0, R дважды непрерывно дифференцируемы по u для всех x, t .

Введём необходимые определения.

Определение 1. Концевые ограничения называются *регулярными* в $p = (x_1, x_2, t_1, t_2)$: $K_1(p) = 0, K_2(p) \leq 0$, если

$$\dim \frac{\partial K_1}{\partial p}(p) = d(K_1), \exists d \in \ker \frac{\partial K_1}{\partial p}(p) : \left\langle \frac{\partial K_2^j}{\partial p}(p), d \right\rangle > 0 \quad \forall j : K_2^j(p) = 0.$$

Определение 2. Смешанные ограничения называются *регулярными*, если для любых (x, u, t) : $R(x, u, t) \leq 0$ существует вектор $q = q(x, u, t)$ такой, что

$$\left\langle \frac{\partial R^j}{\partial u}(x, u, t), q \right\rangle > 0 \quad \forall j \in I(x, u, t) := \{j : R^j(x, u, t) = 0\}. \quad (2)$$

Определение 3. Фазовое ограничение называется *регулярным*, если для любых (x, t) : $\varphi(x, t) = 0$ имеет место $\frac{\partial \varphi}{\partial x}(x, t) \neq 0$.

Определение 4. Фазовое ограничение называется *согласованным* с концевыми ограничениями в точке p^* , если существует число $\varepsilon > 0$ такое, что

$$\{p \in R^{2n+2} : |p^* - p| \leq \varepsilon, K_1(p) = 0, K_2(p) \leq 0\} \subseteq \{p : \varphi(x_1, t_1) \leq 0, \varphi(x_2, t_2) \leq 0\}.$$

Пусть $\xi(t) : R \rightarrow R^m$ заданная измеримая ограниченная функция.

Определение 5. Замыканием справа по мере функции $\xi(t)$ в точке τ называется множество $\Xi^+(\tau)$ таких векторов $u \in R^m$, что

$$\ell(\{t \in [\tau, \tau + \varepsilon] : \xi(t) \in B_\varepsilon(u)\}) > 0 \quad \forall \varepsilon > 0.$$

Здесь, $B_\varepsilon(u) = \{v \in R^m : |v - u| \leq \varepsilon\}$, и ℓ – мера Лебега на R . Соответственно, замыкание слева – это множество $\Xi^-(\tau)$ таких векторов $u \in R^m$, что

$$\ell(\{t \in [\tau - \varepsilon, \tau] : \xi(t) \in B_\varepsilon(u)\}) > 0 \quad \forall \varepsilon > 0.$$

Многозначное отображение $\Xi(t) := \Xi^-(t) \cup \Xi^+(t)$, где $t \in R$, называется замыканием $\xi(t)$ по мере Лебега.

О некоторых свойствах замыкания по мере можно прочесть в [1]. Ниже будем рассматривать оптимальный в задаче (1) процесс (p^*, x^*, u^*) . Замыкание по мере оптимального управления $u^*(t)$ обозначим через $U(t)$.

Определение 6. Процесс (p^*, x^*, u^*) называется *регулярным*, если для всех $t \in T, u \in U(t)$, векторы $\frac{\partial R^j}{\partial u}(u, t), j \in I(u, t)$, линейно независимы, и существует вектор $d = d(u, t) \in R^m$ такой, что $d \in \ker \frac{\partial R^j}{\partial u}(u, t) \quad \forall j \in I(u, t)$,

$$\left\langle \frac{\partial \Gamma}{\partial u}(u, t), d \right\rangle \neq 0 \quad (3)$$

как только $\varphi(x^*(t), t) = \Gamma(u, t) = 0$, где

$$\Gamma(x, u, t) = \left\langle \frac{\partial \varphi}{\partial x}(x, t), f(x, u, t) \right\rangle + \frac{\partial \varphi}{\partial t}(x, t).$$

В общем случае, когда множество $U(t)$ сложно вычислить, достаточно проверить условие (3) для всех $u \in U(t)$. Более того, из условия b) предложения 3.2 из [1] вытекает следующее простое утверждение.

Замечание 1. Любой допустимый процесс в задаче (1) регулярен, как только при всех x, t и всех $u \in U(x, t)$, векторы $\frac{\partial R^j}{\partial u}(x, u, t)$, $j \in I(x, u, t)$, линейно независимы, и существует вектор $d = d(x, u, t) \in R^m$, такой, что $d \in \ker \frac{\partial R^j}{\partial u}(x, u, t)$, $\forall j \in I(x, u, t)$, и

$$\left\langle \frac{\partial \Gamma}{\partial u}(x, u, t), d \right\rangle \neq 0$$

как только $\varphi(x, t) = \Gamma(x, u, t) = 0$.

Напомним определение из [2].

Определение 7. Говорят, что выполняются условия управляемости в конечных точках относительно фазового ограничения, если для $s = 1, 2$,

$$\begin{aligned} \exists f_s \in f(x_s^*, U(x_s^*, t_s^*), t_s^*) : \\ (-1)^s \left[\left\langle \frac{\partial \varphi}{\partial x}(x_s^*, t_s^*), f_s \right\rangle + \frac{\partial \varphi}{\partial t}(x_s^*, t_s^*) \right] > 0 \end{aligned}$$

как только $\varphi(x_s^*, t_s^*) = 0$.

Следующее утверждение есть аналог Леммы 2.1 из [3].

Предложение 1. Пусть процесс (p^*, x^*, u^*) является регулярным. Тогда выполнены условия управляемости.

2. Принцип максимума

Рассмотрим функцию

$$\Gamma(x, u, t) = \left\langle \frac{\partial \varphi}{\partial x}(x, t), f(x, u, t) \right\rangle + \frac{\partial \varphi}{\partial t}(x, t),$$

расширенную функцию Гамильтона–Понтрягина

$$\bar{H}(x, u, \psi, \mu, \lambda^0, t) = \langle \psi, f(x, u, t) \rangle - \mu \Gamma(x, u, t) - \lambda^0 f_0(x, u, t),$$

и малый лагранжиан

$$l(p, \lambda) = \lambda^0 K_0(p) + \langle \lambda^1, K_1(p) \rangle + \langle \lambda^2, K_2(p) \rangle,$$

где $\lambda = (\lambda^0, \lambda^1, \lambda^2)$.

В работе [1] была получена следующая теорема.

Теорема 1. Пусть регулярный процесс (p^*, x^*, u^*) является оптимальным в задаче (1). Предположим, что конечные, фазовые и смешанные ограничения регулярны, и фазовое ограничение согласовано с конечными ограничениями в точке p^* .

Тогда процесс (p^*, x^*, u^*) удовлетворяет принципу максимума, т.е. найдутся вектор $\lambda = (\lambda^0, \lambda^1, \lambda^2) : \lambda^0 \in R, \lambda^1 \in R^{d(K_1)}, \lambda^2 \in R^{d(K_2)}, \lambda^0 \geq 0, \lambda^2 \geq 0, \langle \lambda^2, K_2(p^*) \rangle = 0$, абсолютно непрерывная функция $\psi : [t_1^*, t_2^*] \rightarrow R^n$, функция $\mu : [t_1^*, t_2^*] \rightarrow R^{d(G)}$, и измеримая ограниченная функция $r : [t_1^*, t_2^*] \rightarrow R^{d(R)}$, такие, что

$$\lambda^0 + \left| \psi(t) - \mu(t) \frac{\partial \varphi}{\partial x}(t) \right| > 0 \quad \forall t \in [t_1^*, t_2^*], \quad (4)$$

$$\dot{\psi} = -\frac{\partial \bar{H}}{\partial x}(t) + r(t) \frac{\partial R}{\partial x}(t) \quad \text{н.в. } t, \quad (5)$$

$$\psi(t_s^*) = (-1)^{s+1} \frac{\partial l}{\partial x_s}(p^*, \lambda) + \mu(t_s^*) \frac{\partial \varphi}{\partial x}(t_s^*), \quad s = 1, 2, \quad (6)$$

$$\max_{u \in U(t)} \bar{H}(u, t) = \bar{H}(t) \quad \text{н.в. } t, \quad (7)$$

$$\dot{h} = \frac{\partial \bar{H}}{\partial t}(t) - r(t) \frac{\partial R}{\partial t}(t) \quad \text{н.в. } t, \quad (8)$$

$$h(t_s^*) = (-1)^s \frac{\partial l}{\partial t_s}(p^*, \lambda) - \mu(t_s^*) \frac{\partial \varphi}{\partial t}(t_s^*), \quad s = 1, 2, \quad (9)$$

$$\frac{\partial \bar{H}}{\partial u}(t) = r(t) \frac{\partial R}{\partial u}(t) \quad \text{н.в. } t, \quad (10)$$

$$\langle r(t), R(t) \rangle = 0, \quad r(t) \geq 0 \quad \text{н.в. } t, \quad (11)$$

где $h(t) := \max_{u \in U(t)} \bar{H}(u, t)$.

Более того, функция $h(t)$ абсолютно непрерывна на $[t_1^*, t_2^*]$, а функция μ удовлетворяет следующим свойствам:

- $\mu(t)$ постоянна на каждом интервале времени $S = [s_1, s_2]$, где траектория x^* целиком лежит во внутренней части фазового множества, т.е. когда $\varphi(s) < 0 \quad \forall s \in S$;
- $\mu(t)$ убывает, и $\mu(t_2^*) = 0$;
- $\mu(t)$ непрерывна всюду на $[t_1^*, t_2^*]$ и более того, даже является гельдеровой с показателем $1/2$, т.е.:

$$|\mu(t) - \mu(s)| \leq \text{const} \sqrt{|t - s|} \quad \forall t, s \in [t_1^*, t_2^*]. \quad (12)$$

Выше приняты обозначения и конвенции из [2], Глава 2. Важно отметить, что условие регулярности здесь существенно. Без него $\mu(t)$ может претерпевать разрывы, что показывают соответствующие примеры из [4].

3. Несколько классов задач с фазовыми ограничениями

Нас будут интересовать такие классы задач с фазовыми ограничениями, для которых а priori, т.е. заранее не зная и не вычисляя оптимальный процесс, тем не менее можно гарантировать, что мера множитель Лагранжа $\mu(t)$ из принципа максимума непрерывна всюду на оптимальном отрезке времени. Ясно, что в силу Теоремы 1 так и будет, если гарантировать, что любой допустимый процесс задачи (1) является регулярным. Это зависит от вида ограничений и правой части. Ниже подберём специальные классы задач так, чтобы гарантировать выполнение всех предположений сформулированной теоремы.

Пример 1. Пусть $n = m$, r_1, r_2 – заданные положительные числа, $\phi : R^{2n} \rightarrow R^1$, $\theta : R^n \rightarrow R^n$ – заданные гладкие функции. Рассмотрим задачу:

$$\begin{cases} \int_0^1 \phi(x, u) dt \rightarrow \min, \\ \dot{x} = \theta(x) + u, \\ |u|^2 \leq r_1, |x|^2 \leq r_2, \\ x(0) = x_A, x(1) = x_B. \end{cases}$$

Лемма 1. *Предположим, что*

$$|\langle \theta(x), x \rangle| < \sqrt{r_1 r_2} \quad \forall x : |x|^2 = r_2. \quad (13)$$

Тогда любой допустимый процесс примера 1 регулярен.

Доказательство. Для доказательства воспользуемся Замечанием 1. Покажем, что в примере выполнены все предположения, сформулированные в этом замечании. Тогда в его силу любой допустимый процесс будет регулярным. Действительно, имеем:

$$\varphi(x) = |x|^2 - r_2, \quad R(u) = |u|^2 - r_1, \quad \Gamma(x, u) = \langle 2x, \theta(x) + u \rangle.$$

Покажем, что векторы $\frac{\partial R}{\partial u}(u)$ и $\frac{\partial \Gamma}{\partial u}(x, u)$ линейно независимы на множестве

$$(x, u) : R(u) = 0, \Gamma(x, u) = 0, \varphi(x) = 0.$$

Действительно, имеем, что

$$\frac{\partial R}{\partial u}(u) = 2u, \quad \frac{\partial \Gamma}{\partial u}(x, u) = 2x.$$

Поскольку $\Gamma(x, u) = 0$, то $\langle x, u \rangle = -\langle \theta(x), x \rangle$. Из (13) имеем, что $|\langle x, u \rangle| < \sqrt{r_1 r_2}$. Но $|x| = \sqrt{r_2}$, $|u| = \sqrt{r_1}$, и поэтому последнее неравенство влечёт линейную независимость векторов $\frac{\partial R}{\partial u}(u)$ и $\frac{\partial \Gamma}{\partial u}(x, u)$.

Легко видеть, что выполнены все предположения, сформулированные в Замечании 1. Поэтому любой допустимый процесс является регулярным.

Лемма доказана. \square

Таким образом, в силу Теоремы 1 и Леммы 1 для любого оптимального процесса найдётся гельдерова функция-множитель Лагранжа $\mu(t)$.

Пример 2. Пусть $n = m$, r_1 – заданное положительное число, g – заданный единичный вектор, $\phi : R^{2n} \rightarrow R^1$, $\theta : R^n \rightarrow R^n$ – заданные гладкие функции. Рассмотрим задачу:

$$\begin{cases} \int_0^1 \phi(x, u) dt \rightarrow \min, \\ \dot{x} = \theta(x) + u, \\ |u|^2 \leq r_1, \langle g, x \rangle \leq 0, \\ x(0) = x_A, x(1) = x_B. \end{cases}$$

Лемма 2. *Предположим, что*

$$|\langle \theta(x), g \rangle| < \sqrt{r_1} \quad \forall x : \langle g, x \rangle = 0. \quad (14)$$

Тогда любой допустимый процесс примера 2 регулярен.

Доказательство. Для доказательства воспользуемся Замечанием 1. Покажем, что в примере выполнены все предположения, сформулированные в этом замечании. Тогда в его силу любой допустимый процесс будет регулярен. Действительно, имеем:

$$\varphi(x) = \langle g, x \rangle, \quad R(u) = |u|^2 - r_1, \quad \Gamma(x, u) = \langle g, \theta(x) + u \rangle.$$

Покажем, что векторы $\frac{\partial R}{\partial u}(u)$ и $\frac{\partial \Gamma}{\partial u}(x, u)$ линейно независимы на множестве

$$(x, u) : R(u) = 0, \quad \Gamma(x, u) = 0, \quad \varphi(x) = 0.$$

Действительно, $\frac{\partial R}{\partial u}(u) = 2u$, $\frac{\partial \Gamma}{\partial u}(x, u) = g$. Поскольку $\Gamma(x, u) = 0$, то $\langle g, u \rangle = -\langle \theta(x), g \rangle$. Из (14) имеем, что $|\langle g, u \rangle| < \sqrt{r_1}$, а поскольку $|g| = 1$ и $|u| = \sqrt{r_1}$, то векторы $\frac{\partial R}{\partial u}(u)$ и $\frac{\partial \Gamma}{\partial u}(x, u)$ неколлинеарны.

Легко видеть, что выполнены все предположения, сформулированные в Замечании 1. Поэтому любой допустимый процесс является регулярным.

Лемма доказана. \square

Таким образом, в силу Теоремы 1 и Леммы 2, для любого оптимального процесса найдётся гельдерова функция-множитель Лагранжа $\mu(t)$.

Пример 3. Пусть $n = m$, $k < m$, a – заданное положительное число, g – заданный единичный вектор, $\phi : R^{2n} \rightarrow R^1$, $\theta : R^n \rightarrow R^n$ – заданные гладкие функции. Рассмотрим задачу:

$$\begin{cases} \int_0^1 \phi(x, u) dt \rightarrow \min, \\ \dot{x} = \theta(x) + u, \\ |u^j| \leq a, \quad j = 1, \dots, k, \\ \langle g, x \rangle \leq 0, \\ x(0) = x_A, \quad x(1) = x_B. \end{cases}$$

Лемма 3. Предположим, что

$$\exists j_* > k : g^{j_*} \neq 0. \quad (15)$$

Тогда любой допустимый процесс примера 3 регулярен.

Доказательство. Для доказательства воспользуемся Замечанием 1. Покажем, что в примере выполнены все предположения, сформулированные в этом замечании. Тогда в его силу любой допустимый процесс будет регулярным. Действительно, имеем:

$$\varphi(x) = \langle g, x \rangle, \quad R^j(u) = u^j - a, \quad j = 1, \dots, k,$$

$$R^j(u) = -u^{j-k} - a, \quad j = k+1, \dots, 2k,$$

$$\Gamma(x, u) = \langle g, \theta(x) + u \rangle.$$

Легко видеть, что векторы $\frac{\partial R^j}{\partial u}(u)$, $j \in I(u)$, и $\frac{\partial \Gamma}{\partial u}(x, u)$ линейно независимы на множестве

$$(x, u) : \Gamma(x, u) = 0, \quad \varphi(x) = 0.$$

Действительно, $\frac{\partial R^j}{\partial u}(u)$ есть соответствующий единичный вектор (взятый с плюсом или минусом), у которого все координаты с номером выше, чем k , равны нулю, а $\frac{\partial \Gamma}{\partial u}(x, u) = g$. Поэтому условие (15) сразу влечёт, что векторы неколлинеарны. Значит, выполнены все предположения Замечания 1. Поэтому любой допустимый процесс является регулярным.

Лемма доказана. \square

4. Заключение

В работе были рассмотрены некоторые классы задач управления с фазовыми ограничениями, для которых можно гарантировать непрерывность меры-множителя Лагранжа $\mu(t)$ из принципа максимума. Непрерывность меры является важным фактором, который может быть полезен при отыскании решения задачи с помощью принципа максимума.

Литература

1. Arutyunov A. V., Karamzin D. Y. On Some Continuity Properties of the Measure Lagrange Multiplier from the Maximum Principle for State Constrained Problems // SIAM Journal on Control and Optimization. — 2015. — Vol. 53, No 4. — Pp. 2514–2540.
2. Arutyunov A. V. Optimality Conditions: Abnormal and Degenerate Problems. — Dordrecht/Boston/London: Kluwer Academic Publisher, 2000.
3. Arutyunov A. V., Karamzin D. Y., Pereira F. L. The Maximum Principle for Optimal Control Problems with State Constraints by R. V. Gamkrelidze: Revisited // J. Optim. Theory Appl. — 2011. — Vol. 149. — Pp. 474–493.
4. Zakharov E. V., Karamzin D. Y. On the Study of Conditions for the Continuity of the Lagrange Multiplier Measure in Problems with State Constraints // Differential Equations. — 2015. — Vol. 51, No 3. — Pp. 399–405.

UDC 517.977.52

On Some Classes of Optimal Control Problem with State Constraints

A. V. Gorbacheva^{*†}, D. Yu. Karamzin[‡]

^{*} Department of nonlinear analysis and optimization
Peoples' Friendship University of Russia
6, Miklukho-Maklaya str., Moscow, Russia, 117198

[†] Department of applied mathematics
Russian state social university
4, Wilhelm Pieck str., Moscow, Russia, 129226

[‡] Dorodnicyn Computing Centre of the Russian Academy of Science
40, Vavilova str., Moscow, Russia, 119333

A Borel measure Lagrange multiplier appears in the maximum principle for state constrained problems. The question of continuity or absolute continuity of the measure-multiplier is highly relevant for various applications in particular for some problems of kinematic control. The velocity in such problems is considered as a state variable. As soon as the magnitude of the velocity is bounded, for instance above, (which is quite natural in problems of kinematic control), this leads to the state constraints and to a measure Lagrange multiplier in the necessary optimality conditions. In Control Theory, the methods that are used to solve these conditions often require the continuity of the measure. In this paper, we consider some examples of optimal control problems with state constraints for which one can ensure that this measure is continuous, without a calculation of extremal process.

Key words and phrases: optimal control, maximum principle, state constraints, Borel measure, Hölder condition.

References

1. A. V. Arutyunov, D. Y. Karamzin, On Some Continuity Properties of the Measure Lagrange Multiplier from the Maximum Principle for State Constrained Problems, *SIAM Journal on Control and Optimization* 53 (4) (2015) 2514–2540.
2. A. V. Arutyunov, *Optimality Conditions: Abnormal and Degenerate Problems*, Kluwer Academic Publisher, Dordrecht/Boston/London, 2000.
3. A. V. Arutyunov, D. Y. Karamzin, F. L. Pereira, The Maximum Principle for Optimal Control Problems with State Constraints by R. V. Gamkrelidze: Revisited, *J. Optim. Theory Appl.* 149 (2011) 474–493.
4. E. V. Zakharov, D. Y. Karamzin, On the Study of Conditions for the Continuity of the Lagrange Multiplier Measure in Problems with State Constraints, *Differential Equations* 51 (3) (2015) 399–405.

Математическое моделирование

УДК 004.021, 519.2, 519.6

Численное и имитационное моделирование дисциплин обслуживания очередей типа RED на маршрутизаторе

Д. С. Кулябов, М. Н. Геворкян, Х. Р. Мачука,
К. Диаррассуба, Д. Т. Г. Дали

*Кафедра прикладной информатики и теории вероятностей
Российский университет дружбы народов
ул. Миклухо-Маклая, д. 6, Москва, Россия, 117198*

Целью данной работы является обобщение стохастической модели RED (Random Early Detection) на случай дисциплин обслуживания AURED, SARED и GRED, а также верификация результатов численного моделирования с помощью имитационного моделирования. Стохастическая модель строится на основе системы из трёх стохастических уравнений Ито. Численное решение проводится с помощью стохастических методов Рунге–Кутты со слабой сходимостью второго порядка. Комплекс программ для численного моделирования написан авторами статьи на языке Python версии 3 с использованием библиотек NumPy и SciPy. В статье достаточно подробно описываются компоненты программного комплекса. Для имитационного моделирования авторы использовали открытый пакет программного обеспечения для моделирования компьютерных сетей NS2. В данной статье авторы лишь кратко описывают те моменты, которые касаются дисциплин обслуживания очередей, не останавливаясь на общем описании NS2. Результаты моделирования представлены в виде графиков зависимостей средней и текущей длин очереди от времени. На основе анализа полученных графиков показано, что имитационное и численное моделирование дало качественно соответствующие друг другу результаты.

Ключевые слова: RED, маршрутизация, стохастические дифференциальные уравнения, стохастические численные методы, ns2.

1. Введение

В современных компьютерных сетях большую роль играют системы маршрутизации. Часто именно неэффективная работа маршрутизатора является причиной низкой производительности сети. К современным маршрутизаторам предъявляются строгие требования по обеспечению стабильной работы и должного качества обслуживания сетевого трафика.

Одним из инструментов по управлению трафиком на маршрутизаторе является протокол обслуживания очереди (буфера). В настоящий момент наиболее широко используются дисциплины обслуживания типа RED (Random Early Detection). RED был впервые представлен в статье [1] и в дальнейшем появилось множество усовершенствований исходного алгоритма, в частности алгоритмы ARED, AURED, GRED и SARED [2–4]. Базовые принципы работы всех модификаций алгоритма RED весьма сходны, поэтому имеет смысл говорить о семействе алгоритмов типа RED.

В данной статье представлен комплексный анализ нескольких дисциплин обслуживания типа RED, а именно: ARED [5], AURED [2], GRED [6] и SARED [4]. Для анализа используются методы численного и имитационного моделирования. В первой части статьи изложены основные сведения по изучаемым дисциплинам обслуживания очередей, описываются алгоритмы их функционирования. Далее кратко излагаются стохастическая и детерминированная модели RED, представленные в статьях [7, 8]. Данные модели распространены авторами на изучаемые алгоритмы.

Статья поступила в редакцию 21 февраля 2016 г.

Работа частично поддержана грантами РФФИ № 14-01-00628, 15-07-08795 и 16-07-00556.

Во второй части статьи описывается созданный авторами программный комплекс для численного моделирования дисциплин обслуживания типа RED [9, 10], даётся описание его архитектуры и вычислительных возможностей. Также кратко изложены основные моменты, касающиеся имитационного моделирования дисциплин RED, ARED и GRED в NS2.

В третьей части описываются численный и имитационный эксперименты. Численное моделирование проводится с помощью комплекса программ, описанных во второй части. Имитационное моделирование проводится с помощью пакета программ NS2 (Network Simulator) [11]. На основании проведённых численного и имитационного эксперимента делается вывод о качественном соответствии численной и имитационной моделей.

2. Дисциплины обслуживания очередей

2.1. Алгоритм RED

Рассмотрим алгоритм работы дисциплины Random Early Detection (RED) [1].

Алгоритм 1 Алгоритм RED

```

 $q_{avg} \leftarrow 0$ 
 $count \leftarrow -1$ 

while  $packet$  do
     $q_{avg} \leftarrow QueueAvg( )$ 
    if  $q > 0$  then
         $q_{avg} \leftarrow (1 - w_q)q_{avg} + w_q q$ 
    else
         $m \leftarrow f(time - q_{time})$ 
         $q_{avg} \leftarrow (1 - w_q)^m q_{avg}$ 
    end if

    if  $q_{min} < q_{avg} < q_{max}$  then
         $count \leftarrow count + 1$ 
         $p \leftarrow p_{max}(q_{avg} - q_{min}) / (q_{max} - q_{min})$ 
         $p \leftarrow p / (1 - p \cdot count)$ 

         $Mark(packet, p)$ 
         $count \leftarrow 0$ 
    else if  $q_{avg} \geq q_{max}$  then
         $Mark(packet, 1)$ 
         $count \leftarrow 0$ 
    else
         $count \leftarrow -1$ 
    end if

    if  $q = 0$  then  $q_{time} \leftarrow time$ 
    end if
end while

```

Переменные:

- Q и q — максимальная вместительность очереди и текущая длина очереди;
- \hat{Q} или q_{avg} — средняя длина очереди (avg — average queue size);
- q_{time} — время, с которого очередь находится в пустом состоянии;
- $count$ — число пакетов, начиная с первого помеченного;

Константы:

- w_q — вес очереди;
- q_{min} — минимальный порог заполнения очереди;
- q_{max} — максимальная порог заполнения очереди;
- p_{max} — максимальное значение вероятности сброса p ;

Остальные параметры:

- p — текущая вероятность того, что пакет будет помечен;
- $time$ — текущее время;
- $f(t)$ — линейная функция от времени t .

При использовании RED маршрутизатор вычисляет средний размер очереди, используя низкоуровневый фильтр с экспоненциальным весовым коэффициентом (*экспоненциально взвешенное скользящее среднее*). Средний размер очереди сравнивается с двумя пороговыми значениями — минимальным и максимальным порогом. Если средний размер очереди меньше, чем минимальный порог, то пакеты не маркируются для сброса. Если средний размер очереди превышает максимальный порог, то все прибывающие пакеты маркируются для последующего сброса. При этом гарантируется, что при увеличении числа сброшенных пакетов, средний размер очереди не превысит максимального порога.

Когда средний размер очереди находится в пределах минимального и максимального порогов, то каждый пакет помечается для последующего сброса с вероятностью p_a , где p_a — функция от среднего размера очереди avg . Каждый раз, когда пакет помечается, вероятность того, что он пришёл от конкретного соединения, прямо пропорциональна полосе пропускания этого соединения.

2.2. Алгоритм AURED

Алгоритм AURED (Autonomus RED) [2] вместо фиксированного значения p_{max} изменяет p_{max} периодически, основываясь на оценке уровня производительности.

Алгоритм 2 Алгоритм AURED

```

 $P_k \leftarrow q_k/Q$ 
if  $P_k \leq P_{k-1}$  then
   $op_k \leftarrow -op_{k-1}$ 
else
  if  $op_k > 0$  then
     $p_{max} \leftarrow \alpha + p_{max}$ 
  else
     $p_{max} \leftarrow p_{max} \cdot \beta$ 
  end if
end if
 $P_k \leftarrow P_k$ 
 $op_{k-1} \leftarrow op_k$ 

```

Специфические параметры алгоритма

- P_k — текущий уровень производительности, вычисляется как $P_k = q_k/Q$;
- P_{k-1} — уровень производительности на предыдущем шаге;
- op_k, op_{k-1} — оператор настройки (adjustment operator), предлагается взять $op_0 = 1$;
- α, β — параметры, позволяющие настроить изменение p_{max} . Авторы алгоритма предлагают $\alpha = 0.02$, а $\beta = 0.8$.

2.3. Алгоритм SARED

Задача алгоритма SARED [10] — удерживать очередь в стабильном состоянии и близко к заданному уровню $[q_{min}, q_{max}]$. В SARED для вычисления экспоненциального взвешенного скользящего среднего используется два весовых коэффициента w_{qn} (normal) и w_{qh} (high). На каждом шаге алгоритма корректируется величина \hat{q} по следующим формулам:

$$\hat{q} = \begin{cases} (1 - w_q \text{ normal})\hat{q} + w_q \text{ normal}q, & \hat{q} \in I, \\ (1 - w_q \text{ high})\hat{q} + w_q \text{ high}q, & \hat{q} \notin I, \end{cases}$$

где I обозначает интервал от $q_{min} + 0.4 \cdot (q_{max} - q_{min})$ до $q_{min} + 0.6 \cdot (q_{max} - q_{min})$.

2.4. Алгоритм GRED

В алгоритме GRED (Gentle RED) [6] заданные границы очереди разбиваются на два интервала: $[q_{min}, q_{max}]$ и $[q_{max}, 2q_{max}]$, а также изменяется алгоритм вычисления вероятности сброса.

Алгоритм 3 Алгоритм GRED

```

if  $q_{avg} < q_{min}$  then
   $p = 0$ 
else if  $q_{min} \leq q_{avg} < q_{max}$  then
   $p \leftarrow p_{max}(q_{avg} - q_{min}) / (q_{max} - q_{min})$ 
   $p \leftarrow p / (1 - p \cdot count)$ 
else if  $q_{max} \leq q_{avg} < 2q_{max}$  then
   $p \leftarrow p_{max} + (1 - p_{max}) \cdot (q_{avg} - q_{max}) / q_{max}$ 
   $p \leftarrow p / (1 - p \cdot count)$ 
else if  $q_{avg} \geq 2q_{max}$  then
   $p = 1$ 
end if

```

2.5. Стохастическая и детерминированная модели модуля управления на маршрутизаторе с дисциплиной обслуживания типа RED

Кратко опишем стохастическую модель модуля управления на маршрутизаторе с дисциплиной обслуживания типа RED.

Стохастическая модель представляет собой систему из трёх стохастических дифференциальных уравнений Ито:

$$\begin{cases} dW(t) = \left(\frac{1}{T(t)} - \frac{W^2(t)P(\hat{Q})}{T(t)} \right) dt + \sqrt{\frac{1}{T(t)} + \frac{W^2(t)P(\hat{Q})}{T(t)}} dV^1, \\ dQ(t) = \left(\frac{W(t)}{T(t)} - C \right) dt + \sqrt{\frac{W(t)}{T(t)} - C} dV^2, \\ d\hat{Q}(t) = w_q C(Q(t) - \hat{Q}(t)). \end{cases}$$

Здесь $W(t)$ — размер TCP окна, $Q(t)$ — мгновенная длина очереди, $\hat{Q}(t)$ — экспоненциально взвешенное скользящее среднее длины очереди, $T(t)$ — время двойного оборота (за время двойного оборота приходят все подтверждения на отправленное окно TCP), $P(\hat{Q})$ — функция вычисления вероятности сброса пакета,

C — интенсивность обслуживания, dV^1, dV^2 — винеровские процессы, соответствующие случайным процессам $W(t)$ и $Q(t)$. Вывод и обоснование модели см. в работе [8].

Детерминированная модель, подробно рассмотренная в работе [7], получается из стохастической системы путём отбрасывания стохастического члена.

3. Программное обеспечение для моделирования алгоритмов обслуживания очереди типа RED

3.1. Численное моделирование алгоритмов типа RED

Программный комплекс для численного моделирования дисциплин обслуживания очереди на маршрутизаторе типа RED был написан на языке программирования Python версии 3 с использованием следующих библиотек: `numpy`, `matplotlib` и `scipy`. Весь программный код открыт и доступен по ссылке <https://bitbucket.org/mngev/red-modeling-public>.

3.1.1. Структура программного комплекса

Модуль `red` состоит из нескольких файлов.

- `red.py` — основной файл, в котором находятся все классы, реализующие дисциплины обслуживания очереди типа RED;
- `rungekutta.py` — несколько методов, реализующих детерминированные численные схемы Рунге-Кутты 2,3,4,5 и 6 порядков (используются в `red.py` для решения системы обыкновенных дифференциальных уравнений);
- `sde.py` — набор функций, реализующих стохастические численные схемы типа Рунге-Кутты (также используются в `red.py` для решения системы стохастических дифференциальных уравнений);
- `plot.py` — набор функций для графического представления результатов вычисления;
- `main.py` — сценарии стандартных вычислений (вычислить и нарисовать графики стохастического и детерминированного фазового портрета, решений уравнений и параметров автоколебаний и т.д.);
- `settings.py` — в данном файле несколько функций для обработки конфигурационных `ini` файлов, на основе которых создаются словари python с настройками вычислений.

3.1.2. Описание классов дисциплин обслуживания

При разработке программы использовался объектно-ориентированный подход. В качестве объектов были выбраны дисциплины обслуживания очередей. Так как все они основаны на дисциплине обслуживания RED и отличаются лишь в деталях, то естественным способом их программной реализации было выделение объекта RED в родительский класс и наследование его всеми остальными объектами (ARED, AURED, DSRED, EFRED, POWARED, RARED, SARED, WRED, GRED). При этом подклассы наследуют все методы и атрибуты родительского класса и переопределяют лишь те методы, которые отличают конкретную дисциплину обслуживания от RED. Ни в одном подклассе не вводится новых методов, поэтому достаточно дать *описание родительского класса RED*.

Метод `__init__(self, q_min=0.2, q_max=0.4)` — играет роль конструктора класса RED и инициализирует следующие атрибуты:

- `name` — имя дисциплины обслуживания,
- `Tr` — время прохождения пакета от источника до узла и обратно,
- `wq` — вес очереди,
- `w_max` — максимальный размер TCP-окна,
- `q_min` — минимальное пороговое значение пакетов для алгоритма RED,

- q_max — максимальное пороговое значение пакетов для алгоритма RED,
- c_small — количество обслуживаемых за 1 секунду пакетов,
- R — размер буфера,
- p_max — максимальная вероятность сброса,
- N — количество узлов,
- $delta$ — вычисляется на основе c_small ,
- $time_interval$ — временной интервал интегрирования, кортеж (t_0, T) ,
- $with_pto$ — включить/выключить сброс по тайм-ауту.

Методы класса:

- $C(x)$ — вычисление интенсивности обслуженной нагрузки;
- $P(x)$ — функция вероятности сброса RED применяемая при решении дифференциальных уравнений;
- $P2(x)$ — функция вероятности сброса RED применяемая при решении уравнений автоколебания;
- $Q(x)$ — правая часть уравнения для вычисления мгновенного размера очереди;
- $Qe(x)$ — правая часть третьего уравнения системы, вычисляющая экспоненциально взвешенное скользящее среднее значение мгновенной длины очереди;
- $Ttot(x)$ — время потери по тайм-ауту T_{TO} ;
- $W(x)$ — правая часть первого уравнения системы ОДУ. Функции для вычисления размера TCP-окна;
- $W_sde(x)$ — правая часть первого уравнения системы СДУ. Функции для вычисления размера TCP-окна. Отличие от W в отсутствии знака минус у выражения $x[0]/2.0$;
- $ode(t, x, p)$ — правая часть ОДУ. Три аргумента добавлены для совместности и имеют следующий смысл:
 - t — время,
 - x — переменная в виде массива из трёх элементов $[W(t), Q(t), Qe(t)]$,
 - p — параметры.
- $p_to(x)$ — вероятность сброса пр тайм-ауту;
- $sde_G(x)$ — матрица диффузии (в данном случае она диагональна);
- $sde_f(x)$ — вектор сноса;
- $self_oscillation_equation(x)$ — система из двух уравнений для вычисления параметров автоколебания;
- $solve_ode(step=0.01, x_0=[1.0, 0.0, 0.0])$ — метод решает систему ОДУ, ассоциированную с дисциплиной обслуживания. Для решения используется метод Рунге-Кутты из файла `rungekutta.py`; аргументы:
 - $step$ — шаг,
 - x_0 — начальные значения в виде списка из трёх элементов;
- $solve_sde(step=0.01, x_0=[1.0, 0.0, 0.0])$ — метод решает систему СДУ, ассоциированную с дисциплиной обслуживания. Для решения используется метод из файла `sde.py`; аргументы те же что и у предыдущего метода;
- $solve_self_oscillation_equation_2d(equation_type='ode')$ — решает систему алгебраических уравнений для параметров автоколебания в случае изменения только q_max ;
- $solve_self_oscillation_equation_3d()$ — решает систему алгебраических уравнений для параметров автоколебания в случае изменения и q_max и q_min .

3.1.3. Описание конфигурационных файлов

Конфигурационные файлы позволяют задавать параметры численного моделирования для различных дисциплин обслуживания очереди. Также имеется возможность задать различные параметры для одной и той же дисциплины обслуживания в случае необходимости изучения данной дисциплины при различных параметрах.

Все конфигурационные файлы должны иметь формат `ini` и находиться в директории `settings`. Файл `параметры вычислений.ini` обязателен, он состоит из трёх разделов.

В разделе [Параметры метода] заданы следующие параметры:

- `h` — шаг сетки численного метода,
- `W_0` — начальный размер TCP окна (начальное значение переменной W),
- `Q_0` — начальный размер очереди (начальное значение переменной Q),
- `Qe_0` — начальный средневзвешенный размер очереди (начальное значение переменной \hat{Q}).

В разделе [Прочие настройки] на данный момент указывается формат файла изображения графиков, который необходимо создать. Параметр `формат файла` может принимать значения `pdf` и `png`.

В разделе [Типы вычислений] можно включить или отключить тот или иной тип вычисления:

- Решение детерминированное
- Решение стохастическое
- Спектр детерминированный
- Спектр стохастический
- Автоколебания детерминированные
- Автоколебания стохастические

Также в директории `settings` можно создать файлы настройки параметров дисциплин, для которых предполагается произвести расчёты. Название этого файла может быть произвольным, но его следует обязательно добавить в список `qd_settings`, заданный в файле `main.py`. Все перечисленные в этом списке файлы будут последовательно использованы для вычислений. Если файл в список не включён, то он будет проигнорирован.

Файлы с настройками параметров дисциплин должны обязательно содержать раздел [DEFAULT], в котором перечислены все возможные параметры и им присвоены значения по умолчанию. Все дальнейшие разделы должны начинаться с названия поддерживаемых расчётной программой дисциплин обслуживания (в `red.py` должен присутствовать соответствующий класс). После названия дисциплины должен стоять пробел, а далее может идти любой поясняющий текст.

Если в том или ином разделе не указать вообще ни одного параметра, то будут использованы значения по умолчанию. Название секции при этом убирать из файла не следует, так как в противном случае вычислений для этой дисциплины обслуживания проведено не будет.

Есть возможность для одной и той же дисциплины задать несколько разных наборов параметров, для этого в названии секции надо *через пробел* добавить любое дополнительное слово или число. Например, [RED вариант 1] и [RED вариант 2]. Оно будет использовано в названии файлов с результатами вычислений.

3.2. Имитационное моделирование дисциплины RED, ARED и GRED с помощью NS2

Вся необходимая информация по работе с `ns2` изложена в официальной документации [11] и в примерах, доступных в дистрибутиве. Поэтому здесь мы лишь кратко опишем те моменты, которые касаются дисциплин обслуживания очередей.

В `ns2` существуют реализации трёх разновидностей дисциплины RED: оригинальный алгоритм RED, адаптивный алгоритм ARED и Gentle RED. Выбор конкретного алгоритма осуществляется путём изменения параметров скрипта. Рассмотрим эти параметры подробнее.

- `bytes_`: включает (`true`) или выключает (`false`) режим «byte mode», в котором размер пакетов влияет на вероятность их пометки на сброс;
- `Queue-in-bytes_`: если значение параметра установлено как `true`, то средняя длина очереди будет измеряться в битах. Также при этом параметры `thresh_`

- и `maxthres_` будут измеряться по вычисленному среднему размеру пакетов — `- mean_pktsize_`. По умолчанию устанавливается значение `false`;
- `thres_`: минимальный порог длины очереди q_{\min} ;
- `maxthres_`: максимальный порог длины очереди q_{\max} ;
- `mean_pktsize_`: приблизительная оценка размера пакета в битах. Значение по умолчанию — 500;
- `q_weight_`: весовой фактор w_q , используемый при вычислении средней длины очереди;
- `wate_`: этот параметр позволяет выдерживать интервал между отбрасываемыми пакетами, если установить его значение как `true`;
- `linterm_`: обратное значение параметра p_{\max} . По умолчанию 10;
- `setbit_`: принимает значение `false`, если RED отбрасывает помеченные пакеты. В случае установки значения `true`, в помеченные пакеты добавляется бит перегрузки (congestion bit) — некоторые реализации TCP реагируют на этот бит;
- `drop-tail_`: если значение истина, то при переполнении буфера или при превышении числа пакетов в очереди значения q_{\max} переключатся на алгоритм Drop Tail.

Значения, установленные по умолчанию для параметров `q_weight_`, `maxthres_` и `thres_`, равны соответственно 0,002, 15 и 5. В более поздних реализациях ns2 они вычисляются автоматически.

Параметры ARED и GRED. Чтобы задействовать ARED или GRED необходимо дополнительно задать следующие параметры:

- `adaptive_` — включение или отключение адаптивного алгоритма ARED;
- `alpha_` и `beta_` — значение параметров α и β дисциплины ARED;
- `gentle_` — включение или отключение Gentle RED.

Мониторинг очереди

Одним из наиболее важных объектов в ns-2 является монитор очереди. Он позволяет собирать информацию о длине очереди, о прибывших, уходящих и отброшенных пакетах. Для внедрения монитора между двумя узлами, необходимо добавить следующие строки:

```
#Включаем монитор очереди
set qmon [$ns monitor-queue $R1 $R2 [open qm.tr w] 0.01]
[$ns link $R1 $R2] queue-sample-timeout
```

Объект `monitor-queue` имеет 4 аргумента: первые два определяют соединение, на котором находится очередь, третий — выходной файл, куда будут записываться данные, а четвёртый — частота занесения данных в файл.

Выходной текстовый файл с результатами мониторинга состоит из 11 колонок: время, узел источник, узел приёмник (2 и 3 соответственно узлы, которые определяют очередь), размер очереди в битах, размер очереди в пакетах, число прибывших пакетов, число пакетов, покинувших очередь, число пакетов, отброшенных очередью, число прибывших битов, число битов вышедших из очереди, число отброшенных битов.

Мониторинг очереди RED

Для мониторинга параметров очереди RED (например, между узлами $n2$ и $n3$) необходимо добавить следующие строки кода:

```
set redq [[$ns link $n2 $n3] queue]
set traceq [open red-queue.tr w]
$redq trace curq_
$redq trace ave_
$redq attach $traceq
```

Здесь `curq_` — текущий размер очереди, а `ave_` — средний размер очереди. В результате получим выходной файл, состоящий из трёх колонок. Первая колонка содержит флаг Q (текущий размер очереди) или a (средний размер очереди). Далее следуют время и значение наблюдаемого параметра.

Мониторинг потоков

Файл «монитор потоков» включает более детальную информацию о типе отбрасывания. Он учитывает разницу между ранними отбрасываниями (Early Drops,

ED), т.е. такими, которые произошли благодаря работе алгоритма RED, и отбрасываниями, произошедшими в результате переполнения буфера. Файл имеет следующий формат.

- Колонка 1: время записи информации (данной строки) в файл.
- Колонка 2 и 5: обе колонки дают id потока.
- Колонка 3: `null` (нулевой показатель).
- Колонка 4: вид потока.
- Колонка 6 и 7: источник и пункт назначения потока.
- Колонка 8 и 9: полное число прибывших данных в конкретном потоке в пакетах и в битах.
- Колонка 10 и 11: число ранних отбрасываний из конкретного потока в пакетах и битах.
- Колонка 12 и 13: полное число прибывших данных во всех потоках в пакетах и битах.
- колонка 14 и 15: число ранних отбрасываний во всех потоках в пакетах и байтах.
- Колонки 16 и 17: число отбрасываний обоих типов во всех потоках в пакетах и в битах.
- Колонки 18 и 19: число отбрасываний обоих типов в конкретном потоке в пакетах и битах.

Добавление модулей AURED и SARED

В ns2 отсутствует реализация как AURED, так и SARED. Соответствующие модули были добавлены авторами путём модификации модуля `red.c`. Процесс добавления собственного модуля в ns2 сравнительно прост. Опишем его по шагам (используется дистрибутив `ns-allinone-2.35`):

- в каталоге `ns-allinone-2.35/ns-2.35/queue` создаём файлы с описанием алгоритмов: `aured.cc`, `aured.h`, `sared.cc`, `sared.h` (в нашем случае это были модифицированные копии файлов `red.cc` и `red.h`);
- после этого добавленные модули вносим в соответствующий список в файле `Makefile.in`, находящемся в каталоге `ns-allinone-2.35/ns-2.35`;
- запускаем компиляцию ns путём последовательного запуска двух скриптов `./configure` и `make` из директории `ns-allinone-2.35/ns-2.35`.

4. Численное и имитационное моделирование алгоритмов RED, AURED, SARED и GRED

Для численного моделирования алгоритмов AURED, SARED и GRED был использован вышеописанный программный комплекс. Стохастическая и детерминированная численные модели были распространены на случай алгоритмов AURED, SARED и GRED. Был написан модуль (подкласс), расширяющий программный комплекс и вводящий в него модели вышеперечисленных алгоритмов [9, 10].

Был проведён численный эксперимент с помощью этого комплекса программ, реализующих решение системы стохастических дифференциальных уравнений стохастическим методом Рунге-Кутты слабого порядка 2.0 и сильного порядка 1.0. Полученные результаты были представлены в графическом виде (см. рис. 1, 3, 5, 7, 9).

Также было проведено имитационное моделирование с использованием симулятора NS2. Полученные результаты были представлены в виде графиков зависимости средней и мгновенной длины очереди от времени (см. рис. 2, 4, 6, 8, 10).

- Анализ графиков позволяет утверждать, что численная и имитационная модели дают сходный качественный результат. Средняя длина очереди в начале работы алгоритма резко повышается, а затем убывает и стабилизируется, осциллируя около фиксированного значения.
- Протоколы AURED, SARED и GRED отличаются от RED существенно меньшим колебанием мгновенной длины очереди, что соответствует данным из литературы, посвящённой этим протоколам. Протокол GRED отличается от

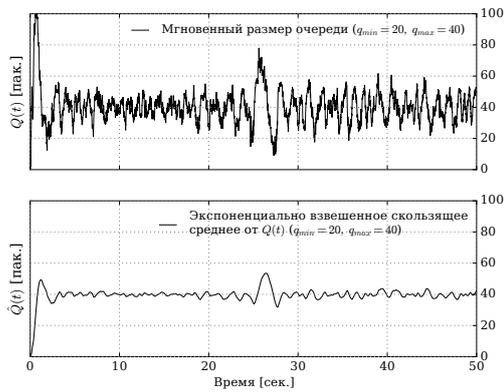


Рис. 1. Результаты численного моделирования алгоритма RED

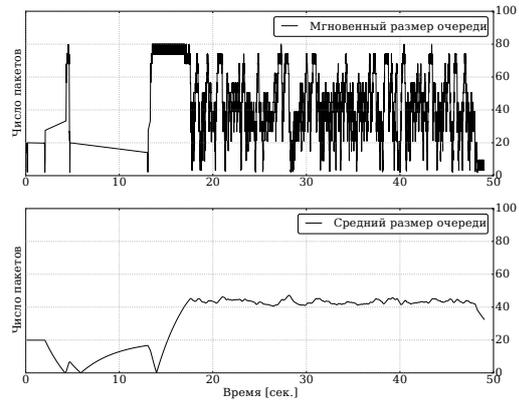


Рис. 2. Результаты имитационного моделирования алгоритма RED

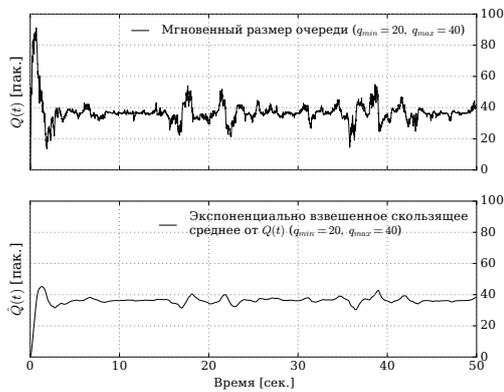


Рис. 3. Результаты численного моделирования алгоритма ARED

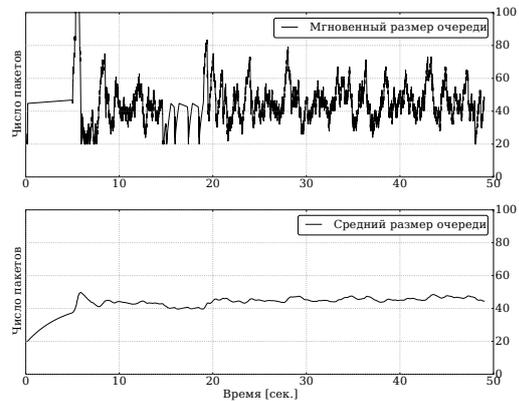


Рис. 4. Результаты имитационного моделирования алгоритма ARED

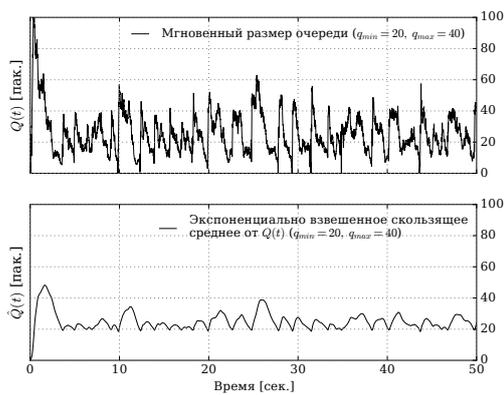


Рис. 5. Результаты численного моделирования алгоритма GRED

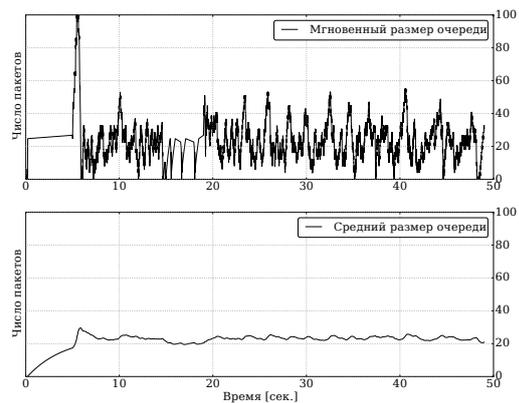


Рис. 6. Результаты имитационного моделирования алгоритма GRED

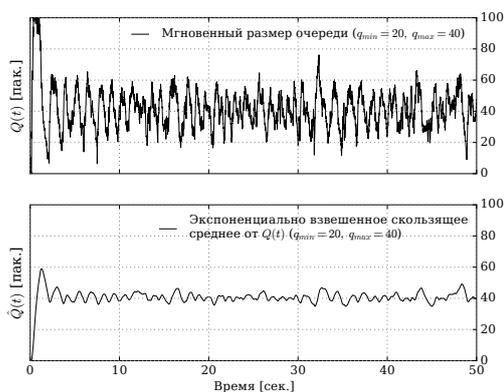


Рис. 7. Результаты численного моделирования алгоритма AURED

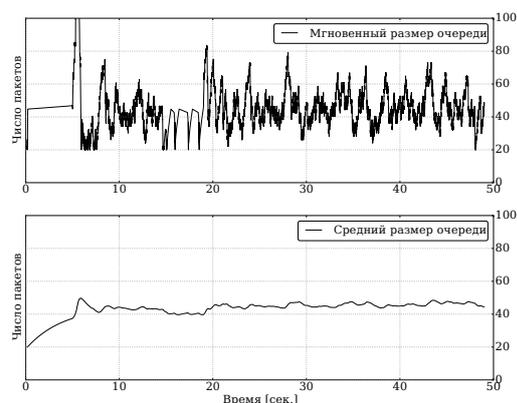


Рис. 8. Результаты имитационного моделирования алгоритма AURED

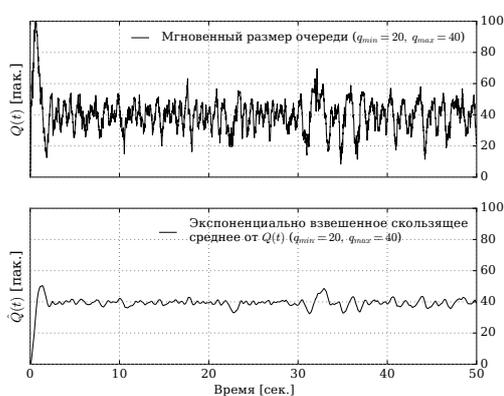


Рис. 9. Результаты численного моделирования алгоритма SARED

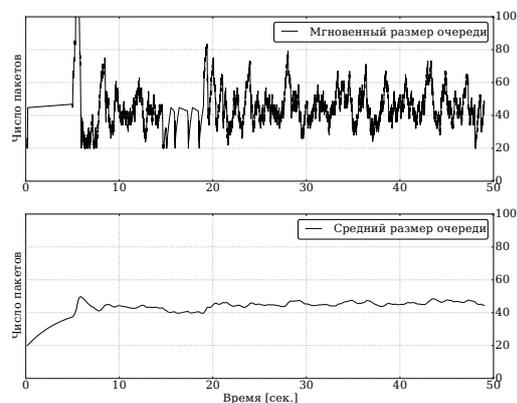


Рис. 10. Результаты имитационного моделирования алгоритма SARED

RED меньшей амплитудой колебания текущей длины очереди в начале своей работы. Меньшие колебания мгновенной длины очереди положительно сказываются на производительности маршрутизатора, так как уменьшается вероятность спонтанного переполнения буфера и переключения маршрутизатора в режим Drop Tail.

Литература

1. *Floyd S., Jacobson V.* Random Early Detection Gateways for Congestion Avoidance // IEEE/ACM Transactions on Networking. — 1993. — No 1. — Pp. 397–413. — <http://www.icir.org/floyd/papers/red/red.html>.
2. *Ho H.-J., Lin W.-M.* AURED — Autonomous Random Early Detection for TCP Congestion Control // Systems and Networks Communications, 2008. ICSNC '08. 3rd International Conference on / IEEE. — Sliema: IEEE, 2008. — Pp. 79–84.
3. Performance Investigations of Some Active Queue Management Techniques Using Simulation / A.-J. Hussein, T. Fadi, D. A. M. et al. // International Journal of New Computer Architectures their Appl. — 2012. — Vol. 2, No 1. — Pp. 286–301.
4. *Javam H., Analoui M.* SARED: Stabilized ARED // Communication Technology, 2006. ICCT '06. International Conference on. — 2006. — Pp. 1–4.

5. Xu Y.-D., Wang Z.-Y., Wang H. ARED: A Novel Adaptive Congestion Controller / IEEE. — Guangzhou: IEEE, 2005. — Pp. 708–714.
6. Digital Enterprise and Information Systems: International Conference / H. Abdeljaber, J. Ababneh, F. Thabtah et al. — Berlin, Heidelberg: Springer Berlin Heidelberg, 2011. — Pp. 592–603.
7. Королькова А. В., Кулябов Д. С. Математическая модель динамики поведения параметров систем типа RED // Вестник РУДН, серия «Математика. Информатика. Физика». — 2010. — № 2. — С. 54–64.
8. Модель управления очередями на маршрутизаторах / Т. Р. Велиева, А. В. Королькова, Д. С. Кулябов, Б. А. Сантуш // Вестник РУДН, серия «Математика. Информатика. Физика». — 2014. — № 2. — С. 81–92.
9. Тьерри Ги Д. Д., Королькова А. В., Геворкян М. Н. Расчёт параметров функционирования модуля активного управления трафиком Adaptive Virtual Queue Random Early Detection. — 2015.
10. Калилу Д., Королькова А. В., Геворкян М. Н. Расчёт параметров функционирования модуля активного управления трафиком Stabilized Adaptiv Random Early Detection. — 2015.
11. Advances in Network Simulation / L. Breslau, D. Estrin, K. Fall et al. // IEEE Computer. — 2000. — Vol. 33, No 5. — Pp. 59–67. — <http://www.isi.edu/~johnh/PAPERS/Bajaj99a.html>.

UDC 004.021, 519.2, 519.6

Numerical Modeling and Simulation of RED Active Queue Management Algorithms

**D. S. Kulyabov, M. N. Gevorkyan, J. R. Machuca,
K. Diarrassuba, D. T. G. Dali**

*Department of Applied Probability and Informatics
Peoples' Friendship University of Russia
6, Miklukho-Maklaya str., Moscow, Russian Federation, 117198*

The aim of this paper is to expand the stochastic model of RED (Random Early Detection) for the case of AURED, SARED and GRED queue service disciplines, as well as the verification of numerical simulation results with NS2 software. A stochastic model is based on the a system of three Ito stochastic equations. The numerical solution is carried out using stochastic Runge–Kutta methods with weak convergence of the second order. Software package for the numerical simulation written by authors in Python version 3 using libraries NumPy and SciPy. The article describes in detail the components of software package. For simulation we use open source software package for modeling Computer Networks NS2. In this article, the authors briefly describe those moments that touch queuing disciplines, not dwelling on the general description of NS2. The simulation results are presented as a plot of the average queue length and the current queue length as functions of time. Based on the analysis of the graphs, it is shown that the simulation and Numerical simulation gave qualitatively related to each other results.

Key words and phrases: RED, routing, stochastic differential equations, stochastic numerical methods, ns2.

References

1. S. Floyd, V. Jacobson, Random Early Detection Gateways for Congestion Avoidance, IEEE/ACM Transactions on Networking (1) (1993) 397–413.
2. H.-J. Ho, W.-M. Lin, AURED — Autonomous Random Early Detection for TCP Congestion Control, in: Systems and Networks Communications, 2008. ICSNC '08. 3rd International Conference on, IEEE, IEEE, Sliema, 2008, pp. 79–84.
3. A.-J. Hussein, T. Fadi, D. A. M., A. Jafar, B. Mahmoud, Performance Investigations of Some Active Queue Management Techniques Using Simulation,

- International Journal of New Computer Architectures their Appl 2 (1) (2012) 286–301.
4. H. Javam, M. Analoui, SARED: Stabilized ARED, in: Communication Technology, 2006. ICCT '06. International Conference on, 2006, pp. 1–4. doi:10.1109/ICCT.2006.341669.
 5. Y.-D. Xu, Z.-Y. Wang, H. Wang, Ared: A novel adaptive congestion controller, IEEE, IEEE, Guangzhou, 2005, pp. 708–714.
 6. H. Abdel-jaber, J. Ababneh, F. Thabtah, A. M. Daoud, M. Baklizi, Digital Enterprise and Information Systems: International Conference, Springer Berlin Heidelberg, Berlin, Heidelberg, 2011, pp. 592–603.
 7. A. V. Korolkova, D. S. Kulyabov, Mathematical Model of the Dynamic Behavior of RED-Like System Parameters, Bulletin of Peoples' Friendship University of Russia. Series "Mathematics. Information Sciences. Physics-(2) (2010) 54–64.
 8. T. R. Velieva, A. V. Korolkova, D. S. Kulyabov, B. A. dos Santos, Model Queue Management on Routers, Bulletin of Peoples' Friendship University of Russia. Series "Mathematics. Information Sciences. Physics-(2) (2014) 81–92.
 9. D. D. G. Thierry, A. V. Korolkova, M. N. Gevorkyan, The Calculation of Operating Parameters of the Module Active Traffic Management Adaptive Virtual Queue Random Early Detection, in russian (2015).
 10. K. Diarrassuba, A. V. Korolkova, M. N. Gevorkyan, The Calculation of Operating Parameters of the Module Active Traffic Management Stabilized Adaptiv Random Early Detection, in russian (2015).
 11. L. Breslau, D. Estrin, K. Fall, S. Floyd, J. Heidemann, A. Helmy, P. Huang, S. McCanne, K. Varadhan, Y. Xu, H. Yu, Advances in Network Simulation, IEEE Computer 33 (5) (2000) 59–67.
URL <http://www.isi.edu/~johnh/PAPERS/Bajaj99a.html>

УДК 537.876+004.94

Моделирование распространения электромагнитных волн методом конечных разностей с помощью openEMS

А. А. Шарапова, Д. С. Кулябов

*Кафедра прикладной информатики и теории вероятностей
Российский университет дружбы народов
ул. Миклухо-Маклая, д. 6, Москва, Россия, 117198*

Существующие методы численного моделирования электромагнитного поля в среде обладают, к сожалению, каждый своими недостатками. Авторы поставили себе задачу проанализировать наиболее популярные методы. В качестве модельной задачи авторами рассматривается линза Люнеберга.

В данной работе авторы рассматривают метод конечных разностей во временной области, программное средство openEMS и его применимость к задачам численного моделирования распространения электромагнитных волн в среде на примере сферической линзы Люнеберга.

Благодаря своей простоте и широким возможностям метод конечных разностей во временной области (Finite-Difference Time-Domain method, FDTD) применяется для решения широкого спектра задач. Существует достаточно большое количество программных инструментов, как с открытым исходным кодом, так и проприетарных, позволяющих производить расчёт этим методом. Программный комплекс openEMS является набором функций для MATLAB или Octave, с помощью которых можно произвести расчёт характеристик электромагнитного поля методом EC-FDTD в декартовых или цилиндрических координатах. Программный комплекс openEMS является бесплатным и имеет открытый исходный код. Поддерживаются параллельные технологии вычисления (MPI).

В данной работе на примере моделирования прохождения электромагнитных волн сквозь сферическую линзу Люнеберга показан процесс работы с openEMS, его установка и настройка, а также даны общие сведения о работе метода FDTD и алгоритма Йи. Приведён пример работы алгоритма Йи. Показан способ анализа и визуализации результатов моделирования с помощью программы ParaView. Приведён исходный код скрипта для моделирования.

Исследованы возможности openEMS и метода FDTD при моделировании распространения электромагнитных волн в среде.

Ключевые слова: Метод конечных разностей во временной области, openEMS, линза Люнеберга.

1. Введение

Метод конечных разностей во временной области (далее FDTD) является, возможно, самым простым и популярным методом численной электродинамики. Как и всякий численный метод, он имеет свои достоинства и недостатки. К его достоинствам помимо простоты относится возможность получить результат для широкого спектра длин волн за один расчёт, возможность задать свойства материала в любой точке расчётной сетки, что позволяет задавать анизотропные, дисперсные и нелинейные среды, возможность наблюдать реальное поведение полей во времени и высокая параллельная эффективность. В тоже время FDTD может быть очень ресурсозатратным методом, особенно при моделировании длинных предметов, таких как, например, провода.

В узком смысле под FDTD понимается использование базового алгоритма Йи для численного решения уравнений Максвелла. Разумеется, с момента выхода статьи Йи он был значительно расширен. Сейчас FDTD включает в себя множество возможностей: моделирование сред с дисперсными и нелинейными свойствами, применение различных типов сеток, постропроцессорная обработка результатов и так далее [1, 2].

Статья поступила в редакцию 1 января 2016 г.

Работа частично поддержана грантами РФФИ № 14-01-00628, 15-07-08795 и 16-07-00556.

Существует достаточно много программных средств, позволяющих производить расчёт методом FDTD. Одним из них является симулятор openEMS [3]. openEMS не является отдельной программой, а представляет набор функций, которые можно использовать в MATLAB или Octave скриптах. В данной работе будет рассмотрен расчёт электромагнитного поля методом EC-FDTD (Equivalent Circuit FDTD) при помощи openEMS.

Работа имеет следующую структуру. В разделе 2 рассматривается классический алгоритм Йи и его применение для расчёта электромагнитного поля в трёхмерном пространстве, а также приведён стандартный алгоритм расчёта методом FDTD. В разделе 3 описываются некоторые особенности работы с пакетом openEMS. В разделе 4 описывается собственно процесс моделирования сферической линзы Лüneберга с помощью электромагнитного симулятора openEMS.

2. Алгоритм Йи

Алгоритм Йи представляет собой схему дискретизации уравнений Максвелла, записанных в дифференциальной форме. Сетки для электрического и магнитного полей смещены по отношению друг к другу на половину шага дискретизации по каждой из пространственных переменных и по времени. В результате узлы, соответствующие компонентам \mathbf{E} , расположены таким образом, что каждый из них окружён четырьмя компонентами \mathbf{H} , и наоборот. Для расчёта значений \mathbf{E} на временном шаге $n + 1/2$ используются значения \mathbf{H} на шаге n . Аналогичным образом значения \mathbf{H} на шаге $n + 1$ рассчитываются с использованием значений \mathbf{E} на шаге $n + 1/2$. Так последовательно рассчитываются все значения полей.

Таким образом, конечно-разностные уравнения позволяют определить электрические и магнитные поля на данном временном шаге на основании известных значений полей на предыдущем, и при заданных начальных условиях вычислительная процедура даёт решение во времени от начала отсчёта с заданным временным шагом.

Ниже описан алгоритм Йи по шагам:

1. Заменить производные в законе Ампера и законе Фарадея конечными разностями.
2. Выразить из полученных дифференциальных уравнений неизвестные значения полей текущего временного шага через уже известные значения полей предыдущего временного шага.
3. Найти значения напряжённости магнитного поля для текущего временного шага.
4. Найти значения напряжённости электрического поля для текущего временного шага.
5. Повторять последние два пункта до тех пор, пока не будут получены значения полей на всём искомом временном промежутке.

2.1. Пример работы алгоритма Йи

Рассмотрим работу метода FDTD поэтапно.

Как уже говорилось, на первом шаге необходимо заменить производные в законах Ампера и Фарадея конечными разностями. Пусть эти уравнения записаны в следующем виде:

$$-\sigma_m \mathbf{H} - \mu \frac{\partial \mathbf{H}}{\partial t} = \nabla \times \mathbf{E}, \quad (1)$$

$$\sigma \mathbf{E} + \varepsilon \frac{\partial \mathbf{E}}{\partial t} = \nabla \times \mathbf{H}. \quad (2)$$

Здесь \mathbf{H} и \mathbf{E} – напряжённости электрического и магнитного полей, μ и ε – диэлектрическая и магнитная проницаемости среды, $\sigma_m \mathbf{H}$ и $\sigma \mathbf{E}$ – плотности электрического тока и её магнитный аналог.

Обозначим компоненты векторов \mathbf{H} и \mathbf{E} следующим образом:

$$H_x(x, y, z, t) = H_x(m\Delta_x, n\Delta_y, p\Delta_z, q\Delta_t) = H_x^q[m, n, p], \quad (3)$$

$$H_y(x, y, z, t) = H_y(m\Delta_x, n\Delta_y, p\Delta_z, q\Delta_t) = H_y^q[m, n, p], \quad (4)$$

$$H_z(x, y, z, t) = H_z(m\Delta_x, n\Delta_y, p\Delta_z, q\Delta_t) = H_z^q[m, n, p], \quad (5)$$

$$E_x(x, y, z, t) = H_x(m\Delta_x, n\Delta_y, p\Delta_z, q\Delta_t) = E_x^q[m, n, p], \quad (6)$$

$$E_y(x, y, z, t) = H_y(m\Delta_x, n\Delta_y, p\Delta_z, q\Delta_t) = E_y^q[m, n, p], \quad (7)$$

$$E_z(x, y, z, t) = H_z(m\Delta_x, n\Delta_y, p\Delta_z, q\Delta_t) = E_z^q[m, n, p]. \quad (8)$$

Здесь x, y, z — координаты узлов в пространстве, а t — координата по времени. Соответственно $\Delta_x, \Delta_y, \Delta_z$ и Δ_t — сеточные шаги по соответствующим направлениям.

Как уже было сказано выше, существуют различные типы расчётных сеток, но мы будем пользоваться сеткой \mathbb{Y} и как самой распространённой. Расположение узлов в сетке \mathbb{Y} и показано на рис. 1.

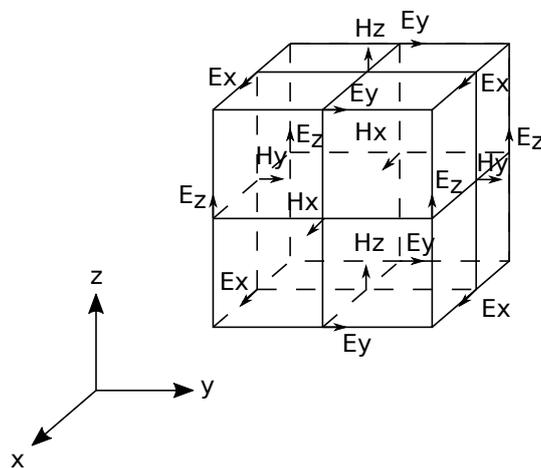


Рис. 1. Расположение узлов в трёхмерной сетке \mathbb{Y} и

Распишем теперь уравнения (1) и (2) покомпонентно:

$$-\sigma_m H_x - \mu \frac{\partial H_x}{\partial t} = \frac{\partial E_z}{\partial y} - \frac{\partial E_y}{\partial z}, \quad (9)$$

$$-\sigma_m H_y - \mu \frac{\partial H_y}{\partial t} = \frac{\partial E_x}{\partial z} - \frac{\partial E_z}{\partial x}, \quad (10)$$

$$-\sigma_m H_z - \mu \frac{\partial H_z}{\partial t} = \frac{\partial E_y}{\partial x} - \frac{\partial E_x}{\partial y}, \quad (11)$$

$$\sigma E_x + \varepsilon \frac{\partial E_x}{\partial t} = \frac{\partial H_z}{\partial y} - \frac{\partial H_y}{\partial z}, \quad (12)$$

$$\sigma E_y + \varepsilon \frac{\partial E_y}{\partial t} = \frac{\partial H_x}{\partial z} - \frac{\partial H_z}{\partial x}, \quad (13)$$

$$\sigma E_z + \varepsilon \frac{\partial E_z}{\partial t} = \frac{\partial H_y}{\partial x} - \frac{\partial H_x}{\partial y}. \quad (14)$$

Как можно видеть из этих уравнений, производная по времени каждого компонента выражена через производные по направлениям другого поля.

Теперь можно заменить частные производные конечными разностями. Рассмотрим уравнение (9):

$$\begin{aligned}
& \frac{H_x^{q+\frac{1}{2}} \left[m, n + \frac{1}{2}, p + \frac{1}{2} \right] + H_x^{q-\frac{1}{2}} \left[m, n + \frac{1}{2}, p + \frac{1}{2} \right]}{2} - \sigma_m \\
& - \mu \frac{H_x^{q+\frac{1}{2}} \left[m, n + \frac{1}{2}, p + \frac{1}{2} \right] - H_x^{q-\frac{1}{2}} \left[m, n + \frac{1}{2}, p + \frac{1}{2} \right]}{\Delta_t} = \\
& = \frac{E_z^q \left[m, n + 1, p + \frac{1}{2} \right] - E_z^q \left[m, n, p + \frac{1}{2} \right]}{\Delta_y} - \\
& - \frac{E_y^q \left[m, n + \frac{1}{2}, p + 1 \right] - E_y^q \left[m, n + \frac{1}{2}, p \right]}{\Delta_z}. \quad (15)
\end{aligned}$$

Отсюда выразим искомое значение $H_x^{q+\frac{1}{2}} \left[m, n + \frac{1}{2}, p + \frac{1}{2} \right]$:

$$\begin{aligned}
H_x^{q+\frac{1}{2}} \left[m, n + \frac{1}{2}, p + \frac{1}{2} \right] &= \frac{1 - \frac{\sigma_m \Delta_t}{2\mu}}{1 + \frac{\sigma_m \Delta_t}{2\mu}} H_x^{q-\frac{1}{2}} \left[m, n + \frac{1}{2}, p + \frac{1}{2} \right] + \\
& + \frac{\Delta_t}{\mu \Delta_y} \left(E_y^q \left[m, n + \frac{1}{2}, p + 1 \right] - E_y^q \left[m, n + \frac{1}{2}, p \right] \right) - \\
& - \frac{1}{1 + \frac{\sigma_m \Delta_t}{2\mu}} \frac{\Delta_t}{\mu \Delta_z} \left(E_z^q \left[m, n + 1, p + \frac{1}{2} \right] - E_z^q \left[m, n, p + \frac{1}{2} \right] \right). \quad (16)
\end{aligned}$$

Точно также следует поступить с уравнениями (10) и (11). После того, как новые значения напряжённости магнитного поля будут найдены, их можно будет использовать для нахождения следующих значений напряжённости электрического поля (уравнения (12)–(14)). Процесс продолжается до тех пор, пока не будут найдены значения \mathbf{H} и \mathbf{E} на всем искомом временном промежутке.

2.2. Порядок расчёта методом FDTD

Порядок расчёта методом FDTD следующий:

1. Задание расчётной области, сетки и граничных условий.
2. В расчётной области задаются материальные тела, составляющие интересующую нас структуру и их свойства.
3. Задание источника излучения.
4. Задание «детектора», отслеживающего изменения полей в нужных точках. Детектор — это не обязательно материальное тело, им может служить просто набор точек.
5. Источник генерирует электромагнитную волну нужного диапазона. Волна падает на тела и рассеивается на них. История распространения волны фиксируется детекторами.

3. Программный пакет моделирования распространения электромагнитных волн openEMS

Программный комплекс openEMS — это электромагнитный симулятор с открытым исходным кодом, использующий для расчёта электромагнитного поля метод EC-FDTD. В отличие от многих других средств оптического моделирования, например, рассмотренных нами в [4], openEMS не является отдельной программой, но набором функций для MATLAB или Octave, написанных на C++.

Рассмотрим примеры использования openEMS для решения задачи моделирования распространения электромагнитных волн в различных структурах.

3.1. Конфигурирование openEMS

Поскольку для вычислений пакет openEMS использует программные комплексы MATLAB или Octave, необходимо сконфигурировать эти комплексы так, чтобы они могли использовать ресурсы openEMS. Для этого надо добавить путь к библиотекам openEMS (см. [5]):

```
addpath('~/opt/openEMS/share/openEMS/matlab');
addpath('~/opt/openEMS/share/CSXCAD/matlab');
```

В случае MATLAB следует добавить строчки в файл `startup.m`, а в случае Octave — в файл `.octaverc`.

3.2. Визуализация с помощью ParaView

Визуализировать результаты симуляции можно с помощью программы ParaView. Её можно установить из стандартных репозиториях. В ParaView необходимо открыть папку, в которой сохранен результат работы скрипта, и выбрать `.vtr` файл (на самом деле их несколько, но в обозревателе они представлены как дерево). Затем нажать кнопку Apply под списком Properties. Чтобы визуализировать амплитуду электромагнитной волны (далее для краткости ЭМВ), нужно в выпадающем списке Coloring выбрать E-field (по умолчанию там стоит SolidColor). Чтобы просмотреть анимацию, нужно нажать кнопку Play на панели инструментов. Также ParaView позволяет анализировать результаты при помощи графиков. Например, зависимость амплитуды ЭМВ от времени в определённой точке расчётной сетки или в зависимости от расстояния [6].

4. Моделирование распространения электромагнитных волн через сферическую линзу Люнеберга

4.1. Постановка задачи

Теперь рассмотрим задачу моделирования сферической линзы Люнеберга в openEMS. Ранее мы уже рассматривали некоторые теоретические аспекты проектирования сферической линзы Люнеберга [7], а также моделирование такой линзы из кубиков, но в рамках геометрической оптики [8, 9]. Как известно, относительная диэлектрическая проницаемость в линзе Люнеберга зависит от расстояния до центра линзы и имеет вид

$$\varepsilon_r(r) = 2 - \left(\frac{r}{R}\right)^2 \quad (0 \leq r \leq R). \quad (17)$$

Свойство линзы таково, что она преобразовывает сферический фронт волны, расходящийся из точки на её поверхности, в плоский фронт [10–12].

Пусть линза имеет диаметр 10λ , где λ — длина волны. Таким образом, чтобы рассчитать диаметр, нужно знать длину волны, но в openEMS можно задать

только частоту излучения. Для нахождения оставшегося параметра воспользуемся формулой

$$\lambda = \frac{c}{\nu}. \quad (18)$$

При задании параметров излучения в openEMS необходимо иметь ввиду следующее ограничение:

$$\max(\Delta_{x,y,z}) < \frac{\lambda_{\min}}{15}. \quad (19)$$

Зададим точечный источник излучения на поверхности линзы. Пусть излучение имеет частоту 10 МГц и длина волны равна 30 м (радиоволны).

Значения электрического поля зафиксированы в плоскости XZ .

4.2. Анализ скрипта openEMS

Ранее мы уже рассматривали синтаксис openEMS на примере более простой модели, волновода с двумя металлическими стенками [13]. Несмотря на то, что линза Лüneберга является более сложной структурой, схема описания модели остаётся той же. Рассмотрим скрипт построчно. Задание пространства FDTD с 300 временными шагами:

```
FDTD = InitFDTD('NrTS',300, 'EndCriteria',0,'OverSampling',50);
```

Задание частоты излучения (10 МГц):

```
freq = 10e6;
```

Задание скорости света (м/с) для расчёта длины волны:

```
c0 = 299792458;
```

Расчёт длины волны (м):

```
lambda = c0/freq;
```

Задание излучения:

```
FDTD = SetSinusExcite(FDTD,freq);
```

Задание граничных условий (идеально согласованные слои — Perfectly Matched Layer, PML):

```
BC = {'PML_8' 'PML_8' 'PML_8' 'PML_8' 'PML_8' 'PML_8'};
```

```
FDTD = SetBoundaryCond(FDTD,BC);
```

```
CSX = InitCSX();
```

Задание размеров расчётной сетки. В данном случае используется прямоугольная система координат:

```
start_mesh=-200;
```

```
end_mesh=200;
```

```
unit = 5;
```

```
mesh.x = SmoothMeshLines([start_mesh end_mesh], unit);
```

```
mesh.y = SmoothMeshLines([start_mesh end_mesh], unit);
```

```
mesh.z = SmoothMeshLines([start_mesh end_mesh], unit);
```

```
CSX = DefineRectGrid(CSX, 5, mesh);
```

Задание параметров линзы:

```
CSX = AddMaterial( CSX, 'sphere_material');
```

Сначала диэлектрическая проницаемость материала линзы устанавливается равной единице:

```
CSX = SetMaterialProperty(CSX,'sphere_material', 'Epsilon', 1);
```

Здесь устанавливается настоящее распределение диэлектрической проницаемости в линзе. К сожалению, радиус линзы нужно задавать в твёрдых числах. Здесь ρ — расстояние до оси z :

```
CSX = SetMaterialWeight(CSX, 'sphere_material', 'Epsilon',  
['2-(rho*rho)/22470']);
```

Расчёт радиуса сферы:

```
sphere_radius = lambda/2*10;
```

Добавление примитива сферы с заданным ранее материалом и найденным радиусом:

```
CSX = AddSphere(CSX,'sphere_material',2,[0 0 0],sphere_radius);
```

Добавление «точечного источника», т.е. очень маленькой площадки, с которой распространяется электромагнитная волна:

```
CSX = AddExcitation(CSX, 'excitation', 1, [0 20 0]);
CSX = AddBox(CSX, 'excitation', 0, [-unit*2 -unit*2 -sphere_radius-unit],
[unit*2 unit*2 -sphere_radius]);
```

Задание плоскости, в которой наблюдается распространение электромагнитной волны:

```
CSX = AddDump(CSX, 'Et', 'DumpType', 0, 'DumpMode', 0);
CSX = AddBox(CSX, 'Et', 0, [start_mesh 0 start_mesh], [end_mesh 0 end_mesh]);
```

Запись результатов в xml-файл:

```
Sim_Path = 'point_source';
Sim_CSX = 'point_source.xml';
[status, message, messageid] = rmdir(Sim_Path, 's');
[status, message, messageid] = mkdir(Sim_Path);
WriteOpenEMS([Sim_Path '/' Sim_CSX], FDTD, CSX);
CSXGeomPlot([Sim_Path '/' Sim_CSX], '--RenderDiscMaterial');
```

```
RunOpenEMS(Sim_Path, Sim_CSX);
```

Результатом расчётов, кроме численных значений, является также визуализация исследуемого процесса, фрагмент которой можно видеть на рис. 2.

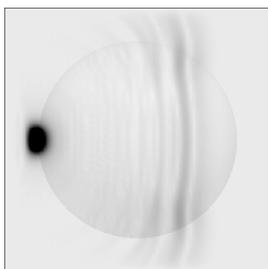


Рис. 2. Прохождение электромагнитной волны через сферическую линзу Люнеберга

5. Заключение

В данной работе даны сведения о работе метода FDTD: описание классического алгоритма Йи, а также общий порядок расчёта методом FDTD. Осуществлено моделирование прохождения электромагнитной волны через сферическую линзу Люнеберга с помощью программы openEMS. Был написан скрипт, решающий данную задачу. Показана применимость openEMS в моделировании градиентной оптики.

Литература

1. Taflove A., Hagness S. C. Computational Electrodynamics: The Finite-difference Time-domain Method. — Artech House, 2005. — ISBN 978-1-58053-832-9.
2. Schneider J. B. Understanding the Finite-Difference Time-Domain Method. — School of Electrical Engineering and Computer Science Washington State University, 2010. — www.eecs.wsu.edu/schneidj/ufdtd.
3. Liebig T. openEMS – Open Electromagnetic Field Solver. — <http://openEMS.de>.
4. Шаранова А. А. Обзор бесплатного программного обеспечения для моделирования линз и оптических систем // Программа 57-й научной конференции МФТИ. — МФТИ, 2015.

5. openEMS Online Manual. — <http://openems.de/index.php/Introduction>.
6. Moreland K. The ParaView Tutorial. — 2015.
7. Шарапова А. А. Нахождение распределения коэффициента преломления для линзы Люнеберга произвольной конфигурации // Информационно-телекоммуникационные технологии и математическое моделирование высокотехнологичных систем: материалы Всероссийской конференции с международным участием. Москва, РУДН, 20–24 апреля 2015 г. — М.: РУДН, 2015. — С. 323–325.
8. Шарапова А. А. Математическое моделирование линзы Люнеберга из кубиков // Труды 55-й научной конференции МФТИ. — МФТИ, 2012. — С. 164.
9. Шарапова А. А. Математическое моделирование сферической линзы Люнеберга из кубиков // Информационно-телекоммуникационные технологии и математическое моделирование высокотехнологичных систем: материалы Всероссийской конференции с международным участием. Москва, РУДН, 22–26 апреля 2013 г. — РУДН, 2013. — С. 221–222.
10. Устройства СВЧ и антенны / Д. И. Воскресенский, В. Л. Гостюхин, В. М. Максимов, Л. И. Пономарев; под ред. Д. И. Воскресенский. — Радиотехника, 2006.
11. Greenwood A. D., Jin J.-M. A Field Picture of Wave Propagation in Inhomogeneous Dielectric Lenses // IEEE Antennas and Propagation Magazine. — 1999.
12. Modeling and simulations of Luneburg lens antennas for communication purposes / S. R. Baev, S. M. Gechev, B. N. Hadjistamov, P. I. Dankov // 16th Telecommunications forum FOR, Serbia, Belgrad. — 2008. — Pp. 488–491. — http://2008.telfor.rs/files/radovi/07_04.pdf.
13. Шарапова А. А. Моделирование распространения электромагнитных волн с помощью openEMS // 58-я научная конференция МФТИ с международным участием. Москва, МФТИ, 23–28 ноября 2015 г. — МФТИ, 2015. — С. 68. — http://conf58.mipt.ru/static/reports_pdf/274.pdf.

UDC 537.876+004.94

Simulation of Wave Propagation with openEMS

A. A. Sharapova, D. S. Kulyabov

*Department of Applied Probability and Informatics
Peoples' Friendship University of Russia
6, Miklukho-Maklaya str., Moscow, Russian Federation, 117198*

Unfortunately, all the existing methods used for modeling computational electrodynamics have their weaknesses. The authors' goal is to analyze the most popular methods. We use spherical Luneburg lens as an illustration.

In this paper authors review the Finite-difference time-domain (FDTD for short), electromagnetic field solver openEMS and its applicability for simulation of wave propagation through medium with the spherical Luneburg lens as an example.

Thanks to its simplicity and broad capabilities, the FDTD method is widely used in various fields. There are quite a lot of simulation tools that implement FDTD, both open-source and proprietary. OpenEMS is an extension for MATLAB and Octave for solving electromagnetic field using the EC-FDTD method. It supports Cartesian and cylindrical coordinates. OpenEMS is free and open-source. It also supports multi-threading, SIMD (SSE) and MPI.

In this paper we simulate wave propagation through spherical Luneburg lens via openEMS thus showing the capabilities of the tool, such as a method to simulate GRIN-optics. We also show how to install and configure it, and how to visualize and analyze the results using ParaView, the application for scientific visualization. Source code of the simulation is presented with appropriate commentaries. In this paper we also briefly review the classic FDTD method and Yee algorithm.

We examined the capabilities of openEMS and FDTD method to simulate wave propagation in a medium.

Key words and phrases: Finite difference time domain method, openEMS, Luneburg lens.

References

1. A. Taflov, S. C. Hagness, Computational Electrodynamics: The Finite-difference Time-domain Method, Artech House, 2005.
2. J. B. Schneider, Understanding the Finite-Difference Time-Domain Method, School of Electrical Engineering and Computer Science Washington State University, 2010.
URL www.eecs.wsu.edu/schneidj/ufdtd
3. T. Liebig, openEMS - Open Electromagnetic Field Solver.
URL <http://openEMS.de>
4. A. A. Sharapova, Review of Free Software for Modeling of Lenses and Optical Systems, in: 57th MFTI Scientific Conference, MFTI, 2015, in Russian.
5. openEMS Online Manual.
URL <http://openems.de/index.php/Introduction>
6. K. Moreland, The ParaView Tutorial (2015).
7. A. A. Sharapova, Calculation of Distribution the Refractive Index in a Luneburg Lens of an Arbitrary Configuration, in: Information and Telecommunication Technologies and Mathematical Modeling of High-Tech Systems, Moscow, PFUR, 20–24 April 2015, PFUR, 2015, in Russian.
8. A. A. Sharapova, Numerical Modeling of a Luneburg Lens Formed from Cubes, in: 55th MFTI Scientific Conference, MFTI, 2012, p. 164, in Russian.
9. A. A. Sharapova, Numerical modeling of a spherical luneburg lens formed from cubes, in: Information and Telecommunication Technologies and Mathematical Modeling of High-Tech Systems, PFUR, 2013, in Russian.
10. D. I. Voskresenskij, V. L. Gostyuhin, V. M. Maksimov, L. I. Ponomarev, Microwave Devices and Antennas, Radio Engineering, 2006, in Russian.
11. A. D. Greenwood, J.-M. Jin, A Field Picture of Wave Propagation in Inhomogeneous Dielectric Lenses, IEEE Antennas and Propagation Magazine.
12. S. R. Baev, S. M. Gechev, B. N. Hadjistamov, P. I. Dankov, Modeling and Simulations of Luneburg Lens Antennas for Communication Purposes, in: 16th Telecommunications forum FOR, Serbia, Belgrad, 2008, pp. 488–491.
URL http://2008.telfor.rs/files/radovi/07_04.pdf
13. A. A. Sharapova, Simulation of Wave Propagation with openEMS, in: 58th MFTI Scientific Conference. Moscow, MFTI, 23–28 November 2015, MFTI, 2016, p. 68, in Russian.

Процедура построения симплектических численных схем для решения гамильтоновых систем уравнений

Б. Батгэрэл^{*†}, Э. Г. Никонов^{*}, И. В. Пузынин^{*}

^{*} Лаборатория информационных технологий
Объединённый институт ядерных исследований

ул. Жолио-Кюри, д. 6, г. Дубна, Московская область, Россия, 141980

[†] Монгольский государственный университет науки и технологии
Улан-Батор, Монголия

Применяемые в настоящее время схемы численного решения систем уравнений динамики многих частиц могут иметь ограничения по шагу и интервалу интегрирования, поскольку при их увеличении численные схемы становятся неустойчивыми и не сохраняют существующие интегралы движения.

В результате при моделировании поведения системы частиц на достаточно больших интервалах времени приходится уменьшать шаг интегрирования, что приводит к существенному увеличению объёма вычислений. В работе предложена новая процедура построения симплектических численных схем для решения гамильтоновых систем уравнений. Предложен подход к симметризации полученных симплектических разностных схем. Численные схемы, построенные при помощи предложенной в работе процедуры, сохраняют энергию системы на большом интервале численного интегрирования при относительно больших величинах шага интегрирования по сравнению с методом Верле, который обычно используется при решении уравнений движения молекулярной динамики. Приведены результаты численных экспериментов, демонстрирующие основные преимущества полученных в работе симметричных симплектических численных схем третьего порядка точности по шагу интегрирования гамильтоновых систем уравнений по сравнению с численными схемами метода Верле второго порядка.

Ключевые слова: гамильтоновы системы уравнений, симплектические разностные схемы, производящие функции, молекулярная динамика.

1. Введение

В классической молекулярной динамике для описания движения частиц используются уравнения Ньютона. При этом для получения координат и скоростей системы, состоящей из N частиц, требуется на каждом временном шаге в трёхмерном случае решить $6N$ уравнений. С практической точки зрения для моделирования поведения системы из N частиц при решении прикладных задач исследования эволюции физических систем необходимо использовать достаточно большой интервал интегрирования по времени. С учётом того, что традиционно используемые для решения уравнений движения молекулярной динамики численные схемы Верле [1] устойчивы лишь при достаточно малых значениях шага интегрирования по времени, на большом интервале интегрирование приводит к большому количеству временных шагов. В этом случае, чтобы увеличить значение шага интегрирования для уменьшения количества шагов при заданном интервале интегрирования, необходимо обеспечить устойчивость схемы по шагу интегрирования. Кроме того, численные методы должны сохранять основные свойства точных решений, в частности должны сохранять первые интегралы, такие как энергия системы и суммарный момент импульса.

Существующие пакеты программ молекулярной динамики для численного интегрирования уравнений движения используют схемы метода Верле. В них интегрирование ведётся с достаточно малым шагом по времени и контролем сохранения гамильтониана (энергии системы), поскольку уже при значении приведённой величины шага порядка 0.1 и количестве шагов порядка 10^3 накопленная вычислительная ошибка приводит к неустойчивости численной схемы и потере сохранения гамильтониана.

Для решения возникающих при молекулярно-динамических расчётах проблем размерности и быстродействия используются следующие пути:

- усовершенствования существующих пакетов:
 - векторизация, как например, в пакете DL-POLY [2];
 - распараллеливание и ускорение межпроцессорных обменов на многопроцессорных системах;
 - разработка спецпроцессоров [3];
- разработка эффективных новых и оптимизация существующих методов численного интегрирования уравнений движения.

Авторами предлагается новый подход к построению численных схем для решения уравнений движения молекулярной динамики. Будет показано, что построенные численные схемы сохраняют в приближенном решении основные свойства точного решения и обладают более высоким порядком аппроксимации, чем схемы метода Верле, а также являются устойчивыми по отношению к шагу аппроксимации.

Предложенный подход состоит в следующем:

- использование гамильтоновой формулировки уравнений движения молекулярной динамики;
- разложение точного решения в ряд Тейлора [4]. При этом возможно применение аппарата компьютерного алгебры для получения аналитических выражений для производных ряда Тейлора;
- использование для вывода численных схем аппарата производящих функций для сохранения геометрических свойств точного решения [1].

В данной работе решаются следующие задачи:

- построить симметричные симплектические численные схемы интегрирования гамильтоновых систем уравнений более высокого порядка точности, чем схемы метода Верле, что позволит увеличить шаг при фиксированном интервале интегрирования и, соответственно, уменьшить общий объем вычислений;
- выполнить численные эксперименты для сравнительного анализа свойств схем метода Верле и построенных численных схем;
- оценить перспективы применения построенных численных схем в существующих пакетах, например в LPMD [5].

2. Постановка задачи

Движение системы N материальных точек в поле с потенциалом $V(\mathbf{q})$, где $\mathbf{q} = (q_1, \dots, q_d)^T$ — координата частицы, $d = 3N$ — размерность пространства координат, может быть описано с использованием системы гамильтоновых уравнений

$$\dot{\mathbf{p}} = -\frac{\partial H(\mathbf{p}, \mathbf{q})}{\partial \mathbf{q}}, \quad (1)$$

$$\dot{\mathbf{q}} = \frac{\partial H(\mathbf{p}, \mathbf{q})}{\partial \mathbf{p}} \quad (2)$$

с начальными условиями

$$\mathbf{p}(0) = \mathbf{p}^0, \quad \mathbf{q}(0) = \mathbf{q}^0. \quad (3)$$

Здесь $\mathbf{p} = (p_1, \dots, p_d)^T$ — импульс частицы, $H(\mathbf{p}, \mathbf{q})$ — гамильтониан системы:

$$H(\mathbf{p}, \mathbf{q}) = \frac{1}{2} \mathbf{p}^T M(\mathbf{q}) \mathbf{p} + V(\mathbf{q}), \quad (4)$$

$M(\mathbf{q})$ — симметричная и положительно определённая матрица масс.

Гамильтонова система уравнений движения эквивалентна уравнениям, полученным в рамках ньютоновского формализма, если силы, действующие на материальные точки, представить в виде $\mathbf{f}(\mathbf{q}) = \nabla V(\mathbf{q})$ и стандартным образом

перейти от дифференциального уравнения Ньютона второго порядка заменой переменных к системе уравнений первого порядка

$$\begin{aligned}\dot{\mathbf{p}} &= -\mathbf{f}(\mathbf{q}), \\ \dot{\mathbf{q}} &= \mathbf{p}.\end{aligned}\tag{5}$$

3. Численные методы

3.1. Геометрические методы

Геометрическим называют такой метод, который сохраняет некоторые геометрические свойства точного решения системы (1)–(3) [1].

Можно выделить следующие геометрические свойства решений:

- отображение $\varphi : (p(t_0), q(t_0)) \mapsto (p(t), q(t))$, реализующее решение гамильтоновой системы уравнений (1)–(3), является симплектическим;
- решение обратимо во времени;
- решение сохраняет значение гамильтониана (4) для любого момента времени.

3.2. Симплектические методы

Дифференцируемое отображение

$$\varphi : U \rightarrow \mathbb{R}^{2d} \quad (U \in \mathbb{R}^{2d})$$

называется симплектическим [1], если якобиан $\varphi'(\mathbf{p}, \mathbf{q})$ удовлетворяет тождеству

$$\varphi'(\mathbf{p}, \mathbf{q})^T J \varphi'(\mathbf{p}, \mathbf{q}) = J,\tag{6}$$

где

$$J = \begin{pmatrix} \mathbf{0}_d & \mathbf{I}_d \\ -\mathbf{I}_d & \mathbf{0}_d \end{pmatrix},$$

$\mathbf{0}_d, \mathbf{I}_d$ – нулевая и единичная матрица размерности d .

3.3. Симплектические численные методы

На дискретном множестве $\{t_k : k = 0, 1, \dots; t_{k+1} - t_k = h\}$ одношаговый метод решения системы (1)–(3), при постоянном шаге по времени h , можно представить в виде

$$(\mathbf{p}^{k+1}, \mathbf{q}^{k+1}) = \Phi_h(\mathbf{p}^k, \mathbf{q}^k),\tag{7}$$

где $\Phi_h(\mathbf{p}, \mathbf{q})$ – преобразование приближенного решения $(\mathbf{p}^k, \mathbf{q}^k)$ при $t = t_k$ в приближенное решение $(\mathbf{p}^{k+1}, \mathbf{q}^{k+1})$ при $t = t_{k+1}$.

Одношаговый метод (7) называется **симплектическим**, если преобразование (7), реализующее приближенное решение гамильтоновой системы (1)–(2), является симплектическим. В качестве примеров симплектических численных методов 1-го и 2-го порядка аппроксимации можно привести явный и неявный методы Эйлера и метод Верле.

Метод Верле:

$$\mathbf{q}^{k+1} = \mathbf{q}^k + h\mathbf{p}^k - \frac{h^2}{2}\mathbf{f}(\mathbf{q}^k),\tag{8}$$

$$\mathbf{p}^{k+1} = \mathbf{p}^k - \frac{h}{2} [\mathbf{f}(\mathbf{q}^k) + \mathbf{f}(\mathbf{q}^{k+1})].\tag{9}$$

Алгоритм вычисления метода Верле:

$$\mathbf{p}^{k+1/2} = \mathbf{p}^k - \frac{h}{2}\mathbf{f}(\mathbf{q}^k), \quad (10)$$

$$\mathbf{q}^{k+1} = \mathbf{q}^k + h\mathbf{p}^{k+1/2}, \quad (11)$$

$$\mathbf{p}^{k+1} = \mathbf{p}^{k+1/2} - \frac{h}{2}\mathbf{f}(\mathbf{q}^{k+1}). \quad (12)$$

Как уже упоминалось, метод Верле является основным методом численного интегрирования уравнений движения в молекулярной динамике.

Если переписать систему уравнений (5) в виде обыкновенного дифференциального уравнения

$$\ddot{\mathbf{q}} = -\mathbf{f}(\mathbf{q}), \quad (13)$$

то метод Верле может быть записан в форме стандартной разностной схемы второго порядка

$$\frac{\mathbf{q}^{k+1} - 2\mathbf{q}^k + \mathbf{q}^{k-1}}{h^2} = -\mathbf{f}(\mathbf{q}^k), \quad k = 0, 1, \dots \quad (14)$$

Большинство стандартных численных методов изначально не являются симплектическими, в частности методы Рунге–Кутты. Для того, чтобы метод стал симплектическим, необходимо проделать ряд нетривиальных шагов для модификации метода в сторону симплектичности.

3.4. Симметричные численные методы

Одношаговый метод (7) называется симметричным, если удовлетворяет условию

$$\Phi_h = \Phi_{-h}^{-1}. \quad (15)$$

В качестве примера симметричного симплектического метода можно привести метод средних точек, который для гамильтоновых систем вида (1), (2) может быть записан следующим образом

$$\mathbf{p}^{k+1} = \mathbf{p}^k - h \frac{\partial H}{\partial \mathbf{q}} \left(\frac{\mathbf{p}^k + \mathbf{p}^{k+1}}{2}, \frac{\mathbf{q}^k + \mathbf{q}^{k+1}}{2} \right), \quad (16)$$

$$\mathbf{q}^{k+1} = \mathbf{q}^k + h \frac{\partial H}{\partial \mathbf{p}} \left(\frac{\mathbf{p}^k + \mathbf{p}^{k+1}}{2}, \frac{\mathbf{q}^k + \mathbf{q}^{k+1}}{2} \right). \quad (17)$$

4. Симплектические преобразования и производящие функции

4.1. Канонические преобразования

В гамильтоновой механике каноническое преобразование – это преобразование канонических переменных и гамильтониана

$$\varphi : (\mathbf{p}, \mathbf{q}) \mapsto (\mathbf{P}, \mathbf{Q}) \quad (18)$$

не меняющее общий вид уравнений Гамильтона [6].

Канонические преобразования взаимнооднозначно определяются производящей функцией $S(\mathbf{p}, \mathbf{q})$, которая является решением уравнения Гамильтона–Якоби и полный дифференциал которой равен

$$dS = \mathbf{P}^\top d\mathbf{Q} - \mathbf{p}^\top d\mathbf{q}. \quad (19)$$

4.2. Производящие функции

Производящая функция может быть выражена через любую пару из четырёх переменных \mathbf{p} , \mathbf{q} , \mathbf{P} , \mathbf{Q} . Возможны четыре варианта выбора пар переменных. Получаемые при этом функции принято называть производящими функциями 1-го, 2-го, 3-го или 4-го типа соответственно.

Производящие функции	Производные
$S = S_1(\mathbf{q}, \mathbf{Q}, t)$	$\mathbf{p} = -\frac{\partial S_1}{\partial \mathbf{q}} \quad \mathbf{P} = \frac{\partial S_1}{\partial \mathbf{Q}}$
$S = S_2(\mathbf{q}, \mathbf{P}, t)$	$\mathbf{p} = \frac{\partial S_2}{\partial \mathbf{q}} \quad \mathbf{Q} = \frac{\partial S_2}{\partial \mathbf{P}}$
$S = S_3(\mathbf{p}, \mathbf{Q}, t)$	$\mathbf{q} = -\frac{\partial S_3}{\partial \mathbf{p}} \quad \mathbf{P} = -\frac{\partial S_3}{\partial \mathbf{Q}}$
$S = S_4(\mathbf{p}, \mathbf{P}, t)$	$\mathbf{q} = \frac{\partial S_4}{\partial \mathbf{p}} \quad \mathbf{Q} = -\frac{\partial S_4}{\partial \mathbf{P}}$

4.3. Связь симплектических преобразований и производящих функций

В соответствии с известной теоремой Пуанкаре [1] для любого симплектического преобразования φ существует производящая функция $S(\mathbf{q}, \mathbf{Q})$, и наоборот, для конкретной производящей функции $S(\mathbf{q}, \mathbf{Q})$ существует симплектическое преобразование φ , которое может быть реконструировано при помощи следующих формул

$$\mathbf{p} = -\frac{\partial S}{\partial \mathbf{q}}(\mathbf{q}, \mathbf{Q}), \quad \mathbf{P} = \frac{\partial S}{\partial \mathbf{Q}}(\mathbf{q}, \mathbf{Q}). \quad (20)$$

5. Процедура построения симплектических разностных схем на основе производящей функции первого типа

Для получения процедуры построения симплектических разностных схем была выбрана производящая функция первого типа $S = S_1(\mathbf{q}, \mathbf{Q}, t)$. При этом могут быть получены два варианта симплектических численных схем. Первый вариант получается при разложении канонической переменной \mathbf{q} при $t = t_{k+1} = t_k + h$ в ряд Тейлора по переменным \mathbf{q} и \mathbf{p} при $t = t_k$, так называемом разложении в ряд Тейлора «вперёд». Второй вариант получается при разложении канонической переменной \mathbf{q} при $t = t_k$ в ряд Тейлора по переменным \mathbf{q} и \mathbf{p} при $t = t_{k+1}$, так называемом разложении в ряд Тейлора «назад». При этом получаются соответственно явная и неявная схема до 3-го порядка включительно. Явно-неявная схема получается при $m = 3$ только в случае разложения в ряд Тейлора «назад». Причём все эти схемы изначально симплектические. В случае же $m > 3$ при использовании разложения в ряд Тейлора и «вперёд» и «назад» получаются явно-неявные симплектические разностные схемы.

В дальнейшем при построении численных схем использованы обозначения:

$$t = t_k, \quad \mathbf{p} = \mathbf{p}^k, \quad \mathbf{q} = \mathbf{q}^k, \\ t = t_{k+1} = t_k + h, \quad \mathbf{P} = \mathbf{p}^{k+1}, \quad \mathbf{Q} = \mathbf{q}^{k+1}.$$

Для производящей функции используются следующие обозначения: $S_{m,1}$ — в случае разложения в ряд Тейлора «вперёд» и $S_{m,2}$ — в случае разложения в ряд Тейлора «назад». Здесь m — порядок аппроксимации точного решения.

5.1. Схема с разложением в ряд Тейлора «вперёд»

1. Представить \mathbf{q}^{k+1} в виде разложения в ряд Тейлора в точке t_k с точностью до $\mathcal{O}(h^{m+1})$.

2. Заменить \mathbf{p}^k в разложении разностной производной от \mathbf{q} , кроме второго члена разложения. Получим уравнение относительно \mathbf{q}^{k+1} .
3. Решить полученное в предыдущем пункте, вообще говоря, нелинейное, для схем выше 3-го порядка аппроксимации, уравнение относительно \mathbf{q}^{k+1} . В случае нелинейного уравнения можно использовать метод Ньютона.
4. Выразить \mathbf{p}^k через \mathbf{q}^k и \mathbf{q}^{k+1} с помощью полученного в первом пункте разложения в ряд Тейлора.
5. Найти производящую функцию $S_{m,1} = S(\mathbf{q}^k, \mathbf{q}^{k+1})$ путём интегрирования полученного в предыдущем пункте выражения по \mathbf{q}^k .
6. Найти \mathbf{p}^{k+1} с помощью равенства

$$\mathbf{p}^{k+1} = \frac{\partial S_{m,1}}{\partial \mathbf{q}^{k+1}} \quad (21)$$

путём дифференцирования.

5.2. Схема с разложением в ряд Тейлора «назад»

1. Представить \mathbf{q}^k в виде разложения в ряд Тейлора в точке t_{k+1} с точностью до $\mathcal{O}(h^{m+1})$.
2. Заменить \mathbf{p}^{k+1} в разложении разностной производной от \mathbf{q} , кроме второго члена разложения.
3. Выразить \mathbf{p}^{k+1} через \mathbf{q}^k и \mathbf{q}^{k+1} с помощью полученного выше разложения в ряд Тейлора.
4. Найти производящую функцию $S_{m,2} = S(\mathbf{q}^k, \mathbf{q}^{k+1})$ путём интегрирования полученного в предыдущем пункте выражения по \mathbf{q}^{k+1} .
5. Найти \mathbf{p}^k помощью равенства

$$\mathbf{p}^k = -\frac{\partial S_{m,2}}{\partial \mathbf{q}^k} \quad (22)$$

путём дифференцирования.

6. Решить, вообще говоря, нелинейное относительно \mathbf{q}^{k+1} уравнение (22). Везде в дальнейшем используется метод Ньютона.
7. Получить \mathbf{p}^{k+1} при помощи полученного во втором пункте выражения, используя известное значение \mathbf{q}^k и полученное в предыдущем пункте значение \mathbf{q}^{k+1} .

5.3. Симплектическая разностная схема 2-го порядка

Для вычисления величины \mathbf{p}^k используем разложение для \mathbf{q} в ряд Тейлора «вперёд» до второго порядка

$$\mathbf{q}^{k+1} = \mathbf{q}^k + h\mathbf{p}^k - \frac{h^2}{2}\mathbf{f}(\mathbf{q}^k) + \mathcal{O}(h^3).$$

В результате получим явную симплектическую разностную схему 2-го порядка

$$\mathbf{p}^k = \frac{\mathbf{q}^{k+1} - \mathbf{q}^k}{h} + \frac{h}{2}\mathbf{f}(\mathbf{q}^k), \quad (23)$$

$$S_2 = \frac{h}{2} \left(\frac{\mathbf{q}^{k+1} - \mathbf{q}^k}{h} \right)^2 - \frac{h}{2} [V(\mathbf{q}^k) + V(\mathbf{q}^{k+1})], \quad (24)$$

$$\mathbf{p}^{k+1} = \frac{\mathbf{q}^{k+1} - \mathbf{q}^k}{h} - \frac{h}{2} \mathbf{f}(\mathbf{q}^{k+1}). \quad (25)$$

Если для вычисления величины \mathbf{p}^{k+1} использовать разложение для \mathbf{q} в ряд Тейлора «назад» также до второго порядка

$$\mathbf{q}^k = \mathbf{q}^{k+1} - h\mathbf{p}^{k+1} - \frac{h^2}{2} \mathbf{f}(\mathbf{q}^{k+1}) + \mathcal{O}(h^3),$$

то получим ту же самую схему (23),(25).

Если же ввести обозначение

$$\frac{\mathbf{q}^{k+1} - \mathbf{q}^k}{h} = \mathbf{p}^{k+1/2},$$

то получим известную схему Верле (10)–(12), которая является симплектической и симметричной.

5.4. Явная и неявная симплектические разностные схемы 3-го порядка

Для построения явных и неявных схем третьего порядка, с использованием производящей функции первого типа, для нахождения величин \mathbf{p}^k и \mathbf{p}^{k+1} , используем разложения точного решения для канонической переменной \mathbf{q} в ряд Тейлора для явной схемы «вперёд», для неявной – «назад».

Явная схема.

$$\mathbf{q}^{k+1} = \mathbf{q}^k + h\mathbf{p}^k - \frac{h^2}{2} \mathbf{f}(\mathbf{q}^k) - \frac{h^3}{6} \nabla \mathbf{f}(\mathbf{q}^k) \cdot \boxed{\mathbf{p}^k} + \mathcal{O}(h^4)$$

Если заменить \mathbf{p}^k в последнем члене ряда разностным отношением

$$\boxed{\mathbf{p}^k} = \frac{\mathbf{q}^{k+1} - \mathbf{q}^k}{h} + \mathcal{O}(h),$$

то получим явную схему 3-го порядка

$$\mathbf{p}^k = \frac{\mathbf{q}^{k+1} - \mathbf{q}^k}{h} + \frac{h}{2} \mathbf{f}(\mathbf{q}^k) + \frac{h^2}{6} \nabla \mathbf{f}(\mathbf{q}^k) \cdot \frac{\mathbf{q}^{k+1} - \mathbf{q}^k}{h} \quad (26)$$

$$S_{3,1} = \frac{h}{2} \left(\frac{\mathbf{q}^{k+1} - \mathbf{q}^k}{h} \right)^2 - \frac{h}{3} [2V(\mathbf{q}^k) + V(\mathbf{q}^{k+1})] - \frac{h^2}{6} \mathbf{f}(\mathbf{q}^k) \cdot \frac{\mathbf{q}^{k+1} - \mathbf{q}^k}{h}, \quad (27)$$

$$\mathbf{p}^{k+1} = \frac{\mathbf{q}^{k+1} - \mathbf{q}^k}{h} - \frac{h}{6} [\mathbf{f}(\mathbf{q}^k) + 2\mathbf{f}(\mathbf{q}^{k+1})]. \quad (28)$$

Введя обозначение

$$\Theta^k = \left[\mathbf{I} + \frac{h^2}{6} \nabla \mathbf{f}(\mathbf{q}^k) \right]^{-1},$$

получим формулы для вычисления \mathbf{p} и \mathbf{q} на следующем временном шаге

$$\mathbf{p}^{k+1/2} = \mathbf{p}^k - \frac{h}{2} \mathbf{f}(\mathbf{q}^k),$$

$$\tilde{\mathbf{p}}^{k+1/2} = \Theta^k \mathbf{p}^{k+1/2},$$

$$\begin{aligned}\mathbf{q}^{k+1} &= \mathbf{q}^k + h\tilde{\mathbf{p}}^{k+1/2}, \\ \mathbf{p}^{k+1} &= \tilde{\mathbf{p}}^{k+1/2} - \frac{h}{6} [\mathbf{f}(\mathbf{q}^k) + 2\mathbf{f}(\mathbf{q}^{k+1})].\end{aligned}$$

Неявная схема.

Для получения неявной схемы представим значение для переменной \mathbf{q} на k -ом шаге в виде разложения в ряд Тейлора «назад»

$$\mathbf{q}^k = \mathbf{q}^{k+1} - h\mathbf{p}^{k+1} - \frac{h^2}{2}\mathbf{f}(\mathbf{q}^{k+1}) + \frac{h^3}{6}\nabla\mathbf{f}(\mathbf{q}^{k+1}) \cdot \boxed{\mathbf{p}^{k+1}} + \mathcal{O}(h^4).$$

Представим \mathbf{p}^{k+1} в последнем члене разложения в виде разностной производной

$$\boxed{\mathbf{p}^{k+1}} = \frac{\mathbf{q}^{k+1} - \mathbf{q}^k}{h} + \mathcal{O}(h).$$

В результате получим следующее выражение для \mathbf{p}^{k+1} :

$$\mathbf{p}^{k+1} = \frac{\mathbf{q}^{k+1} - \mathbf{q}^k}{h} - \frac{h}{2}\mathbf{f}(\mathbf{q}^{k+1}) + \frac{h^2}{6}\nabla\mathbf{f}(\mathbf{q}^{k+1}) \cdot \frac{\mathbf{q}^{k+1} - \mathbf{q}^k}{h}. \quad (29)$$

После интегрирования по \mathbf{q}^{k+1} получаем производящую функцию $S_{3,2}$

$$S_{3,2} = \frac{h}{2} \left(\frac{\mathbf{q}^{k+1} - \mathbf{q}^k}{h} \right)^2 - \frac{h}{3} [V(\mathbf{q}^k) + 2V(\mathbf{q}^{k+1})] + \frac{h^2}{6}\mathbf{f}(\mathbf{q}^{k+1}) \cdot \frac{\mathbf{q}^{k+1} - \mathbf{q}^k}{h} \quad (30)$$

и дифференцированием по \mathbf{q}^k с учётом знака находим выражение для \mathbf{p}^k :

$$\mathbf{p}^k = \frac{\mathbf{q}^{k+1} - \mathbf{q}^k}{h} + \frac{h}{6} [2\mathbf{f}(\mathbf{q}^k) + \mathbf{f}(\mathbf{q}^{k+1})]. \quad (31)$$

Далее аналогично явной схеме получаем формулы для неявной схемы третьего порядка

$$\mathbf{p}^{k+1/2} = \mathbf{p}^k - \frac{h}{2}\mathbf{f}(\mathbf{q}^k), \quad (32)$$

$$\tilde{\mathbf{q}}^{k+1} = \mathbf{q}^k + h\mathbf{p}^{k+1/2}, \quad (33)$$

$$\mathbf{F}(\mathbf{q}^{k+1}) = \tilde{\mathbf{q}}^{k+1}, \quad (34)$$

$$\boldsymbol{\mu}^{k+1} = \frac{h}{6} [\mathbf{f}(\mathbf{q}^{k+1}) - \mathbf{f}(\mathbf{q}^k)], \quad (35)$$

$$\boldsymbol{\theta}^{k+1} = \mathbf{I} + \frac{h^2}{6}\nabla\mathbf{f}(\mathbf{q}^{k+1}), \quad (36)$$

$$\tilde{\mathbf{p}}^{k+1/2} = \boldsymbol{\theta}^{k+1}(\mathbf{p}^{k+1/2} - \boldsymbol{\mu}^{k+1}), \quad (37)$$

$$\mathbf{p}^{k+1} = \tilde{\mathbf{p}}^{k+1/2} - \frac{h}{2}\mathbf{f}(\mathbf{q}^{k+1}), \quad (38)$$

где функция $\mathbf{F}(\mathbf{q}^{k+1})$ определяется по формуле

$$\mathbf{F}(\mathbf{q}^{k+1}) = \mathbf{q}^{k+1} + \frac{h^2}{6} [\mathbf{f}(\mathbf{q}^{k+1}) - \mathbf{f}(\mathbf{q}^k)]. \quad (39)$$

6. Симметричные симплектические разностные схемы

Для построения симметричных симплектических разностных схем рассмотрим одно параметрическое семейство производящих функций следующего вида.

$$S_m(\alpha) = \alpha S_{m,1} + (1 - \alpha) S_{m,2}. \quad (40)$$

Здесь $\alpha : 0 \leq \alpha \leq 1$, $S_{m,1}$ и $S_{m,2}$ – производящие функции, полученные при помощи процедур (21) и (22) соответственно. Каждая из этих производящих функций $S_m(\alpha)$ порождает симплектическую разностную схему m -го порядка. Если $\alpha = 1/2$, то мы получаем процедуры построения симметричных симплектических разностных схем.

Получим процедуры построения симметричных симплектических разностных схем 3-го и 4-го порядка, поскольку схемы 3-го и 4-го порядка представляют наибольший практический интерес при проведении численных расчётов. Для простоты представления схема 4-го порядка приведена для одномерного случая. Схемы более высокого порядка аппроксимации имеет смысл выводить с использованием средств компьютерной алгебры.

6.1. Симметричная симплектическая схема для $m = 3$

$$S_{3,3} = \frac{h}{2} \left(\frac{\mathbf{q}^{k+1} - \mathbf{q}^k}{h} \right)^2 - \frac{h}{2} [V(\mathbf{q}^k) + V(\mathbf{q}^{k+1})] + \frac{h^2}{12} [\mathbf{f}(\mathbf{q}^{k+1}) - \mathbf{f}(\mathbf{q}^k)] \cdot \frac{\mathbf{q}^{k+1} - \mathbf{q}^k}{h}, \quad (41)$$

$$\mathbf{p}^k = \frac{\mathbf{q}^{k+1} - \mathbf{q}^k}{h} + \frac{h}{12} [5\mathbf{f}(\mathbf{q}^k) + \mathbf{f}(\mathbf{q}^{k+1})] + \frac{h^2}{12} \nabla \mathbf{f}(\mathbf{q}^k) \cdot \frac{\mathbf{q}^{k+1} - \mathbf{q}^k}{h}, \quad (42)$$

$$\mathbf{p}^{k+1} = \frac{\mathbf{q}^{k+1} - \mathbf{q}^k}{h} - \frac{h}{12} [\mathbf{f}(\mathbf{q}^k) + 5\mathbf{f}(\mathbf{q}^{k+1})] + \frac{h^2}{12} \nabla \mathbf{f}(\mathbf{q}^{k+1}) \cdot \frac{\mathbf{q}^{k+1} - \mathbf{q}^k}{h}. \quad (43)$$

6.2. Симметричная симплектическая схема для $m = 4$

Для построения симметричной симплектической численной схемы 4-го порядка сначала проведём вывод **схемы «вперёд»**. Сначала представим q_{k+1} в виде разложения в ряд Тейлора «вперёд»

$$q_{k+1} = q_k + hp_k - \frac{h^2}{2} f(q_k) - \frac{h^3}{6} f'(q_k) p_k - \frac{h^4}{24} [f''(q_k) p_k^2 - f'(q_k) f(q_k)] + \mathcal{O}(h^5).$$

После замены p_k в члене разложения с h^3

$$p_k = \frac{q_{k+1} - q_k}{h} + \frac{h}{2} f(q_k) + \mathcal{O}(h^2)$$

и замены квадрата p_k в члене разложения с h^4

$$p_k = \frac{q_{k+1} - q_k}{h} + \mathcal{O}(h)$$

получим следующее выражение для p_k , зависящее только от q_k и q_{k+1} :

$$p_k = \frac{q_{k+1} - q_k}{h} + \frac{h}{2} f(q_k) + \frac{h^2}{6} f'(q_k) \left[\frac{q_{k+1} - q_k}{h} + \frac{h}{2} f(q_k) \right] +$$

$$+ \frac{h^3}{24} \left[f''(q_k) \left(\frac{q_{k+1} - q_k}{h} \right)^2 - f'(q_k) f(q_k) \right].$$

Далее интегрированием получаем производящую функцию

$$S_{4,1} = \frac{h}{2} \left(\frac{q_{k+1} - q_k}{h} \right)^2 - \frac{h}{4} [3V(q_k) + V(q_{k+1})] - \\ - \frac{h^2}{4} f(q_k) \left(\frac{q_{k+1} - q_k}{h} \right) - \frac{h^3}{24} f'(q_k) \left(\frac{q_{k+1} - q_k}{h} \right)^2 - \frac{h^3}{48} [f^2(q_{k+1}) - f^2(q_k)],$$

и дифференцированием выражение для p_{k+1}

$$p_{k+1} = \frac{q_{k+1} - q_k}{h} - \frac{h}{4} [f(q_k) + f(q_{k+1})] - \\ - \frac{h^2}{12} f'(q_k) \left(\frac{q_{k+1} - q_k}{h} \right) + \frac{h^3}{24} f'(q_{k+1}) f(q_{k+1}).$$

Аналогично, в соответствии с процедурой (22), строится **схема «назад»** 4-го порядка. Сначала представим q_k в виде разложения в ряд Тейлора «назад»

$$q_k = q_{k+1} - hp_{k+1} - \frac{h^2}{2} f(q_{k+1}) + \frac{h^3}{6} f'(q_{k+1}) p_{k+1} - \\ - \frac{h^4}{24} [f''(q_{k+1}) p_{k+1}^2 - f'(q_{k+1}) f(q_{k+1})] + \mathcal{O}(h^5).$$

Затем проведём замену p_{k+1} в члене разложения с h^3

$$p_{k+1} = \frac{q_{k+1} - q_k}{h} - \frac{h}{2} f(q_k) + \mathcal{O}(h^2),$$

и замену квадрата p_{k+1} в члене разложения с h^4

$$p_{k+1} = \frac{q_{k+1} - q_k}{h} + \mathcal{O}(h).$$

В результате получим выражение для p_{k+1} , зависящее только от q_k и q_{k+1}

$$p_{k+1} = \frac{q_{k+1} - q_k}{h} - \frac{h}{2} f(q_{k+1}) + \frac{h^2}{6} f'(q_{k+1}) \left(\frac{q_{k+1} - q_k}{h} \right) + \\ + \frac{h^3}{24} \left[f''(q_{k+1}) \left(\frac{q_{k+1} - q_k}{h} \right)^2 + f'(q_{k+1}) f(q_{k+1}) \right].$$

Далее интегрированием получаем соответствующую производящую функцию

$$S_{4,2} = \frac{h}{2} \left(\frac{q_{k+1} - q_k}{h} \right)^2 - \frac{h}{4} [V(q_k) + 3V(q_{k+1})] + \\ + \frac{h^2}{4} f(q_{k+1}) \left(\frac{q_{k+1} - q_k}{h} \right) - \frac{h^3}{24} f'(q_{k+1}) \left(\frac{q_{k+1} - q_k}{h} \right)^2 + \\ + \frac{h^3}{48} [f^2(q_{k+1}) - f^2(q_k)]$$

и дифференцированием выражение для p_k :

$$p_k = \frac{q_{k+1} - q_k}{h} + \frac{h}{4} [f(q_k) + f(q_{k+1})] - \frac{h^2}{12} f'(q_{k+1}) \left(\frac{q_{k+1} - q_k}{h} \right) + \frac{h^3}{24} f'(q_k) f(q_k).$$

Далее с помощью формулы (40) при $\alpha = 0.5$

$$S_{4,3} = \frac{1}{2} (S_{4,1} + S_{4,2})$$

получаем симметричную производящую функцию $S_{4,3}$

$$S_{4,3} = \frac{h}{2} \left(\frac{q_{k+1} - q_k}{h} \right)^2 - \frac{h}{2} [V(q_k) + V(q_{k+1})] + \frac{h^2}{8} [f(q_{k+1}) - f(q_k)] \left(\frac{q_{k+1} - q_k}{h} \right) - \frac{h^3}{48} [f'(q_{k+1}) + f'(q_k)] \left(\frac{q_{k+1} - q_k}{h} \right)^2 + \frac{h^3}{48} [f^2(q_{k+1}) - f^2(q_k)].$$

В результате после соответствующей процедуры дифференцирования получим симметричную симплектическую схему 4-го порядка:

$$p_k = \frac{q_{k+1} - q_k}{h} + \frac{h}{8} [3f(q_k) + f(q_{k+1})] - \frac{h^2}{24} [2f'(q_k) - f'(q_{k+1})] \left(\frac{q_{k+1} - q_k}{h} \right) + \frac{h^2}{48} f''(q_k) \left(\frac{q_{k+1} - q_k}{h} \right)^2 - \frac{h^3}{24} f'(q_k) f(q_k),$$

$$p_{k+1} = \frac{q_{k+1} - q_k}{h} - \frac{h}{8} [f(q_k) + 3f(q_{k+1})] - \frac{h^2}{24} [f'(q_k) - 2f'(q_{k+1})] \left(\frac{q_{k+1} - q_k}{h} \right) - \frac{h^2}{48} f''(q_{k+1}) \left(\frac{q_{k+1} - q_k}{h} \right)^2 - \frac{h^3}{24} f'(q_{k+1}) f(q_{k+1}).$$

7. Численные эксперименты

Численные эксперименты выполнены для сравнительного анализа численной схемы метода Верле и численных схем более высокого порядка аппроксимации при $m = 3$, полученных с использованием разработанной авторами процедуры, на примере задачи Кеплера [1]. В ней движение двух материальных точек описывается гамильтонианом следующего вида

$$H(\mathbf{p}, \mathbf{q}) = \frac{1}{2} (p_1^2 + p_2^2) - \frac{1}{\sqrt{q_1^2 + q_2^2}}. \quad (44)$$

Численные расчёты проводились при следующих параметрах:

- 1) $h = 0,01$, $T = 1000$;
- 2) $h = 0,05$, $T = 1000$;
- 3) $h = 0,10$, $T = 1000$;
- 4) $h = 0,20$, $T = 5000$.

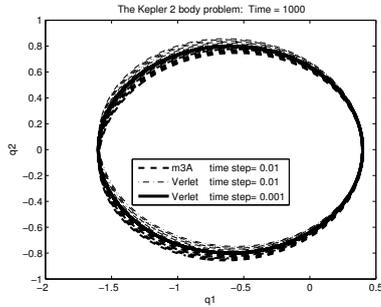
Здесь h – шаг по времени, T – граница интервала времени $[0, T]$ эволюции системы.

Для каждого фиксированного шага по времени h сравнивались результаты расчётов, полученные с использованием метода Верле, и результаты, полученные

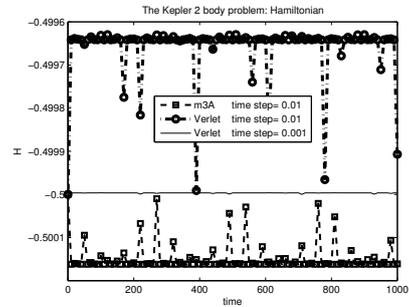
с использованием построенных в работе явной, неявной и симметричной численной схемы 3-го порядка. В качестве «точного» решения использовались результаты расчётов по методу Верле с шагом по времени, равным $h/10$.

Результаты моделирования представлены на рис. 1–14.

Задача Кеплера. Явная схема.

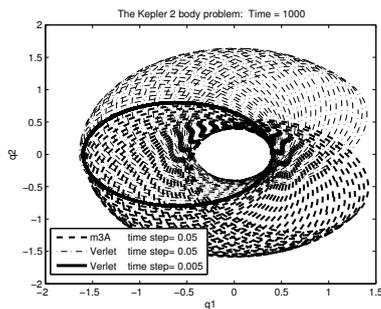


а) Фазовые траектории

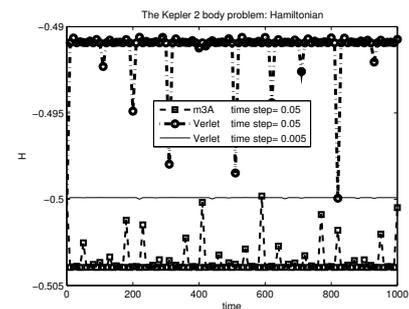


б) Зависимость гамильтониана от времени

Рис. 1. Явная схема, схема Верле при $h = 0.01$ и схема Верле при $h = 0.001$

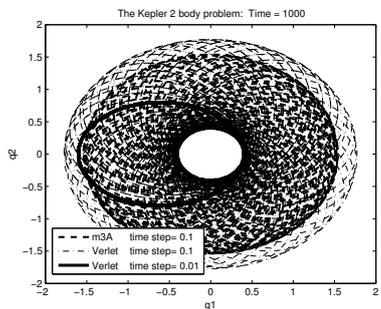


а) Фазовые траектории

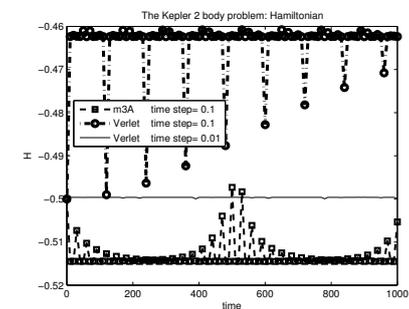


б) Зависимость гамильтониана от времени

Рис. 2. Явная схема, схема Верле при $h = 0.05$ и схема Верле при $h = 0.005$



а) Фазовые траектории



б) Зависимость гамильтониана от времени

Рис. 3. Явная схема, схема Верле при $h = 0.1$ и схема Верле при $h = 0.01$

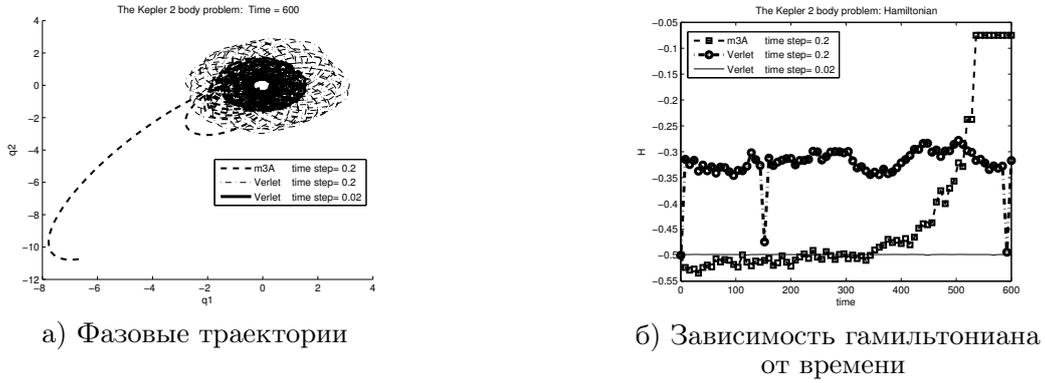


Рис. 4. Явная схема, схема Верле при $h = 0.2$ и схема Верле при $h = 0.02$

Как показывают численные эксперименты, при величинах шага по времени до $h = 0.1$ включительно схема Верле и явная схема третьего порядка очень близки и по расчётным фазовым траекториям и по значениям гамильтониана на всем интервале интегрирования (рис. 1–3). Начиная с величины шага $h = 0.2$ схема Верле оказалась более устойчивой, хотя и менее точной по сравнению с явной схемой третьего порядка. Кроме того, начиная примерно с 500-го шага она становится неустойчивой (рис. 4) по шагу интегрирования. **Задача Кеплера. Неявная схема.**

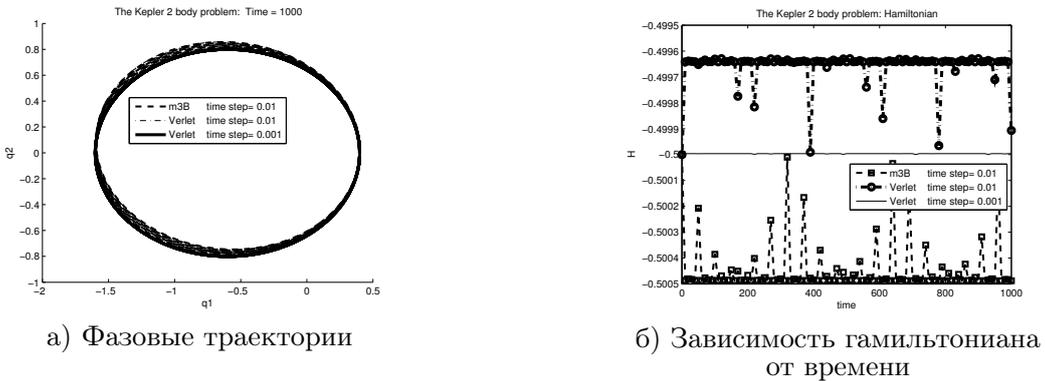


Рис. 5. Неявная схема, схема Верле при $h = 0.01$ и схема Верле при $h = 0.001$

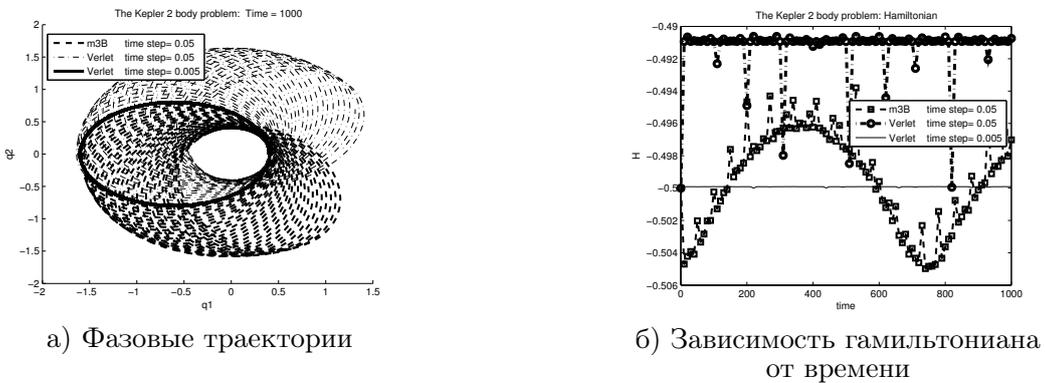


Рис. 6. Неявная схема, схема Верле при $h = 0.05$ и схема Верле при $h = 0.005$

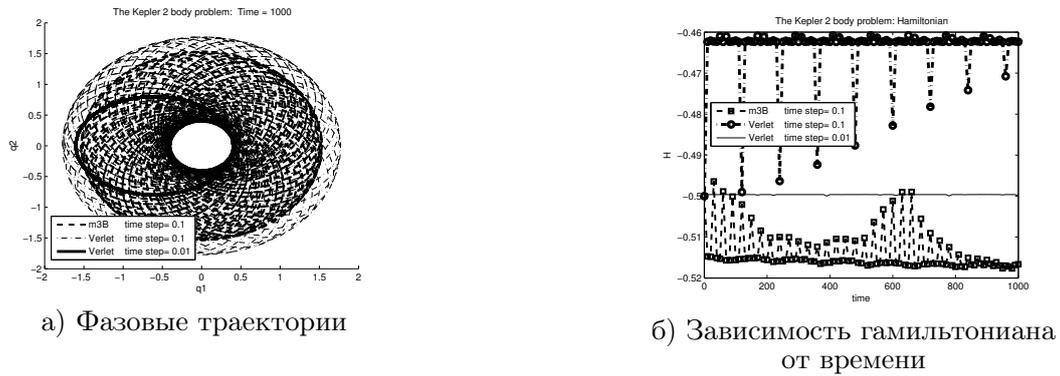


Рис. 7. Неявная схема, схема Верле при $h = 0.1$ и схема Верле при $h = 0.01$

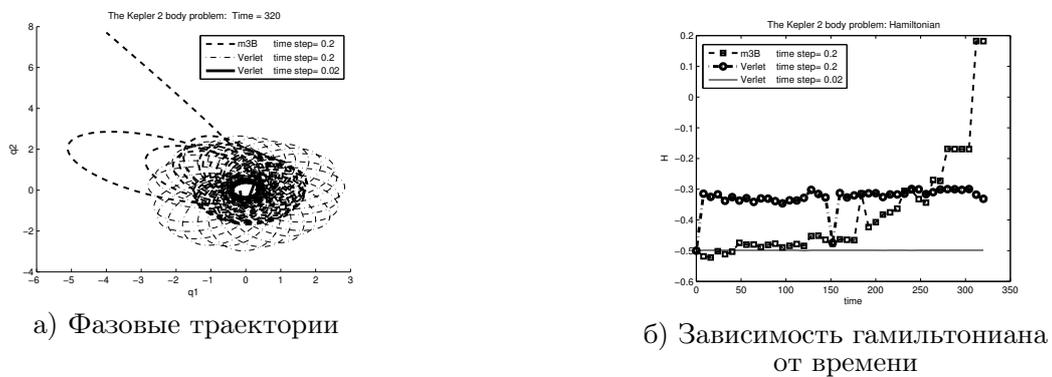


Рис. 8. Неявная схема, схема Верле при $h = 0.2$ и схема Верле при $h = 0.02$

Аналогично численным экспериментам с явной схемой третьего порядка, в случае неявной схемы, при величинах шага по времени до $h = 0.1$ включительно, схема Верле и неявная схема третьего порядка очень близки и по расчётным фазовым траекториям и по значениям гамильтониана на всем интервале интегрирования (рис. 5–7). Начиная с величины шага $h = 0.2$ схема Верле снова оказывается более устойчивой по сравнению с неявной схемой третьего порядка, однако примерно с 250-го шага становится неустойчивой по шагу интегрирования (рис. 8).

Задача Кеплера. Симметричная схема.

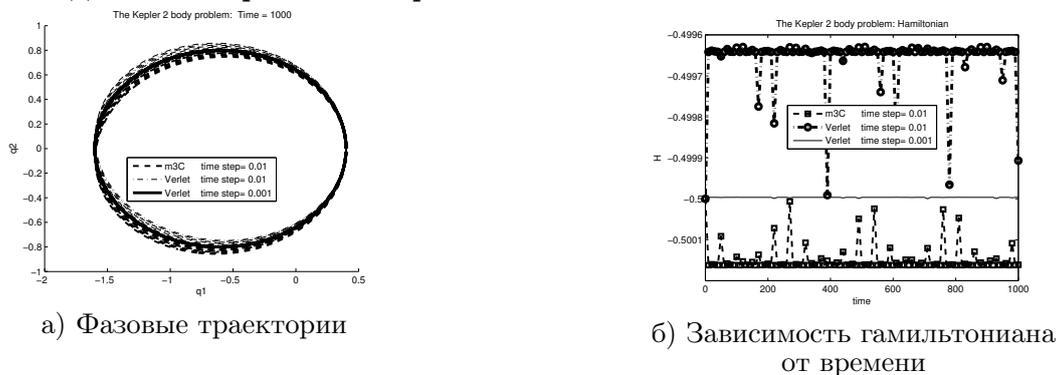


Рис. 9. Симметричная схема, схема Верле при $h = 0.01$ и схема Верле при $h = 0.001$

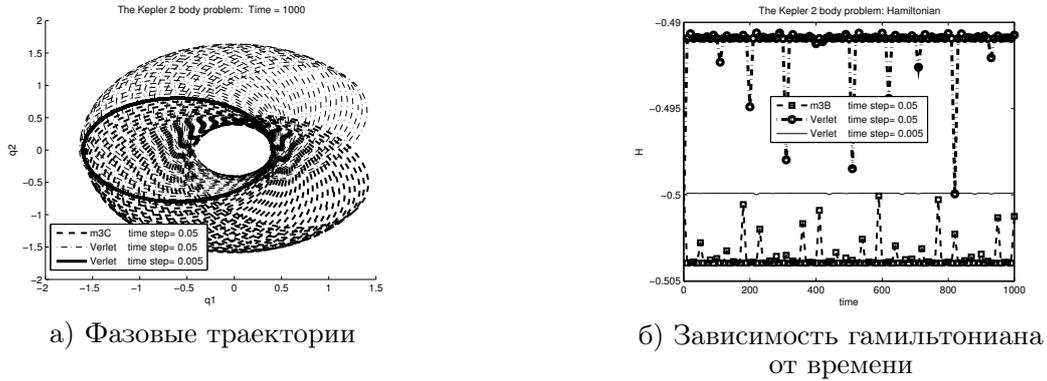


Рис. 10. Симметричная схема, схема Верле при $h = 0.05$ и схема Верле при $h = 0.005$

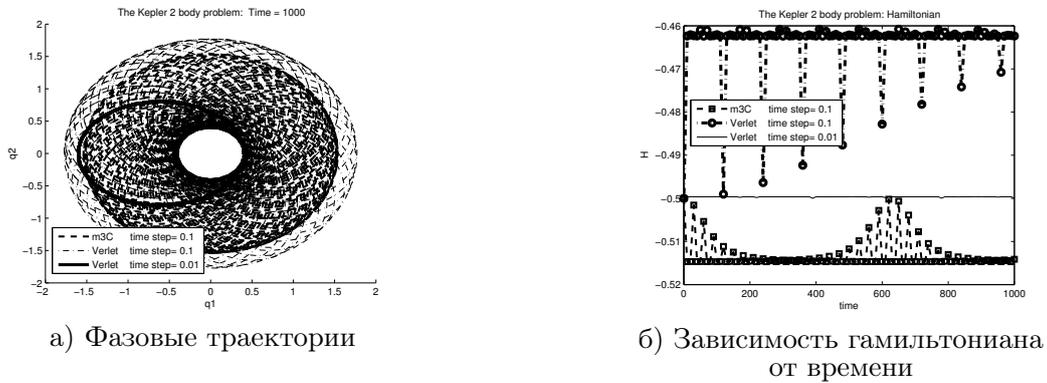


Рис. 11. Симметричная схема, схема Верле при $h = 0.1$ и схема Верле при $h = 0.01$

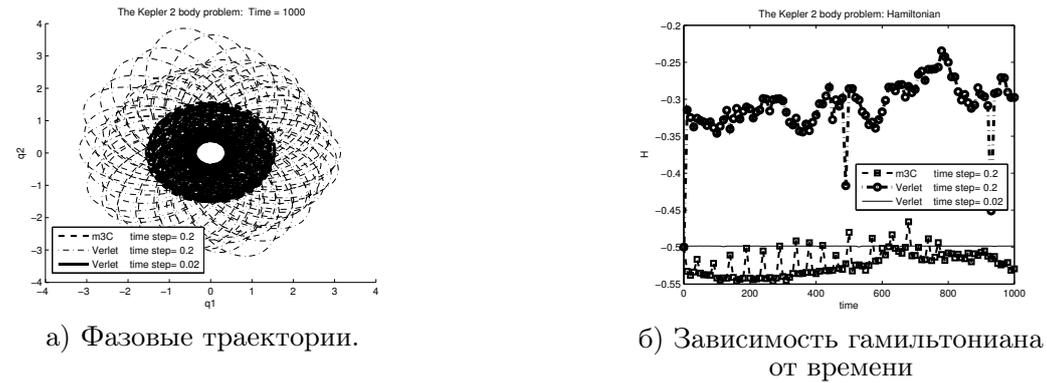


Рис. 12. Симметричная схема, схема Верле при $h = 0.2$ и схема Верле при $h = 0.02$

Совсем иная картина открывается при сравнении результатов расчётов с использованием симметричной симплектической численной схемы третьего порядка. Вплоть до величины шага численного интегрирования по времени $h = 0.2$ включительно симметричная симплектическая схема третьего порядка существенно более точна по расчётным фазовым траекториям и по отклонениям значений

гамильтониана от начального значения на всем интервале интегрирования (рис. 9–12).

Задача Кеплера. Симметричная схема. $T = 5000$.

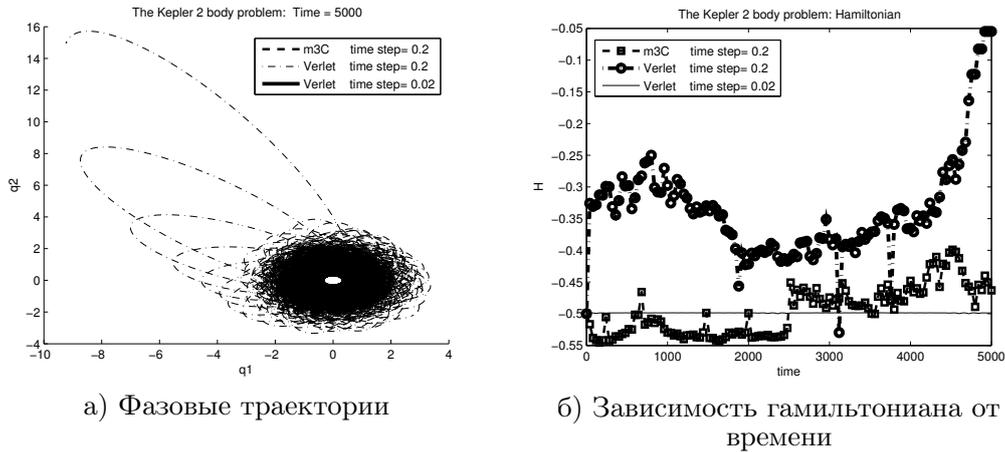


Рис. 13. Симметричная схема, схема Верле при $h = 0.2$ и схема Верле при $h = 0.02$

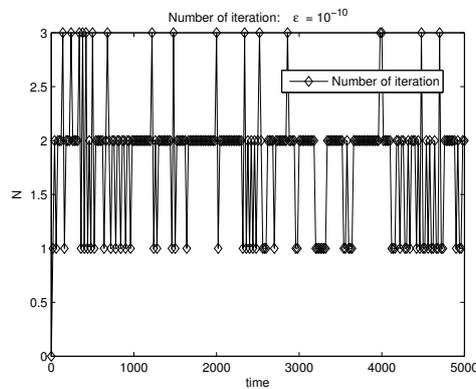


Рис. 14. Количество итераций метода Ньютона на каждом шаге по времени при $\epsilon = 10^{-10}$

И, наконец, в численных расчётах, которые проводились на большем интервале эволюции системы $T = 5000$, симметричная симплектическая численная схема третьего порядка продемонстрировала существенно большую точность и устойчивость при величине шага $h = 0.2$. При этом схема Верле становится неустойчивой уже начиная с интервалов порядка $T = 4000$ (рис. 13). При это вычислительные затраты в методе Ньютона при вычислениях для неявной части симметричной схемы не превышают 3 итераций при точности вычислений $\epsilon = 10^{-10}$ (рис. 14).

8. Заключение

1. Предложен подход к построению симметричных симплектических численных схем интегрирования уравнений движения метода молекулярной динамики в гамильтоновой формулировке.
2. Описан алгоритм получения симметричных симплектических разностных схем заданного порядка аппроксимации с использованием аппарата производящих функций.
3. Приведены симметричные симплектические разностные схемы до 4-го порядка аппроксимации включительно. Схемы более высокого порядка аппроксимации требуют более сложного вывода и в дальнейшем могут быть получены с использованием аналитических вычислений.
4. Проведены численные эксперименты, показавшие, что полученные на основе разработанного подхода разностные схемы третьего порядка аппроксимации сохраняют устойчивость для достаточно больших значений шага и интервала интегрирования в отличие от метода Верле. Они с более высокой точностью сохраняют гамильтониан системы на всем интервале интегрирования.

Литература

1. Hairer E., Lubich C., Wanner G. Geometric Numerical Integration. — Berlin, Heidelberg: Springer, 2006. — P. 644.
2. Методы молекулярной динамики для моделирования физических и биологических процессов / Х. Т. Холмуродов, М. В. Алтайский, И. В. Пузынин и др. // Физика элементарных частиц и атомного ядра. — 2003. — Т. 34, № 2. — С. 472–515.
3. MDGRAPE-4: a Special-Purpose Computer System for Molecular Dynamics Simulations / I. Ohmura, G. Morimoto, Y. Ohno et al. // Philos Trans A Math Phys Eng Sci. — 2014. — Vol. 372, No 2021. — Pp. 1–16.
4. Akishin P. G., Puzyrin I. V., Vinitsky S. I. A Hybrid Numerical Method for Analysis of Dynamics of the Classical Hamiltonian Systems // Comp. Math. Applic. — 1997. — Vol. 34, No 2-4. — Pp. 45–73.
5. Las Palmeras Molecular Dynamics: Flexible and modular molecular dynamics / S. Davis, C. Loyola, F. Gonzalez, J. Peralta // Computer Physics Communications. — 2010. — Vol. 181, No 12. — Pp. 2126–2139.
6. Гантмаагер Ф. Р. Лекции по аналитической механике. — М.: Физматлит, 2001. — С. 262.

UDC 519.622, 004.021, 004.942

A Procedure for Constructing Symplectic Numerical Schemes for Solving of Hamiltonian Systems of Equations

B. Batgerel^{*†}, E. G. Nikonov^{*}, I. V. Puzyrin[†]

^{*} Laboratory of Information Technologies
Joint Institute for Nuclear Research
6, Joliot-Curie, Dubna, Moscow region, Russia, 141980

[†] The Mongolian University of Science and Technology
8th khoroo, Baga toiruu 34, Sukhbaatar district
Ulaanbaatar, Mongolia, 14191

Numerical schemes which is used for solving of many-particle dynamics systems of equations can have restrictions on a step and an interval of integration because if its increase the numerical schemes became unstable and don't conserve existing integrals of motion. As a result when we simulate many-particle system behavior on the sufficiently large time interval we should decrease an integration step which leads to considerable increasing of computation quantity. In this paper a new procedure for constructing symplectic numerical schemes for solving of Hamiltonian systems of equations is proposed. A method for symmetrization of received symplectics numerical schemes is proposed too. Constructed by proposed in the

paper procedure numerical schemes conserve energy of a system on the large interval of numerical integration for relatively large integration step in comparison with Verlet method which is usually used for solving of equations of motion in molecular dynamics. Results of numerical experiments are given in the paper. These results show main advantages of received symmetric symplectic numerical schemes of third order of accuracy for the integration step for the Hamiltonian systems of equations in comparison with numerical schemes of Verlet method of second order of accuracy.

Key words and phrases: Hamiltonian systems of equations, symplectic difference schemes, generating functions, molecular dynamics.

References

1. E. Hairer, C. Lubich, G. Wanner, Geometric Numerical Integration, Springer, Berlin, Heidelberg, 2006.
2. H. T. Holmurodov, M. V. Altajskij, I. V. Puzynin, et al., Molecular Dynamics Methods for Simulation of Physical and Biological Processes, Physics of Elementary Particles and Atomic Nuclei 34 (2) (2003) 472–515, in Russian.
3. I. Ohmura, G. Morimoto, Y. Ohno, A. Hasegawa, M. Taiji, MDGRAPE-4: a Special-Purpose Computer System for Molecular Dynamics Simulations, Philos Trans a Math Phys Eng Sci. 372 (2021) (2014) 1–16.
4. P. G. Akishin, I. V. Puzynin, S. I. Vinitzky, A Hybrid Numerical Method for Analysis of Dynamics of the Classical Hamiltonian Systems, Comp. Math. Applic. 34 (2-4) (1997) 45–73.
5. S. Davis, C. Loyola, F. Gonzalez, J. Peralta, Las palmeras molecular dynamics: Flexible and modular molecular dynamics, Computer Physics Communications 181 (12) (2010) 2126–2139.
6. F. R. Gantmakher, Lectons on Analytical Mechanics, PhysMathLit, Moscow, 2001, in Russian.

Информатика и вычислительная техника

УДК 621.39

Оценка времени установления сессии между пользователями при наличии межсетевого экрана

К. Е. Самуйлов, А. Ю. Ботвинко, Э. Р. Зарипова

*Кафедра прикладной информатики и теории вероятностей
Российский университет дружбы народов
ул. Миклухо-Маклая, д. 6, Москва, Россия, 117198*

Для эффективной разработки, дальнейшего внедрения и эксплуатации информационных систем связи необходимо предусмотреть своевременную защиту программ и баз данных, средств хранения, обработки и передачи информации. Для анализа производительности, надёжности и безопасности телекоммуникационных систем, а также для получения первичных оценок временных характеристик широко используются методы теории массового обслуживания. В статье для исследования метода оценки времени установления сессии при наличии межсетевого экрана выбрана процедура установления сессии между двумя конечными пользователями по протоколу установления сессий (Session Initiation Protocol, SIP) между двумя пользователями с одним межсетевым экраном по пути следования сигнальных сообщений. Под межсетевым экраном (firewall) подразумевается программный или аппаратный комплекс, реализующий функции фильтрации сетевого трафика между отправляющей и принимающей стороной по некоторому набору правил, определяемых политикой безопасности. Предложенный метод оценки основан на применении модели открытой экспоненциальной сети массового обслуживания. Приведён пример расчёта среднего времени установления сессии и средней задержки запроса сессии. Расчёт проведён для технических характеристик, соответствующих прокси-серверу CiscoASA 5500-Xc обслуживающим модулем SSP-10. Расчёт временных характеристик показывает приемлемость применения данного оборудования в сети связи и его малое влияние на временные характеристики даже при высокой нагрузке.

Ключевые слова: Session Initiation Protocol, межсетевой экран, время установления сессии, открытая сеть массового обслуживания, качество восприятия.

Введение

Для обеспечения надёжной информационной инфраструктуры необходимо своевременно разрабатывать и внедрять аппаратные или программные комплексы, обеспечивающие фильтрацию входящего потока по некоторым правилам. При внедрении таких комплексов, обеспечивающих контроль и безопасность, необходимо учитывать временные задержки для информационных потоков, проходящих проверку на безопасность. В данной работе исследуется время установления сессии между двумя пользователями, где информационные потоки проходят через межсетевой экран (МЭ) в прямом и обратном направлении.

В качестве примера выбрана одна из наиболее чувствительных к временным задержкам услуг — установление сессии между двумя абонентами при наличии двух прокси-серверов и одного межсетевого экрана по пути следования сигнальных сообщений. Время установления сессии включает время передачи сигнальных сообщений протокола установления сессий (Session Initiation Protocol, SIP) между двумя пользователями, время фильтрации сообщений в межсетевом экране, время ожидания и время обслуживания на каждом из прокси-серверов. Наличие МЭ увеличивает время установления сессии, однако обеспечивает контроль и фильтрацию проходящих через него сетевых сообщений. МЭ не избавляет от угроз утечки информации или от загрузки пользователями вредоносных программ, например, вирусов. Использование межсетевых экранов для защиты сетей вносит

Статья поступила в редакцию 27 ноября 2015 г.

Исследование выполнено при частичной финансовой поддержке РФФИ в рамках научных проектов № 15-07-03608, 15-07-03051.

задержку при передаче сообщений и, соответственно, вызывает необходимость учёта дополнительных задержек при установлении сессии, что, в свою очередь, приводит к повышению требований к производительности межсетевых экранов по сравнению с фильтрацией не мультимедийного трафика. В статье ставится задача оценки времени установления сессии при внедрении в сеть МЭ.

Статья имеет следующую структуру. Первый раздел содержит описание процедуры установления сессии между двумя пользователями при наличии двух серверов SIP и одного МЭ. Приведён порядок обмена сигнальными сообщениями по протоколу установления сессий SIP, перечислены логические и функциональные блоки, принимающие участие в процедуре установления сессии. После детального изучения рекомендованных международными стандартами показателей качества восприятия было отмечено, что время установления сессии не может превышать 2 с [1].

Во втором разделе статьи исследованы временные характеристики, которые необходимо учитывать при установлении сессии: время установления сессии и задержка запроса установления сессии (англ. Session Request Delay, SRD), эту характеристику в дальнейшем будем называть задержкой запроса сессии. Время установления сессии является характеристикой качества восприятия (англ., Quality of Equipment, QoE). В разделе предложен метод оценки временных характеристик.

В третьем разделе предложенный метод используется для расчета времени установления сессии для оборудования Cisco. Оценено влияние времени фильтрации трафика в межсетевом экране на среднее время установления сессии и среднюю задержку запроса сессии. Заключение содержит основные выводы исследования.

1. Процедура установления сессии при наличии межсетевого экрана

В спецификациях протокола SIP, разработанных группой MMUSIC (Multiparty Multimedia Session Control) комитета IETF (Internet Engineering Task Force) [2,3], предусмотрены 3 основных сценария установления соединения: простое взаимодействие между двумя клиентами пользователей, соединение двух пользователей с участием сервера переадресации и соединение двух пользователей с использованием прокси-сервера.

Данные сценарии отличаются способом поиска адреса и приглашения вызываемого пользователя. При простом взаимодействии клиентов вызывающему пользователю для установления соединения необходимо знать текущий адрес вызываемого пользователя. При использовании сервера переадресации вызывающий пользователь передаёт серверу переадресации сообщение с известным ему адресом вызываемого пользователя, а сервер переадресации обеспечивает переадресацию вызова на текущий адрес этого пользователя. В этом случае сервер переадресации определяет текущий адрес вызываемого пользователя у сервера определения местоположения. В сценарии с использованием прокси-сервера функции поиска и приглашения вызываемого пользователя возлагаются на прокси-сервер.

На практике сценарии установления соединения могут состоять из более сложных цепочек – например, вызов может пройти сервер переадресации и несколько прокси-серверов. Более того, запросы могут быть размножены и переданы по разным маршрутам и т. п. В статье исследован один из возможных простых сценариев – установление сессии между двумя пользователями с участием цепочки из двух прокси-серверов и одного межсетевого экрана. Прокси-сервер Proxy-1 выполняет посреднические функции для первого пользователя, а прокси-сервер Proxy-2 – для второго пользователя. Предположим, что оборудование первого пользователя расположено внутри сети, защищённой межсетевым экраном, обеспечивающим контроль проходящих сетевых пакетов с целью выявления и предотвращения попыток несанкционированного доступа или злоупотребления сетевыми ресурсами. В этих условиях при фильтрации межсетевым экраном исходящих пакетов от

первого пользователя и входящих пакетов от второго пользователя появляется дополнительная задержка. Схематичное расположение пользователей и функциональных элементов сети представлены на рис. 1.

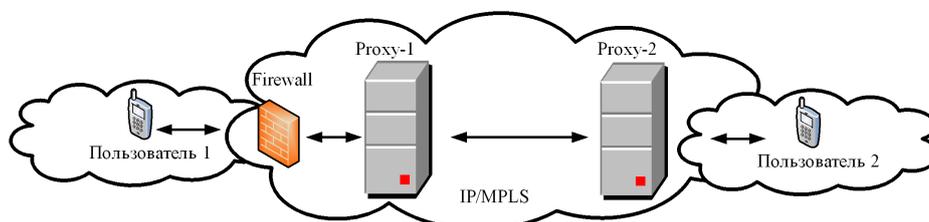


Рис. 1. Взаимодействие функциональных элементов при установлении сессии при наличии межсетевого экрана

Функциональными элементами, принимающими участие в процедуре установления сессии, являются прокси-серверы (Proxy-1 и Proxy-2), межсетевого экрана (Firewall), магистральная сеть IP/MPLS и оборудование пользователей. Заметим, что некоторые из этих элементов могут быть реализованы на практике совместно, например, межсетевого экрана может быть объединён с прокси-сервером в одном оборудовании. Для предварительной оценки времени установления сессии выбрана процедура, где каждый функциональный элемент реализован отдельно.

Сессия инициируется первым пользователем сигнальным сообщением Invite в момент нажатия клавиши на своём оборудовании. Процедура установления сессии учитывает запросы и ответы, пересылаемые от одного функционального блока к другому. Запрос Invite содержит информацию об адресате, ответом на запрос Invite является сообщение 100 Trying, которое высылается на предыдущий функциональный блок при успешной обработке сообщения Invite [1, 3]. Заметим, что в рамках данной работы исследуется случай успешного установления сессии без ретрансляций сообщений. С помощью данного сценария можно оценить среднее время установления сессии T_S от отправления инициирующего сообщения до начала предоставления данных и среднюю задержку запроса сессии T_{SRD} от момента инициации сессии до получения первым абонентом ответа 180 Ringing, которое указывает, что все функциональные блоки готовы к обслуживанию абонентов [3]. Процедура установления сессии учитывает время обслуживания в сети, магистральная сеть IP/MPLS представлена в виде отдельного функционального блока, время обслуживания в сети при движении в одну сторону от одного абонента к другому будет равно половине времени передачи в петле связи (англ., Round Trip Time, RTT). Сигнальное сообщение протокола установления сессии 180 Ringing передаётся от вызываемого абонента к вызывающему абоненту при успешном получении инициирующего сообщения Invite. Далее в процедуре установления сессии предусмотрен обмен сообщениями 200 Ok и Ack, что подтверждает приём ответа на запрос Invite. После обмена этими сигнальными сообщениями сессия считается установленной, начинается обмен данными между двумя абонентами.

Последовательность сигнальных сообщений в виде процедуры установления сессии при наличии одного межсетевого экрана представлена на рис. 2. Необходимо заметить, что внедрение межсетевого экрана в процедуру установления сессии увеличивает время установления сессии, т.к. происходит фильтрация трафика, в том числе сигнального, проходящего через межсетевого экрана [4].

2. Расчёт времени установления сессии при наличии межсетевого экрана

Для оценки влияния времени фильтрации в межсетевого экране на время установления сессии между двумя абонентами предлагается математическая модель

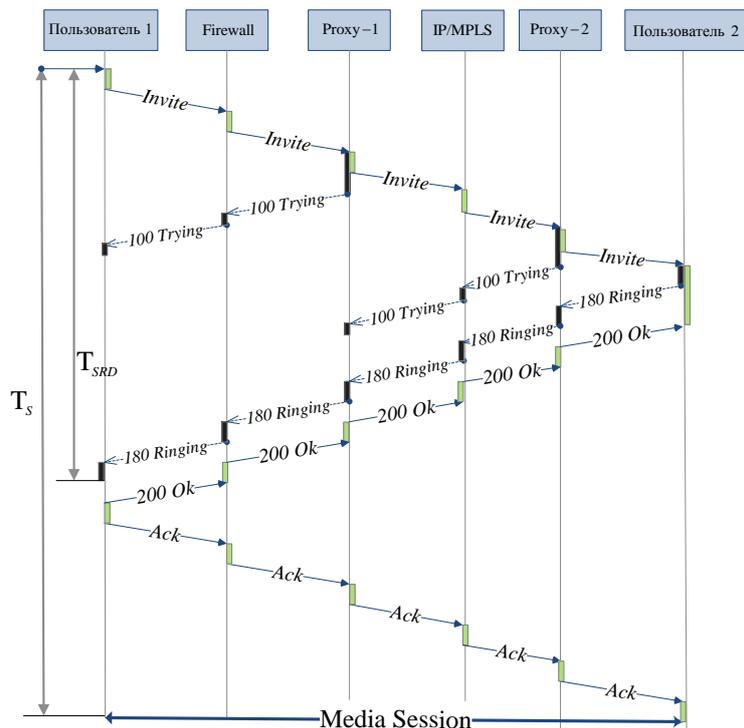


Рис. 2. Процедура установления сессии при наличии межсетевого экрана

в виде открытой сети массового обслуживания (СеМО), в которой время пребывания в сети соответствует времени установления сессии. Подобный подход исследовался в работах [5, 6]. При данном подходе необходимо оценить время ожидания и время обслуживания последовательно по каждому функциональному блоку, затем, просуммировав все интервалы времени, получить время установления сессии. Таким образом, СеМО будет состоять из шести узлов в виде систем массового обслуживания (СМО), каждая СМО будет представлять отдельный функциональный блок, соответствующий процедуре установления сессии. Блоки, соответствующие оборудованию обоих пользователей и сети IP/MPLS, будут представлены с помощью СМО $M|M|\infty$, остальные блоки — в виде СМО $M|M|1|\infty$.

Обозначив интенсивность входящих сообщений λ_0 , определим условие существования стационарного режима в виде (1), где μ_i^{-1} — время обслуживания сообщений прибором в СМО, соответствующей i -му блоку в процедуре установления сессии [6]

$$\lambda_0 < \min\left(\frac{\mu_2}{5}; \frac{\mu_3}{5}; \frac{\mu_5}{4}\right). \quad (1)$$

Данное неравенство (1) следует из того, что в узлах 2, 3 и 5 могут возникать очереди, что нарушает стационарный режим.

Учитывая подход, предложенный в работах [5, 6], средняя задержка запроса сессии T_{SRD} и среднее время установления сессии T_S определяется в виде (2) и (3) соответственно.

$$T_{SRD} = 2\mu_1^{-1} + \frac{2}{\mu_2 - 5\lambda_0} + \frac{2}{\mu_3 - 5\lambda_0} + 2\mu_4^{-1} + \frac{2}{\mu_5 - 4\lambda_0} + \mu_6^{-1}; \quad (2)$$

$$T_S = 2\mu_1^{-1} + \frac{3}{\mu_2 - 5\lambda_0} + \frac{3}{\mu_3 - 5\lambda_0} + 3\mu_4^{-1} + \frac{3}{\mu_5 - 4\lambda_0} + 2\mu_6^{-1}. \quad (3)$$

Время фильтрации межсетевым экраном сигнального сообщения равно времени пребывания сигнального сообщения во втором функциональном блоке (Firewall). Время фильтрации одного сообщения при установлении сессии представлено формулой (4)

$$T_F = \frac{1}{\mu_2 - 5\lambda_0}. \quad (4)$$

Проведём расчёт показателей влияния наличия межсетевого экрана на среднюю задержку запроса сессии и на среднее время установления сессии. Эти показатели отражены в формулах (5) и (6) соответственно.

$$N_{T_F_T_{SRD}} = \frac{2T_F}{T_{SRD}}; \quad (5)$$

$$N_{T_F_T_S} = \frac{3T_F}{T_S}. \quad (6)$$

3. Пример расчета времени установления сессии

Для предварительной оценки временных характеристик были выбраны исходные данные, характерные для оборудования Cisco. Технические характеристики прокси-сервера соответствуют прокси-серверу Cisco на платформе Sun Fire V120. Для характеристик межсетевого экрана выбран Cisco ASA 5500-X с обслуживающим модулем (Security Service Processor - SSP) SSP-10. В таблице 1 перечислены исходные данные и их соответствие введённым в статье обозначениям.

Таблица 1

Исходные данные

Функциональный элемент	Пользователь-1	Firewall	Прокси-1	IP/MPLS	Прокси-2	Пользователь-2
Обозначение	μ_1^{-1}	μ_2^{-1}	μ_3^{-1}	μ_4^{-1}	μ_5^{-1}	μ_6^{-1}
Время обслуживания, мс	0,1	0,5	0,4	50	0,4	0,1

Результаты расчета средней задержки запроса сессии T_{SRD} и среднего времени установления сессии T_S приведены в виде графиков зависимости от интенсивности поступающих запросов (рис. 3).

Условие существования стационарного режима (1) позволяет провести расчёт для значений интенсивности входящего потока вплоть до 400 запросов в секунду. Технические характеристики соответствуют прокси-серверу Cisco на платформе Sun Fire V120 и межсетевому экрану Cisco ASA 5500-X с обслуживающим модулем SSP-10. Предварительная оценка временных характеристик при наличии одного межсетевого экрана показывает приемлемые результаты, которые удовлетворяют рекомендованным международным стандартам показателей качества восприятия, среднее время установления сессии гораздо меньше 2 с и даже при интенсивности 380 запросов/с достигает значения 0,2 с. Средняя задержка запроса сессии SRD всегда меньше среднего времени установления сессии и при интенсивности входящего потока 380 запросов/с достигает 0,15 с.

Интересна оценка доли времени фильтрации сигнальных сообщений в межсетевом экране при расчёте временных характеристик. На рис. 4 изображены

показатели влияния времени фильтрации в межсетевом экране на время установления сессии. Отмечено, что время пребывания сигнальных сообщений в межсетевом экране не превосходит 10 процентов при интенсивности входящих сообщений для значений равных 370 запросов/с.

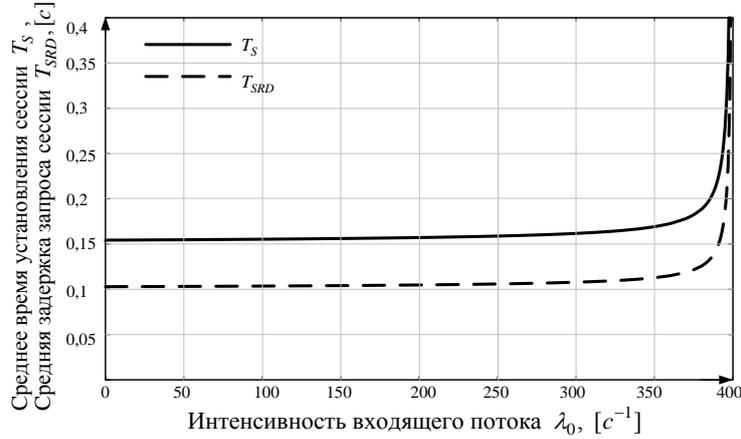


Рис. 3. Временные характеристики при установлении сессии с одним межсетевым экраном

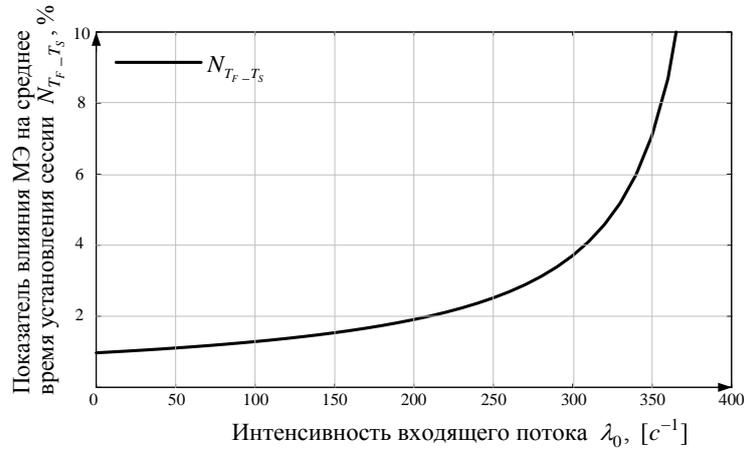


Рис. 4. Оценка доли времени фильтрации в межсетевом экране при установлении сессии

Заключение

В статье предложена математическая модель установления сессии между двумя пользователями с участием двух прокси-серверов и одного межсетевого экрана. Проведена оценка времени установления сессии и средней задержки запроса сессии. Расчёт временных характеристик показал приемлемость использования оборудования Cisco ASA 5500-X с обслуживающим модулем SSP-10, использование межсетевого экрана необходимо в целях обеспечения информационной безопасности.

Полученные результаты могут найти применение при решении задач оценивания влияния МЭ на качество установления соединения протокола SIP. При дальнейших исследованиях производительности систем связи область применения построенной математической модели и предложенного метода может быть расширена на случай оценки показателей качества других услуг связи, чувствительных к длительности задержки передачи данных.

Литература

1. Triple-play Services Quality of Experience (QoE) Requirements Cluster: Techrep Technical Report-126 / DSL Forum. — 2006.
2. Rosenberg J., Schulzrinne H., Camarillo G. et al. RFC 3261, SIP: Session Initiation Protocol. — 2002. — <http://www.ietf.org/rfc/rfc3261.txt>.
3. Malas D., Morton A. RFC 6076, Basic Telephony SIP End-to-End Performance Metrics. — 2011. — <http://www.ietf.org/rfc/rfc6076.txt>.
4. Иванов К. В., Тутубалин П. И. Марковские модели защиты автоматизированных систем управления специального назначения. — Казань: ГБУ «Республиканский центр мониторинга качества образования», 2012. — С. 216.
5. Самуйлов К. Е., Лузгачев М. В., Плаксина О. Н. Разработка вероятностной модели для анализа показателей качества протокола инициирования сеансов связи. — Вестник РУДН. Серия: «Математика. Информатика. Физика», 2007. — С. 53–63.
6. Gaidamaka Y., Zaripova E. Session Setup Delay Estimation Methods for IMS-Based IPTV Services / Ed. by S. Balandin, S. Andreev, Y. Koucheryavy. — Springer International Publishing Switzerland, 2014. — Vol. 8638. — Pp. 408–418.

UDC 621.39

Session Setup Time Estimation in the Network with a Firewall

К. Е. Samouylov, A. Yu. Botvinko, E. R. Zaripova

*Department of Applied Probability and Informatics
Peoples' Friendship University of Russia
6, Miklukho-Maklaya str., Moscow, Russian Federation, 117198*

In modern telecommunication networks is implemented information transmitting security for protection programs and databases. We propose the approach of signaling messages processing estimation in case when the network includes one firewall and two proxy servers. Firewall is based on security rules and define security police. We choose a session setup procedure between two users using the Session Initiation Protocol as an illustration of the method. Signaling messages are moving from the first user to the second through IP/MPLS network, proxy servers and firewall. The chain of signaling messages must successfully service before media traffic starts. We get preliminary estimation of performance measures for modern firewall Cisco ASA 5500-X with Security Service Processor SSP-10. Numerical example shows, that firewall Cisco ASA 5500-X serves signaling messages can be implemented. Mean waiting time and mean service time in this firewall has little effect on the session setup time and the session request delay.

Key words and phrases: Session Initiation Protocol, firewall, session setup time, open queuing network, Quality of Experience.

References

1. Triple-Play Services Quality of Experience (QoE) Requirements Cluster, Tech. Rep. Technical Report-126, DSL Forum (2006).

2. J. Rosenberg, H. Schulzrinne, G. Camarillo, A. Johnston, J. Peterson, R. Sparks, M. Handley, E. Schooler, RFC 3261, SIP: Session Initiation Protocol (June 2002). URL <http://www.ietf.org/rfc/rfc3261.txt>
3. D. Malas, A. Morton, RFC 6076, Basic Telephony SIP End-to-End Performance Metrics (Jan. 2011). URL <http://www.ietf.org/rfc/rfc6076.txt>
4. K. V. Ivanov, P. I. Tutubalin, Markov Models of Protection of Automated Control Systems for Special Purposes, Kazan: “Republican Center of Education Quality Monitoring”, 2012, in Russian.
5. K. E. Samouylov, M. V. Luzgachev, O. N. Plaksina, Modelling SIP Connections with Open Multiclass Queueing Networks, Bulletin of Peoples’ Friendship University of Russia. Series “Mathematics. Information Sciences. Physics”, 2007, in Russian.
6. Y. Gaidamaka, E. Zaripova, Session Setup Delay Estimation Methods for IMS-based IPTV Services, Vol. 8638, Springer International Publishing Switzerland, 2014, pp. 408–418.

Fuzzy Conceptual Graphs for Knowledge Representation in Process-Oriented Organizations

E. J. Azofeifa, G. M. Novikova

*Department of Information Technologies
Peoples' Friendship University of Russia
6, Miklukho-Maklaya str., Moscow, Russia, 117198*

The use of fuzzy conceptual graphs for representing knowledge in process-oriented organizations is considered. Two types of knowledge, procedural and declarative, are discussed; the difference between them is shown along with representation and usage details in knowledge bases. A formal definition of fuzzy conceptual graphs is given, and their ability to represent declarative and procedural domain knowledge in a simple and understandable way is shown. The structure of the knowledge base includes three levels: an ontological layer, which contains concepts and integrates declarative and procedural knowledge; a middle or interface layer, which describes business processes based on real data; and the ground layer, the layer of real (historical) data, which collects primary information about the current state of objects and the relationships between them. The difference between the types of information on each of the layers of the knowledge base is shown, as well as the application method of fuzzy conceptual graphs on the interface layer. A description of the mechanisms of interaction and rules for reflecting data from the ground layer to the interface layer is provided. Mathematical methods for analysis of primary data and fuzzy knowledge indicators are described, which aid in decision-making, optimization and refinement of procedural knowledge systems.

Key words and phrases: Fuzzy conceptual graphs, knowledge representation, business process, knowledge base, ontology, intelligent agent.

1. Introduction

The span of Business Intelligence (BI) tools and similar approaches is due to grow steadily in the next years, following the current expansion of information in organizations. Many companies already consider BI as a necessity, more than a competitive advantage. Consequently, specializations or branches of BI technologies have been emerging during the last years. One of them is Business Process Intelligence (BPI), or the application of BI techniques to business processes [1]. BPI relates to the emerging data inside of a company; namely, the information concerning processes and the resources associated to them. However, BPI is prone to encounter issues that, most likely, have been dragged through the lifecycle of the company. Some of these problems are related to technological limitations, like absence of event logging, while others are related to Knowledge Representation (KR), like incompatibilities in the ontological knowledge among business units or lack of standardized process documentation.

Conceptual Graphs (CGs) are an intuitive and easily understandable means to represent knowledge [2, p. v], and they can be combined with Fuzzy Logic (FL) to create Fuzzy Conceptual Graphs (FCG), which are focused on the approximate representation of imprecise knowledge [3, p. 1]. The aim of this work is to present FCGs as a suitable formalism for KR and reasoning in process-oriented organizations. For this purpose, a Knowledge Base (KB) structure is proposed based on the FCG formalism. Mapping between declarative and procedural knowledge is discussed, and an interface between ontological and factual knowledge is presented. Fuzzy approaches to quantify the characteristics of the KB are proposed, including a pseudometric based on the flow of information through entities.

2. Background

Knowledge representation can be thought of as an internal representation of reality inside an intelligent agent. It consists of symbols that represent a limited collection of propositions about the world, and as such, is only able to approximate reality. Choosing a specific type of KR will determine how and what to perceive from reality.

Ontologies are abstract models consisting of concepts that are relevant for describing the real world [4, p. 30]. They provide a set of terms used to represent reality, and for this reason, they are considered the heart of every knowledge representation. They are typically built on top of a taxonomy, which is a hierarchical structure of concepts using the relation "is a kind of" among them. However, they vary according to which kind of knowledge they hold. Two types of knowledge can be distinguished [5]. Declarative knowledge describes facts or the understanding that something is true. Procedural knowledge, on the other hand, refers to the ability to perform a task in an efficient way [4, p. 48].

To establish a basis for reasoning, ontologies must be embedded into a suitable framework, which includes a formal system of logic and an efficient computational environment. Conceptual graphs are a logically precise formalism which can represent knowledge in a humanly readable, and computationally manageable form [6]. Fuzzy logic is an extension of traditional logical systems, which offers the framework for dealing with uncertainty and imprecision. In contrast to exact KR, FL is a much more similar way to represent the human mind [7].

A CG is composed of concept nodes representing entities and relation nodes representing relationships between these entities. A concept node refers to a specific entity by adding an individual marker to its label. Otherwise, the concept node refers to an unspecified entity. In fuzzy logic, an object is said to be of a type with an uncertainty and/or truth degree. The main difference between CGs and FCGs is only that CGs are based on basic concept types and basic relation types, whereas FCGs are based on fuzzy concept types and fuzzy relation types [3, p. 36].

A basic CG [2, p. 26] is composed of a vocabulary $V = (TC, TR, I)$ and a 4-tuple $G = (C, R, E, l)$ where:

- TC is the set of concept types.
- TR is the set of relation symbols, which is partitioned into subsets T_{1R}, \dots, T_{kR} of relation symbols of arity $1, \dots, k$, respectively.
- I is the set of individual markers, $*$ denotes the generic marker and $M = I \cup \{*\}$ is the set of markers.
- (C, R, E) is a finite, undirected and bipartite multigraph denoted $graph(G)$.
- C is the concept node set.
- R is the relation node set.
- E is the family of edges.
- l is a labeling function that satisfies:
 - A concept node c is labeled by a pair $(type(c), marker(c))$, where $type(c) \in TC$ and $marker(c) \in I \cup \{*\}$.
 - A relation node r is labeled by $l(r)$ or $type(r) \in TR$.
 - Edges incident to a relation node r are totally ordered and labeled from 1 to $arity(type(r))$. An edge labeled i between a relation r and a concept c is denoted (r, i, c) .

Computer systems that make use of human knowledge are referred to as knowledge-based systems (KBS). A typical KBS consists at least of a knowledge base containing organized knowledge represented by KR formalisms, and a reasoning engine with mechanisms to imitate the problem solving process of a human expert [4, p. 54]. Organized knowledge in a KB refers to ontologies, facts, rules and constraints [2, p. 2]. Factual knowledge makes assertions about a concrete situation.

The simplest KR model consists of a KB with ontologies and facts [2, p. 312]. When working with complex systems, however, other kinds of components have to be added to the model in order to correctly describe the state of the world and its evolution in time. Inference rules represent implicit or general knowledge about a domain, which

helps to fully describe a world. Evolution rules transform factual knowledge: they represent possible actions leading from one world to another.

3. A KB using the FCG formalism

The following is an approach to structuring declarative and procedural knowledge in a KB using the FCG formalism. On this proposal, ontological and factual knowledge are bound together by CGs structured in layers. There is a middle layer (interface) between the ontological and factual (ground) layers. Rules and constraints are context-dependent, and they are assumed to be developed aside from the ontological structure.

3.1. Ontological layer

Consider an enterprise *XYZ* divided in different departments or business units. These units are organized in a hierarchical or non-hierarchical way, or a combination of both, with varying levels of interdependence. This interdependence between departments can be expressed in several ways, depending on the task at hand: for example, interdependence measures could be the number of shared resources between business units, or the distributed weight of the different channels of communication between departments (See Figure 1).

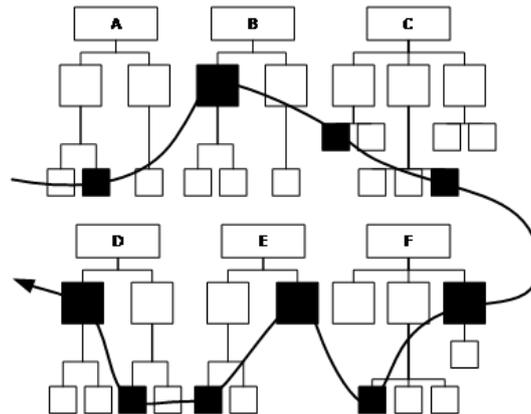


Figure 1. Interdependent process

Basically, physical or abstract entities, together with the relations between them, are commonly shared (up to a certain degree) among the units of a company. However, these dependencies are often ignored or hidden for several reasons, which contributes to problems in corporate visibility. In companies with large operations, for example, multiple users share and manipulate information or assets, even if they do not know about each other's actions. This can give rise to inconsistencies or issues, which require prompt and efficient solutions in order to resume normal operation. In the ideal scenario of a company with full documentation of its ontologies and visibility of its processes, finding solutions to this kind of problems is straightforward. In a real scenario, on the contrary, there can be flaws and inconsistencies in the form or content of the knowledge representation, which could prevent or limit problem solving.

Consider the same company *XYZ*, with a number n of different business units. Each one of these departments possesses a set of ontologies of type o in the form of CGs, where type o could be either declarative or procedural. The set of n departments is part of a CG representing the hierarchical structure of the company, with the top of the hierarchy being the company itself. If at least one of the business units possesses an ontology and, more specifically, its set of vertices (concepts) is not empty, then the

maximal join [2, p. 218], [8] of the ontologies of each unit, together with the taxonomy, results in an upper or top-level ontology, denoted by T , which is itself a CG.

Notice that process analysis relies on the existence of documented business processes or their generation from event logs. In the proposed KB, existing business processes have to be embedded into the ontological layer as ontologies of the procedural type, maintaining the overall CG structure. Naturally, newly discovered business processes should update the ontological knowledge. Several approaches exist to merge declarative and procedural knowledge. For example, the Static and Dynamic Knowledge Representation Framework (SD-KRF) [9] expresses processes, nodes, tasks, events, transitions, actions and decisions as ontology classes. A method for creating ontological knowledge is based on providing compatibility to a set of tools for modeling processes and corporate infrastructure [10].

3.2. Interface layer

Factual knowledge asserts that some entities exist and that they are related by some relationships [2, p. 22]. It is obtained after transforming raw historical (or real-time) data into useful information for analysis and decision-making. In this KB, a way of expressing factual knowledge is proposed. It consists in merging the obtained information from historical data into the upper ontology T , creating a new layer on the CG structure. The newly created layer, which is denoted interface layer, is the boundary between ontologies and historical data and serves as the ground on which business processes take place.

Three general rules are proposed for the generation of this interface. First, relations on the interface layer cannot exist between any arbitrary kind of concepts or relations, but only between instances of concepts or relations. It is important to understand the difference between an instance (instantiated type) and a grounded instance [11, p. 92]. An instance refers to a more precise rendering of a certain concept or relation by means of a specific configuration of its attributes, with the exception of those that express quantification or multiplicity (quantifiers) and grounding (determiners). An instance is an unspecified entity in a CG. However, instances cannot come into existence without first being grounded. An example of an instance would be a concept with a special configuration of its aspect attributes (e.g. size, color, shape). A grounded instance would be the same instance with added attributes of unique identification, quantity and a specific position in space and time. It corresponds to a specific entity in a CG with an individual marker.

The second property constrains the representations on the interface layer to procedures or processes. Specifically, relations between concept instances are expected to represent activities or events of a process execution. This property can be considered as a qualitative projection of the observed world into an ontology, because it determines which concept instances interrelate and how. Thereby, the interface layer is the ground on which most process analyses take place.

The third proposed rule states that the relations between instances on the interface layer should be weighted, which means that some relations will be stronger than others with magnitudes that range over an interval (i.e. from 0 to 1), representing a chosen value (e.g. cost, benefit, implication or causation). Notice that this property enforces a transformation of the layer from a normal CG into a FCG. This property can be considered as a quantitative projection of the observed world into an ontology, because it assigns magnitudes to the qualitative relations defined on the second property. Hence, a structure for the interface layer is proposed as follows.

Consider a basic CG G defined over a vocabulary V , where $graph(G)$ is constrained to be an undirected bipartite graph and the set of individuals $I = \emptyset$; then every node in G corresponds to an instance that is not grounded. Furthermore, the set R can be generalized to contain other kinds of nodes, e.g. actors, *demons*, constraint overlays [12], in order to enable procedural expression on the CG. A restriction is set so that all relations are binary (n-ary relations can be transformed polynomially into binary ones [2, p. 360]). Edges labeled 1 correspond to arcs exiting a node heading

to a relation, and edges labeled 2 are directed the opposite way [2, p. 114]. Thus, G corresponds to the interface layer.

Let c be a concept node and r a relation node on the interface layer, $i \in \{1, 2\}$ and $E_c = \{(r^*, i^*, c^*) \in E | c^* = c\}$ the set of edges adjacent to the node c . The number of edges incident to c that are labeled i is expressed as follows:

$$q(c) = \sum_{b \in E_c} n_b * u(b),$$

where n_b is the number of occurrences of the grounded edge b . The function u is then defined as follows:

$$u(n) = \begin{cases} 1 & \text{if } l(b) = i, \\ 0 & \text{if } l(b) \neq i, \end{cases}$$

where l is the labeling function. A function $p(c, b)$ denoting the probability of entering or exiting node c through edge b is defined as follows:

$$p(c, b) = \frac{n_b}{q(c)}.$$

Thus, the probability of exiting or entering node c is expressed as follows:

$$P = \sum_{b \in E_c} p(c, b) * u(b),$$

which evaluates to 0 or 1 depending on the existence of exiting or entering edges. When $i = 1$, the function p is denoted as the node routing through b . Updating the function with $i = 2$ provides the probability of entering the node c from b , which is denoted as the node inflow from b . In presence of n-ary relations, the same can be extended to a relation node c by applying p with $i = 1$ to obtain the inflow and $i = 2$ for the routing.

In the interface layer, however, relations are constrained to be binary. This means that both edges incident to a relation will have the same value of p . Thereby, the function $p(c, b)$, where b is incident to a relation r , will be denoted as the *fuzzy membership* of r on the interface layer. Thus, the underlying CG is transformed to a FCG, which can be visualized by a directed graph on which relations are represented by fuzzy edges (See Figure 2).

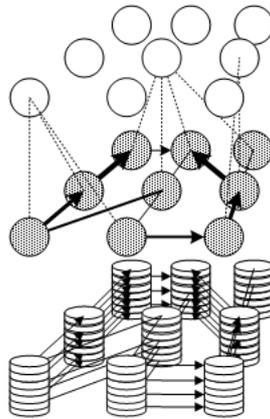


Figure 2. Layers on the KB

A quantification of the relation of a node to the entire layer is also proposed. Consider the function $f(c)$:

$$f(c) = \frac{q(c)}{2 \sum_{r \in R} n_r},$$

where n_r is the cardinality of the set of grounded instances of the relation r . When $i = 1$, f represents the exit factor or $ext(c)$, which is the weight of the exiting grounded edges from c in relation to the totality of grounded relations. When $i = 2$, the entering edges to c are considered, and this is denoted as the entrance factor or $ent(c)$. From both exit and entrance factors a correspondent pseudometric can be obtained. Consider the following expression:

$$d(x, y) = |f(y) - f(x)|,$$

where x, y are concept nodes. This corresponds to a percentage of the totality of grounded relations. When $i = 1$, function d is denoted as exit distance or d_{ext} . When $i = 2$, the correspondent entrance distance or d_{ent} is obtained. Consider the following relation:

$$h(c) = \frac{ext(c)}{ext(c) + ent(c)}.$$

The function h corresponds to the flow factor. If $h(c) > 1/2$, c is considered a source, with a degree of truth or fuzzy membership of $h(c)$. It is considered a sink if $h(c) < 1/2$, with a degree of truth of $1 - h(c)$. The flow factor can be assigned to nodes as the fuzzy membership value, depending on the task.

3.3. Ground layer

The proposed ground layer is composed of grounded instances of concepts and relations between them. This layer contains the historical (or real-time) data formatted as CG structures. However, this layer is unconstrained in relation to any ontology. In other words, grounded instances, which have been verified not to belong to any ontology, can be preserved on the layer. This is a crucial requirement of any process-oriented system, because it allows process discovery by means of process mining [13, p. 10].

4. Support for process analysis in the KB

4.1. Multiple contexts

Possible interdependence between work units and their ontologies or business processes is not the only source of complexity for analysis: the kind of analysis that is needed can be different every time, depending on a variety of factors. For example, analysis approaches can be generalized for the whole company while considering only a specific kind of resource, or they can be specific for a defined process while considering all the interrelations with adjacent processes. A multiplicity of scenarios can be brought to mind in which a static kind of knowledge is not suitable for the analysis of the task at hand. Therefore, it is necessary to consider the context of the analysis that will take place, so that the appropriate instantiations can be performed (See Figure 3).

A new instantiation of the interface layer can be obtained by defining a subset of all the concepts or relations that satisfy a certain condition (a filtering operation) and generating the interface layer in the ordinary way over the resulting node subset. New instantiations of the layer can also be obtained by extracting all the grounded instances that satisfy a certain condition, and thus obtaining fuzzy values for the interface nodes by means of performing a certain operation over the resulting grounded subsets. Combinations of these approaches can also be suitable for complex analyses.

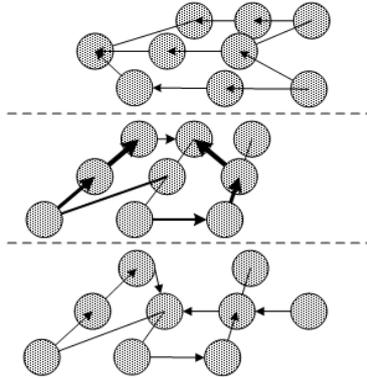


Figure 3. Interface instantiations

4.2. Process analysis

Graph search over the interface layer provides an interesting array of possibilities for process analysis. Business metrics can be reflected on concept instances on the layer, making it possible to visualize the correlations between them. Abnormal behavior of a metric can be determined by critical factor analysis (CFA) over the layer [1]. A straightforward approach to CFA is to translate a subgraph of the interface layer into a decision tree, in which degrees of membership and fuzzy relations of the graph can represent uncertainty. This creates a visual representation of every outcome from a starting node. If a visual approach is complicated to obtain, a path to a certain goal state (metric) from an initial state on the interface layer can be created by means of graph search techniques. A search can further be delimited by applying a set of constraints over the interface layer, e.g. preventing it to enter certain nodes where a condition is valid. The probability of ending in the goal state is obtained by the product of the fuzzy values of the edges of the path, assuming that they are normalized.

If only a specific business metric needs to be analyzed, an approach is to compare the relevant nodes with the proposed exit and entrance distances, as a measure for impact or priority. Furthermore, if the extracted process data comprises information about allocation of resources, it is possible to perform capacity planning. Groups of instances can be considered as separate entities subsumed under a common concept. By adding to the layer groups of instances, together with the correspondent constraints, a normal search can provide appropriate resource allocations to achieve the desired performance. The proposed flow factor can provide insight about existing bottlenecks and expose possible focus points for improvement or optimization.

5. Conclusion

The present work proposes the use of Conceptual Graphs with Fuzzy Logic for representing knowledge in process-oriented environments. Conceptual Graphs were discussed as a suitable and straightforward approach for representing and unifying knowledge. For this reason, a Knowledge Base structure based on Fuzzy Conceptual Graphs was proposed, with a high-level of abstraction. The KB features three layers: an ontological layer, merging declarative and procedural knowledge; a middle or interface layer, where real business processes are projected from historical or real-time data; a ground layer, where concrete instances and relations among them are stored for retrieval. The following coefficients are proposed to support process analysis on the interface layer: entrance factor, exit factor, entrance distance, exit distance, flow factor. Possible methods to analyze processes over the KB are exposed to provide evidence for the suitability of the proposed KB for process-oriented organizations.

References

1. J. Cardoso, W. Aalst, Handbook of Research on Business Process Modeling, Information Science Reference, Hershey, PA, 2009.
2. M. Chein, M. Mugnier, Graph-Based Knowledge Representation, Springer, London, 2009.
3. T. Cao, Conceptual Graphs and Fuzzy Logic, Springer-Verlag, Berlin, 2010.
4. G. Jakus, V. Milutinović, S. Omerović, S. Tomažič, Concepts, Ontologies, and Knowledge Representation, Springer, Heidelberg, 2013.
5. G. M. Novikova, An integral knowledge representation model for an intelligent management system, in: E. V. Popov (Ed.), Dynamic intelligent systems, Russian Central House of Knowledge, pp. 110–112.
6. J. F. Sowa, Logic: Graphical and algebraic, Manuscript (1997).
7. L. Zadeh, Toward extended fuzzy logic - a first step, Fuzzy Sets and Systems 160 (21) (2009) 3175–3181.
8. C. Laudy, J. Ganascia, C. Sedogbo, High-level fusion based on conceptual graphs, in: 10th International Conference on Information Fusion, 2007.
9. F. Smith, M. Proietti, Behavioral reasoning on semantic business processes in a rule-based framework, in: Communications in Computer and Information Science, 2014, pp. 293–313.
10. N. Silega, T. Loureiro, M. Noguera, Model-driven and ontology-based framework for semantic description and validation of business processes, IEEE Latin America Transactions 12 (2) (2014) 292–299.
11. R. Langacker, Foundations of Cognitive Grammar: descriptive application, Vol. 2, Stanford University Press, California, USA, 1991.
12. J. Lee, L. F. Lai, W. T. Huang, Task-based specifications through conceptual graphs, IEEE Expert 11 (4) (1996) 60–70.
13. W. Aalst, Process mining, Springer, Berlin, 2011.

УДК 004.82:519.17

Нечеткие концептуальные графы для представления знаний в процессно-ориентированной организации

Э. Х. Азофейфа, Г. М. Новикова

*Кафедра информационных технологий
Российский университет дружбы народов
ул. Миклухо-Маклая, д. 6, Москва, Россия, 117198*

Рассматривается использование нечетких концептуальных графов для представления знаний в процессно-ориентированных организациях. Рассмотрены два типа знаний — процедурные и декларативные, показано их различие и особенности представления и использования в базах знаний. Дано формальное определение нечетких концептуальных графов. Показаны их возможности для представления в простой и понятной форме как декларативных, так и процедурных знаний предметной области. Предложена структура Базы Знаний, которая включает три уровня: онтологический слой, содержащий концепты понятий и интегрирующий декларативные и процедурные знания; средний слой, описывающий схемы бизнес-процессов на основе исторических данных, формирующихся на базовом уровне; нижний слой, слой реальных данных, где собирается первичная информация о текущих состояниях объектов и отношений между ними. Показано отличие типов информации на каждом из слоев базы знаний, а также способ использования нечеткого концептуального графа на средне-интерфейсном слое. Рассмотрены механизмы взаимодействия и правила преобразования информации между средне-интерфейсным слоем и базовым слоем. Описан математический аппарат и методы анализа первичной информации для поддержки принятия решений по оптимизации и уточнению процедурных знаний системы, а также показатели, используемые в процессе анализа нечетких знаний.

Ключевые слова: нечёткие концептуальные графы, представления знаний, бизнес-процесс, базы знаний, онтологии, интеллектуальный агент.

Литература

1. *Cardoso J., Aalst W.* Handbook of Research on Business Process Modeling. — Hershey, PA: Information Science Reference, 2009. — Pp. 456–480.
2. *Chein M., Mugnier M.* Graph-Based Knowledge Representation. — London: Springer, 2009.
3. *Cao T.* Conceptual Graphs and Fuzzy Logic. — Berlin: Springer-Verlag, 2010.
4. Concepts, Ontologies, and Knowledge Representation / G. Jakus, V. Milutinović, S. Omerović, S. Tomažič. — Heidelberg: Springer, 2013.
5. *Новикова Г. М.* Интегрированная модель представления знаний в интеллектуальной системе управления // Динамические интеллектуальные системы / под ред. Э. В. Попов. — ЦРДЗ. — С. 110–112.
6. *Sowa J. F.* Logic: Graphical and Algebraic. — Manuscript. — 1997.
7. *Zadeh L.* Toward Extended Fuzzy Logic - a First Step // Fuzzy Sets and Systems. — 2009. — Vol. 160, No 21. — Pp. 3175–3181.
8. *Laudy C., Ganascia J., Sedogbo C.* High-Level Fusion Based on Conceptual Graphs // 10th International Conference on Information Fusion. — 2007.
9. *Smith F., Proietti M.* Behavioral Reasoning on Semantic Business Processes in a Rule-Based Framework // Communications in Computer and Information Science. — 2014. — Pp. 293–313.
10. *Silega N., Loureiro T., Noguera M.* Model-Driven and Ontology-Based Framework for Semantic Description and Validation of Business Processes // IEEE Latin America Transactions. — 2014. — Vol. 12, No 2. — Pp. 292–299.
11. *Langacker R.* Foundations of Cognitive Grammar: descriptive application. — California, USA: Stanford University Press, 1991. — Vol. 2.
12. *Lee J., Lai L. F., Huang W. T.* Task-Based Specifications Through Conceptual Graphs // IEEE Expert. — 1996. — Vol. 11, No 4. — Pp. 60–70.
13. *Aalst W.* Process mining. — Berlin: Springer, 2011.

Рассматривается задача анализа и выбора информативных признаков, являющихся яркостными и геометрическими инвариантами. Предложен аналитический обзор работ, в которых поднималась проблема построения инвариантов и решались практические задачи. В числе инвариантов к мультипликативным и аддитивным преобразованиям яркости рассмотрены коэффициент корреляции, нормализованная гистограмма, некоторые отношения особых отрезков линий и углы. Рассмотрены как простейшие геометрические, так и интегральные инварианты, устойчивые к аффинным преобразованиям, а также к преобразованиям растяжения. Среди всех геометрических инвариантов выделяются инварианты H_c , интерес к которым не ослабевает и в настоящее время. Поскольку в первоисточниках отсутствуют доказательства утверждений об инвариантности, то большое внимание уделяется проверке на некоторых моментах инвариантности к часто употребляемой операции поворота. Рассмотренные признаки можно использовать для решения задач распознавания графических образов, измерения расстояний, анализа сложных сцен. Предполагается, что в дальнейшем на основе яркостных и геометрических инвариантов будут решены задачи распознавания лиц человека. Статья предназначена для широкого круга аспирантов и инженеров, занимающихся задачами распознавания графических образов в различных приложениях.

Ключевые слова: изображение, графический образ, распознавание, яркостные инварианты, инвариантные моменты, аффинные преобразования, растяжения.

1. Введение

В задачах распознавания графических образов желательно иметь такие характеристики изображения, которые не зависели бы от масштаба, ориентации и качества анализируемого снимка. Такие характеристики называют инвариантами. Широкое распространение получили яркостные и геометрические инварианты изображений [1, 2]. Инварианты полезны тем, что способны при определённых условиях выступать в роли существенных признаков, передающих характерные особенности изображения. Они позволяют проводить корректное сравнение изображений подвёрнутых, преобразованиями, что увеличивает вероятность получения правильного решения. Это могут быть, например, аффинные преобразования, растяжения, изменения яркости и контрастности.

Инвариантный подход к распознаванию является одним из немногих подходов, допускающих точную математическую постановку задачи и позволяющих выработать понятия классов объектов. Инварианты обладают способностью к обобщению, что является необходимым атрибутом искусственного интеллекта, например искусственных нейронных сетей [3].

Разобьём множество всех изображений объектов на снимке или множестве снимков, представляющих интерес, на непересекающиеся классы таким образом, чтобы каждый класс состоял из всех таких изображений, которые отличаются друг от друга лишь преобразованием указанной группы. Задача распознавания заключается в том, чтобы для произвольного изображения объекта на снимке определить, к какому классу он относится.

Задачу математически можно было бы решать путём построения полной системы инвариантов от изображений объектов относительно преобразований группы, значения которых постоянны в пределах каждого класса и различны для любых двух разных классов. Однако, следует отметить, что реальное распознавание объектов усложняется тем, что инварианты объектов, на самом деле, вследствие влияния различных факторов, например помех, могут флуктуировать, они являются

«субинвариантами». В этой связи в качестве методов распознавания приходится привлекать вероятностные подходы и специальные метрики сравнения инвариантов, например метрику Евклида-Махаланобиса [4].

В некоторых случаях авторы оригинальных работ не приводят необходимых доказательств, подтверждающих инвариантность предложенных характерных признаков. Таким образом, утверждение об инвариантности носит характер гипотезы и требует перепроверки.

2. Яркостные инварианты изображений

Цифровое изображение обычно является результатом дискретизации непрерывной функции $f(x, y)$ и хранится в компьютере в виде двумерного массива $I(i, j)$, где $i = 0, 1, 2, \dots, N_x - 1$ и $j = 0, 1, 2, \dots, N_y - 1$. Каждый элемент этого массива представляет собой пиксель с интенсивностью $I(i, j)$, изменяющейся в диапазоне от 0 до $L - 1$. Величина L обычно является степенью двойки (например, 64, 256) и называется глубиной изображения [5].

Пусть исходные изображения представлены в формате JPEG и цветовой модели RGB. Для дальнейшей работы часто достаточно воспользоваться полутоновым изображением (изображением в градациях серого) [6]. Согласно стандарту ITU-R BT.601 яркость отдельного пикселя в таком случае определяется по формуле $z = 0.56G + 0.33R + 0.11B$, где R , G и B представляют собой компоненты вектора в пространстве RGB.

По отношению к мультипликативным изменениям яркости инвариантами выступают следующие соотношения [1]:

$$\frac{B_{min}}{B_{max}}, \quad \frac{B_{max}}{B_{cp}}, \quad \frac{\mu(B)}{B_{cp}},$$

где B_{min} , B_{max} , B_{cp} — соответственно минимальное, максимальное и среднее значения яркостей изображения, $\mu(B) = \sum_{i,j} |B_{ij} - B_{cp}|$.

В случае аддитивного преобразования, инвариантами являются

$$B_{max} - B_{min}, B_{max} - B_{cp}.$$

Инварианты по отношению к аддитивному и мультипликативному преобразованию можно получить следующим образом:

$$B_{max} - B_{min}, B_{max} + B_{min} - 2B_{cp}.$$

Коэффициент корреляции, часто используемый для сравнения изображений, имеет вид

$$\gamma_R(B_1, B_2) = \frac{\langle B_1, B_2 \rangle}{\|B_1\| \|B_2\|},$$

где $B = |B_{ij}|$ — матрица яркостей пикселей, $(i, j = 0, 1, \dots, n)$, $\langle B_1, B_2 \rangle = \sum_i \sum_j B_1(i, j) B_2(i, j)$ — скалярное произведение двух матриц $(i, j \in R)$, $\|B\|$ —

норма матрицы B , $\|B\| = \sqrt{\sum_{i,j \in R} B^2(i, j)}$.

Для того, чтобы корреляционный коэффициент был инвариантом к аддитивным и мультипликативным преобразованиям, используют формулу

$$\gamma_R(\widehat{B}_1, \widehat{B}_2) = \frac{\langle \widehat{B}_1, \widehat{B}_2 \rangle}{\|\widehat{B}_1\| \|\widehat{B}_2\|},$$

где $\widehat{B}_k(i, j) = B_k(i, j) - \overline{B}_k$, где \overline{B}_k — среднее арифметическое яркостей изображения.

Коэффициент корреляции вычисляют при наложении одного изображения на другое, причём сравнение происходит по соответствующим пикселям.

3. Геометрические инварианты изображений

Как отмечалось ранее, геометрические инварианты — это такие характеристики изображения, которые не зависят от масштаба, ориентации, качества снимка. Рассмотрим задачу получения инвариантов по отношению к аффинным преобразованиям.

3.1. Простейшие геометрические инварианты

Инвариантами к операции плоского поворота изображения объекта являются площадь S и длина контура изображения L .

Пусть (\bar{x}, \bar{y}) — центр изображения, d_{max}, d_{min} — максимальное и минимальное значения расстояний до границы изображения, как показано на рис. 1.

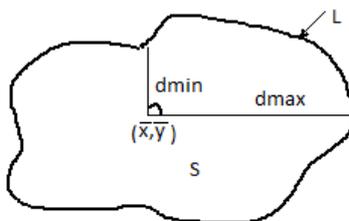


Рис. 1. Пояснения к задаче построения геометрических инвариантов

Тогда инвариантами к операции масштабирования и поворота являются: $\frac{d_{min}}{d_{max}}$;
 α — угол между d_{min} и d_{max} ; $\frac{L}{2S}$.

Опишем вокруг отдельного изображения окружность, тогда инвариантом к повороту и масштабированию будет отношение $\frac{S}{S_R}$, где S — площадь изображения, S_R — площадь описанного круга.

На свойстве подобия окружностей может быть построен метод сравнения и распознавания сложного изображения. Для этого отдельные изображения представляются совокупностью вписанных в него кругов. Таким образом, паре (эталонное и тестируемое изображение) сопоставляется вектор, компонентами которого являются радиусы и координаты центров соответствующих кругов. Исходя из этого можно найти расстояние между изображениями.

Яркостные и геометрические характеристики могут быть связанными. Будем рассматривать функцию яркости изображения как стационарный случайный процесс [7]. В этом случае искомыми признаками для каждого снимка будут служить числовые характеристики случайного процесса. К наиболее часто используемым признакам относятся яркостные характеристики, такие как гистограмма распределения значений яркости на изображении. Составим нормализованную гистограмму для каждого исходного изображения:

$$p(z_i) = \frac{n_i}{n},$$

где n_i — количество пикселей яркости (интенсивности) z_i ($i = 0, \dots, L - 1$), n — общее число пикселей в изображении, L — количество градаций. Величина $p(z_i)$ является оценкой вероятности появления пикселя с интенсивностью z_i . Следовательно, выполняется условие нормировки:

$$\sum_{k=0}^{L-1} p(z_k) = 1.$$

Нормализация гистограммы позволяет избежать зависимости результатов распознавания от масштабирования.

3.2. Коэффициенты тейлоровского разложения

В работе [8] определяются и исследуются четыре дифференциальных инварианта точки изображения относительно изменения масштаба и поворота. Будем предполагать, что функция яркости $f(x, y)$ дважды дифференцируема, тогда справедливо разложение Тейлора второго порядка с центром в произвольной точке области (рис. 2). Предположим, что снимок подвергся преобразованию (повороту и растяжению).

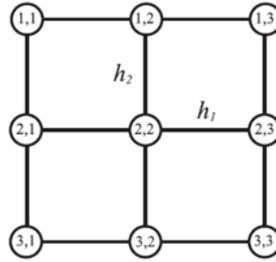


Рис. 2. Прямоугольная сетка дискретного изображения

Рассмотрим разложение функции $f(x, y)$ с центром в точке $(2, 2)$. Если найти значение данного разложения в узлах сети согласно рис. 2 и приравнять их данным $f_{i,j}$, то в результате получим систему из 9 уравнений на 6 неизвестных коэффициентов тейлоровского разложения. В матричной форме можно записать: $A * X = F$, где X — столбец неизвестных коэффициентов $a, p_1, p_2, b_{11}, b_{12}, b_{22}$; F — столбец значений функции в узлах $f_{22}, f_{12}, f_{21}, f_{23}, f_{32}, f_{11}, f_{13}, f_{33}, f_{31}$.

Матрица A равна

$$A = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ 1 & -h_1 & 0 & \frac{h_1^2}{2} & 0 & 0 \\ 1 & 0 & -h_2 & 0 & \frac{h_2^2}{2} & 0 \\ 1 & 0 & h_2 & 0 & \frac{h_2^2}{2} & 0 \\ 1 & h_1 & 0 & \frac{h_1^2}{2} & 0 & 0 \\ 1 & -h_1 & -h_2 & \frac{h_1^2}{2} & \frac{h_2^2}{2} & h_1 h_2 \\ 1 & -h_1 & h_2 & \frac{h_1^2}{2} & \frac{h_2^2}{2} & -h_1 h_2 \\ 1 & h_1 & h_2 & \frac{h_1^2}{2} & \frac{h_2^2}{2} & h_1 h_2 \\ 1 & h_1 & -h_2 & \frac{h_1^2}{2} & \frac{h_2^2}{2} & -h_1 h_2 \end{pmatrix}.$$

Решая систему по методу наименьших квадратов, получим

$$X = A^{-1} * F,$$

где псевдообратная матрица A^{-1} равна

$$A^{-1} = \begin{pmatrix} \frac{5}{9} & \frac{2}{9} & \frac{2}{9} & \frac{2}{9} & \frac{2}{9} & -\frac{1}{9} & -\frac{1}{9} & -\frac{1}{9} & -\frac{1}{9} \\ 0 & -\frac{1}{6h_1} & 0 & 0 & \frac{1}{6h_1} & -\frac{1}{6h_1} & -\frac{1}{6h_1} & \frac{1}{6h_1} & \frac{1}{6h_1} \\ 0 & 0 & -\frac{1}{6h_2} & \frac{1}{6h_2} & 0 & -\frac{1}{6h_2} & \frac{1}{6h_2} & \frac{1}{6h_2} & -\frac{1}{6h_2} \\ -\frac{2}{3h_1^2} & \frac{1}{3h_1^2} & -\frac{2}{3h_1^2} & -\frac{2}{3h_1^2} & \frac{1}{3h_1^2} & \frac{1}{3h_1^2} & \frac{1}{3h_1^2} & \frac{1}{3h_1^2} & \frac{1}{3h_1^2} \\ -\frac{2}{3h_2^2} & -\frac{2}{3h_2^2} & \frac{1}{3h_2^2} & \frac{1}{3h_2^2} & -\frac{2}{3h_2^2} & \frac{1}{3h_2^2} & \frac{1}{3h_2^2} & \frac{1}{3h_2^2} & \frac{1}{3h_2^2} \\ 0 & 0 & 0 & 0 & 0 & \frac{1}{4h_1h_2} & -\frac{1}{4h_1h_2} & \frac{1}{4h_1h_2} & -\frac{1}{4h_1h_2} \end{pmatrix}.$$

Теорема. Для квадратичной функции $f(x, y)$ формула $X = A^{-1} * F$ даёт точные значения коэффициентов тейлоровского разложения второго порядка с центром в данной точке.

Следствие. Для квадратичной функции $f(x, y)$ эта формула позволяет вычислить точные значения инвариантов. Аналогичные исследования выполнены в работах [9–11].

3.3. Инвариантные моменты

Значительных успехов достигла теория инвариантных моментов, основанная на поиске интегральных инвариантов. Математиком Ну в работах [2, 12] были предложены семь моментов, инвариантных к группе аффинных преобразований. Можно показать методом прямой подстановки соответствующих преобразований и коэффициентов в выражения для моментов и эквивалентных преобразований, что выражения, полученные Ну, действительно являются инвариантами к операции сдвига и масштабированию.

Доказательство инвариантности к операциям сдвига не представляет затруднений. Проверка инвариантности моментов к операции масштабирования содержится в работе [13].

Покажем на нескольких примерах, что моменты $Hu(m_1, m_2, \dots, m_7)$ являются инвариантами к операции поворота изображения вокруг начала координат. Для этого будем пользоваться выражением для центральных моментов бинарного изображения:

$$m_{pq} = \frac{1}{N} \sum_{i=1}^N (x_i - \bar{x})^p (y_i - \bar{y})^q,$$

где $p + q \leq 3$, (x_i, y_i) — искомая точка изображения, (\bar{x}, \bar{y}) — центр изображения, N — число точек изображения.

а. Проверка инварианта m_1 :

$$m_1 = m_{02} + m_{20} = \frac{1}{N} \sum_{i=1}^N [(x_i - \bar{x})^2 (y_i - \bar{y})^2].$$

Подставив соответствующее преобразование поворота точек изображения вокруг начала координат на угол, получим

$$m_1 = \frac{1}{N} \sum_{i=1}^N \left[\left((x_i \ y_i) \begin{pmatrix} \cos \alpha \\ -\sin \alpha \end{pmatrix} - (\bar{x} \ \bar{y}) \begin{pmatrix} \cos \alpha \\ -\sin \alpha \end{pmatrix} \right)^2 + \right.$$

$$+ \left((x_i \ y_i) \begin{pmatrix} \sin \alpha \\ \cos \alpha \end{pmatrix} - (\bar{x} \ \bar{y}) \begin{pmatrix} \sin \alpha \\ \cos \alpha \end{pmatrix} \right)^2 \Big].$$

Достаточно воспользоваться следующим утверждением,

$$\left((a \ b) \begin{pmatrix} \cos \alpha \\ -\sin \alpha \end{pmatrix} \right)^2 + \left((a \ b) \begin{pmatrix} \sin \alpha \\ \cos \alpha \end{pmatrix} \right)^2 = a^2 + b^2$$

и рассмотреть выражение в квадратных скобках

$$\begin{aligned} & \left((x_i \ y_i) \begin{pmatrix} \cos \alpha \\ -\sin \alpha \end{pmatrix} - (\bar{x} \ \bar{y}) \begin{pmatrix} \cos \alpha \\ -\sin \alpha \end{pmatrix} \right)^2 + \\ & \quad + \left((x_i \ y_i) \begin{pmatrix} \sin \alpha \\ \cos \alpha \end{pmatrix} - (\bar{x} \ \bar{y}) \begin{pmatrix} \sin \alpha \\ \cos \alpha \end{pmatrix} \right)^2 = \\ & = \left((x_i - \bar{x} \ y_i - \bar{y}) \begin{pmatrix} \cos \alpha \\ -\sin \alpha \end{pmatrix} \right)^2 + \left((x_i - \bar{x} \ y_i - \bar{y}) \begin{pmatrix} \sin \alpha \\ \cos \alpha \end{pmatrix} \right)^2 = \\ & \qquad \qquad \qquad = (x_i - \bar{x})^2 + (y_i - \bar{y})^2. \end{aligned}$$

Таким образом, выражение в скобках является инвариантом, что и требовалось доказать.

b. Проверка инварианта m_2 :

$$\begin{aligned} m_2 &= (m_{20} - m_{02})^2 + 4m_{11}^2 = \\ &= \left(\frac{1}{N} \sum_{i=1}^N [(x_i - \bar{x})^2 - (y_i - \bar{y})^2] \right)^2 + 4 \left(\frac{1}{N} \sum_{i=1}^N [(x_i - \bar{x})(y_i - \bar{y})] \right)^2. \end{aligned}$$

Утверждение 1.

$$\left((a \ b) \begin{pmatrix} \cos \alpha \\ -\sin \alpha \end{pmatrix} \right)^2 - \left((a \ b) \begin{pmatrix} \sin \alpha \\ \cos \alpha \end{pmatrix} \right)^2 = (a^2 - b^2)(\cos^2 \alpha - \sin^2 \alpha) - 4ab \sin \alpha \cos \alpha.$$

Рассмотрим выражение по частям.

Первая часть выражения:

$$\begin{aligned} m_{20} - m_{02} &= \frac{1}{N} \sum_{i=1}^N \left[\left((x_i \ y_i) \begin{pmatrix} \cos \alpha \\ -\sin \alpha \end{pmatrix} - (\bar{x} \ \bar{y}) \begin{pmatrix} \cos \alpha \\ -\sin \alpha \end{pmatrix} \right)^2 - \right. \\ & \quad \left. - \left((x_i \ y_i) \begin{pmatrix} \sin \alpha \\ \cos \alpha \end{pmatrix} - (\bar{x} \ \bar{y}) \begin{pmatrix} \sin \alpha \\ \cos \alpha \end{pmatrix} \right)^2 \right], \end{aligned}$$

$$\begin{aligned} m_{20} - m_{02} &= \\ &= \frac{1}{N} \sum_{i=1}^N \left[\left((x_i - \bar{x} \ y_i - \bar{y}) \begin{pmatrix} \cos \alpha \\ -\sin \alpha \end{pmatrix} \right)^2 - \left((x_i - \bar{x} \ y_i - \bar{y}) \begin{pmatrix} \sin \alpha \\ \cos \alpha \end{pmatrix} \right)^2 \right] = \\ &= \frac{1}{N} \left[\sum_{i=1}^N \left[((x_i - \bar{x})^2 - (y_i - \bar{y})^2)(\cos^2 \alpha - \sin^2 \alpha) - 4(x_i - \bar{x})(y_i - \bar{y}) \sin \alpha \cos \alpha \right] \right], \end{aligned}$$

$$\begin{aligned}
& (m_{20} - m_{02})^2 = \\
& = \frac{1}{N^2} \left[\sum_{i=1}^N \left[((x_i - \bar{x})^2 - (y_i - \bar{y})^2)(\cos^2 \alpha - \sin^2 \alpha) - 4(x_i - \bar{x})(y_i - \bar{y}) \sin \alpha \cos \alpha \right] \right]^2 \times \\
& \quad \times \left[((x_i - \bar{x})^2 - (y_i - \bar{y})^2)(\cos^2 \alpha - \sin^2 \alpha) - 4(x_i - \bar{x})(y_i - \bar{y}) \sin \alpha \cos \alpha \right]^2 = \\
& \quad = \left[(x_i - \bar{x})^2 - (y_i - \bar{y})^2 \right]^2 (\cos^2 \alpha - \sin^2 \alpha)^2 - \\
& \quad - 8 \left[(x_i - \bar{x})^2 - (y_i - \bar{y})^2 \right] (\cos^2 \alpha - \sin^2 \alpha)^2 \cos \alpha \sin \alpha (x_i - \bar{x})(y_i - \bar{y}) + \\
& \quad + 16 \left[(x_i - \bar{x})^2 (y_i - \bar{y})^2 \cos^2 \alpha \sin^2 \alpha \right].
\end{aligned}$$

Рассмотрим вторую часть:

$$\begin{aligned}
4m_{11}^2 & = 4 \left(\frac{1}{N} \sum_{i=1}^N [(x_i - \bar{x})(y_i - \bar{y})] \right)^2 = \\
& = \frac{4}{N^2} \left(\sum_{i=1}^N \left((x_i \ y_i) \begin{pmatrix} \cos \alpha \\ -\sin \alpha \end{pmatrix} - (\bar{x} \ \bar{y}) \begin{pmatrix} \cos \alpha \\ -\sin \alpha \end{pmatrix} \right) \times \right. \\
& \quad \left. \times \left((x_i \ y_i) \begin{pmatrix} \sin \alpha \\ \cos \alpha \end{pmatrix} - (\bar{x} \ \bar{y}) \begin{pmatrix} \sin \alpha \\ \cos \alpha \end{pmatrix} \right) \right)^2 = \\
& = \frac{4}{N^2} \left(\sum_{i=1}^N (x_i \cos \alpha - y_i \sin \alpha - \bar{x} \cos \alpha + \bar{y} \sin \alpha) (x_i \sin \alpha + y_i \cos \alpha - \bar{x} \sin \alpha - \bar{y} \cos \alpha) \right)^2 = \\
& = \frac{4}{N^2} \left(\sum_{i=1}^N (x_i - \bar{x} \cos \alpha - (y_i - \bar{y}) \sin \alpha) (x_i - \bar{x}) \sin \alpha + (y_i - \bar{y}) \cos \alpha \right)^2 = \\
& = \frac{4}{N^2} \left(\sum_{i=1}^N (x_i - \bar{x} \cos \alpha - (y_i - \bar{y}) \sin \alpha) (x_i - \bar{x}) \sin \alpha + (y_i - \bar{y}) \cos \alpha \right)^2 = \\
& = \frac{4}{N^2} \left(\sum_{i=1}^N ((x_i - \bar{x})^2 \cos \alpha \sin \alpha - (x_i - \bar{x})(y_i - \bar{y}) \sin^2 \alpha + \right. \\
& \quad \left. + (x_i - \bar{x})(y_i - \bar{y}) \cos^2 \alpha - (y_i - \bar{y})^2 \cos \alpha \sin \alpha) \right)^2 = \\
& = \frac{4}{N^2} \left(\sum_{i=1}^N ((x_i - \bar{x})^2 - (y_i - \bar{y})^2 \cos \alpha \sin \alpha + (x_i - \bar{x})(y_i - \bar{y})(\cos^2 \alpha - \sin^2 \alpha)) \right)^2 = \\
& = \frac{4}{N^2} \left(\sum_{i=1}^N [((x_i - \bar{x})^2 - (y_i - \bar{y})^2)^2 \cos^2 \alpha \sin^2 \alpha + \right. \\
& \quad \left. + 2((x_i - \bar{x})^2 - (y_i - \bar{y})^2) \cos \alpha \sin \alpha (x_i - \bar{x})(y_i - \bar{y})(\cos^2 \alpha - \sin^2 \alpha) + \right. \\
& \quad \left. + (x_i - \bar{x})^2 (y_i - \bar{y})^2 (\cos^2 \alpha - \sin^2 \alpha)^2 \right].
\end{aligned}$$

Собирая обе части, получим

$$\begin{aligned}
& \frac{1}{N} \sum_{i=1}^N \left[\left[(x_i - \bar{x})^2 - (y_i - \bar{y})^2 \right]^2 (\cos^2 \alpha - \sin^2 \alpha)^2 - \right. \\
& 8 \left[(x_i - \bar{x})^2 - (y_i - \bar{y})^2 \right] (\cos^2 \alpha - \sin^2 \alpha)^2 \cos \alpha \sin \alpha (x_i - \bar{x})(y_i - \bar{y}) + \\
& 16 \left[(x_i - \bar{x})^2 (y_i - \bar{y})^2 \cos^2 \alpha \sin^2 \alpha \right] + \\
& 4(x_i - \bar{x})^2 - (y_i - \bar{y})^2)^2 \cos^2 \alpha \sin^2 \alpha + \\
& 8((x_i - \bar{x})^2 - (y_i - \bar{y})^2) \cos \alpha \sin \alpha (x_i - \bar{x})(y_i - \bar{y})(\cos^2 \alpha - \sin^2 \alpha) + \\
& 4(x_i - \bar{x})^2 (y_i - \bar{y})^2 (\cos^2 \alpha - \sin^2 \alpha)^2 = \\
& \left[(x_i - \bar{x})^2 - (y_i - \bar{y})^2 \right]^2 \left[(\cos^2 \alpha - \sin^2 \alpha)^2 + 4 \cos^2 \alpha \sin^2 \alpha \right] + \\
& + \left[(x_i - \bar{x})^2 (y_i - \bar{y})^2 \right] \left[4(\cos^2 \alpha - \sin^2 \alpha)^2 + 16 \cos^2 \alpha \sin^2 \alpha \right] = \\
& \left(\left[(x_i - \bar{x})^2 - (y_i - \bar{y})^2 \right]^2 + 4 \left[(x_i - \bar{x})^2 (y_i - \bar{y})^2 \right] \right) \times \\
& \times \left[(\cos^2 \alpha - \sin^2 \alpha)^2 + 4 \cos^2 \alpha \sin^2 \alpha \right] = \\
& \left(\left[(x_i - \bar{x})^2 - (y_i - \bar{y})^2 \right]^2 + 4 \left[(x_i - \bar{x})^2 (y_i - \bar{y})^2 \right] \right) \left[(\cos^2 \alpha + \sin^2 \alpha)^2 \right] = \\
& \left[(x_i - \bar{x})^2 - (y_i - \bar{y})^2 \right]^2 + 4 \left[(x_i - \bar{x})^2 (y_i - \bar{y})^2 \right].
\end{aligned}$$

Что и требовалось доказать.

Аналогично проводится проверка для других моментов, что требует несколько большего объёма преобразований. Эти доказательства в настоящей работе не приводятся ввиду их громоздкости.

4. Заключение

В статье выполнен обзор работ, который позволил выявить несколько полезных инвариантов к преобразованиям яркости и аффинным преобразованиям. Показано, что в качестве инвариантов могут выступать как очень простые признаки, так и сложные интегральные признаки, которые являются более устойчивыми к различным помехам. К их числу можно отнести коэффициенты тейлоровского разложения изображения и инварианты Ну. Для двух моментов Ну продемонстрирована их инвариантность к операции поворота. Работа выполнена в рамках проекта Программы фундаментальных исследований ОНИТ 1 РАН «Интеллектуальные информационные технологии, системный анализ и автоматизация».

Литература

1. *Путьгин Е. П., Аверин С. И.* Обработка изображений в робототехнике // Машиностроение. — 1990. — Т. 320. — С. 1–320.
2. *Wong R. Y., Hall E. L.* Scene Matching with Invariant Moments // Computer Graphics and Image Processing. — 1978. — Vol. 9, No 8. — Pp. 16–24.
3. *Шмидт А. А.* Инвариантные методы в теории распознавания изображений // Дисс. на соиск. учен. степени к.ф.-м.н. — 1984. — Т. 164. — С. 1–164.
4. *Хачумов М. В.* Расстояния, метрики и кластерный анализ // Искусственный интеллект и принятие решений. — 2012. — Т. 9, № 1. — С. 81–89.
5. *Theodoridis S., Koutroumbas K.* Pattern Recognition. — 2003.
6. NTSC. — <http://en.wikipedia.org/wiki/NTSC>.

7. *Сойфер В. А.* Методы компьютерной обработки изображений // Физматлит. — 2003.
8. *Самарина О. В., Славский В. В.* Инварианты изображения относительно поворотов и растяжений // Вестник СамГУ. — Естественнонаучная серия. — 2007. — Т. 10, № 9/1(59). — С. 128–137.
9. *Peter J. O.* Equivalence, Invariants, and Symmetry // Cambridge University Press. — 1995.
10. *Walker K. N., Cootes T. F., Taylor C. J.* Locating Salient Facial Features Using Image Invariants. Medical Biophysics. — UK, 1998.
11. *Gouet M. V., Montesinos P., Pele D.* Stereo Matching of Color Images Using Differential Invariants. — International Conference on Image Processing, 1999.
12. *Hu M. K.* Visual Pattern Recognition by Moment Invariants // IRE Transactions on Information Theory. — 1962. — Vol. 8. — Pp. 179–187.
13. *Абрамов Н. С., Фраленко В. П.* Определение расстояний на основе системы технического зрения и метода инвариантных моментов // Информационные технологии и вычислительные системы. — 2012. — № 4. — С. 32–39.

UDC 004.931

Invariant in the Pattern Recognition

Duy Thanh Nguyen

*Department of Information Technology
Russian Peoples' Friendship University
6, Miklukho-Maklaya str., Moscow, Russia, 117198*

Consider the problem analysis and selection of informative features, which are the luminance and geometric invariants. An analytical overview of the works, which raises the problem of constructing invariants and solve practical problems. Among the invariants to the multiplicative and additive transformations brightness discussed the correlation coefficient, the normalized histogram of some special relationship of line segments and angles. It is considered as the simplest geometric and integral invariants resistant to affine transformations, as well as the transformation strain. Among all geometric invariants allocated invariants H_u , in which interest is not waning and now. Since there is no evidence in the original sources of the invariance of the allegations, the great attention is paid to checking on some points invariance to rotation of frequently used operations. The above features can be used for solving problems of pattern recognition, distance measurement, analysis of complex scenes. It is expected that in the future on the basis of brightness and geometric invariants will be solved the problem of recognizing individual rights. This article is intended for a wide range of post-graduate students and engineers involved in pattern recognition tasks in various applications.

Key words and phrases: image, graphical image, recognition, brightness invariants, invariant moments, affine transformations, stretching.

References

1. E. P. Putyatin, S. I. Averin, Image Processing in Robotics, Mechanical engineering 320 (1990) 1–320, in Russian.
2. R. Y. Wong, E. L. Hall, Scene Matching with Invariant Moments, Computer Graphics and Image Processing 9 (8) (1978) 16–24.
3. A. A. Schmidt, Invariant Methods in the Theory of Image Recognition, Diss. on soisk. scientists. PhD degree. Physics and Mathematics 164 (1984) 1–164, in Russian.
4. M. V. Khachumov, Distances, Metrics and Cluster Analysis, Artificial Intelligence and Decision Making 9 (1) (2012) 81–89, in Russian.
5. S. Theodoridis, K. Koutroumbas, Pattern Recognition, Elsevier Academic, 2003.
6. Ntsc.
URL <http://en.wikipedia.org/wiki/NTSC>

7. V. A. Soifer, *Methods of Computer Image Processing*, FIZMATLIT, 2003, in Russian.
8. O. V. Samarin, V. V. Slavskii, *Image Invariants with Respect to Rotations and Dilations*, Vestnik of SSU. — Natural Science Series 10 (9).
9. P. J. Olver, *Equivalence, Invariants, and Symmetry*, 1995.
10. K. N. Walker, T. F. Cootes, C. J. Taylor, *Locating Salient Facial Features Using Image Invariants*. Medical Biophysics, Manchester University, UK, 1998.
11. M. V. Gouet, P. Montesinos, D. Pele, *Stereo Matching of Color Images Using Differential Invariants.*, International Conference on Image Processing.
12. M. K. Hu, *Visual Pattern Recognition by Moment Invariants*, IRE Transactions on Information Theory 8 (1962) 179–187.
13. N. S. Abramov, V. P. Pavlenko, *Distances Based Vision System and Method of Invariant Moments*, Information technology and computer systems (4) (2012) 32–39, in Russian.

Физика

UDC 539.12

Lagrangian Density of Lepton and Baryon Phases in Nonlinear 8-Spinor Model

V. I. Molotkov

*Department of Theoretical Physics and Mechanics
Peoples' Friendship University of Russia
6, Miklukho-Maklaya str., Moscow, Russia, 117198*

The Skyrme's idea (1954) for describing baryons as topological solitons was based on the identification of baryon number B with the topological charge of the degree type $B = \text{deg}(S^3 \rightarrow S^3)$. It serves as the generator of the third homotopy group $\pi_3(S^3) = \mathbb{Z}$. The similar idea to describe leptons as topological solitons was announced by Faddeev (1972). He identified the lepton number L with the Hopf invariant Q_H .

The 8-spinor field is suggested to unify Skyrme and Faddeev models describing baryons and leptons as topological solitons. The special 8-spinor Brioschi identity is used to include leptons and baryons as two possible phases of the effective spinor field model, with Higgs potential depending on the $j^\mu j_\mu$ being added to the Lagrangian.

To this end the generalization of the Mie electrodynamics within the scope of the effective 8-spinor field model is suggested. For this field model the quadratic spinor quantities entering the Brioschi identity are constructed. Then the symmetry groups, which generate S^2 - and S^3 -submanifolds in general S^8 biquadratic spinor manifold, are found. For unifying these phases, common vacuum state should conserve only one component in both lepton and baryon cases.

In the present paper we try to construct Lagrange density for homotopy groups $\pi_3(S^2)$ and $\pi_3(S^3)$, which describe lepton and baryon phases.

Key words and phrases: 8-spinor, topological charge, solitons, homotopy groups, Brioschi identity, Skyrme–Faddeev model.

1. Introduction

In the Skyrme model the particles-solitons possess the topological charge of the degree type $B = \text{deg}(S^3 \rightarrow S^3)$, which is interpreted as the baryon number B . In the Faddeev model the particles are endowed with the topological invariant of the Hopf type, which is interpreted as the Lepton number. This invariant serves as the generator of the third homotopy group $\pi_3(S^2) = \mathbb{Z}$. Prof. Yu. P. Rybakov suggested [1] to unify these two approaches describing baryons and leptons as two possible phases of the effective 8-spinor field model. For this goal, the special 8-spinor Brioschi identity is used [2]:

$$j_\mu j^\mu - \tilde{j}_\mu \tilde{j}^\mu = s^2 + p^2 + v^2 + a^2. \quad (1)$$

In this identity the following quadratic spinor quantities are introduced:

$$\begin{aligned} s &= \bar{\psi}\psi, \\ p &= i\bar{\psi}\gamma_5\psi, \\ v &= \bar{\psi}\lambda\psi, \\ a &= i\bar{\psi}\gamma_5\lambda\psi, \\ j_\mu &= \bar{\psi}\gamma_\mu\psi, \\ \tilde{j}_\mu &= \bar{\psi}\gamma_\mu\gamma_5\psi, \end{aligned}$$

Received 8th December, 2015.

The author expresses his gratitude to his scientific supervisor Prof. Yu. P. Rybakov and appreciates the helpful discussions of the problem in question with I. E. Bulyzhenkov.

with $\bar{\psi} = \psi^+ \gamma_0$ and matrices $\lambda = \sigma_i \otimes I_4$ standing for the Pauli matrices in the isotopic space. Here and below we use the Weyl representation for Dirac matrices γ_μ , $\mu = 0, 1, 2, 3$, which may be written in the block form using the Pauli matrices $\sigma_1, \sigma_2, \sigma_3$:

$$\gamma_0 = \begin{pmatrix} 0 & I_2 \\ I_2 & 0 \end{pmatrix}, \quad \gamma_k = \begin{pmatrix} 0 & \sigma_k \\ -\sigma_k & 0 \end{pmatrix}, \quad k = 1, 2, 3.$$

Hence the matrix γ_5 , which is the product of the four gamma matrices, is written as follows:

$$\gamma_5 = \begin{pmatrix} I_2 & 0 \\ 0 & -I_2 \end{pmatrix}.$$

We use the denotation I_n for the unit matrix of size n , $n \in \mathbb{N}$.

By analogy with [3, 4] we consider the 8-spinor space, in which 8-spinor is defined as column:

$$\psi = \text{col}(\varphi, \chi, \xi, \theta), \quad (2)$$

with $\varphi = \text{col}(\varphi_1, \varphi_2)$, $\chi = \text{col}(\chi_1, \chi_2)$, $\xi = \text{col}(\xi_1, \xi_2)$, $\theta = \text{col}(\theta_1, \theta_2)$ being 2-spinors.

In this 8-spinor space S^2 - and S^3 - submanifolds, which can describe the lepton and baryon sectors, were introduced [5]. The symmetry group $SU(2)$ was proposed to describe leptons:

$$\begin{pmatrix} \varphi \\ \chi \end{pmatrix} = U_L \begin{pmatrix} \xi \\ \theta \end{pmatrix} = \begin{pmatrix} 0 & i \\ i & 0 \end{pmatrix} \begin{pmatrix} \xi \\ \theta \end{pmatrix}, \quad (3)$$

Simple calculations show that we have the homotopy group $\pi_3(S^2)$, which generates the topological Hopf-like charge $s^2 + a_1^2 + a_3^2 \neq 0$:

$$\begin{aligned} s &= 2(\xi^+ \theta + \theta^+ \xi), \\ a_1 &= 2(\theta \theta^+ - \xi \xi^+), \\ a_3 &= -2i(\theta^+ \xi - \xi^+ \theta), \\ s^2 + a_1^2 + a_3^2 &= 4(\theta^+ \theta + \xi^+ \xi)^2. \end{aligned}$$

The symmetry group $SU(2)$ was proposed for the description of baryons:

$$\begin{pmatrix} \varphi \\ \chi \end{pmatrix} = U_B \begin{pmatrix} \xi^* \\ \theta^* \end{pmatrix} = \begin{pmatrix} 0 & i \\ i & 0 \end{pmatrix} \begin{pmatrix} \xi^* \\ \theta^* \end{pmatrix}, \quad (4)$$

This S^3 -manifold is generated by the structure $p^2 + s^2 + a_1^2 + a_2^2$:

$$\begin{aligned} s &= 2(\xi^+ \theta + \theta^+ \xi), \\ p &= 2i(\xi^+ \theta - \theta^+ \xi), \\ a_1 &= \left((\theta^+)^2 + \theta^2 - \xi^2 - (\xi^+)^2 \right), \\ a_2 &= i \left((\theta^+)^2 - \theta^2 + \xi^2 - (\xi^+)^2 \right), \\ s^2 + p^2 + a_1^2 + a_2^2 &= 4 \left[(\theta^+ \theta + \xi^+ \xi)^2 - (\xi^+ \theta - \theta^+ \xi)^2 \right]. \end{aligned}$$

To add to this, common vacuum state was found [6], which conserves only one component $s = 2(\xi^+\theta + \theta^+\xi)$ in both the lepton and the baryon phases:

$$\psi_V = \begin{pmatrix} iC \\ iC \\ C \\ C \end{pmatrix}, \tag{5}$$

where $\theta = \xi = C = \begin{pmatrix} C_1 \\ C_2 \end{pmatrix}$, $C_1, C_2 \in \mathbb{C}$.

2. Current in 8-spinor model

We can write down the components of current $j_\mu = \bar{\psi}\gamma_\mu\psi = \psi^+\gamma_0\gamma_\mu\psi$ entering the Brioschi identity (1):

$$\begin{aligned} j_0 &= \psi^+\gamma_0\gamma_0\psi = \varphi^+\varphi + \chi^+\chi + \xi^+\xi + \theta^+\theta, \\ j_1 &= \psi^+\gamma_0\gamma_1\psi = -\varphi^+\varphi + \chi^+\chi - \xi^+\xi + \theta^+\theta, \\ j_2 &= \psi^+\gamma_0\gamma_2\psi = i[\varphi_1^+\varphi_2 - \varphi_2^+\varphi_1 + \chi_2^+\chi_1 - \chi_1^+\chi_2 + \xi_1^+\xi_2 - \xi_2^+\xi_1 + \theta_2^+\theta_1 - \theta_1^+\theta_2], \\ j_3 &= \psi^+\gamma_0\gamma_3\psi = -\varphi_1^+\varphi_1 + \varphi_2^+\varphi_2 + \chi_1^+\chi_1 - \chi_2^+\chi_2 - \xi_1^+\xi_1 + \xi_2^+\xi_2 + \theta_1^+\theta_1 - \theta_2^+\theta_2. \end{aligned}$$

According to (3) and (4), only one component of the current converses in lepton and baryon phases:

$$j_0 = \psi_L^+\gamma_0\gamma_0\psi_L = \psi_B^+\gamma_0\gamma_0\psi_B = 2(\xi^+\xi + \theta^+\theta), \quad j_1 = j_2 = j_3 = 0. \tag{6}$$

The same situation is in the vacuum state:

$$j_0 = \psi_V^+\gamma_0\gamma_0\psi_V = 4(\psi_1^*\psi_1 + \psi_2^*\psi_2), \quad j_1 = j_2 = j_3 = 0. \tag{7}$$

3. Lagrangian density

By analogy with [1], let us choose the following Lagrangian density for the 8-spinor field model:

$$\mathcal{L} = \frac{1}{2\lambda^2} \overline{D_\mu\psi}\gamma^\alpha j_\alpha D^\mu\psi + \frac{\epsilon^2}{4} f_{\mu\nu}f^{\mu\nu} - V(j_\mu j^\mu) - \frac{1}{16\pi} F_{\mu\nu}^2 \tag{8}$$

with $D^\mu = \partial_\mu\psi - ie_0A_\mu\Gamma_e\psi$ — gauge covariant derivative, $V(j_\mu j^\mu) = \frac{\sigma^2}{8} (j_\mu j^\mu - \kappa^2)^2$ — Higgs potential and $f_{\mu\nu}$ stands for the antisymmetric tensor of the Faddeev–Skyrme type:

$$f_{\mu\nu} = (\overline{\psi}\gamma^\alpha D_{[\mu}\psi) (\overline{D_{\nu]}\psi}\gamma_\alpha\psi) \tag{9}$$

with λ and ϵ being constant parameters of the model. We use the denotation \mathcal{L}_i for terms in this Lagrangian density. Let us write the first σ -model term \mathcal{L}_1 , which is a product of the covariant derivatives:

$$\begin{aligned} \mathcal{L}_1 &= \overline{D_\mu\psi}\gamma^\alpha j_\alpha D^\mu\psi = D_\mu\psi^+\gamma_0\gamma^\alpha j_\alpha D^\mu\psi = \\ &= (\partial_\mu\psi^+ + ie_0A_\mu\Gamma_e\psi^+) \gamma_0\gamma^\alpha (\psi^+\gamma_0\gamma_\alpha\psi) (\partial^\mu\psi - ie_0A^\mu\Gamma_e\psi) = \\ &= \partial_\mu\psi^+\gamma_0\gamma^\alpha (\psi^+\gamma_0\gamma_\alpha\psi) \partial^\mu\psi - \partial_\mu\psi^+\gamma_0\gamma^\alpha (\psi^+\gamma_0\gamma_\alpha\psi) ie_0A^\mu\Gamma_e\psi + \\ &\quad + ie_0A_\mu\Gamma_e\psi^+\gamma_0\gamma^\alpha (\psi^+\gamma_0\gamma_\alpha\psi) \partial^\mu\psi - |ie_0A_\mu\Gamma_e|^2 (\psi^+\gamma_0\gamma_\alpha\psi)^2. \end{aligned} \tag{10}$$

Our model (8) supposes that at space infinity $|r| \rightarrow \infty$ the following boundary conditions hold:

$$\psi \rightarrow \psi_V, \quad j_\mu j^\mu \rightarrow \kappa^2, \quad A_\mu \rightarrow 0. \quad (11)$$

To satisfy this asymptotic behaviour of localized configurations charge generator Γ_e should satisfy the following conditions:

$$\Gamma_e \psi_V = 0, \quad \Gamma_e^2 = \Gamma_e. \quad (12)$$

Using (5) operator Γ_e has the following form:

$$\Gamma_e = I_4 \otimes \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}. \quad (13)$$

We can calculate the spinor term $\overline{D_\mu \psi} \gamma^\alpha j_\alpha D^\mu \psi$ for two phases and the vacuum. For leptons and baryons it has the following form:

$$\begin{aligned} \mathcal{L}_1 &= \overline{D_\mu \psi} \gamma^\alpha j_\alpha D^\mu \psi = D_\mu \psi^+ \gamma^0 \gamma^\alpha j_\alpha D^\mu \psi = \\ &= 4 (\xi^+ \xi + \theta^+ \theta) [D_\mu \xi^+ D^\mu \xi + D_\mu \theta^+ D^\mu \theta]. \end{aligned} \quad (14)$$

Similarly, for the vacuum state one gets:

$$\mathcal{L}_{V1} = \overline{D_\mu \psi_V} \gamma^\alpha j_\alpha D^\mu \psi_V = 16 (C_1^* C_1 + C_2^* C_2) [D_\mu C_1^* D^\mu C_1 + D_\mu C_2^* D^\mu C_2], \quad (15)$$

where C_1 and C_2 take complex values, which describe the vacuum state in (5). Since C_1 and C_2 are constant and $\Gamma_e \psi_V = 0$ (12), the covariant derivative is identically zero in the vacuum state $D^\mu \psi_V = 0$. As a result the σ -model term of the Lagrangian density (15) is also identically zero:

$$\mathcal{L}_{V1} \equiv 0. \quad (16)$$

4. Conclusion

Finally we give the summary of the results obtained:

- The effective 8-spinor field model is suggested. For this field model the components of current quantities are constructed.
- For suggested lepton and baryon sectors σ -model term of Lagrange function is found.

References

1. Y. P. Rybakov, Soliton Configurations in Generalized Mie Electrodynamics, *Phys. of Nuclei* 74 (7) (2011) 1102–1105.
2. Y. P. Rybakov, Topological Solitons in 8-Spinor Mie Electrodynamics, *Physics of Atomic Nuclei* 76 (10) (2013) 1284–1288.
3. V. I. Molotkov, The Effective 8-Spinor Model of Elementary Particles, in: Conference Abstracts of International Student Conference “Science and Progress”, Petergof, 2011.
4. V. I. Molotkov, Structure of nonlinear 16-spinor model of elementary particles, *journal = Physics in Higher Education* 76 (10) (2014) 1284–1288.
5. V. I. Molotkov, Description of Lepton and Baryon Phases in Skyrme–Faddeev Spinor Model, *Bulletin of Peoples’ Friendship University of Russia. Series Mathematics. Information Sciences. Physics* (2) (2015) 73–77.
6. V. I. Molotkov, Y. P. Rybakov, Description of Lepton and Baryon Phases in 8-spinor Model of Elementary Particles, *Physics in Higher Education*(in Russian).

УДК 539.12

Лагранжиан лептонного и барионного секторов в нелинейной восьмиспинорной модели

В. И. Молотков

*Кафедра теоретической физики и механики
Российский университет дружбы народов
ул. Миклуто-Маклая, д. 6, Москва, Россия, 117198*

Идея Скирма (1954) состоит в том, что барион интерпретируется как частица-солитон, которая имеет топологический заряд $B = \text{deg}(S^3 \rightarrow S^3)$. Этот заряд B служит генератором гомотопической группы $\pi_3(S^3) = \mathbb{Z}$. Аналогичная идея для описания лептонов используется в модели Фаддеева (1972). В ней в роли лептонного числа L используется инвариант типа Хопфа Q_H .

Для объединения двух подходов, описывающих лептоны и барионы как топологические заряды, предлагается использовать 8-спинорное поле. Использование специального 8-спинорного тождества Бриоски позволяет рассматривать лептоны и барионы как секторы в общей спинорной модели с потенциалом Хиггса, зависящего от $j^\mu j_\mu$, входящего в лагранжиан. С этой целью рассматривается обобщение электродинамики Ми в рамках эффективной 8-спинорной полевой модели. Кроме того были обнаружены группы симметрий, образующие S^2 и S^3 подмногообразия в общем биквадратном спинорном S^8 -многообразии. Для объединения двух секторов было построено общее вакуумное состояние, сохраняющее лишь одну компоненту в каждом секторе.

В настоящей работе предлагается попытка выписать функцию Лагранжа для гомотопических групп $\pi_3(S^2)$ и $\pi_3(S^3)$, которые описывают барионы и лептоны.

Ключевые слова: 8-спинор, топологический заряд, солитоны, гомотопические группы, тождество Бриоски, модель Скирма-Фаддеева.

Литература

1. *Rybakov Y. P.* Soliton Configurations in Generalized Mie Electrodynamics // *Phys. of Nuclei*. — 2011. — Vol. 74, No 7. — Pp. 1102–1105.
2. *Rybakov Y. P.* Topological Solitons in 8-Spinor Mie Electrodynamics // *Physics of Atomic Nuclei*. — 2013. — Vol. 76, No 10. — Pp. 1284–1288.
3. *Molotkov V. I.* The Effective 8-Spinor Model of Elementary Particles // *Conference Abstracts of International Student Conference “Science and Progress”*. — Petergof: 2011.
4. *Молотков В. И.* Структура нелинейной спинорной модели элементарных частиц // *Физическое образование в вузах*. — 2014. — Т. 76, № 10. — С. 1284–1288.
5. *Molotkov V. I.* Description of Lepton and Baryon Phases in Skyrme–Faddeev Spinor Model // *Bulletin of Peoples’ Friendship University of Russia. Series Mathematics. Information Sciences. Physics*. — 2015. — No 2. — Pp. 73–77.
6. *Молотков В. И., Рыбаков Ю. П.* Описание барионного и лептонного секторов в 8-спинорной модели элементарных частиц // *Физическое образование в вузах. Труды конференции-конкурса молодых физиков*. — 2015. — Т. 21, № 1С.

УДК 621.321

Сравнительный анализ спектрально-фотометрических характеристик бытовых источников света

И. Василеска, М. А. Корнеева, В. П. Стёпин

*Кафедра прикладной физики
Российский университет дружбы народов
ул. Миклуто-Маклая, д. 6, Москва, Россия, 117198*

Правильная организация искусственного освещения имеет большое значение для создания комфортных бытовых условий в местах, предназначенных для выполнения работ со значительным зрительным напряжением, а также для решения задачи эффективного использования электроэнергии. Целью данной работы является проведение комплексных экспериментальных исследований спектральных, фотометрических и электрических характеристик источников света бытового назначения различных типов. Были выявлены наиболее и наименее энергоэффективные источники, показаны особенности применения в различных условиях, а также проведена проверка соответствия характеристик образцов нормам СанПиН.

Для исследования были выбраны лампы накаливания, металло-галогенные, светодиодные, люминесцентные и компактные люминесцентные лампы. По результатам прямых измерений были определены диаграммы направленности, световой поток и класс энергопотребления отобранных источников света.

Все экспериментальные данные были сведены в единую таблицу, что позволяет наглядно проследить особенности исследованных образцов. Было обнаружено, что измеренные величины для некоторых образцов отличаются от заявленных производителем, причём наибольшее расхождение отмечено у компактных люминесцентных ламп.

Анализ полученных результатов позволил определить целесообразность конкретного применения исследованных источников света. Полученные результаты будут полезны как для специалистов и производителей в области светотехники, так и конечных потребителей.

Ключевые слова: источники света, лампы накаливания, металло-галогенные лампы, светодиодные лампы, люминесцентные лампы, компактные люминесцентные лампы, спектр излучения, цветовая температура, освещённость, диаграмма направленности.

1. Введение

Свет является одним из важнейших явлений природы, влияющих на жизнедеятельность человека, так как более чем 80% информации об окружающем мире и происходящих процессах человек получает через зрительные ощущения. Разработка искусственных источников света (ИС) являлась неотъемлемой частью развития техногенной цивилизации. Современные ИС являются электрическими приборами, в которых электрическая энергия преобразуется в излучение в различных спектральных диапазонах и с различной энергетической эффективностью. Естественно, что в различных областях целевого использования (освещение помещений, освещение рабочего места, мест кратковременной зрительной работы в жилых помещениях и т.п.) к ИС предъявляются различные требования. В этой связи разработка эффективных бытовых ИС по-прежнему является актуальной задачей [1].

Для соблюдения условий оптимальной рабочей среды разработаны общие нормы [2, 3], регламентирующие параметры естественного и искусственного освещения. Однако в области бытового применения ИС эти правила достаточно часто не соблюдаются. Кроме того, в них не отражено, что освещение воспринимается естественным только в случае, если регистрируемое глазом излучение имеет непрерывный спектр в видимой области спектра с распределением энергии, близким к распределению энергии в солнечном излучении [1]. Отмечена особенность,

что глаза меньше всего утомляются при жёлто-зелёном либо белом освещении, а больше всего при красном и фиолетовом [1].

Целью данной работы являлось проведение систематизированных измерений спектральных, фотометрических и энергетических характеристик, наиболее распространённых ИС бытового назначения для сравнительного анализа и определения их соответствия условиям наилучшего зрения, целесообразности их конкретного применения в бытовых условиях, требованиям СанПиН и СНиП [2, 3] и справедливости заявленных производителем характеристик.

2. Постановка и проведение экспериментов

Для исследований были отобраны серийно выпускаемые для цепей с напряжением 220 В лампы бытового назначения: для ламп накаливания (ЛН) — до 100 Вт, для металло-галогенных ламп (МГЛ) — до 46 Вт, для светодиодных ламп (LED) — до 7 Вт, для компактные люминесцентные лампы (КЛЛ) — до 32 Вт, для люминесцентные лампы дневного света (ЛД) — до 18 Вт.

Регистрация спектров производилась с помощью монохроматора-спектрографа MS3500I с дифракционной решёткой 1200 шт/мм, входная и выходная щели составляли 40 мкм. В качестве детектора использовался ФЭУ с фиксированным напряжением 1000 В.

Спектрометрический тракт был откалиброван по длине волны с помощью ртутной лампы (ДРС 50-1), а по абсолютной интенсивности с помощью лампы СИРШ 6-40 (аттестована во ВНИИОФИ). Регистрация спектров испытуемых образцов проводилась на фиксированном расстоянии от световода — 50 см. Обработка полученных спектров проводилась с помощью графического пакета Origin.

Освещённость регистрировалась при помощи люксметра ТКА-ПКМ С051 (государственный реестр средств измерений №24248-09), предназначенного для измерения освещённости в видимой области спектра (380–760 нм) с диапазоном от 10 до 200000 лк (погрешность $\pm 8.0\%$). Измерение распределения освещённости на поверхности 70×70 см проводилось для каждого ИС на расстоянии 50 см.

Коэффициент пульсации определялся при помощи прецизионного фотодиода (быстрый кремниевый фотоприёмника APDF, рабочий диапазон — 320–1000 нм, максимальная чувствительность в диапазоне 730–830 нм). Расстояние между ИС и диодом составляло 50 см.

Измерение электрических характеристик (напряжения питания, потребляемый ток) проводилось с помощью осциллографа (TDS 2024 В). Величина тока определялась по падению напряжения на измерительном шунте ($0,47 \pm 0,05$ Ом).

Тепловые потери определялись с помощью термопары цифрового мультиметра MASTECH MY-62. Измерения проводились по всей поверхности колбы ИС через 10 мин после включения. Было выявлено, что нагрев происходит равномерно по всей поверхности (отклонения не более 10% от максимального значения), в дальнейшем приводится значение лишь максимальной температуры, если это не оговорено дополнительно.

Параметры ИС, заявленные производителями, и результаты экспериментальных измерений сведены в единую таблицу (см. ниже). Существенные отклонения и нарушения санитарных норм выделены особо.

3. Экспериментальные результаты и их обсуждение

Основной характеристикой любого источника света является его спектр излучения. Естественно, что ближе спектр ИС к солнечному, тем более он благоприятен для человека. Как известно [4], солнечный спектр является спектром излучения чёрного тела, представляющим из себя непрерывные континуум. Спектральная излучательная способность в данном случае описывается формулой Планка

и характеризуется цветовой температурой с типичным значением ~ 6000 К. Источники, имеющие аналогичный тип распределения излучения (серое тело), но меньшую яркость, так же характеризуются цветовой температурой. Важно отметить, что если спектральное распределение не описывается формулой Планка (является либо линейчатым, либо сплошным, но с распределением другого типа), то характеризовать его цветовой температурой некорректно. Все существующие ИС бытового назначения имеют спектр, отличный от солнечного. Говорить о «степени соответствия» спектра солнечному сложно, так как нет стандартизованного критерия, который бы его определял. Тем не менее, можно выделить одно принципиальное условие: спектр солнца является сплошным и, следовательно, сравнивать с ним можно только ИС со сплошным спектром. На рис. 1 представлены спектры ЛН и МГЛ разных мощностей, на рис. 2 — LED-спектры.

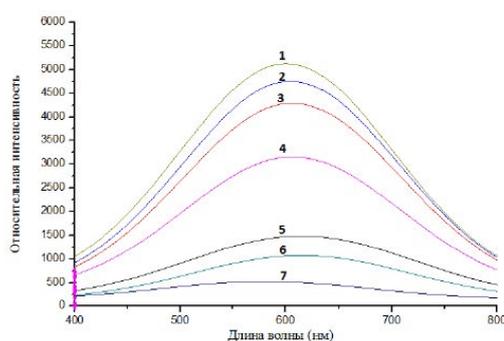


Рис. 1. Спектры ЛН и МГЛ:
 1 — МГЛ 42 Вт; 2 — ЛН 100 Вт;
 3 — ЛН 60 Вт; 4 — МГЛ 42 Вт;
 5 — ЛН 40 Вт;
 6 — ЛН (матовая) 60 Вт;
 7 — МГЛ (рефлекторная) 46 Вт

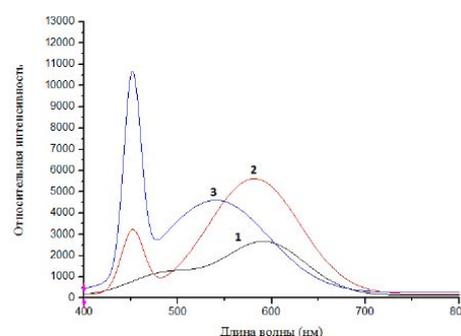


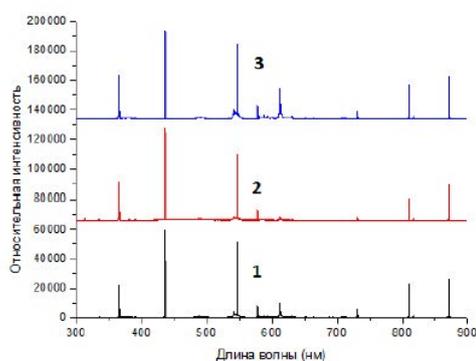
Рис. 2. Спектры LED:
 1 — 5,5 Вт (2700 К); 2 — 7 Вт (2700 К);
 3 — 7 Вт (6400 К)

Из рис. 1 и 2 следует, что спектры ЛН, МГЛ и LED ламп имеют сплошной спектр, но форма их существенно отличается. Лампы накаливания и металлогалогенные лампы имеют одинаковый принцип действия: в обоих случаях излучающим элементом является вольфрамовая нить накала, но колбы МГЛ заполнены буферным газом (парами галогенов), увеличивающими срок службы. Спектр излучения нагретой вольфрамовой нити является спектром серого тела, перекрывающим видимый диапазон и уходящим в инфракрасную область. Цветовая температура серого тела рассчитывается по положению максимума спектра. Спектральные измерения цветových температур ЛН и МГЛ показали их соответствие значениям, заявленным производителями. В данном случае цветовая температура хорошо отражает степень отличия спектра ИС от солнечного: отношение температуры Солнца к температуре ламп ± 2 . В остальных случаях простое сравнение спектров является не корректным.

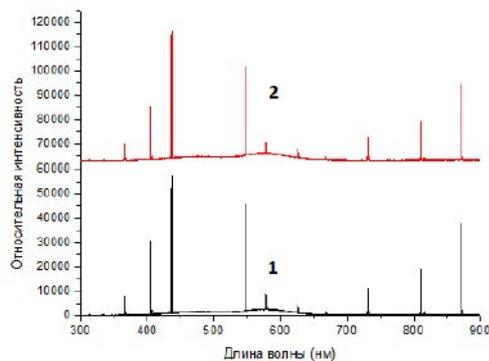
Существенные отличия в спектрах светодиодных ламп и ламп накаливания (см. рис. 1 и рис. 2) свидетельствуют о том, что светодиодные лампы не обладают спектром серого тела, следовательно, говорить в данном случае о цветовой температуре, несмотря на заявления производителей, нельзя. В отличие от ламп накаливания, имеющих существенный участок спектра в инфракрасной области, спектр светодиодных ламп лежит исключительно в области видимого света, что повышает их эффективность. По принципу формирования цветопередачи светодиодные лампы могут быть двух типов [5] — многокристальные и люминофорные. Спектр многокристальных светодиодов получается за счёт наложения узких спектров каждого отдельного кристалла. Спектр люминофорных светодиодов формирует излучение люминофора, который в свою очередь возбуждается

за счёт излучения светодиода в синей или ближней ультрафиолетовой области. По форме спектра на рис. 2 можно определить, что исследованные лампы являются люминофорными светодиодными лампами. Говорить о соответствии солнечному спектру в данном случае сложно. Солнечный спектр имеет максимум интенсивности при длине волны — 550 нм. Одна из исследованных ламп — 3 (рис. 2, температура 6400 К) также имеет максимум на этой длине волны, но при этом у неё есть второй пик в области синего диапазона длин волн (450 нм). Два других исследованных образца — 1, 2 (рис. 2, температуры 2700 К) имеют иное распределение интенсивности в спектральной области излучения. У изученных образцов 1-й максимум соответствует области красного диапазона и приходится на длину волны 600 нм, а второй максимум сосредоточен в области синего диапазона: у одной из ламп он приходится на длину волны 450 нм, а у второй вообще слабо выражен. Таким образом, при покупке светодиодных ламп предсказать форму спектра невозможно, так как указанная производителем цветовая температура не несёт информации о качестве спектра.

На рис. 3а представлены спектры КЛЛ. Несмотря на заявленную производителем цветовую температуру, данные лампы в пределах видимого диапазона имеют обеднённый линейчатый спектр. Видно, что все лампы имеют одинаковый спектральный состав, а разница в качественном ощущении света (тёплый, холодный свет) достигается за счёт перераспределения энергии в спектре. Следует отметить, данный спектр соответствует спектру ртутной газоразрядной лампы, хотя некоторые производители заявляют об отсутствии паров ртути в изделии. Несмотря на название, компактные люминесцентные лампы люминофора не содержат, в отличие от люминесцентных ламп дневного света, в спектре которых присутствует континуум, соответствующий излучению люминофора, на фоне которого выступают линии ртути (рис. 3б). Важно отметить присутствие в спектре излучения линий в ближнем ультрафиолете (365 нм). Говорить о соответствии спектров КЛЛ солнечному невозможно, отличия в данном случае принципиальны.



(а) 1 — КЛЛ 13 Вт (2700 К); 2 — КЛЛ 13 Вт (6500 К); 3 — КЛЛ 32 Вт (2700 К)



(б) 1 — ЛД 18 Вт (2700 К); 2 — ЛД 18 Вт (4000 К)

Рис. 3. Спектры люминесцентных ламп

При использовании ИС важную роль играет поле допустимой освещённости. Согласно требованиям ГОСТ [6] по освещению рабочих поверхностей для кратковременной работы на поверхности $0,4 \times 0,3$ м необходимо > 200 лк, а на рабочем столе на площади $0,3 \times 0,2$ м необходимо > 300 лк. Как показали измерения, данным требованиям удовлетворяют практически все лампы (фиксированное расстояние от ИС до рабочей поверхности — 50 см), за исключением некоторых маломощных ламп (лампа накаливания 40 Вт).

Согласно ГОСТ [7] были определены диаграммы направленности силы света для каждого ИС. Вид диаграмм направленностей большинства ламп (рис. 4) соответствует равномерно направленному типу кривой силы света. Нарушение сферической симметрии диаграммы направленности (рис. 4) в нижней её части обусловлено наличием цоколя ИС. У ламп накаливания, в отличие от КЛЛ и светодиодных ламп, часть излучения распространяется в заднее полупространство, т. е. в сторону цоколя.

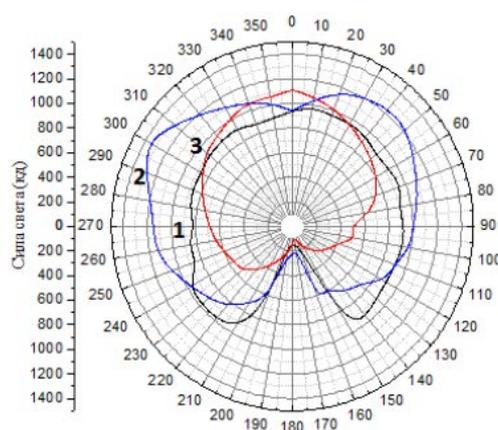
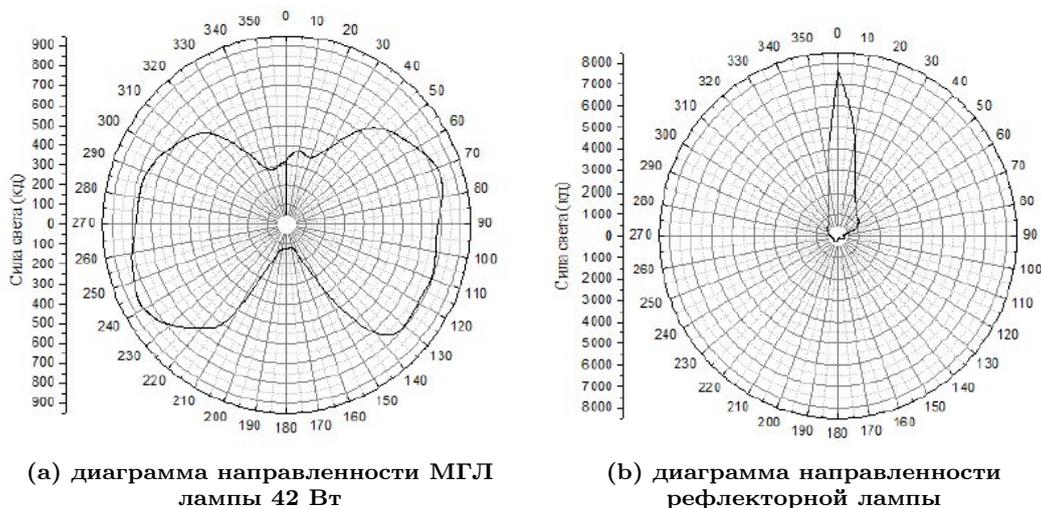


Рис. 4. Диаграммы направленности (для фиксированного расстояния — 50 см): 1 — лампа накаливания 60 Вт, 2 — КЛЛ 12 Вт, 3 — светодиодная лампа 7,7 Вт

Диаграммы направленности МГЛ (рис. 5а) характеризуются широким типом кривой распределения силы света. Диаграмма имеет чёткие боковые максимумы, что резко отличает их от ламп накаливания, КЛЛ и светодиодных ламп.



(а) диаграмма направленности МГЛ лампы 42 Вт

(б) диаграмма направленности рефлекторной лампы

Рис. 5. Диаграммы направленности МГЛ лампы (для фиксированного расстояния — 50 см):

Диаграмма направленности рефлекторной лампы существенно отличается от ИС рассмотренных выше (рис. 5b). Данный случай относится к концентрированному типу кривой диаграммы направленности силы света. Эта особенность отмечена производителем на упаковке. Подобная диаграмма обусловлена конструктивными особенностями и преимущественно определяется форм-фактором данной лампы, что обеспечивает узко направленный пучок света (угол рассеяния $\sim 20^\circ$ при указанных производителем — 30°).

Немаловажным параметром качества освещенности, создаваемой ИС, является коэффициент пульсации, который в соответствии с СНиП [3] не должен превышать 20%. Измерения показали превышение допустимой нормы пульсаций у КЛЛ и люминесцентных ламп дневного света (см. табл. 1).

Таблица 1
Экспериментальные и заявленные производителями характеристики бытовых ламп

Тип ламп	Заявленная мощность, Вт		Измеренная мощность, Вт		Площадь с освещенностью >20лк * см ²		Площадь с освещенностью >40лк ** ±8% см ²		Заявленный световой поток, лм		Измеренный световой поток, лм		Коэффициент пульсации ±3%, %		Угол рассеяния, ±5%, °		Эффективность, лм/Вт	Заявленный класс энергопотребления		Измеренный класс энергопотребления		Тип спектра	Заявленная цветовая температура, °К		Измеренная цветовая температура, ±3%, °К		Присутствие линий в УФ (365 нм)	Наличие ртутных линий в спектре		Тепловые потери, ±10%, %
	40	42,6	855	0	400	469	7	290	11	Е	Е	сплошной	2700	2705	нет	нет		160												
Лампа накаливания	60	63,8	2733	707	710	760	6,2	300	12	Е	Е	сплошной	2700	2706	нет	нет	176													
	100	102	4900	4900	1340	1609	4,3	310	16	С	Д	сплошной	2700	2705	нет	нет	181													
	60	63,8	2826	707	630	621	3,5	290	10	Д	Е	сплошной	2700	2704	нет	нет	156													
Матовая лампа накаливания	42	44,7	2206	660	630	681	5,9	211	15	А	Е	сплошной	3000	2700	нет	нет	115													
	42	44,7	1735	572	630	682	5,8	280	15	А	Е	сплошной	2700	2700	нет	нет	120													
Рефлекторная металлогалогенная лампа	46	48,1	3117	1886	300	320	6,8	20	7	Д	Е	сплошной	2700	2630	нет	нет	192													
	5,5	6,38	2734	700	470	498	5,5	250	78	А+	А++	сплошной	2700	×	нет	нет	32													
Светодиодная лампа	7	8,51	4900	4900	560	591	5	260	69	А++	А++	сплошной	2700	×	нет	нет	33													
	7	8,51	4900	4900	600	643	8	250	76	А++	А++	сплошной	6500	×	нет	нет	32													
	9	11,6	1963	380	522	580	20	250	50	А++	А+	линейчатый	2700	×	нет	нет	66													
КЛЛ	11	-	4900	1585	600	-	25	-	-	А++	-	линейчатый	2700	×	нет	нет	58													
	12	14,1	4900	1663	650	675	20	270	48	А++	А	линейчатый	2700	×	нет	нет	68													
	13	15,2	4900	1886	754	787	23	260	52	А++	А+	линейчатый	2700	×	нет	нет	74													
	13	15,2	4900	1725	650	670	22	260	44	А	А	линейчатый	6400	×	нет	нет	71													
	32	39,9	4900	4900	1780	1860	16	260	47	А++	А	линейчатый	2700	×	нет	нет	78													
	14	16,6	4900	3770	740	760	25	260	46	А	А	линейчатый	2700	×	нет	нет	62													
Лампа дневного света	18	17,7	4900	2463	1200	1359	27,6	230	77	А++	А++	линейчатый с континуумом	2700	×	нет	нет	30													
	18	17,7	4900	1256	1200	1359	27,6	230	77	А++	А++	линейчатый с континуумом	4000	×	нет	нет	31													

* освещение мест кратковременной зрительной работы по ГОСТ 8607-82: не менее 200 лк на площади 1200 см² "—" измерения не проводились

** освещение письменных столов по ГОСТ 8607-82: не менее 300 лк на площади 600 см² "×" определение не возможно

Для определения эффективности (лм/Вт) и класса энергопотребления [8] ИС необходимо определить световой поток и потребляемую мощность. Световой поток был рассчитан по результатам измерений диаграмм направленности в соответствии с ГОСТ [9]. Потребляемая мощность для конкретного ИС была рассчитана исходя из осциллографических измерений и представлена в таблице. Полученные результаты также свидетельствуют о расхождениях с данными, заявленными производителями. В частности, из таблицы видно, что МГЛ относятся к наихудшему классу энергопотребления (Е), а лампы типа КЛЛ не могут быть отнесены к ИС с наилучшим типом энергопотребления (А++), а только к типу А/А+.

Тепловые потери ламп хорошо согласуются с классом энергопотребления. Как и предполагалось, самым высоким тепловыделением обладают лампы накаливания ($\geq 150^\circ\text{C}$), за ними идут металло-галогенные ($\geq 100^\circ\text{C}$), КЛЛ ($\geq 50^\circ\text{C}$) и светодиодные и лампы дневного света ($\leq 50^\circ\text{C}$). Исключением является лампа рефлекторного типа, имеющая очень высокую температуру в центре ($\approx 200^\circ\text{C}$), при этом боковая поверхность имеет температуру $\leq 50^\circ\text{C}$. Так же стоит отметить, что матовая поверхность колбы лампы снижает температуру на 10–15% по сравнению с лампами аналогичного типа и мощности.

4. Заключение

Полученные результаты свидетельствуют о том, что большая часть отобранных ИС имеет отклонения от заявленных производителем характеристик, что свидетельствует о низком выходном контроле изделий на предприятиях. Полученные результаты измерений можно обобщить по каждому типу ИС, выделив основные достоинства и недостатки.

Светодиодные лампы обладают наилучшими характеристиками:

- ✓ сплошной спектр;
- ✓ низкий коэффициент пульсации;
- ✓ высокая эффективность (A++);
- ✓ низкое тепловыделение;
- × упоминание производителем цветовой температуры является некорректным.

Лампы накаливания и металло-галогенные лампы уступают светодиодным лампам по энергопотреблению и тепловыделению:

- ✓ сплошной спектр;
- ✓ низкий коэффициент пульсации;
- ✓ соответствие заявленной цветовой температуре;
- × низкая эффективность и высокое тепловыделение.

Люминесцентные лампы дневного света имеют много недостатков, но два существенных достоинства:

- ✓ высокая эффективность (A++);
- ✓ низкое тепловыделение;
- × линейчатый спектр с континуумом;
- × присутствие линии в УФ;
- × высокий коэффициент пульсации (превышает нормы [2, 3]).

Компактные люминесцентные лампы имеют наихудшие характеристики, несмотря на заявления производителей, и обладают только одним достоинством:

- ✓ высокая эффективность (A, A+);
- × среднее тепловыделение;
- × линейчатый спектр;
- × присутствие линии в УФ;
- × высокий коэффициент пульсации (превышает нормы [2, 3]).

Отдельно следует отметить лампы с колбой типа «**рефлектор**». Такая форма влияет только на угол раствора диаграммы направленности и тем самым на площадь освещённости. Все остальные характеристики не отличаются от характеристик ламп соответствующего типа, в нашем случае это МГЛ. Лампы с таким цоколем хорошо использовать для освещения рабочей поверхности. Несмотря на то, что три исследованные МГЛ имели близкую потребляемую мощность, рефлекторная лампа обеспечивает большую площадь поверхности с освещённостью > 300 лк.

Литература

1. *Варфоломеева Л. П.* Энергоэффективное электрическое освещение. — М.: Издательский дом МЭИ, 2013.

2. СанПиН. 2.2.1/2.1.1.1278-03. — 2003.
3. СНиП. 23-05-2010. — 2010.
4. Сивухин Д. В. Общий курс физики. Т. 4. Оптика. — М.: Издательский дом МЭИ, 2005.
5. Елисеев Н. П., Решенов С. П. О предельных световых и цветовых характеристиках белых светодиодов. — М.: Светотехника, 2012.
6. ГОСТ. 8607-82. — 1982.
7. ГОСТ. 55702-2013 года (пункт 7.2.5). — 2013.
8. Murphy T. Maximum Efficiency of White Light. — Department of physics UC San Diego, 2011.
9. ГОСТ. 54350-2011 года (формула (5) в пункте 11.3 и формула (Д.4) в пункте Приложения Д). — 2011.

UDC 621.321

Comparative Analysis of Spectral-Photometric Characteristics of Domestic Sources of Light

I. Vasileska, M. A. Korneeva, V. P. Stepin

*Department of Applied Physics
Peoples' Friendship University of Russia
6, Miklukho-Maklaya str., Moscow, Russia, 117198*

Correct organization of artificial lighting is very important for creation comfortable living conditions in work places with significant eyestrain, as well as solving the problem of energy efficiency. The purpose of this study is to conduct a comprehensive experimental study of spectral, photometric and electrical characteristics of domestic light sources. Were identified the most and least energy-efficient different light sources in specific circumstances, as well as verification their conformity with the characteristics of the standards SanPiN.

In this work were studied incandescent lamps, metal halide, LED, fluorescent and compact fluorescent lamps. According to the results of direct measurements were determined: the directional patterns, the luminous fluxes and energy efficient classes of domestic light sources.

All experimental resultants have been summarized in a single table which allows comparing the characteristics of the studied domestic light sources. It was founded that measured values for some of the light sources differ from those stated by the manufacturer, which was noted in compact fluorescent lamps.

Analysis of the results allowed us to determine the feasibility of the specific light sources. The results will be useful for the specialists and manufacturers in lighting technology, as well as ordinary users.

Key words and phrases: the light sources, incandescent lamps, metal halide lamps, LED lamps, fluorescent lamps, compact fluorescent lamps, the emission spectrum, color temperature, illumination, directional patterns.

References

1. L. P. Valfrolomeeva, Energetic Efficiency of Electric Illumination, Publishing House MEI, Moscow, 2013, in Russian.
2. SanPin, 2.2.1/2.1.1.1278-03, 2003, in Russian.
3. SNiP, 23-05-2010, 2010, in Russian.
4. D. V. Sivukhin, General Course of Physics, Vol. 4. Optics, Fizmatlit, Moscow, 2005, in Russian.
5. N. P. Eliseev, S. P. Reshenov, The limit of light and color characteristics of white LEDs, Svetotekhnika, Moscow, 2012.
6. GOST, 8607-82, 1982, in Russian.
7. GOST, 55702-2013 (# 7.2.5), 2013, in Russian.
8. T. Murphy, Maximum Efficiency of White Light, Department of physics UC San Diego, 2011.
9. GOST, 54350-2011 (formula (5) in # 11.3 and formula (D.4) in Appendix D), 2011, in Russian.

СВЕДЕНИЯ ОБ АВТОРАХ

Азофейфа Гомез Эстебан Хосуэ — студент кафедры информационных технологий РУДН (e-mail: esteban.azofeifa@gmail.com)

Балт Батгэрэл — кандидат физико-математических наук, старший научный сотрудник научного отдела вычислительной физики лаборатории информационных технологий Объединённого института ядерных исследований (e-mail: batgerel@jinr.ru, тел.: +7 (49621) 64-598)

Ботвинко Анатолий Юрьевич — аспирант, инженер в/ч 68242 (e-mail: botviay@sci.pfu.edu.ru)

Василеска Ивана — студент кафедры прикладной физики РУДН (Македония) (e-mail: ivonavasileska@yahoo.com, тел.: +7 (495) 9550924)

Геворкян Мигран Нельсонович — кандидат физико-математических наук, доцент кафедры прикладной информатики и теории вероятностей РУДН (e-mail: mngevorkyan@sci.pfu.edu.ru, тел.: +7(495) 9520250)

Горбачёва Анна Викторовна — аспирант кафедры нелинейного анализа и оптимизации РУДН, преподаватель кафедры прикладной математики РГСУ (e-mail: avgorbacheva@inbox.ru, тел.: +79151144000)

Дали Джерейе Тьерри Ги — студент кафедры прикладной информатики и теории вероятностей РУДН, (e-mail: dthierryguy@yahoo.fr)

Диаррассуба Калилу — студент кафедры прикладной информатики и теории вероятностей РУДН, (e-mail: kalheslebelge@yahoo.fr)

Зарипова Эльвира Ринатовна — кандидат физико-математических наук, старший преподаватель кафедры прикладной информатики и теории вероятностей РУДН (e-mail: ezarip@sci.pfu.edu.ru, тел.: +7(495) 9550999)

Калениченко Валентина Николаевна — студент кафедры прикладной информатики и теории вероятностей РУДН (e-mail: vnkalenichenko@yandex.ru, тел.: +7(495) 9522823)

Карамзин Дмитрий Юрьевич — доктор физико-математических наук, старший научный сотрудник отдела методов нелинейного анализа ФИЦ ИУ РАН (e-mail: Dmitry_karamzin@mail.ru, тел.: +79154505177)

Корнеева Мария Анатольевна — ассистент кафедры прикладной физики РУДН (e-mail: korneevama@mail.ru, тел.: +7 (495) 9550924)

Кулябов Дмитрий Сергеевич — доцент, кандидат физико-математических наук, доцент кафедры прикладной информатики и теории вероятностей РУДН (e-mail: dharma@mx.pfu.edu.ru, тел.: +7(495) 9520250)

Матюшенко Сергей Иванович — кандидат физико-математических наук, доцент кафедры прикладной информатики и теории вероятностей РУДН (e-mail: matushenko@list.ru, тел.: +7(495) 9522823)

Мачука Осуна Хосе Роландо — студент кафедры прикладной информатики и теории вероятностей РУДН, (e-mail: cheosurf@gmail.com)

Молотков Вячеслав Иванович — аспирант кафедры теоретической физики и механики РУДН (e-mail: molotvi@gmail.com, тел.: +7 (916) 9032028)

Нгуен Зуй Тхань — аспирант кафедры информационных технологий РУДН (e-mail: nguyenduythanh1410@gmail.com)

Никонов Эдуард Германович — старший научный сотрудник, доктор физико-математических наук, начальник сектора научно-технического отдела внешних коммуникаций и распределённых информационных систем лаборатории информационных технологий Объединённого института ядерных исследований (e-mail: e.nikonov@jinr.ru, тел.: +7 (49621) 64-722)

Новикова Галина Михайловна — доцент, кандидат технических наук, доцент кафедры информационных технологий РУДН (e-mail: novikova_gm@mail.ru)

Пузынин Игорь Викторович — профессор, доктор физико-математических наук, советник дирекции лаборатории информационных технологий Объединённого института ядерных исследований (e-mail: ipuzynin@jinr.ru)

Пяткина Дарья Анатольевна — доцент, кандидат физико-математических наук, доцент кафедры прикладной информатики и теории вероятностей РУДН (e-mail: daria_ru@mail.ru, тел.: +7(495) 9522823)

Самуйлов Константин Евгеньевич — доктор технических наук, заведующий кафедрой прикладной информатики и теории вероятностей РУДН (e-mail: ksam@sci.pfu.edu.ru, тел.: +7(495) 9550999)

Стёпин Вячеслав Петрович — студент кафедры прикладной физики РУДН (e-mail: vjacheslav-stepin@rambler.ru, тел.: +7 (495) 9550924)

Шарапова Анастасия Андреевна — студент кафедры прикладной информатики и теории вероятностей РУДН (e-mail: asharapovarudn@gmail.com)

ПРАВИЛА ОФОРМЛЕНИЯ СТАТЕЙ

Редколлегия журнала «Вестник Российского университета дружбы народов», серия «Математика. Информатика. Физика» просит авторов придерживаться следующих правил при представлении статьи в журнал.

1. Статьи представляются на русском или английском языке.
2. Объём статьи не должен превышать 0,75 авт. печ. л.
3. Автор представляет в редакцию электронную версию рукописи, набранную в системе \LaTeX (используется версия $\text{\LaTeX} 2_{\epsilon}$, для набора формул используется макропакет $\mathcal{A}\mathcal{M}\mathcal{S}\text{-}\text{\LaTeX}$). К электронному варианту прилагается отпечатанный на бумаге экземпляр или файл в формате Postscript или PDF. Для каждой статьи указываются коды УДК и MSC (PACS).
4. Название, аннотация, ключевые слова, фамилии и инициалы авторов, название организации, где работают авторы, предоставляются на русском и английском языках.
5. Аннотация (от 150 до 250 слов) должна описывать основные цели исследования; объяснять, как было проведено исследование, без методологических деталей; суммировать наиболее важные результаты и их значение; быть написана качественным русским/английским языком. В аннотации не допускаются ссылки на цитированную литературу, аббревиатуры и громоздкие формулы.
6. Текст статьи должен включать введение, заключение, список использованных источников (не менее 5). Глубина разбивки текста не должна превышать трёх уровней (разделы, пункты и подпункты).
7. Рисунки принимаются в электронном виде. Каждый рисунок должен быть помещён в отдельный файл. Принимаемые форматы файлов: 1) векторные: PDF, EPS, \TeX ; 2) растровые: PNG, TIFF, GIF (возможна инкапсуляция в EPS).
8. Размер рисунка вместе с подписью не должен превышать 14x19 см. Разрешение растрового рисунка должно находиться в пределах 300–600 dpi.
9. Рисунки должны быть чёрно-белые. Возможность использования полутоновых и фотографических изображений может быть рассмотрена отдельно. Фоны должны быть только штрихованные. Сеточные фоны и полтона не допускаются.
10. Список литературы подготавливается в системе \BibTeX и представляется в двух отдельных файлах — по требованиям ГОСТ 7.1-84 (смешанный список источников на русском и других языках) и по стандарту Harvard (полностью на иностранном языке (англ., франц., нем., исп.); для источников на русском языке требуется перевод на английский язык). Ссылки на неопубликованные работы не допускаются.
11. Рукопись должна быть тщательно выверена. Необходимо указать точные сведения об авторах: фамилию, имя, отчество полностью, учёную степень, учёное звание, должность с указанием подразделения и организации, почтовый и электронный адрес места работы авторов и телефоны. После подготовки редакцией к набору размеченный и исправленный автором текст статьи и исправленная электронная версия возвращаются в редакцию. Корректурa для просмотра высылается по e-mail.
12. Возвращение статьи автору на доработку не означает, что она принята к опубликованию. Доработанный вариант статьи редколлегия рассматривает вновь. В случае отклонения статьи редколлегия оставляет за собой право не возвращать автору один её экземпляр.

13. Все статьи, планируемые к публикации в журнале, проходят обязательную процедуру рецензирования двумя рецензентами, являющимися экспертами в данной области. Один из рецензентов может быть выбран из состава редколлегии (внутреннее рецензирование). В обязательном порядке статья направляется независимому эксперту (на внешнее рецензирование). Процедура рецензирования является анонимной для авторов. По результатам рецензирования статья может быть отклонена, направлена автору на доработку или принята к печати. Требования к рецензиям: оценка существа работы и возможности её публикации в журнале; конкретное перечисление ошибок в методологии и инструментарии (если они есть); предложения по доработке текста.

14. При получении положительного заключения рецензента материалы помещаются в «портфель» редакции для дальнейшего опубликования. При получении отрицательного заключения одного из рецензентов статья рассматривается на заседании редакционной коллегии, которая принимает решение об отклонении статьи или о необходимости получения дополнительной рецензии независимого эксперта.

Контактный адрес для решения технических вопросов по предоставляемым материалам в журнал Вестник РУДН. Серия «Математика. Информатика. Физика»: Россия, Москва, ул. Орджоникидзе, д. 3, комн. 313; e-mail: vestnik-miph@sci.pfu.edu.ru.

Адрес для переписки: 117198, Россия, Москва, ул. Миклухо-Маклая, д. 6, факультет физико-математических и естественных наук, профессору Севастьянову Л. А.

ф. СП-1

ФГУП «ПОЧТА РОССИИ»

АБОНЕМЕНТ на журнал

18235

(индекс издания)

ВЕСТНИК РУДН
Серия «Математика.
Информатика. Физика»

Количество
комплектов:

на 20__ год по месяцам

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

Куда

(почтовый индекс)

(адрес)

Кому

(фамилия, инициалы)

ДОСТАВОЧНАЯ КАРТОЧКА

ПВ	место	литер

на журнал

18235

(индекс издания)

ВЕСТНИК РУДН

Серия «Математика. Информатика. Физика»

Стои- мость	подписки	_____ руб. ___ коп.	Количество комплектов:	
	переадресовки	_____ руб. ___ коп.		

на 20__ год по месяцам

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

Куда

(почтовый индекс)

(адрес)

Кому

(фамилия, инициалы)

ф. СП-1

ФГУП «ПОЧТА РОССИИ»

АБОНЕМЕНТ на журнал

ВЕСТНИК РУДН

(индекс издания)

Серия _____

Количество комплектов:

на 2010 год по месяцам

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

Куда

(почтовый индекс)

(адрес)

Кому

(фамилия, инициалы)

ДОСТАВОЧНАЯ КАРТОЧКА

ПВ	место	литер

на журнал

(индекс издания)

ВЕСТНИК РУДН

Серия _____

Стоимость	подписки	_____ руб. __ коп.	Количество комплектов:	<input type="text"/>
	переадресовки	_____ руб. __ коп.		

на 2010 год по месяцам

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

Куда

(почтовый индекс)

(адрес)

Кому

(фамилия, инициалы)

ВЕСТНИК
Российского университета дружбы народов

Научный журнал. Основан в 1993 г.

Серия
МАТЕМАТИКА. ИНФОРМАТИКА. ФИЗИКА
2016, № 1

Редактор *К. В. Зенкин*
Компьютерная вёрстка: *А. В. Королькова*

Адрес редакции:
Российский университет дружбы народов
ул. Орджоникидзе, 3, г. Москва, Россия, 115419
тел. +7(495)955-07-16

Адрес редакционной коллегии
серии «Математика. Информатика. Физика»:
ул. Орджоникидзе, 3, комн. 313, г. Москва, Россия, 115419
тел. +7(495) 952-02-50, e-mail: vestnik-miph@sci.pfu.edu.ru
сайт: <http://vestnik-miph.sci.pfu.edu.ru>

Подписано в печать 18.03.2016. Формат 60 × 84/8. Бумага офсетная.
Печать офсетная. Гарнитура «Computer Modern».
Усл. печ. л. 12,61 Тираж 500 экз. Заказ № 99

Типография ИПК РУДН
ул. Орджоникидзе, 3, ГСП-1, г. Москва, Россия, 115419, тел. 952-04-41

BULLETIN
of Peoples' Friendship University of Russia

Series
MATHEMATICS. INFORMATION SCIENCES. PHYSICS
2016, № 1

Editor *K. V. Zenkin*
Computer design: *A. V. Korolkova*

Address of editorial board:
Peoples' Friendship University of Russia
Ordzhonikidze str., 3, Moscow, Russia, 115419
Tel. +7 (495) 955-07-16

Address of editorial board
series "Mathematics. Information Sciences. Physics":
Ordzhonikidze str., 3, Moscow, Russia, 115419
Tel. +7 (495) 952-02-50, e-mail: vestnik-miph@sci.pfu.edu.ru
site: <http://vestnik-miph.sci.pfu.edu.ru>

Printing run 500 copies

Address of PFUR publishing house
Ordzhonikidze str., 3, Moscow, Russia, 115419
Tel. +7 (495) 952-02-50