



Том 4. № 3
2018

научный журнал

ТРУДЫ УЧЕБНЫХ ЗАВЕДЕНИЙ СВЯЗИ

Темы номера:

- ✓ Стохастическая модель метеорного канала
- ✓ Оценка восприятия нелинейных искажений сигнала
- ✓ Анонимные коммуникационные системы с низкой латентностью

■ ТЕЛЕКОММУНИКАЦИИ ■ ТЕЛЕВИДЕНИЕ И РАДИОВЕЩАНИЕ ■ СЕТИ СВЯЗИ И СИСТЕМЫ КОММУТАЦИИ
■ СИСТЕМЫ СВЯЗИ С ПОДВИЖНЫМИ ОБЪЕКТАМИ ■ АУДИОВИЗУАЛЬНАЯ ТЕХНИКА ■ ПОЧТОВАЯ СВЯЗЬ И ИНФОРМАТИКА

Федеральное государственное бюджетное образовательное
учреждение высшего образования
«Санкт-Петербургский государственный университет телекоммуникаций
им. проф. М.А. Бонч-Бруевича» (СПбГУТ)

Научный журнал

ТРУДЫ
УЧЕБНЫХ ЗАВЕДЕНИЙ СВЯЗИ

Том 4

№ 3

Санкт-Петербург

2018

Описание журнала

Научный журнал. Публикуются материалы научных работ, выполненных в области телекоммуникаций, телевидения и радиовещания, сетей связи и систем коммутации, систем связи с подвижными объектами, аудиовизуальной техники, почтовой связи, информатизации отрасли связи.

Выпускается с 1960 года. Выходит 4 раза в год.

Редакционный совет

- Розанов Н.Н.** д.ф.-м.н., проф., чл.-корр. РАН, начальник отдела теоретических исследований Института лазерной физики «ГОИ им. С.И. Вавилова»
Кучерявый Е. PhD, Технологический университет Тампере, Финляндия
Гошек И. PhD, Технологический университет г. Брно, Чешская республика
Тиамийу О.А. PhD, Университет Илорина, Нигерия
Козин И.Д. д.ф.-м.н., проф., профессор кафедры телекоммуникационных систем Алматинского университета энергетики и связи, Республика Казахстан
Самуйлов К.Е. д.т.н., проф., заведующий кафедрой прикладной информатики и теории вероятностей РУДН
Степанов С.Н. д.т.н., проф., заведующий кафедрой «Сети связи и системы коммутации» МТУСИ
Росляков А.В. д.т.н., проф., заведующий кафедрой автоматической электросвязи ПГУТИ
Кучерявый А.Е. д.т.н., проф., заведующий кафедрой сетей связи и передачи данных СПбГУТ
Канаев А.К. д.т.н., проф., заведующий кафедрой «Электрическая связь» ПГУПС
Новиков С.Н. д.т.н., проф., заведующий кафедрой безопасности и управления в телекоммуникациях СибГУТИ
Дворников С.В. д.т.н., проф., профессор кафедры радиосвязи ВАС
Коржик В.И. д.т.н., проф., профессор кафедры защищенных систем связи СПбГУТ
Ковалгин Ю.А. д.т.н., проф., профессор кафедры радиосвязи и вещания СПбГУТ
Владыко А.Г. к.т.н., директор НИИ «Технологии связи» СПбГУТ

Редакционная коллегия

Главный редактор **Бачевский С.В.**, д.т.н., проф.
Зам. главного редактора **Буйневич М.В.**, д.т.н., проф.
Ответственный редактор **Аникевич Е.А.**, к.т.н.
Выпускающие редакторы **Татарникова И.М.**, **Яшугин Д.Н.**

Регистрационная информация

Журнал зарегистрирован Федеральной службой по надзору в сфере связи, информационных технологий и массовых коммуникаций: № 77-17986 от 07.04.2004

Подписной индекс по каталогу МАП: 35664

Размещение в РИНЦ (elibrary.ru) по договору: № 59-02/2013R от 20.02.2013

Контактная информация

Учредитель и издатель:	Федеральное государственное бюджетное образовательное учреждение высшего образования «Санкт-Петербургский государственный университет телекоммуникаций им. проф. М.А. Бонч-Бруевича» (СПбГУТ)	Адрес: 193232, Санкт-Петербург, пр. Большевиков, 22/1, к. 334/2 редакции: Тел.: +7 (812) 326-31-63, м. т. 2022, +79643759970 E-mail: tuzs@spbgut.ru Web: http://tuzs.sut.ru ВК: http://vk.com/spbtuzs
-------------------------------	---	---

Выходные данные

Подписано в печать	Усл.-печ. л.	Формат	Тираж	Заказ
27.09.2018	14,13	60×84 _{1/8}	1000 экз.	№ 2463
Отпечатано	Санкт-Петербургский университет ГПС МЧС России			
01.10.2018	196105, Санкт-Петербург, Московский пр., д. 149			

Description

Scientific journal. The materials of scientific papers carried out in the field of telecommunications, television and broadcasting, communication networks and switching systems, communication systems with mobile objects, audiovisual equipment, mail services, informatization of communication industry are published.

Since 1960. Published 4 times per year.

Editorial Council

- Rozanov N.N.** DSc, prof., member-corr. RAS, Open Joint Stock Company «S.I. Vavilov State Optical Institute»
- Koucheryavy Y.** PhD, Tampere University of Technology, Finland
- Hošek I.** PhD, Brno University of Technology, Czech Republic
- Tiamiyu O.A.** PhD, University of Ilorin, Nigeria
- Kozin I.D.** DSc, prof., Almaty University of Power Engineering and Telecommunications, the Republic of Kazakhstan
- Samuilov K.E.** DSc, prof., RUDN University
- Stepanov S.N.** DSc, prof., Moscow Technical University of Communication and Informatics
- Roslyakov A.V.** DSc, prof., Povolzhskiy State University of Telecommunications and Informatics
- Koucheryavy A.E.** DSc, prof., The Bonch-Bruевич Saint-Petersburg State University of Telecommunication
- Kanaev A.K.** DSc, prof., Emperor Alexander I-st Petersburg State Transport University
- Novikov S.N.** DSc, prof., Siberian State University of Telecommunications and Information Sciences
- Dvornikov S.V.** DSc, prof., Telecommunications Military Academy
- Korzhik V.I.** DSc, prof., The Bonch-Bruевич Saint-Petersburg State University of Telecommunication
- Kovalgin Yu.A.** DSc, prof., The Bonch-Bruевич Saint-Petersburg State University of Telecommunication
- Vladyko A.G.** PhD, The Bonch-Bruевич Saint-Petersburg State University of Telecommunication

Editorial Board

Editor-in-chief **Bachevsky S.V.**, DSc, prof.
Deputy editor-in-chief **Buinevich M.V.**, DSc, prof.
Executive editor **Anikevich E.A.**, PhD
Managing editors **Tatarnikova I.M.**, **Yashugin D.N.**

Registration Information

Registered by Federal Service for Supervision of Communications, Information Technology and Mass Media
on 07.04.2004 № 77-17986

Subscription index for MAP catalog: 35664

Accommodation in RINC (elibrary.ru) by agreement: № 59-02/2013R on 20.02.2013

Contact Information

Publisher:	Federal State Budget-Financed Educational Institution of Higher Education «The Bonch-Bruевич Saint-Petersburg State University of Telecommunications» (SPbSUT)	Post 193232, Saint-Petersburg, address: Prospekt Bolshevikov, 22/1 Phone: +7 (812) 326-31-63, local 2022, +79643759970 E-mail: tuzs@spbgut.ru Web: http://tuzs.sut.ru VK: http://vk.com/spbtuzs
-------------------	--	--

Publisher's Imprint

Signed for print 27.09.2018	Cond. printed sheet 14,13	Format 60×84 _{1/8}	Circulation 1000 copies	Order № 2463
Printed 01.10.2018	Saint Petersburg University of State Fire Service of Emercom of Russia 196105, Saint-Petersburg, Moskovskiy prospect, 149			

СОДЕРЖАНИЕ

CONTENTS

Андреева Е.И., Былина М.С., Глаголев С.Ф. Доценко С.Э., Чаймарданов П.А. Свойства временных оптических солитонов в оптических волокнах и возможность их использования в телекоммуникациях. Часть 3	5	Andreeva E., Bylina M., Glagolev S., Dotsenko S., Chaimardanov P. Properties of temporary optical solitons in optical fibers and the possibility of their use in telecommunications. Part 3
Ануфренко А.В., Канаев А.К., Логин Э.В. Модель воздействия злоумышленника на фрагмент транспортной сети связи на основе технологии Carrier Ethernet	17	Anufrenko A., Kanaev A., Login E. Model of the attacker's influence on a fragment of transport communication network based on the Carrier Ethernet technology
Дворников С.В. Билинейные масштабно-временные распределения энергии аффинного класса в частотно-временном пространстве (продолжение обзора)	26	Dvornikov S. Bilinear scale-temporary distributions of energy of the affine class in the frequency- temporary space (continued review)
Жиров В.А., Орлов А.Е., Смирнов А.А. Модель радиолинии спутниковой связи в составе высокоскоростной спутниковой системы	45	Zhirov V., Orlov A., Smirnov A. The model of the satellite communication radio link in the high-speed satellite system
Капралов Д.Д., Кирик Д.И. Стохастическая модель метеорного радиоканала	54	Kapralov D., Kirik D. Stochastic model of meteor-burst radio channel
Макаренко С.И. Усовершенствованный протокол маршрутизации EIGRP, обеспечивающий повышенную устойчивость сетей связи	65	Makarenko S. The improved EIGRP protocol for high network stability
Макаров Л.М., Поздняков А.В., Протасеня С.В. Эргодическая модель атомарных конструкций	74	Makarov L., Pozdnyakov A., Protasenya S. Ergodic model of atomic designs
Муртала И., Тиамийу О.А. Сравнительный анализ анонимных коммуникационных систем с низкой латентностью	85	Murtala I., Tiamiyu O.A. Comparative analysis of low latency anonymous communication systems
Пахомов М.О., Рожнов В.С. Метод количественной оценки слухового восприятия нелинейных искажений. Часть 2. Выбор меры тональности сигнала и учет ее влияния на субъективное восприятие искажений	98	Pakhomov M., Rozhnov V. Method for quantitative evaluation of auditory perception of nonlinear distortion. Part 2. Metric for music signal tonality and its impact on subjective perception of distortions
СВЕДЕНИЯ ОБ АВТОРАХ	106	AUTHORS INDEX

СВОЙСТВА ВРЕМЕННЫХ ОПТИЧЕСКИХ СОЛИТОНОВ В ОПТИЧЕСКИХ ВОЛОКНАХ И ВОЗМОЖНОСТЬ ИХ ИСПОЛЬЗОВАНИЯ В ТЕЛЕКОММУНИКАЦИЯХ. ЧАСТЬ 3

Андреева Е.И.¹, Былина М.С.^{1*}, Глаголев С.Ф.¹, Доценко С.Э.¹, Чаймарданов П.А.¹

¹Санкт-Петербургский государственный университет телекоммуникаций им. проф. М.А. Бонч-Бруевича,

Санкт-Петербург, 193232, Российская Федерация

*Адрес для переписки: BylinaMaria@mail.ru

Информация о статье

УДК 621.39, 530.182

Язык статьи – русский

Ссылка для цитирования: Андреева Е.И., Былина М.С., Глаголев С.Ф., Доценко С.Э., Чаймарданов П.А. Свойства временных оптических солитонов в оптических волокнах и возможность их использования в телекоммуникациях. Часть 3 // Труды учебных заведений связи. 2018. Т. 4. № 3. С. 5-16.

Аннотация: В третьей части работы выполнено теоретическое исследование процессов распространения импульсов, имеющих форму гиперболического секанса, по специальным одномодовым оптическим волокнам с уменьшающейся хроматической дисперсией. Моделирование показало, что в волокнах с уменьшающейся дисперсией поддерживается солитонный режим на значительных расстояниях даже при наличии в ней потерь, которые компенсируются сосредоточенными оптическими усилителями. С помощью моделирования также показано, что вместо волокон с уменьшающейся дисперсией можно успешно использовать несколько последовательно соединенных оптических волокон с дискретно уменьшающейся дисперсией групповых скоростей. Также теоретически рассмотрены квазисолитонные волоконно-оптические системы связи, в волоконно-оптических линейных трактах которых используются секции, состоящие из двух оптических волокон с разными знаками дисперсий групповых скоростей. Причем остаточная дисперсия групповых скоростей в секции имеет отрицательный знак. Это позволяет в сочетании с оптическими усилителями реализовать в волоконно-оптических линейных трактах квазисолитонный режим работы. Квазисолитонные импульсы называют также солитонами с управляемой дисперсией. Проведенные расчеты и моделирование с использованием программы OptiSystem позволили получить инженерные формулы и рекомендации для проектирования квазисолитонных волоконно-оптических систем связи.

Ключевые слова: оптические солитоны, одномодовое оптическое волокно, волокна с уменьшающейся дисперсией, волоконно-оптическая система связи, дисперсия групповых скоростей, фазовая самомодуляция, оптическое усиление, солитоны с управляемой дисперсией, моделирование.

Введение

Предлагаемая Вашему вниманию статья является продолжением ранее опубликованных в предыдущих номерах журнала статей с таким же названием [1, 2]. В данной части рассматриваются возможности использования квазисолитонных волоконно-оптических систем связи (ВОСС) с бинарной амплитудной модуляцией и волоконно-оптическими линейными трактами (ВОЛТ), в которых осуществляется управление хроматической дисперсией (ХД). В работе исследуются процессы распространения квазисолитонных импульсов, имеющих форму гиперболического секанса, в различных схемах ВОЛТ [3, 4]: со специальным одно-

модовым оптическим волокном (ОВ) с уменьшающейся хроматической дисперсией; с секциями, состоящими из нескольких последовательно соединенных одномодовых ОВ с дискретно уменьшающейся дисперсией групповых скоростей (ДГС); с секциями, состоящими из двух одномодовых ОВ с разными знаками ДГС, величина которых выбрана так, чтобы остаточная ДГС была отрицательной. Проведенные расчеты и моделирование с использованием программы OptiSystem [5] показали, что в рассмотренных ВОСС можно реализовать квазисолитонный режим работы. При этом необходимо компенсировать потери в ОВ с помощью оптических усилителей.

Результаты теоретических исследований и моделирования позволили сформулировать условия существования и возможности передачи на большие расстояния квазисолитонов с управлением ХД. Предлагаются инженерные формулы для проектирования солитонных ВОСС.

Теоретический анализ, также как в предыдущих статьях [1, 2], проводился на основе решений нелинейного уравнения Шредингера (НУШ).

1. Общие сведения

Для удобства чтения третьей части статьи напомним основные обозначения и кратко резюмируем полученные во второй части [2] результаты.

Фундаментальным солитоном называется секансный импульс (импульс, имеющий форму гиперболического секанса), который при распространении в ОБ без потерь сохраняет свою форму, амплитуду и длительность. Квазисолитонным импульсом называется реальный оптический импульс, близкий по форме к фундаментальному солитону, но в какой-то степени изменяющий и периодически восстанавливающий свои амплитуду и длительность при распространении по ОБ.

Периодическое дискретное усиление квазисолитонных импульсов можно математически описать, используя НУШ в следующей форме [2]:

$$i \frac{\partial u}{\partial z} + \frac{1}{2} \cdot \frac{\partial^2 u}{\partial \tau^2} + |u|^2 \cdot u = -i \cdot \Gamma \cdot u + \frac{i}{2} \cdot g(z) \cdot L_D \cdot u, \quad (1)$$

где $\tau = T / T_0(0)$ – нормированное время T , отсчитываемое от середины импульса; $T_0(Z)$ – полуширина секансного импульса, которая в общем случае зависит от расстояния Z вдоль ОБ; $T_0(0)$ – полуширина импульса на входе в ОБ; $u(z, \tau)$ – комплексная нормированная амплитуда напряженности электрического поля – огибающая квазисолитонного импульса, распространяющегося в направлении оси Z по ОБ, которая для секансного импульса имеет вид $u(0, \tau) = \text{sech}(\tau)$; $z = Z / L_D$ – нормированное расстояние вдоль ОБ; $L_D = T_0^2(0) / |\beta_2|$ – дисперсионная длина ОБ; β_2 – дисперсия групповых скоростей (ДГС); $\Gamma = \alpha \cdot L_D / 2 = \alpha \cdot T_0^2(0) / (2 \cdot |\beta_2|)$ – затухание в ОБ длиной L_D для напряженности электрического поля; α – коэффициент затухания ОБ; $g(z)$ – функция дискретного усиления в точках расположения ОУ:

$$g(z) = \sum_{m=1}^{N_A} G_m \cdot \delta(z - z_m). \quad (2)$$

В (2) N_A – общее число дискретных ОУ; G_m – усиление сосредоточенного ОУ, расположенного на расстоянии $z = z_m$. Если предположить, что одинаковые ОУ с коэффициентом усиления G_m размещены равномерно с нормированным периодом $z_A = L_A / L_D$, то $z_m = m \cdot z_A$. В приведенных выражениях L_A – длина усилительного участка

(УУ). В качестве сосредоточенных ОУ обычно используют усилители на основе ОБ, легированных эрбием (EDFA, от англ. Erbium Doped Fiber Amplifier).

Расчеты [2] показывают, что квазисолитонный режим на УУ можно поддержать, если:

$$z_A < 1 \text{ или } (L_A < L_D). \quad (3)$$

Физическая причина необходимости ограничения длины УУ L_A (3) заключается в том, что EDFA повышают энергию солитона до входного уровня в легированном эрбием ОБ, имеющем длину несколько метров, а не обеспечивают непрерывное восстановление солитона на протяжении всего УУ. При этом длительность усиленного солитона динамически восстанавливается на выходе ОУ, но в процессе восстановления солитон теряет часть энергии в виде рассеянного излучения. Рассеянное излучение после прохождения многих УУ может накапливаться до заметного уровня и постепенно разрушает солитонный режим.

Для анализа процессов распространения квазисолитонных импульсов в ВОСС, состоящей из нескольких УУ, при выполнении условия (3), огибающую импульсов $u(z, \tau)$ при поиске решения (1) удобно представить в виде произведения [2, 4]:

$$u(z, \tau) = \sqrt{p(z)} \cdot v(z, \tau), \quad (4)$$

где $p(z)$ – быстро меняющаяся и $v(z, \tau)$ – медленно меняющаяся функции. Функция $v(z, \tau)$ удовлетворяет следующему НУШ:

$$i \frac{\partial v}{\partial z} + \frac{1}{2} \cdot \frac{\partial^2 v}{\partial \tau^2} + p(z) \cdot |v|^2 \cdot v = 0. \quad (5)$$

Функция $p(z)$ является решением обыкновенного дифференциального уравнения:

$$\frac{dp}{dz} = [G(z) \cdot L_D - \Gamma] \cdot p. \quad (6)$$

Решение (6) при выполнении (3) имеет вид (рисунк 1):

$$p(z) = p_{0m} \cdot \sum_{m=0}^{N_A-1} \{ \Phi(z - m \cdot z_A) - \Phi[z - (m+1) \cdot z_A] \} \times \exp[-2\Gamma \cdot (z - z_A \cdot m)], \quad (7)$$

где $\Phi(x)$ – функция Хевисайда, которая принимает значение равное 1 при $x \geq 0$ и 0 при $x < 0$.

Условием существования солитонов порядка N в ОБ без потерь является равенство [2, 3]:

$$N^2 = L_D / L_{NL} = \gamma \cdot P_{0m} \cdot T_0^2(0) / |\beta_2| = \gamma \cdot W_0 \cdot T_0(0) / (2 \cdot |\beta_2|), \quad (8)$$

где $L_{NL} = 1 / (\gamma \cdot P_{0m})$ – нелинейная длина ОБ; P_{0m} – пиковая мощность импульса на входе в ОБ; γ – коэффициент нелинейности; W_0 – энергия солитона N -го порядка. Фундаментальному солитону соответствует $N = 1$.

В общем случае значительное уменьшение пиковой мощности и энергии квазисолитонного импульса с увеличением z (рисунок 1) сопровождаются увеличением его длительности t_u и возникновением рассеянного излучения за счет неполной компенсации ХД. Большие и быстрые изменения $p(z)$ могут разрушить квазисолитон. Однако, форма квазисолитонного импульса $v(z, \tau)$ и его длительность мало меняются при выполнении условия (3). На рисунке 1 качественно показаны эти изменения для относительной амплитуды $v(z, 0)$, которая возвращается к первоначальному значению, равному 1, после прохождения ОУ. Форма квазисолитонного импульса практически восстанавливается.

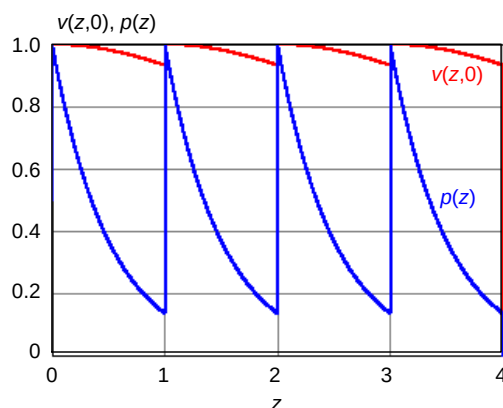


Рис. 1. Зависимости от нормированного расстояния быстро $p(z)$ и медленно $v(z, 0)$ меняющихся функций для ВОСС, состоящей из четырех УУ

Основной идеей поддержания квазисолитонного режима является сохранение средней величины $\bar{p}_m = 1$ нормированной пиковой мощности $p_m(z)$ на УУ на уровне (8), необходимом для формирования фундаментального солитона в ОВ без потерь [2]:

$$\begin{aligned} \bar{p}_m = 1 &= z_A^{-1} \cdot \int_0^{z_A} p_m(z) \cdot dz = \\ &= z_A^{-1} \cdot \int_0^{z_A} \exp(-2 \cdot \Gamma \cdot z) \cdot dz. \end{aligned} \quad (9)$$

Нормированное значение пиковой мощности на входе УУ p_{m0} можно определить из (9):

$$\begin{aligned} p_{m0} = \frac{1}{\bar{p}_m} &= \left(\frac{1}{z_A} \int_0^{z_A} p_m(z) dz \right)^{-1} = \\ &= \frac{2 \cdot \Gamma \cdot z_A}{1 - \exp(-2 \cdot \Gamma \cdot z_A)} = \frac{G \cdot \ln(G)}{G - 1}, \end{aligned} \quad (10)$$

где $G = \exp(2 \cdot \Gamma \cdot z_A)$ – требуемое значение коэффициента усиления ОУ.

Используя (8) и (10), можно определить реальное значение пиковой мощности на входе УУ, необходимое для поддержания квазисолитонного режима [2] в ОВ с потерями:

$$P_{omy} = p_{m0} \cdot P_{om} = \frac{p_{m0} \cdot |\beta_2|}{\gamma \cdot T_0^2(0)}. \quad (11)$$

Такое же значение пиковой мощности требуется для поддержания квазисолитонного режима с распределенными усилителями на основе вынужденного комбинационного рассеяния с встречной накачкой [2]. Наименьшее значение пиковой мощности $P_{omy} = P_{om}$ требуется при использовании распределенных усилителей с двунаправленной накачкой [2].

2. Использование специальных ОВ с изменяющейся ХД

Общим недостатком [2] солитонных ВОСС с управлением затухания является ограничение на длину УУ (2).

Еще в 1987 году [4] была предложена схема построения солитонной ВОСС, использующая дискретные ОУ, и позволяющая увеличить расстояние L_A между ними. Суть предложения состояла в использовании специального ОВ, в котором ДГС, имеющая отрицательное значение, уменьшается по модулю вдоль ОВ [4]. В таких ОВ, называемых в [4] волокнами с уменьшающейся дисперсией (ВУД) убывающая ДГС компенсирует уменьшающуюся фазовую самомодуляцию (ФСМ), испытываемую солитонами, амплитуда которых уменьшается из-за потерь в ОВ. Фактически в этом случае управление дисперсией сочетается с управлением потерями.

Эволюция солитона в ВУД описывается НУШ (4), в котором член, содержащий вторую производную, включает новый параметр d , зависящий от z из-за изменений ДГС вдоль ОВ. Модифицированное НУШ приобретает вид:

$$i \frac{\partial v}{\partial z} + \frac{d(z)}{2} \cdot \frac{\partial^2 v}{\partial \tau^2} + p(z) \cdot |v|^2 \cdot v = 0, \quad (12)$$

где $v = u/\sqrt{p}$, $d(z) = \beta_2(z)/\beta_2(0)$ и $p(z)$ учитывают изменения пиковой мощности, вызванные потерями.

Уравнение (12) отличается от стандартного НУШ без учета потерь в ОВ зависимостью его второго и третьего членов от z . Введем новую переменную:

$$z' = \int_0^z d(z) \cdot dz. \quad (13)$$

Переход к z' (13) перенормирует масштаб длины в зависимости от локального значения ДГС и (12) примет вид:

$$i \frac{\partial v}{\partial z'} + \frac{1}{2} \cdot \frac{\partial^2 v}{\partial \tau^2} + \frac{p(z)}{d(z)} \cdot |v|^2 \cdot v = 0. \quad (14)$$

Если выбрать зависимость ДГС от расстояния так, чтобы $d(z) = p(z) \equiv \exp(-\Gamma \cdot z)$, уравнение (14) сведется к стандартному НУШ, полученному в отсутствии потерь в ОВ (выражение (3) в [1]). В

результате потери при использовании ВУД не влияют на форму фундаментального солитона, несмотря на убывание его энергии с увеличением расстояния. Условие существования квазисолитонного режима (3) в ВОЛТ с дискретными ОУ снимается.

Форма фундаментальных солитонов может поддерживаться в ОВ с потерями, если ДГС убывает в нем экспоненциально:

$$|\beta_2(Z)| = |\beta_2(0)| \cdot \exp(-\alpha \cdot Z), \quad (15a)$$

также как пиковая мощность:

$$|P_0(Z)| = P_0(0) \cdot \exp(-\alpha \cdot Z). \quad (15b)$$

Этот результат можно пояснить, если из (8) записать выражение для отношения пиковой мощности к ДГС для фундаментального солитона на входе ВУД:

$$\frac{P_0(0)}{|\beta_2(0)|} = \frac{P_0(Z)}{|\beta_2(Z)|} = \frac{1}{\gamma \cdot T_0^2}. \quad (16)$$

Из (16) следует, что фундаментальный солитон сохраняет свою форму и длительность даже в ОВ с потерями.

Были изготовлены ВУД с приблизительно экспоненциальным профилем изменения ДГС [4]. Технология изготовления ВУД включала уменьшение диаметра сердцевин по длине ОВ в процессе его вытягивания. Уменьшение диаметра сердцевин ОВ меняет вклад волноводной дисперсии в величину β_2 , что приводит к уменьшению β_2 . В проведенном исследовании ДГС в ВУД изменялась в 10 раз на длине 40 км. В эксперименте, использующем ВУД длиной 40 км, фундаментальные солитоны сохраняли длительность и форму, несмотря на потери, составляющие 8 дБ.

Рассмотрим идеализированную ситуацию, когда ДГС в ВУД с затуханием $\alpha = 0,2$ дБ/км изменяется в соответствии с (15a) и имеет начальное значение ДГС $\beta_{2\max} = \beta_2(0) = -5$ пс²/км. Определим по (16) начальное значение пиковой мощности для ВУД, которое составляет $P_0(0) = 25$ мВт для нелинейного показателя преломления $n_2 = 25 \cdot 10^{-21}$ м²/Вт, коэффициента нелинейности $\gamma = 2$ (Вт·км)⁻¹ и площади модового поля $A_{ef} = 55$ мкм². Тогда для УУ длиной $L_A = 100$ км, конечное значение пиковой мощности $P_0(0) = 0,25$ мВт, а ДГС $\beta_{2\min} = \beta_2(L_A) = -0,05$ пс²/км. Изменения уровня пиковой мощности $P_0(Z)$ и ДГС $\beta_2(Z)$ от расстояния в ОВ для ВУД показаны на рисунке 2. Солитонный режим будет сохраняться на больших расстояниях при установке сосредоточенных ОУ через каждые $L_A = 100$ км.

В настоящее время ВУД серийно не выпускаются. Рассмотрим возможности замены ВУД на несколько последовательно соединенных ОВ с уменьшающейся ДГС. В [4] показано, что большая часть преимуществ ВУД может быть реализована даже при использовании всего четырех отдельных

ОВ. Будем полагать, что используются ОВ одинаковой длины $l_{ob} = 25$ км, имеющие все параметры, кроме ДГС, такие же, как у ВУД. ДГС определяются, как средние значения ДГС для ВУД на каждом 25 километровом участке. Рассмотрим УУ длиной 100 км, состоящий из четырех ОВ. Расчетные значения ДГС для этих ОВ с номерами от 0 до 3 составили: $\beta_{20} = -3,292$; $\beta_{21} = -1,042$; $\beta_{22} = -0,330$; $\beta_{23} = -0,104$. Ступенчатые изменения ДГС вдоль УУ показаны на рисунке 2б.

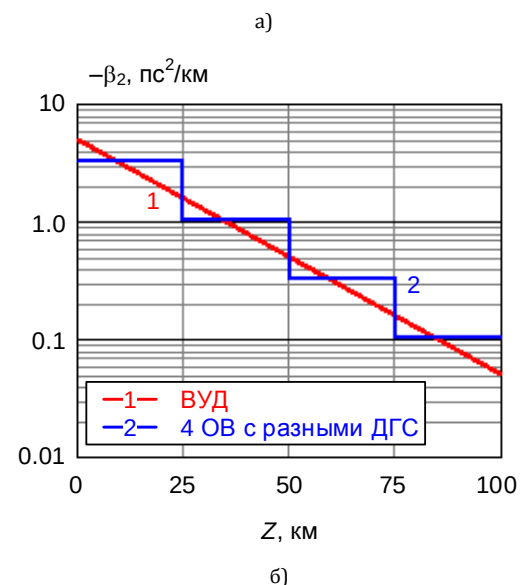
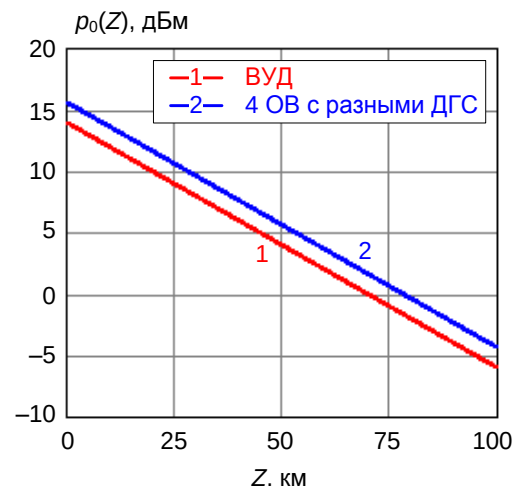


Рис. 2. Зависимости параметров ВОЛТ с использованием ВУД и четырех последовательно включенных ОВ, обеспечивающих квазисолитонный режим, от расстояния: а) пиковая мощность фундаментального солитона; б) ДГС

Для оценки требуемой пиковой мощности для обеспечения квазисолитонного режима на УУ, составленном из четырех ОВ, было проведено моделирование процессов распространения секансных импульсов для одного и нескольких (N_{yy}) составных УУ в схеме, представленной на рисунке 3.

Схема ВОСС содержит:

– генератор периодической битовой последовательности (ГПБ) со скоростью 10 Гбит/с, кото-

рой модулируется оптический генератор секансных импульсов (ГСИ);

- элемент, позволяющий изменять количество УУ $N_{уу}$, каждый из которых состоит из четырех ОВ с дискретно уменьшающейся ДГС, оптического усилителя и оптического фильтра (ОФ);
- фотоприемное устройство (ФПУ).

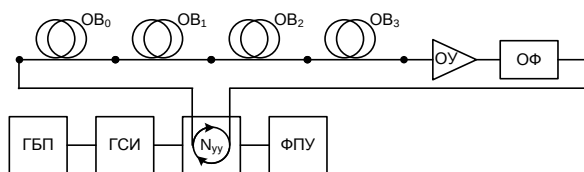


Рис. 3. Схема квазисолитонной ВОСС с составными УУ

Для проведения исследований осуществлялся контроль средней оптической мощности, спектра излучения и осциллограмм сигнала в различных точках ВОСС, а также величины Q -фактора на выходе ФПУ.

Результаты моделирования приведены в таблице 1. Они показывают, что для этой схемы ВОСС требуется пиковая мощность 36,5 мВт, которая превышает значение, необходимое для ВУД (25 мВт). Изменения уровня пиковой мощности на составном УУ показаны на рисунке 2а. Квасисолитонный режим и высокое качество связи обеспечивается на расстояниях более 600 км.

На рисунке 4 показаны результаты моделирования формы квазисолитонных импульсов на вхо-

де и выходе ВОЛТ, составленного из шести УУ $L_A = 100$ км, общей длиной 600 км. Шумы на рисунке 4б не показаны.

Недостатками такого технического решения, реализующего квазисолитонный режим, являются:

- большой перепад в значениях ДГС между началом и концом УУ;
- большое среднее значение ДГС в первом сегменте УУ, увеличивающее необходимое значение пиковой мощности на входе;
- очень маленькое значение ДГС на последнем участке ВОЛТ, которое снижает пороги для нелинейных явлений, например, для четырехволнового смешения.

**ТАБЛИЦА 1. Результаты моделирования процессов
распространения секансных импульсов в квазисолитонной
ВОСС с составными УУ**

Длина ВОСС L , км	Козф. усил. ОУ, G	Вх. мощн. P_0 , мВт	Вых. мощн. P_b , мВт	Длит. имп. $t_{и}$, пс	Q - фактор	P_l / P_0
100	20,7	25	23,5	23	148	0,94
100	20,7	30	29	21	162	0,97
100	20,7	36	37,5	20	175	1,04
100	20,7	36,5	38	20	176	1,04
100	20,7	37	39	19	178	1,05
100	20,53	36,5	36,5	20	176	1,0
200	20,53	36,5	38	18	131	1,04
400	20,53	36,5	36,5	15	95	1,0
600	20,53	36,5	29	19	87	0,79

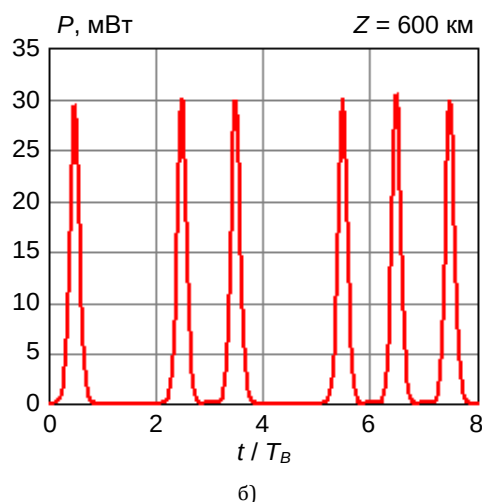
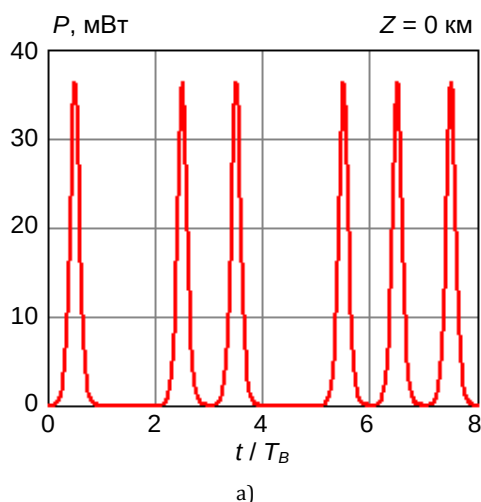


Рис. 4. Формы импульсов: а) на входе ВОЛТ; б) на выходе ВОЛТ длиной 600 км (T_B – длительность битового интервала)

3. Схемы с чередованием оптических волокон с разными знаками ХД

3.1. Анализ без учета ФСМ

В [4] показано, что использование квазисолитонного режима при низкой средней ДГС улучшает характеристики ВОСС. Рассмотрим простейшее техническое решение, позволяющее сочетать низкую среднюю отрицательную ДГС со сравнительно

большими значениями отрицательной и положительной ДГС в чередующихся ОВ. Проанализируем процессы в секции из двух ОВ без учета ФСМ. Первое ОВ длиной l_a должно иметь отрицательную аномальную ДГС β_{2a} , а второе длиной l_n положительную нормальную ДГС β_{2n} . Определим остаточную (среднюю) ДГС секции из двух ОВ:

$$\Delta\bar{\beta}_2 = \frac{\beta_{2n} \cdot l_n + \beta_{2a} \cdot l_a}{l_a + l_n}. \quad (17)$$

Из (17) видно, что остаточная ДГС может иметь очень маленькую отрицательную величину при достаточно большой ДГС каждого ОВ. Это значительно ослабляет эффект четырехволнового смешения и хроматическую дисперсию третьего порядка в каждом ОВ.

Итак, в каждой секции можно обеспечить малую отрицательную величину остаточной ДГС, которую можно затем компенсировать нелинейным эффектом ФСМ. При этом в пределах секции пиковая мощность, длительность, чирпинг и форма квазисолитонного импульса будут периодически осциллировать. Такие квазисолитоны называют солитонами с управляемой дисперсией (УД солитоны) [4].

Наличие в секции ОВ с нормальной положительной ДГС на первый взгляд делает ее неработоспособной, т.к. на таких участках в принципе не поддерживается солитонный режим распространения импульсов. Однако жизнеспособность таких схем известна с 1995 года [4].

Период секции $L_{\text{сек}} = l_a + l_n$ может быть меньше расстояния L_A между ОУ, которое должно быть кратно длине секции:

$$L_A = n \cdot L_{\text{сек}}, \quad (18)$$

где n – количество секций на ОУ.

Для реализации квазисолитонного режима необходимо использовать на входе каждой секции чирпированные импульсы, т.е. импульсы с линейной частотной модуляцией. Рассмотрим влияние чирпинга на процессы распространения квазисолитонных импульсов, учитывая, что их форма близка к гауссовской.

Количественно величину и знак чирпинга C гауссовского импульса с полушириной $T_0(Z)$ можно определить по выражению [3]:

$$C(Z) = T_0^2(Z) \cdot d\omega/dT, \quad (19)$$

где $d\omega/dT$ характеризует изменение частоты внутри импульса.

Напомним некоторые особенности распространения чирпированных импульсов по ОВ [3]. Без учета нелинейных явлений и ХД третьего порядка полуширина чирпированного оптического гауссовского импульса $T_0(Z)$ изменяется с увеличением расстояния Z в ОВ по закону [3]:

$$T_0(Z) = T_0(0) \cdot \sqrt{\left(1 + \frac{C(0) \cdot \beta_2 \cdot Z}{T_0^2(0)}\right)^2 + \left(\frac{\beta_2 \cdot Z}{T_0^2(0)}\right)^2}, \quad (20)$$

где $T_0(0)$ и $C(0)$ – полуширина и чирпинг гауссовского импульса на входе в ОВ (при $Z = 0$).

На рисунке 5 показаны рассчитанные по (20) зависимости относительной ширины гауссовского

импульса $T_0(Z) / T_0(0)$ в ОВ с ДГС $\beta_2 = -5$ пс²/км от расстояния Z .

При проведении расчетов мы ориентировались на ВОСС с амплитудной модуляцией и скоростью передачи 10 Гбит/с. При скважности $Q = 5$ полуширина импульса на уровне половины амплитуды должна составлять 10 пс. Учитывая результаты моделирования, которые будут приведены ниже, и сопоставление их с результатами теоретических расчетов, мы приняли полуширину входного импульса $T_0(0) = 11,5$ пс. Расчеты проведены для различных значений чирпинга $C(0)$ входного импульса.

На рисунке 5 представлены зависимости относительной ширины импульса $T_0(Z) / T_0(0)$ от расстояния при отсутствии начального чирпинга $C(0) = 0$ и при значениях чирпинга $C(0) = +1,2$ и $C(0) = -1,2$. Видно, что при отрицательном значении начального чирпинга $C(0)$, совпадающем по знаку с величиной ДГС β_2 , ширина импульса растет значительно быстрее, чем при отсутствии чирпинга у входного импульса. При разных знаках β_2 и $C(0)$ с увеличением Z ширина импульса сначала уменьшается, а потом начинает возрастать. Это наиболее интересный для нас случай.

На рисунке 5 также представлены графики трех зависимостей полуширины импульса от расстояния Z с начальным положительным чирпингом $C(0) = 1,2, 2,2$ и $2,7$. Видно, что с уменьшением начального чирпинга минимальная полуширина чирпированного импульса также уменьшается. При этом изменяется расстояние, на котором восстанавливается начальная ширина импульса. Можно теоретически и экспериментально подобрать величину $C(0)$ для получения требуемых длин первого l_a и второго l_n ОВ, при которых ширина импульса восстанавливается. В нашем случае для длины ОВ $l_a = 20$ км потребовалось значение начального чирпинга $C(0) = 2,2$.

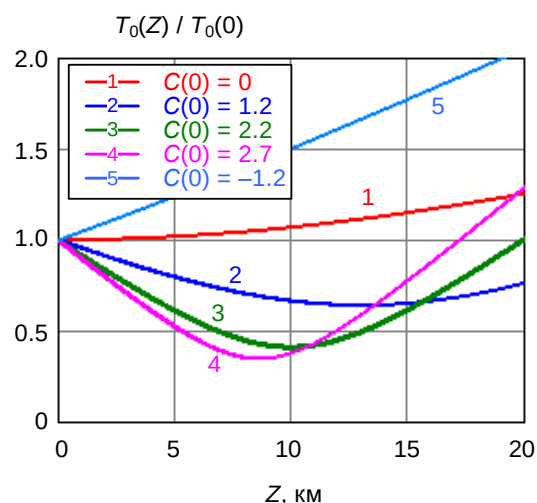


Рис. 5. Зависимости относительной ширины импульса в ОВ от расстояния при различных значениях чирпинга входного импульса

Для того чтобы во втором ОВ ширина импульса также вначале уменьшалась, а потом восстанавливалась, необходимо, чтобы чирпинг входного импульса был отрицательным.

Для анализа выражения (20) определим производную от $T_0(Z)$ по Z :

$$\frac{1}{T_0(0)} \cdot \frac{dT_0(Z)}{dZ} = \frac{T_0(0)}{T_0(Z)} \cdot \left(1 + \frac{C(0) \cdot \beta_2 \cdot Z}{T_0^2(0)} \right) \times \times \frac{C(0) \cdot \beta_2}{T_0^2(0)} + \frac{\beta_2^2 \cdot Z}{T_0^4(0)}. \quad (21)$$

На рисунке 6 показаны зависимости относительной ширины гауссовского импульса (20) и относительной производной ширины импульса (21) от расстояния для первого ОВ при следующих значениях параметров:

$$T_0(0) = 11,5 \text{ пс}, \beta_2 = -5 \text{ пс}^2/\text{км}, C(0) = 2,2.$$

Параметры входного импульса и ОВ подбирались для $l_a = 20$ км.

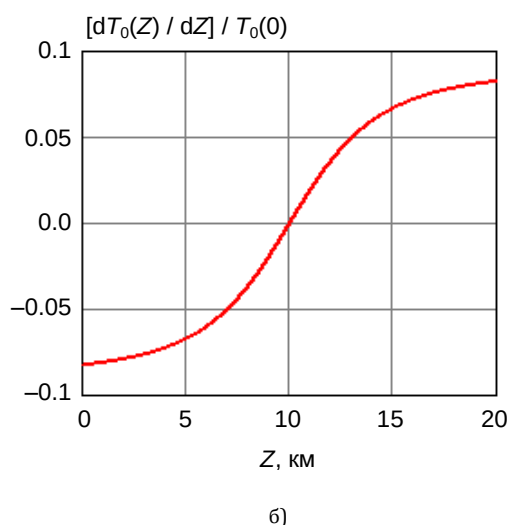
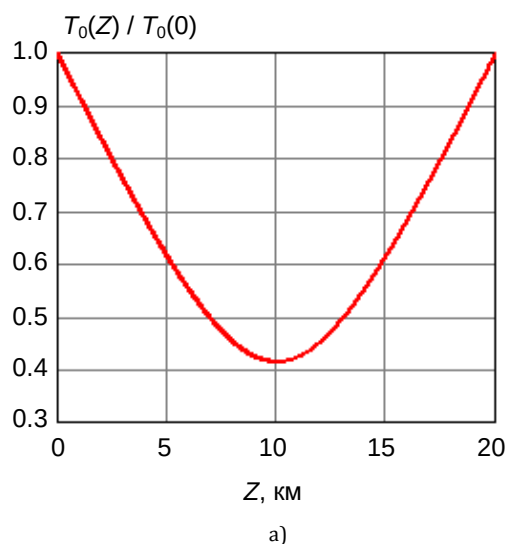


Рис. 6. Зависимости параметров импульса в ОВ длиной 20 км от расстояния в ОВ: а) относительная полуширина; б) производная относительной полуширины

Определим из (21) расстояние $Z_{\min}$, на котором длительность импульса будет минимальной. Для этого приравняем выражение (21) к нулю. Учитывая, что знаменатель всегда больше 0 и не влияет на результат, получим:

$$Z_{\min} = \frac{-C(0) \cdot T_0^2(0)}{\beta_2 \cdot (1 + C^2(0))}. \quad (22)$$

Для $T_0(0) = 11,5$ пс, $C(0) = 2,2$ и $\beta_2 = -5$ пс²/км из (22) получим $Z_{\min} = 10$ км.

На расстоянии $Z_{\min}$ ширина импульса становится минимальной, а чирпинг – равным 0. С дальнейшим увеличением расстояния вновь возникает чирпинг, но другого знака по сравнению с начальным чирпингом $C(0)$. Длительность импульса снова нарастает с увеличением расстояния и на некотором расстоянии Z_{ob} , возвращается к исходной длительности $T_0(Z_{ob}) = T_0(0)$, а чирпинг принимает значение $C(Z_{ob}) = -C(0)$.

Из (21) определим длину волокна Z_{ob} , на которой длительность импульса восстанавливается. Для этого надо выполнить условие:

$$\left(1 + \frac{C(0) \cdot \beta_2 \cdot Z_{ob}}{T_0^2(0)} \right)^2 + \left(\frac{\beta_2 \cdot Z_{ob}}{T_0^2(0)} \right)^2 = 1. \quad (23)$$

Из (23) и (22) следует, что:

$$Z_{ob} = \frac{-2 \cdot C(0) \cdot T_0^2(0)}{\beta_2 \cdot (1 + C^2(0))} = 2 \cdot Z_{\min}. \quad (24)$$

На рисунке 7 показаны зависимости длины ОВ Z_{ob} (24), на которой происходит восстановление ширины импульса, от $C(0)$ при различных β_2 и $T_0(0)$ на входе в ОВ.

Определим минимальное значение полуширины импульса на протяжении одного ОВ. Для этого подставим значение $Z_{\min}$ из (22) в (20):

$$T_0(Z_{\min}) = \frac{T_0(0)}{\sqrt{1 + C^2(0)}} \quad (25)$$

Таким образом, минимальное относительное значение полуширины импульса зависит только от начального значения чирпинга (рисунок 8).

Определим, как изменяется величина чирпинга на протяжении одного ОВ. Напомним, что гауссовский нечирпированный импульс имеет минимальную ширину. В нашем случае это импульс в ОВ на расстоянии $Z_{\min}$, имеющий длительность $T_0(Z_{\min})$. Запишем выражение, связывающее величину чирпинга с расстоянием [3]:

$$T_0(Z) = T_0(Z_{\min}) \cdot \sqrt{1 + C^2(Z)}, \quad (26)$$

из (26) получаем:

$$C(Z) = \sqrt{[T_0(Z)/T_0(Z_{\min})]^2 - 1}. \quad (27)$$

На рисунке 9 приведены результаты расчетов зависимости чирпинга от расстояния в первом ОВ

с отрицательной ДГС по выражениям (20), (25) и (27) при $C(0) = 2,2$, $\beta_2 = -5 \text{ пс}^2/\text{км}$ и $T_0(0) = 11,5 \text{ пс}$.

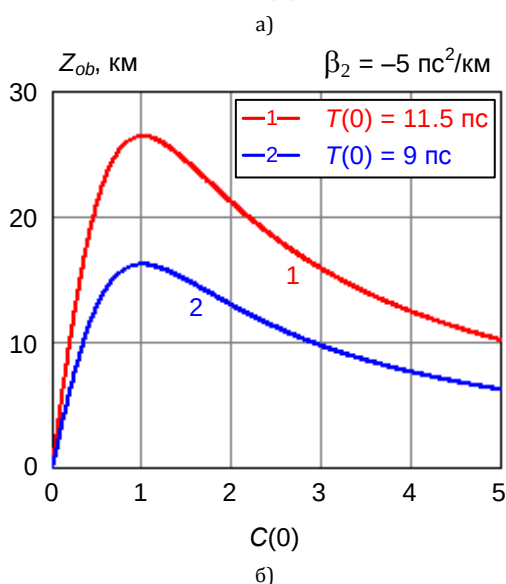
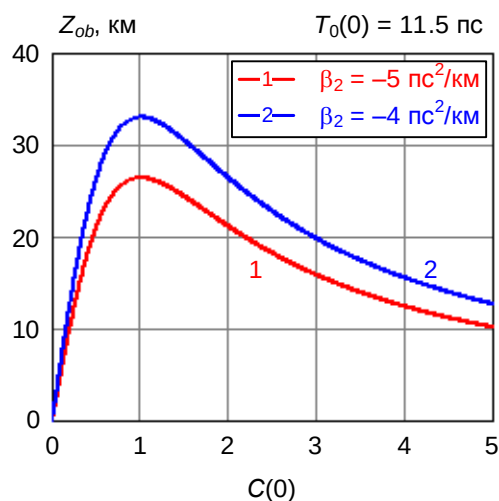


Рис. 7. Зависимости длины ОВ Z_{ob} , на которой происходит восстановление первоначальной длительности импульса, от начального чирпинга: а) при различных ДГС первого ОВ; б) при различных длительностях импульса на входе в первое ОВ

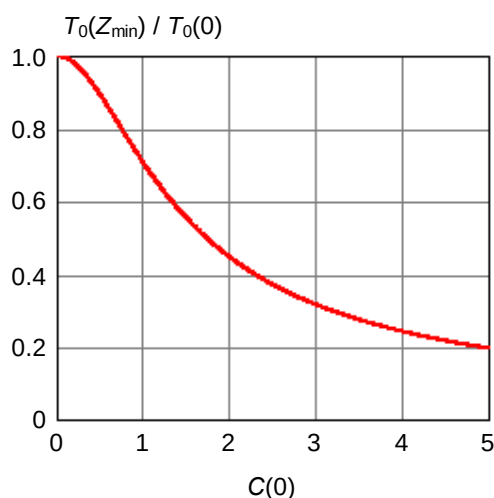


Рис. 8. Зависимость минимальной относительной ширины импульса от начального значения чирпинга на входе в ОВ

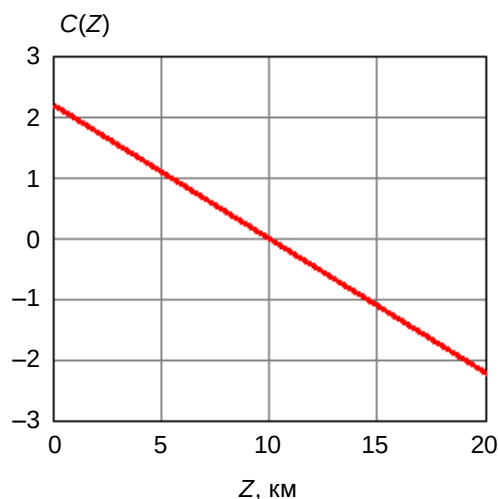


Рис. 9. Зависимость чирпинга от расстояния в первом ОВ

Из рисунка 9 видно, что положительный чирпинг в первом ОВ линейно уменьшается с увеличением расстояния. На расстоянии $Z = Z_{ob} = 20 \text{ км}$ чирпинг изменяет свой знак и по модулю становится равным начальному значению $C(Z_{ob}) = -C(0)$.

Рассмотрим изменения пиковой мощности в первом ОВ с изменением расстояния. Энергию гауссовского нечирпированного импульса с пиковой мощностью $P_m(Z_{min})$ в ОВ без потерь на расстоянии Z_{min} можно определить по выражению [3]:

$$W_0 = \sqrt{\pi} \cdot T_0(Z_{min}) \cdot P_m(Z_{min}). \quad (28)$$

При распространении в ОВ без потерь импульс будет сохранять свою энергию, а его пиковая мощность $P_m(Z)$ будет зависеть от расстояния, ширины импульса и его чирпинга.

Выразим пиковую мощность импульса P_{0m} на входе в ОВ без потерь через его энергию и запишем выражение для зависимости пиковой мощности $P_m(Z)$ от расстояния без учета нелинейных явлений:

$$P_{0m} = W_0 / (\sqrt{\pi} \cdot T_0(0)), \quad (29)$$

$$P_m(Z) = \frac{W_0}{\sqrt{\pi} \cdot T_0(Z)} = \frac{P_{0m}}{\sqrt{1 + \frac{C(0) \cdot \beta_2 \cdot Z}{T_0^2(0)} + \left(\frac{\beta_2 \cdot Z}{T_0^2(0)}\right)^2}}. \quad (30)$$

При учете коэффициента затухания α в ОВ энергия импульса будет зависеть от расстояния Z и выражение для пиковой мощности (30) надо скорректировать:

$$P_m(Z) = \frac{W_0 \cdot \exp(-\alpha \cdot Z)}{\sqrt{\pi} \cdot T_0(Z)} = \frac{P_{0m} \cdot \exp(-\alpha \cdot Z)}{\sqrt{1 + \frac{C(0) \cdot \beta_2 \cdot Z}{T_0^2(0)} + \left(\frac{\beta_2 \cdot Z}{T_0^2(0)}\right)^2}}. \quad (31)$$

Если следом за первым поставить второе ОВ с ДГС другого знака, то длительность импульса в нем будет уменьшаться с увеличением расстояния, так как чирпинг на входе во второе ОВ и его ДГС будут иметь разные знаки. Во втором ОВ процессы будут аналогичны процессам в первом ОВ, т.е. уменьшающаяся ширина импульса после достижения минимального значения возвращается к исходной, а чирпинг будет монотонно увеличиваться и на некотором расстоянии практически вернется к значению $C(0)$.

Таким образом, можно сформировать секцию из двух ОВ с разными знаками ДГС и практически пол-

ностью скомпенсировать в ней ХД, оставив для реализации квазисолитонного режима небольшую некомпенсированную отрицательную ДГС. Ее среднее значение можно определить по выражению (17).

На рисунках 10а–10в показаны зависимости относительной длительности (20), чирпинга (27) и пиковой мощности (30) оптических импульсов от расстояния в секции из двух ОВ без потерь. Расчеты проведены без учета нелинейных явлений при параметрах: $l_a = l_n = 20$ км, $\beta_{2a} = -5$ пс²/км, $\beta_{2n} = 4,9$ пс²/км, $\Delta\beta_2 = -0,05$ пс²/км, $T_0(0) = 11,5$ пс и $C(0) = 2,2$.

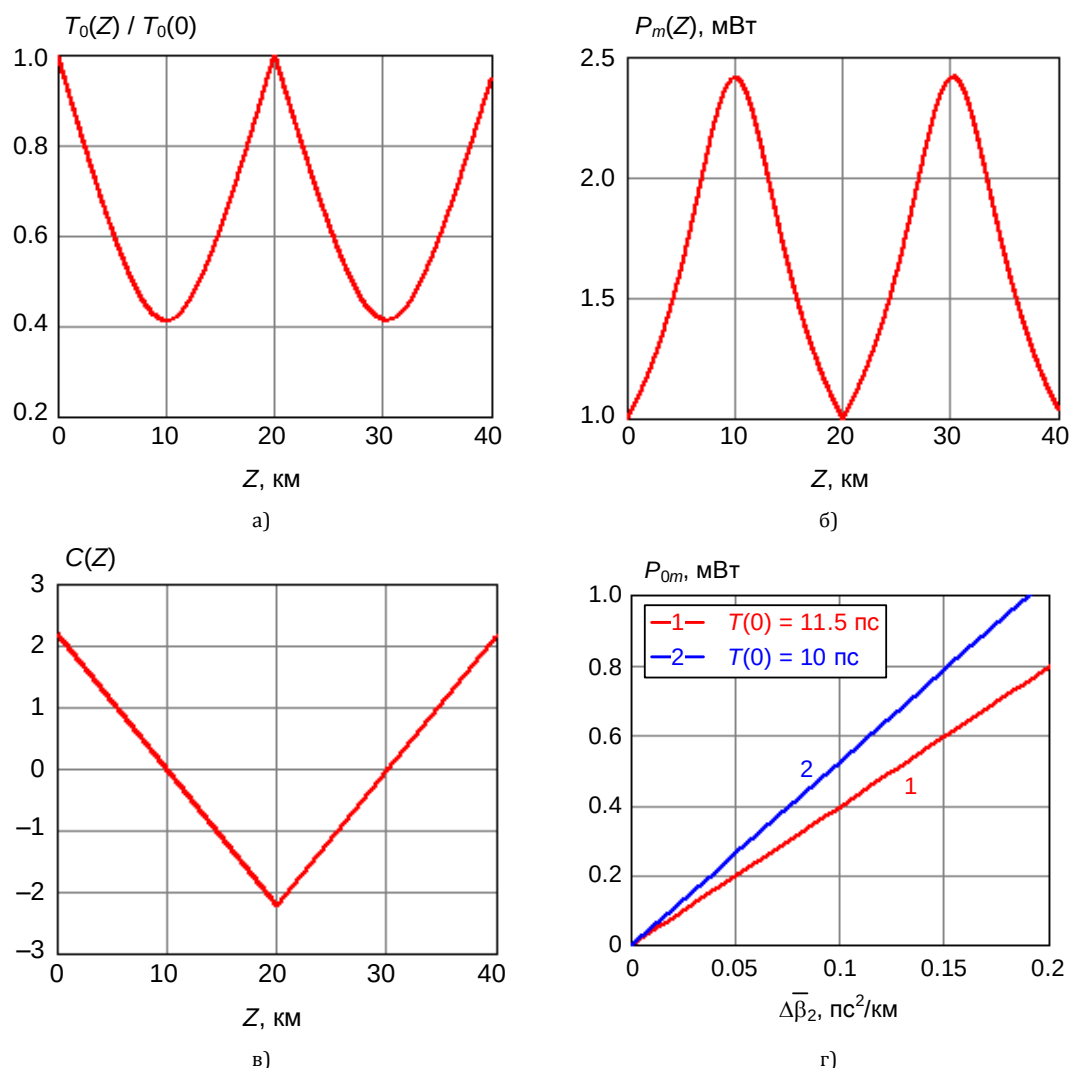


Рис. 10. Параметры УД солитона в секции из двух ОВ без потерь: а) зависимость относительной полуширины от расстояния; б) зависимость пиковой мощности от расстояния; в) зависимость чирпинга от расстояния; г) зависимость пиковой мощности на входе секции от средней ДГС

3.2. Учет ФСМ

Отметим, что в схеме линейного тракта ВОСС с использованием УД солитонов влияние ФСМ сравнительно небольшое по сравнению с влиянием ХД. Оценим величину входной пиковой мощности P_{0m} для компенсации остаточной средней ХД $\Delta\bar{\beta}_2$ в секции из двух ОВ без потерь по выражению:

$$P_{0m} = \Delta\bar{\beta}_2 / (\gamma \cdot T_0^2(0)). \quad (32)$$

На рисунке 10г приведена зависимость (32) необходимой для формирования УД солитона пиковой мощности на входе УУ, состоящего из одной или нескольких секций ОВ без потерь, от остаточной ДГС. Видно, что необходимая мощность при

$\Delta\beta_2 = 0,05$ пс²/км в ОВ без потерь очень мала и составляет всего $P_{0m} = 0,2$ мВт.

Определим по (10) и (11) входную пиковую мощность P_{0my} для поддержания на УУ длиной 120 км средней величины пиковой мощности $\bar{P}_m = 0,2$ мВт (32), необходимой для формирования квазифундаментального солитона в ОВ с потерями: $P_{0my} = G \cdot \ln(G) \cdot P_{0m} / (G - 1) = 1,11$ мВт. Коэффициент усиления ОУ был принят равным $G = 250$ (24 дБ).

Более строго можно учесть влияние ФСМ на процессы распространения УД солитонов с помощью НУШ (12), применявшегося ранее для ВУД. Однако из-за чередования знаков ДГС лучше использовать размерные величины Z и t и переписать (12) в виде [4]:

$$i \frac{\partial B}{\partial Z} - \frac{\beta_2(Z)}{2} \cdot \frac{\partial^2 B}{\partial t^2} + \gamma \cdot p(Z) \cdot |B|^2 \cdot B = 0, \quad (33)$$

где $B(Z, t) = A(Z, t) / \sqrt{p(Z)}$; Z – длина в естественных единицах; $p(Z)$ – решение (7) уравнения (6).

Уравнение (33) можно решить численно методом расщепления и быстрого преобразования Фурье [4]. Можно найти приближительное периодическое решение подгонкой параметров входного импульса (его длительности, чирпинга и пиковой мощности). Форма таких УД солитонов ближе к гауссовской, чем к секансной, как у фундаментальных солитонов.

В [4] обосновано, что УД солитон отвечает периодическому решению системы из двух дифференциальных уравнений первого порядка для полуширины импульса $T_0(Z)$ и чирпинга $C(Z)$:

$$\frac{dT_0(Z)}{dZ} = \beta_2(Z) \cdot \frac{C(Z)}{T_0(Z)}, \quad (34a)$$

$$\frac{dC(Z)}{dZ} = \frac{\gamma \cdot W_0 \cdot p(Z)}{\sqrt{2\pi} \cdot T_0(Z)} + \frac{\beta_2(Z)}{T_0^2(Z)} \cdot (1 + C^2(Z)). \quad (34b)$$

Эти уравнения решаются с периодическими граничными условиями:

$$T_0(0) = T_0(L_A), \quad C(0) = C(L_A), \quad (34в)$$

чтобы обеспечить восстановление начального состояния квазисолитонного импульса после каждого ОУ.

Для решения системы (34) выразим из (34a) $C(Z)$:

$$C(Z) = \frac{T_0(Z)}{\beta_{2a}} \cdot \frac{dT_0(Z)}{dZ}. \quad (35a)$$

Подставим (35a) в (34b) и получим одно дифференциальное уравнение второго порядка для ОВ с отрицательной ДГС:

$$\frac{T_0(Z)}{\beta_{2a}} \cdot \frac{d^2 T_0(Z)}{dZ^2} - \frac{\beta_{2a}}{T_0^2(Z)} - \frac{\gamma \cdot W_0 \cdot \exp(-\alpha \cdot Z)}{\sqrt{2\pi} \cdot T_0(0)} = 0. \quad (35б)$$

Для решения (35б) использовался численный метод Рунге-Кутты. В качестве начальных условий

использовалась длительность чирпированного импульса $T_0(0) = 11,5$ пс и производная от длительности $dT_0(0) / dZ$, которую можно получить из (21). Она составляет $-0,095$ пс/км.

Результаты решения (35) для зависимости относительной мощности и чирпинга от расстояния в ОВ с отрицательной ДГС и ранее используемыми параметрами представлены на рисунке 11. Видно, что результаты более простых расчетов без учета ФСМ совпадают с решением НУШ.

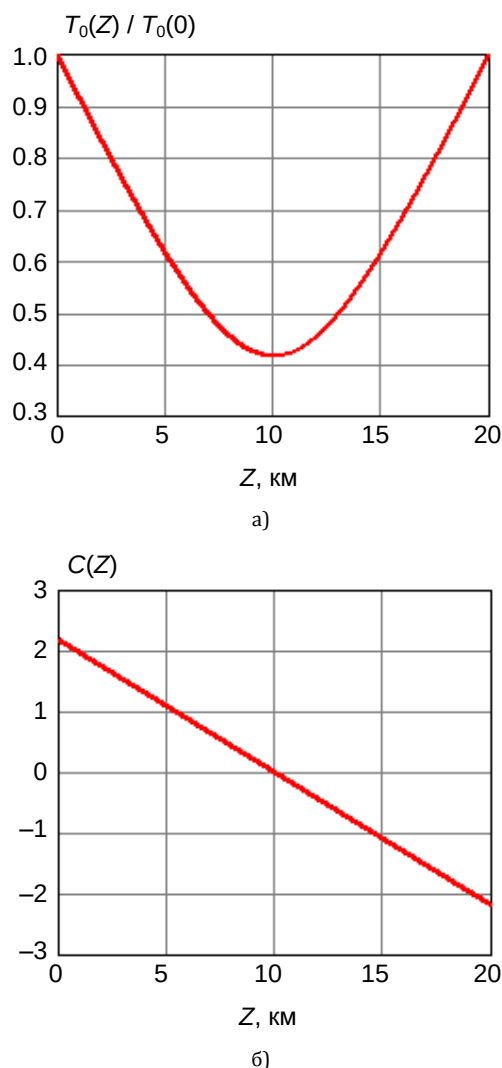


Рис. 11. Зависимости параметров импульса от расстояния в ОВ длиной 20 км: а) относительная полуширина; б) чирпинг

Для проверки теоретических положений проводилось моделирование процессов распространения УД солитонов в схеме ВОСС, приведенной на рисунке 12.

Схема содержит генераторы ГБП и ГСИ, ФПУ, как в схеме на рисунке 3, а также элементы, позволяющие изменять количество УУ N_{yy} и количество секций M на каждом УУ (каждая секция состоит из двух ОВ с разными знаками ДГС), ОУ и ОФ, установленные в конце каждого УУ.

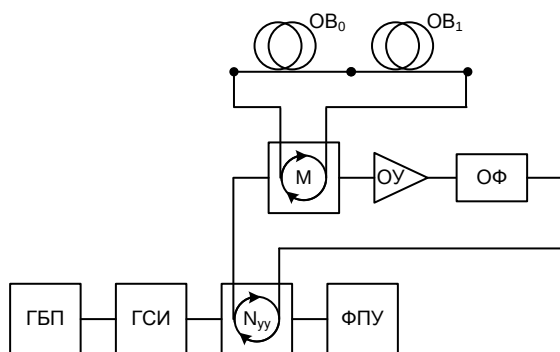


Рис. 12. Схема для моделирования ВОСС с использованием УД-солитонов

На рисунке 13 и в таблице 2 показаны результаты моделирования процессов прохождения импульсов по ВОЛТ, состоящему из нескольких ранее рассмотренных секций ОБ с чередующимися знаками ДГС. Шумы на рисунке 13б не показаны.

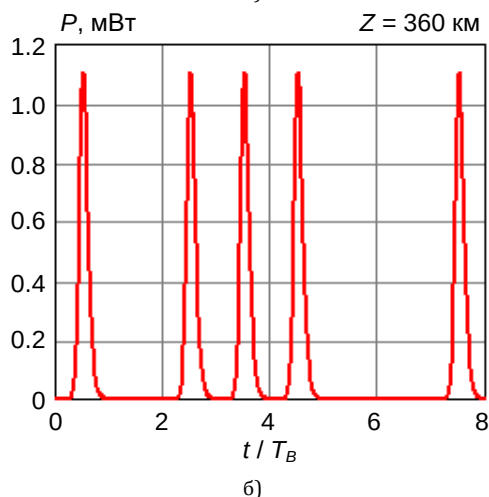
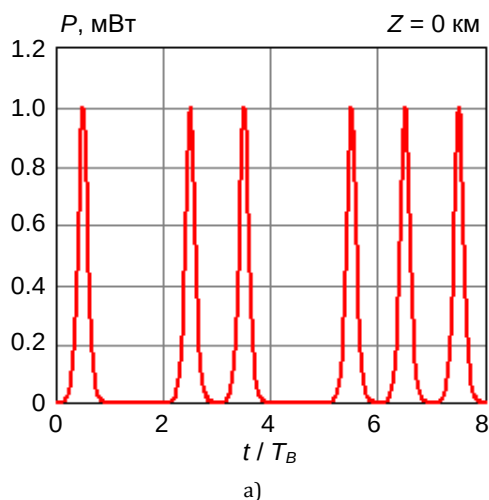


Рис. 13. Формы импульсов: а) на входе ВОЛТ; б) на выходе ВОЛТ длиной 360 км

Моделирование проведено при ранее использованных параметрах, а также при $M = 3$, $N_{yy} = 3$ и $G = 355$ (25,5 дБ).

Результаты моделирования для различных входных пиковых мощностей и расстояний, при-

веденные в таблице 2 показывают, что режим УД солитонов поддерживается в ВОСС на больших расстояниях при расчетных значениях пиковой мощности на входе. Увеличение мощности от 1 до 10 мВт повышает качество связи, т.е. Q -фактор возрастает, а длительность выходного импульса не превышает длительность входного до расстояний 360 км. Видно, что сохранение квазисолитонного режима на больших расстояниях требует тонкой настройки усиления на каждом УУ.

ТАБЛИЦА 2. Результаты моделирования процессов распространения секансных импульсов в квазисолитонной ВОСС с несколькими УУ

Пиковая мощность P_{0m} на входе ВОЛТ, мВт	Длина ВОЛТ, км	Пиковая мощность P_{1m} на выходе ВОЛТ, мВт	Длительность импульса на выходе ВОЛТ, пс	Q -фактор	Отношение пиковых мощностей P_{1m} / P_{0m}
1	120	1,04	17,0	16,1	1,04
1	240	1,10	18,0	11,1	1,10
1	360	1,10	19,0	9,20	1,10
1	480	1,13	21,0	8,68	1,13
1	600	1,26	22,5	9,40	1,26
3	120	3,15	20,0	28,5	1,05
3	240	3,20	16,0	19,8	1,07
3	360	3,20	19,0	16,8	1,07
3	480	3,30	23,0	15,6	1,10
3	600	3,70	25,0	16,9	1,23
10	120	10,2	17,0	51,0	1,02
10	360	10,9	20,0	31,0	1,09
10	480	9,80	24,0	29,1	0,98
10	600	10,4	28,0	28,5	1,04

Заключение

Проведенные исследования показали, что:

1) Поддерживать квазисолитонный режим на значительных расстояниях и увеличить длину УУ позволяет использование специальных ОБ с экспоненциально уменьшающейся с расстоянием ДГС. Однако такие ОБ не выпускаются серийно.

2) Специальное ОБ можно успешно заменить на четыре последовательно соединенных ОБ с уменьшающейся ДГС. Однако это требует увеличения входной пиковой мощности и сопровождается нежелательным увеличением влияния четырехволнового смешения на процессы распространения сигналов.

3) Наилучшие результаты по реализации квазисолитонного режима в ВОСС позволяет получить решение, основанное на использовании на УУ нескольких секций, состоящих из двух ОБ с отрицательной и положительной ДГС, значения которых выбраны так, чтобы остаточная ДГС в секции имела небольшое отрицательное значение.

К сожалению, часть результатов теоретических расчетов и моделирования квазисолитонных ВОСС

с управлением дисперсией не удалось привести в настоящей статье из-за ее ограниченного объема. В результате не были описаны схемы с УД солитонами, использующие чередование ОВ и распределенное оптическое усиление. Не исследованы так-

же предельные возможности по повышению скорости передачи и использованию спектрального уплотнения в квазисолитонных ВОСС. С согласия редакции эти вопросы будут рассмотрены в следующем номере журнала.

Список используемых источников

1. Андреева Е.И., Былина М.С., Глаголев С.Ф., Чаймарданов П.А. Свойства временных оптических солитонов в оптических волокнах и возможность их использования в телекоммуникациях. Часть 1 // Труды учебных заведений связи. 2018. Т. 4. № 1. С. 5–11. DOI:10.31854/1813-324x-2018-1-5-12.
2. Андреева Е.И., Былина М.С., Глаголев С.Ф., Чаймарданов П.А. Свойства временных оптических солитонов в оптических волокнах и возможность их использования в телекоммуникациях. Часть 2 // Труды учебных заведений связи. 2018. Т. 4. № 2. С. 26–35. DOI:10.31854/1813-324x-2018-2-26-35.
3. Агравал Г. Нелинейная волоконная оптика: Пер. с англ. М.: Мир, 1996. 323 с.
4. Кившарь Ю.С., Агравал Г.П. Оптические солитоны. От волоконных световодов до фотонных кристаллов: Пер. с англ. под ред. Н.Н. Розанова. М.: ФИЗМАТЛИТ, 2005. 648 с.
5. User's Reference // Optical Communication System Design Software (OptiSystem). Optiwave. 2014.

* * *

PROPERTIES OF TEMPORARY OPTICAL SOLITONS IN OPTICAL FIBERS AND THE POSSIBILITY OF THEIR USE IN TELECOMMUNICATIONS. PART 3

E. Andreeva¹, M. Bylina¹, S. Glagolev¹, S. Dotsenko¹, P. Chaimardanov¹

¹The Bonch-Bruевич State University of Telecommunications,
St. Petersburg, 193232, Russian Federation

Article info

Article in Russian

For citation: Andreeva E., Bylina M., Glagolev S., Dotsenko S., Chaimardanov P. Properties of Temporary Optical Solitons in Optical Fibers and the Possibility of their Use in Telecommunications. Part 3 // Proceedings of Telecommunication Universities. 2018. Vol. 4. Iss. 3. PP. 5–16.

Abstract: *In the third part of the paper, a theoretical study of the propagation processes of pulses having the form of hyperbolic secant has been carried out using special single-mode optical fibers with decreasing chromatic dispersion. Simulation has shown that the soliton regime is maintained at the longitude at considerable distances even in the presence of losses in the special fibers, which are compensated by discrete optical amplifiers. Using the simulation, it is also shown that instead of special fibers, several series-connected optical fibers can be successfully used with discretely decreasing group velocity dispersion (GVD). Quasi-soliton fiber-optic communication systems are also theoretically considered, in fiber-optic linear paths of which sections consisting of two optical fibers with different signs of GVD are used. And the residual GVD in the section has a negative sign. This allows, in combination with an op-amp, to realize a quasisoliton mode of operation in the fiber-optic linear path. Quasisoliton pulses are also called solitons with a controlled dispersion. Calculations and modeling using the OptiSystem allowed obtaining engineering formulas and recommendations for designing quasi-soliton fiber-optic communication systems.*

Keywords: *optical soliton, temporary optical soliton, single-mode optical fiber, fiber optic communication system, group velocity dispersion, self-phase modulation, optical fiber with decreasing chromatic dispersion, optical amplification, solitons with a controlled dispersion, simulation.*

МОДЕЛЬ ВОЗДЕЙСТВИЯ ЗЛОУМЫШЛЕННИКА НА ФРАГМЕНТ ТРАНСПОРТНОЙ СЕТИ СВЯЗИ НА ОСНОВЕ ТЕХНОЛОГИИ CARRIER ETHERNET

А.В. Ануфренко¹, А.К. Канаев², Э.В. Логин^{2*}

¹Военная академия связи имени Маршала Советского Союза С.М. Буденного,

Санкт-Петербург, 194064, Российская Федерация

²Петербургский государственный университет путей сообщения Императора Александра I,

Санкт-Петербург, 190031, Российская Федерация

*Адрес для переписки: elinabeneta@yandex.ru

Информация о статье

УДК 654.027

Язык статьи – русский

Ссылка для цитирования: Ануфренко А.В., Канаев А.К., Логин Э.В. Модель воздействия злоумышленника на фрагмент транспортной сети связи на основе технологии Carrier Ethernet // Труды учебных заведений связи. 2018. Т. 4. № 3. С. 17–25.

Аннотация: Транспортная сеть связи играет ключевую роль в обеспечении переноса разнородного трафика между узлами доступа. При этом современная пакетная технология Carrier Ethernet, реализующая ряд механизмов для контроля и управления состоянием фрагментов транспортной сети связи, не имеет встроенных механизмов защиты от целенаправленного воздействия злоумышленника. Задача, связанная с оценкой вероятностно-временных характеристик воздействия злоумышленника на сеть, функционирующей на основе упомянутой технологии, является актуальной. В статье рассмотрено ее функционирование в условиях компьютерной атаки типа «Подмена доверенного объекта». Разработана и представлена блок-схема алгоритма функционирования варианта атаки данного типа, а также представлена ее математическая модель, разработанная с помощью метода топологического преобразования стохастических сетей и с учетом профильной модели атаки.

Ключевые слова: транспортная сеть связи, Carrier Ethernet, компьютерная атака, подмена доверенного объекта.

Введение

Транспортная сеть связи (ТрС) является ключевым элементом телекоммуникационной сети, функционирование которой обеспечивает требуемую безопасность управления движением на железнодорожном транспорте. При этом стоит отметить, что современной технологией, которая представляет наибольший интерес относительно применения ее принципов в ТрС, является пакетная технология Carrier Ethernet (CE).

Посредством ТрС обеспечивается перенос трафика (в том числе и служебной информации) между сетями доступа. Структура ТрС имеет распределенный и протяженный характер и может включать в себя множество сетевых элементов. При этом ТрС на основе CE обладает мощными средствами контроля состояния отдельных элементов и сети в целом, а также средствами автоматического и автоматизированного управления. Однако указанные средства управления сетями связи операторского класса не имеют самостоятельных ме-

ханизмов обеспечения безопасности от воздействия компьютерной атаки (КА). Все это говорит о возможной уязвимости ее отдельных фрагментов, поскольку задача, связанная с оценкой вероятностной характеристики воздействия злоумышленника на ТрС CE, остается нерешенной.

Carrier Ethernet как объект воздействия злоумышленником

Сетевая технология Ethernet долгое время полностью отвечала требованиям ведущих операторов связи по качеству и скорости передачи информации. Однако, несмотря на все достоинства Ethernet, ее использованию в сетях операторского класса препятствовал ряд недостатков. Развитие LAN-сетей потребовало усовершенствования технологии Ethernet, что привело к разработке Ethernet операторского класса или Carrier Ethernet. В рамках стандарта IEEE 802.3 международной организацией Metro Ethernet Forum (MEF) разработаны рекомендации для технологии CE и дано

ее определение как универсальной стандартизирующей службы и сети операторского класса.

Техническим комитетом MEF в рамках технологии Carrier Ethernet описаны подходы к реализации сетевого управления и взаимодействия через сеть связи. Процессы, затрагивающие эксплуатацию, управление и обслуживание (OAM, *от англ.* Operation, Administration, Maintenance) в CE позволяют осуществлять сквозное управление неисправностями (Fault Management), управление производительностью (Performance Management), управление конфигурациями (Configuration Management), а также мониторинг параметров работы (Performance Monitoring) в нескольких взаимосвязанных сетях.

На каждом уровне управления имеется свой набор функций, взаимодействие которых позволяет эффективно управлять сетевыми процессами и взаимодействием нескольких сетей.

Одной из целевых задач реализации OAM является Fault Management, которая обеспечивает возможность обнаружения, проверки, локализации неисправностей и оповещения о проблемах с доступностью сервисов. Выполнение перечисленных функций осуществляется в соответствии со стандартами IEEE 802.1ag и IEEE 802.3ah.

Стандарт IEEE 802.1ag ассоциируется как Connectivity Fault Management (CFM) для реализации OAM в сетях с поддержанием технологии пакетных технологий. Основные механизмы CFM включают в себя набор устройств, предоставляющих услугу под управлением определенного оператора, набор элементов для осуществления процесса мониторинга параметров работы сети.

Стандарт ITU-T Y.1731 также, как и стандарт CFM (IEEE 802.1ag) позволяет выполнять функции

управления неисправностями. При этом имеются особенности относительно механизмов, оповещающих об отказе, и способов рассылки служебных сообщений. В данном случае оба стандарта дополняют друг друга.

Стандарт IEEE 802.3ah [1–2] определяет функции, связанные с предоставлением информации о соединениях для передачи сообщений OAM.

Управление конфигурациями осуществляется с помощью стандарта MEF E-LMI [1–2]. Здесь выполняются такие механизмы, как периодический опрос и оповещения о статусе удаленного пользовательского интерфейса, а также оповещения о добавлении, удалении и статусе виртуального соединения.

В регламентирующих документах IEEE 802.1ag [1] и ITU-T Y.1731 [2] для управления соединениями на структуре сети определены границы ответственности (MD, *от англ.* Maintenance domain). Как показано на рисунке 1, границы могут проходить в одном месте, но не могут пересекаться. MD могут быть вложенными (до 8 уровней вложения). Например, такими MD могут быть Клиент/Оператор/Арендуемый сегмент сети. Для мониторинга доступности сервиса в MD определены Maintenance Association (MA). MA в свою очередь определяются как набор Maintenance End Points (MEP) на границах домена. Итак, MD – это группа устройств, предоставляющих услугу под управлением определенного оператора; MA – мониторинг экземпляра сервиса в пределах MD; MP – точка мониторинга, которая генерирует и отвечает на сообщения CFM. Точка MIP (*от англ.* Maintenance Domain Intermediate Point), определяет путь между MEP и место сбоя на этом пути.

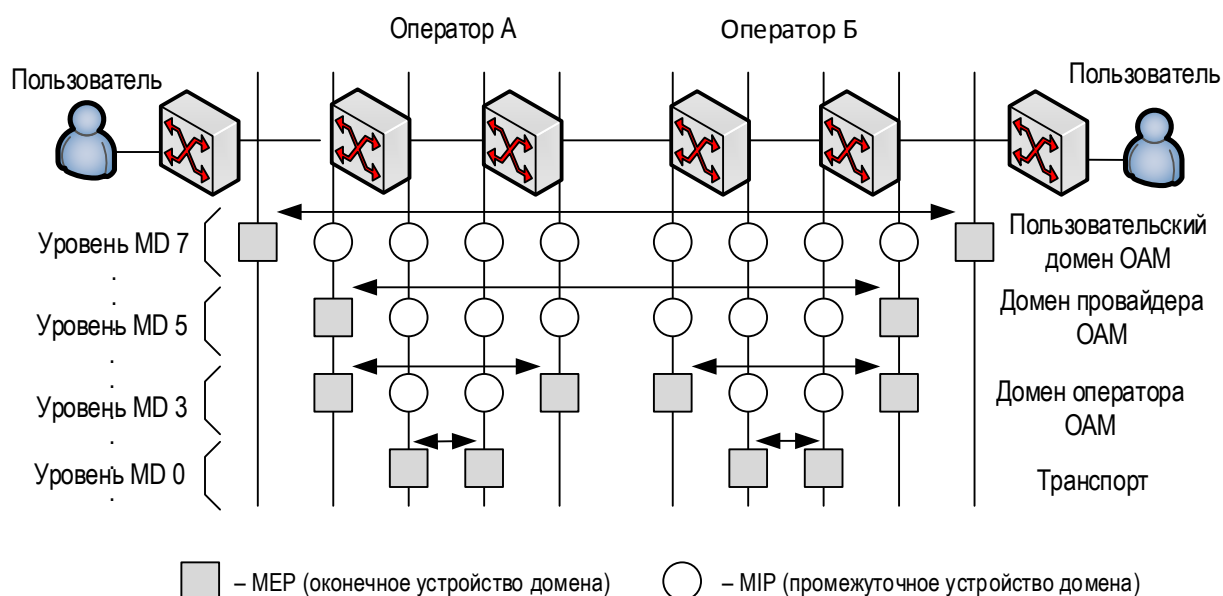


Рис. 1. Архитектура OAM Carrier Ethernet

с помощью средств наблюдения за состоянием ПТС КА и работу системы управления безопасностью связи (СУ), обеспечивающей настройку оптимальных параметров СОВ ТрС с учетом данных, поступающих от средств наблюдения за состоянием ПТС КА.

В этом случае эффективность функционирования ТрС в условиях воздействия на нее ПТС КА обеспечивается более оперативной реализацией цикла по обеспечению ИБ транспортной сети связи по отношению к реализации цикла воздействия программно-технических средств, реализующих компьютерную атаку.

Обобщенно работу ТрС в условиях воздействия системы КА можно представить следующим образом. Реализуется передача диспетчерской информации ($\bar{t}_2$, $\bar{t}_7$, $\bar{t}_{12}$ и т.д.), для чего сначала осуществляются процедуры формирования физического или логического соединения (ФЛС) – $\bar{t}_1$, на что затрачивается среднее время $\bar{t}_{\text{вх}}$.

Затем ($\bar{t}_3$, $\bar{t}_8$ и т.д.) ПТС КА реализует компьютерную атаку за среднее время $\bar{t}_{\text{КА}}$, которую СОВ ТрС сможет обнаружить ($\bar{t}_4$, $\bar{t}_9$ и т.д.) за среднее время $\bar{t}_{\text{ро}}$, определяемое временем реакции системы мониторинга ТрС. Обнаружив воздействие КА, система обнаружения вторжения ТрС будет принимать меры по восстановлению связи ($\bar{t}_5$, $\bar{t}_{10}$ и т.д.) за среднее время $\bar{t}_{\text{пер}}$. После этого СОВ транспортной сети связи обеспечивает формирование ФЛС ($\bar{t}_6$, $\bar{t}_{11}$ и т.д.), на что затрачивается некоторое среднее время $\bar{t}_{\text{вх}}$ и передача диспетчерской информации возобновляется.

Среднее время, затрачиваемое на принятие мер защиты формирования ФЛС ($t_{\text{пмз}}$), характеризует реакцию СУ на воздействие компьютерной атаки, то есть $\bar{t}_{\text{рсу}} = \bar{t}_{\text{пмз}} + \bar{t}_{\text{вх}} = \bar{t}_{\text{пер}} + \bar{t}_{\text{ро}} + \bar{t}_{\text{вх}}$. Среднее время от момента передачи диспетчерской информации до момента воздействия КА является временем реакции ПТС КА ($\bar{t}_{\text{р ТКС}}$) [7–10]. Реакция ПТС КА определяется временем, необходимым для обнаружения безопасного маршрута передачи диспетчерской информации после принятия мер защиты, и временем, необходимым ПТС КА для реализации повторного воздействия на идентифицированный маршрут передачи диспетчерской информации.

Интервал времени $\bar{t}_{\text{КА}}$ назовем циклом воздействия ПТС КА. В реальных условиях один цикл работы программно-технического средства, реализующего компьютерную атаку, отличается по длительности от другого. Это обусловлено тем, что составляющие циклов – время подготовки ПТС КА к работе, принятие мер защиты маршрута от компьютерной атаки, время формирования ФЛС и т.д. – имеют в общем случае различную длительность в каждом цикле.

Среднее время, затрачиваемое на принятие и реализацию мер защиты от КА (время цикла по обеспечению информационной безопасности ТрС), характеризуется реакцией СУ на воздействие КА, то есть $\bar{t}_{\text{рсу}} = \bar{t}_{\text{пмз}} + \bar{t}_{\text{вх}} = \bar{t}_{\text{пер}} + \bar{t}_{\text{ро}} + \bar{t}_{\text{вх}}$ [7–10].

Исследуемая модель функционирования ТрС в условиях целенаправленного воздействия ПТС КА рассматривается с точки зрения функционирования системы управления СЕ в условиях дестабилизирующего воздействия на нее КА. В основе цикла воздействия ПТС КА лежит компьютерная атака типа «Подмена доверенного объекта», которая представляет собой ряд подпроцессов, направленных на передачу кадров Ethernet, предназначенных для реализации некоторых функций OAM (кадр Eth-AIS, кадр CCM с информацией Eth-RDI, кадр Eth-LCK) и обеспечения успешной идентификации этих кадров системой управления СЕ, как поступивших из доверенного источника. Цикл по обеспечению ИБ транспортной сети связи включает в себя задачи по обнаружения ложных Ethernet-кадров и меры по восстановлению безопасной передачи информации по ТрС путем переключения с подвергнувшегося атаке на защищенный от КА резервный маршрут передачи.

Описание компьютерной атаки на ТрС типа «Подмена доверенного объекта»

Реализация компьютерной атаки типа «Подмена доверенного объекта» описывается с применением метода топологического преобразования стохастических сетей (ТПСС), в котором сложный процесс реализации КА декомпозируется на элементарные подпроцессы, каждый из которых характеризуется функцией распределения или средним временем выполнения процесса [10]. Этапы создания математической модели процессов реализации КА с применением ТПСС подробно изложены в [9, 10].

Определение вероятностно-временных характеристик (ВВХ) стохастической сети (СС) процесса реализации КА будем осуществлять при помощи метода двухмоментной аппроксимации.

Тогда среднее время реализации КА, определяемое как начальный момент первого порядка, описывается при помощи выражения [10]:

$$\bar{t}_{\Pi} = -\frac{d}{ds} \left[\frac{Q(s)}{Q(0)} \right]_{s=0}, \quad (1)$$

где $Q(s)$ – эквивалентная функция СС, которая определяется топологическим уравнением Мейсона.

Дисперсия времени реализации КА $D[t_{\Pi}]$, определяемая как второй центральный момент, описывается выражением [8–10]:

$$D[t_{\Pi}] = \frac{d^2}{ds^2} \left[\frac{Q(s)}{Q(0)} \right]_{s=0} - \left\{ -\frac{d}{ds} \left[\frac{Q(s)}{Q(0)} \right]_{s=0} \right\}^2. \quad (2)$$

Вычисление математического ожидания и дисперсии позволяет приближенно определить функцию распределения времени реализации КА как неполную гамма-функцию [8–10]:

$$F(t) = \begin{cases} 0, & t < 0; \\ \int_0^t \frac{\mu^\alpha}{\Gamma(\alpha)} x^{\alpha-1} \exp[-\mu x] dx, & t \geq 0, \end{cases} \quad (3)$$

где $\alpha = \frac{\bar{t}_\Pi^2}{D[t_\Pi]}$, $\mu = \frac{\bar{t}_\Pi}{D[t_\Pi]}$ – параметры формы и масштаба, соответственно.

Для формирования математической модели процесса функционирования КА с применением ТПСС введем следующие ограничения и допущения [7–10]:

- функция распределения случайных величин относятся к классу экспоненциальных;
- вероятности, соответствующие ветвям СС, определяются статистическими методами;
- времена реализации отдельных подпроцессов искомого процесса имеют экспоненциальное распределение;
- модель предполагает отсутствие новых заявок до окончания обработки предыдущей;
- потоки заявок являются неконкурирующими.

С целью сокращения времени реализации КА злоумышленник будет воздействовать несколькими способами. При этом одним из наиболее уязвимых мест, на которое противник будет воздействовать, является система мониторинга и система управления. Исходя из этого в процессе функционирования ТрС учитывается воздействие КА на канальном уровне СЕ.

Блок-схема алгоритма атаки показана на рисунке 4. Злоумышленник реализует компьютерную атаку типа «Подмена доверенного объекта» в следующей последовательности:

- преодолевает систему идентификации и аутентификации сообщений ОАМ, циркулирующих в ТрС за среднее время $\bar{t}_{\text{преод}}$ с функцией распределения времени $W(t)$;
- посылает кадр Ethernet об отключении сигнала аварии (кадр Eth-AIS) за среднее время $\bar{t}_{1.2.1}$ с функцией распределения времени $D(t)$, который с вероятностью P_1 воспринимается система управления ТрС, как кадр, поступивший из доверенного источника;
- посылает кадр Ethernet о состоянии неисправности (кадр ССМ с информацией Eth-RDI) за среднее время $\bar{t}_{1.2.2}$ с функцией распределения времени $L(t)$, который с вероятностью P_2 воспринимается СУ транспортной сети связи, как кадр, поступивший из доверенного источника;
- посылает кадр Ethernet об административном блокировании (кадр Eth-LCK) за среднее время $\bar{t}_{1.2.3}$ с функцией распределения времени $M(t)$, который с вероятностью P_3 воспринимается СУ ТрС, как кадр, поступивший из доверенного источника;

– кроме того, в ходе реализации КА возникают потери и искажения информации или ее блокировка иными средствами защиты, что предполагает их повторного выполнения за среднее время $\bar{t}_{\text{повт}}$ с функцией распределения времени $Z(t)$.

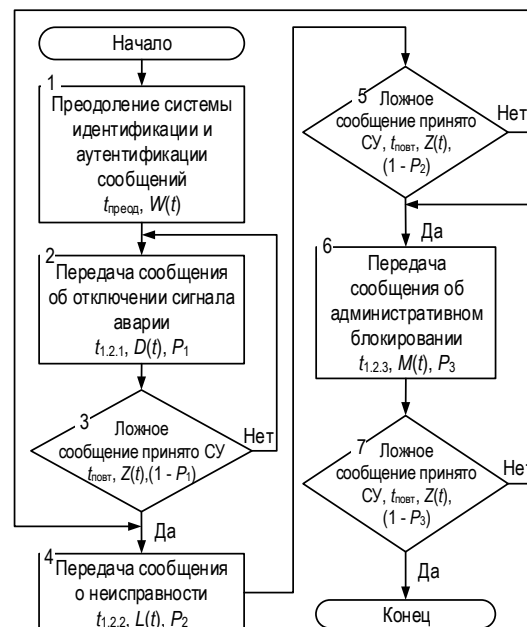


Рис. 4. Блок-схема алгоритма компьютерной атаки типа «Подмена доверенного объекта»

Разработанная математическая модель КА типа «Подмена доверенного объекта» [7, 8] в виде стохастической сети представлена на рисунке 5. Представим подпроцесс, характеризующийся функцией распределения $Q(t)$ и средним временем выполнения процесса функционирования КА, используя преобразование Лапласа [7–10].

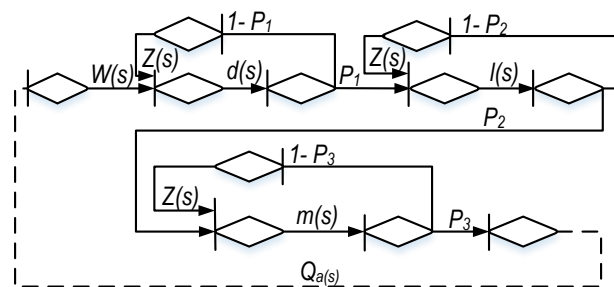


Рис. 5. Стохастическая сеть процесса реализации компьютерной атаки на ТрС через подсистему ОАМ

Ветви разработанной стохастической сети ($W(s)$, $m(s)$, $l(s)$, $d(s)$, $Z(s)$) описывают процессы КА в соответствии с блок-схемой алгоритма компьютерной атаки, отображенной на рисунке 4 [7–10]. На основе данной сети при использовании топологического уравнения Мейсона получена эквивалентная функция $Q(s)$, которая имеет следующий вид [8–10]:

$$F(s) = \frac{\frac{w}{w+s} \cdot \frac{d}{d+s} \cdot \frac{1}{1+s} \cdot \frac{m}{m+s} \cdot P_1 \cdot P_2 \cdot P_3}{1 - \frac{z}{z+s} \cdot \left[(1-P_1) \cdot \frac{d}{d+s} + (1-P_2) \cdot \frac{1}{1+s} + (1-P_3) \cdot \frac{m}{m+s} \right] + \left(\frac{z}{z+s} \right)^2 \cdot \left[(1-P_1) \cdot (1-P_2) \cdot \frac{d}{d+s} \cdot \frac{1}{1+s} + (1-P_1) \cdot (1-P_3) \cdot \frac{1}{1+s} \cdot \frac{m}{m+s} + (1-P_2) \cdot (1-P_3) \cdot \frac{d}{d+s} \cdot \frac{m}{m+s} \right] - \left(\frac{z}{z+s} \right)^3 \cdot (1-P_1) \cdot (1-P_2) \cdot (1-P_3) \cdot \frac{d}{d+s} \cdot \frac{1}{1+s} \cdot \frac{m}{m+s}} \quad (4)$$

Результаты расчетов $F(t)$, $\overline{t_{KA}}$, полученные с помощью выражений (1-3) представлены на рисунках 6-8. В качестве исходных данных используются значения параметров времени реализации КА, приведенных в [8]: $\overline{t_{пред}} = 5$ мин; $\overline{t_{1.2.1}} = 5$ мин; $\overline{t_{1.2.2}} = 5$ мин; $\overline{t_{1.2.3}} = 5$ мин; $P_1 = 0,9$; $P_2 = 0,9$, $P_3 = 0,9$.

На рисунке 6а представлены результаты моделирования в виде функции распределения времени реализации КА для разных значений вероятности P_1 восприятия СУ транспортной сети связи кадра Eth-AIS, посланного ПТС КА, как кадра, поступившего из доверенного источника. На рисунке 6а показано, что с увеличением значения вероятности P_1 значительно возрастает вероятность успешной реализации КА за то же время.

На рисунке 6 представлены функции распределения времени реализации КА для разных значений вероятности P_1 при изменении вероятности P_2 восприятия системы управления ТРС кадра ССМ с информацией Eth-RDI, посланного ПТС КА, как кадра, поступившего из доверенного источника.

Из графиков (рисунок 6) видно, что при осуществлении КА несколькими способами, реализуемыми последовательной цепочкой, рост вероятности успешной реализации каждого из способов компьютерной атаки ведет к значительному росту вероятности принятия СУ ложной информации.

На рисунке 7а представлены результаты моделирования в виде функции распределения времени реализации КА для разных значений времени $\overline{t_{1.2.1}}$ посланного кадра Eth-AIS. Из рисунка видно, что вероятность принятия СУ ложной информации зависит от времени передачи ПТС компьютерной атаки кадра Eth-AIS. С увеличением времени $\overline{t_{1.2.1}}$ вероятность успешной реализации КА уменьшается.

На рисунке 7б представлены результаты моделирования, из которых видно, что рост времени $\overline{t_{повт}}$ повторного выполнения рассматриваемой КА уменьшает вероятность ее успешной реализации, когда атака продолжается более 20 мин и увеличивает, когда атака продолжается менее 20 мин.

На рисунке 7в показаны результаты моделирования в виде функции распределения времени реализации КА для разных значений времени $\overline{t_{пред}}$ преодоления системы идентификации и аутентификации сообщений, циркулирующих в ТРС. Вероятность успешной реализации КА увеличивает уменьшение времени $\overline{t_{пред}}$.

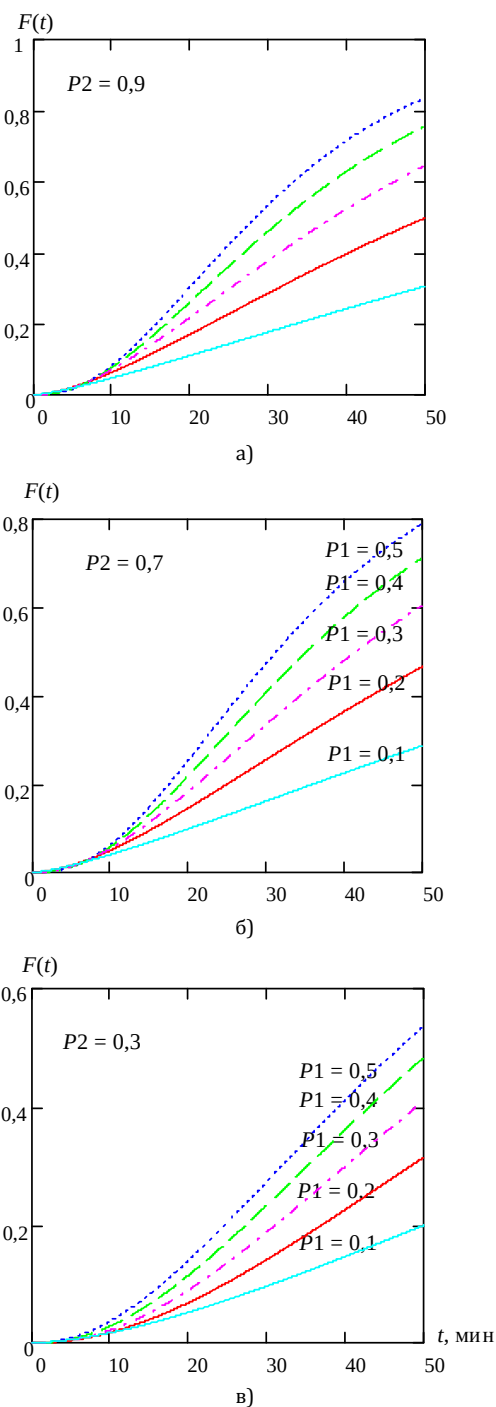


Рис. 6. Зависимость интегральной функции распределения вероятности принятия СУ ложной информации от времени реализации компьютерной атаки при вероятности восприятия СУ ТРС кадра ССМ, как кадра, поступившего из доверенного источника, равна: а) 0,9; б) 0,7; в) 0,3

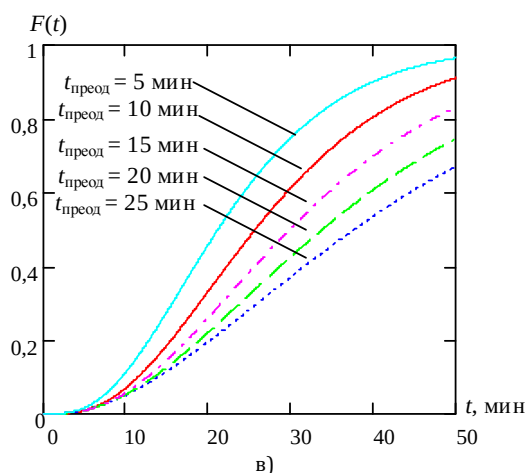
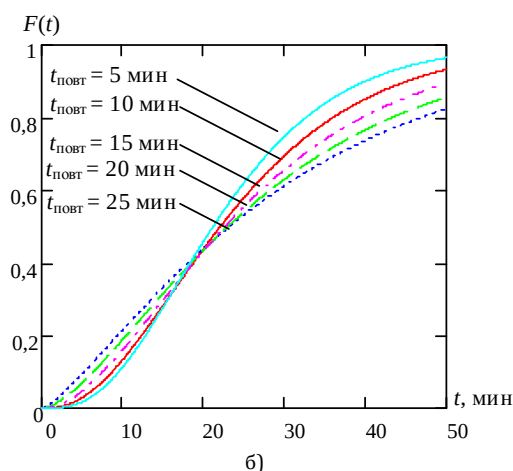
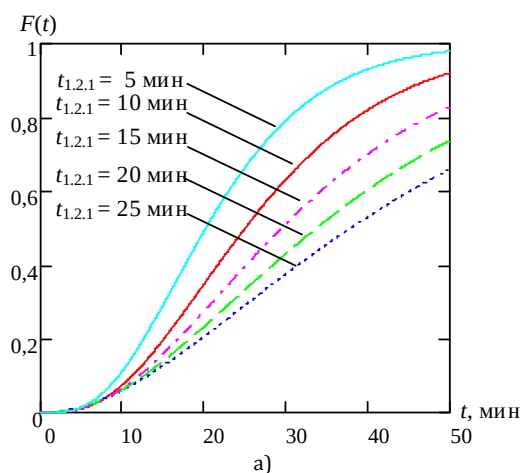


Рис. 7. Зависимость интегральной функции распределения вероятности принятия СУ ложной информации от времени реализации компьютерной атаки при различных значениях времени: а) отправки кадра Eth-AIS; б) повторного выполнения КА; в) преодоления системы идентификации и аутентификации сообщений, циркулирующих в ТрС

Представленные на рисунках 6 и 7 графики иллюстрируют изменение длительности реализации КА на транспортной сети связи СЕ для различных вероятностей установления данного воздействия при изменении интегральной функции (комплексная оценка всех этапов компьютерной атаки), давая возможность определить вероятность $P_n(t \leq T_3)$ решения задачи по подмене доверенного

объекта за время не более заданного T_3 . Это позволяет оценивать эффективность функционирования ТрС СЕ в условиях воздействия КА и обоснованно задавать требуемые временные характеристики реализации КА с целью перспективного моделирования механизмов защиты ТрС СЕ.

Разработанная математическая модель дает возможность определения различных функциональных зависимостей времени реализации КА от изменения ее параметров. Так, например, с помощью модели можно определить среднее время $t(P1)$ реализации КА (рисунки 8 и 9).

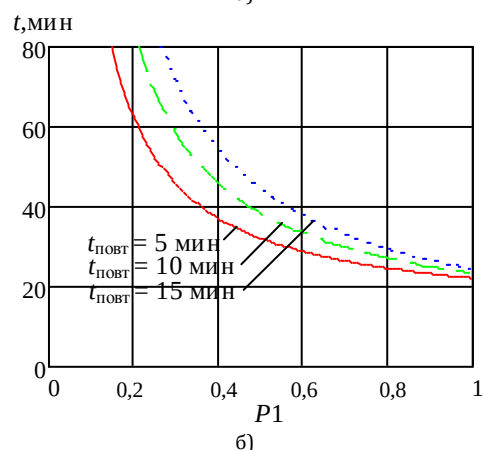
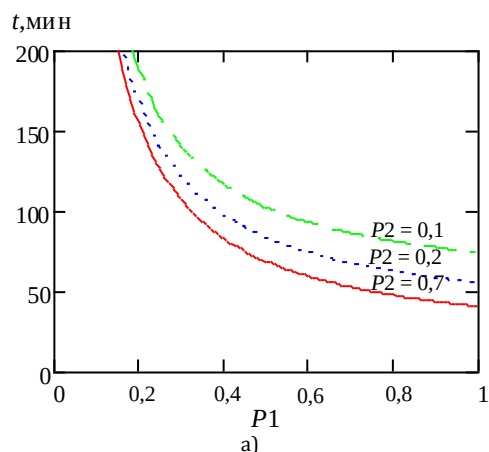


Рис. 8. Зависимость среднего времени реализации компьютерной атаки от вероятности восприятия СУ ТрС кадра Eth-RDI, как кадра, поступившего из доверенного источника: а) при различной вероятности восприятия СУ ТрС кадра ССМ, как кадра, поступившего из доверенного источника; б) при различных значениях времени повторного выполнения КА

Графики на рисунке 8а демонстрируют зависимость среднего времени реализации КА от вероятности восприятия системы управления ТрС кадра Eth-RDI, как кадра, поступившего из доверенного источника при разной вероятности выполнения процесса передачи кадра ССМ системой ПТС КА. Из графиков видно, что рост $P1$ и $P2$ ведет к снижению времени успешной реализации КА. Низкое значение $P1$ (0 – 0,2) снижает чувствительность разработанной модели КА к изменению $P2$ вероятности восприятия СУ ТрС кадра ССМ, как кадра, поступившего из доверенного источника.

Графики на рисунке 8б демонстрируют зависимость среднего времени реализации КА от вероятности восприятия СУ ТрС кадра Eth-RDI, как кадра, поступившего из доверенного источника при разных значениях времени повторного выполнения КА. Из графиков видно, что высокое значение $P1$ (0,8 – 1) ведет к снижению чувствительности разработанной модели КА к времени $\overline{t_{повт}}$ повторного выполнения рассматриваемой КА.

На рисунке 9а показаны результаты моделирования, демонстрирующие зависимость среднего времени реализации компьютерной атаки от вероятности восприятия СУ ТрС кадра Eth-RDI, как кадра, поступившего из доверенного источника при разных значениях времени преодоления системы идентификации и аутентификации сообщений, циркулирующих в ТрС. Низкое значение $P1$ (0 – 0,2) снижает чувствительность разработанной модели КА к изменению времени $t_{преод}$ преодоления системы идентификации и аутентификации сообщений.

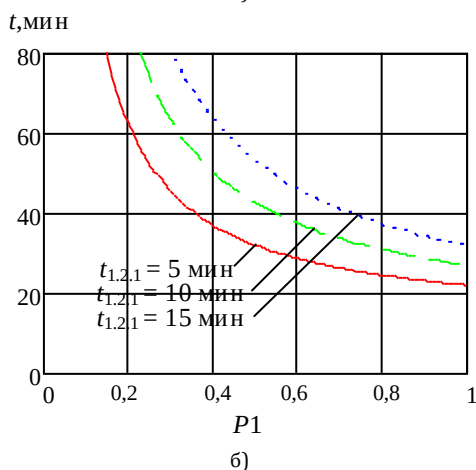
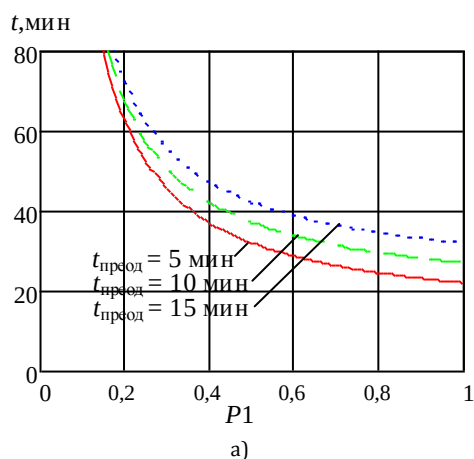


Рис. 9. Зависимость среднего времени реализации компьютерной атаки от вероятности восприятия СУ ТрС кадра Eth-RDI, как кадра, поступившего из доверенного источника при разных значениях времени: а) преодоления системы идентификации и аутентификации сообщений, циркулирующих в ТрС; б) задержки кадра Eth-AIS

На рисунке 9б показаны результаты моделирования, демонстрирующие зависимость среднего

времени реализации компьютерной атаки от вероятности восприятия СУ ТрС кадра Eth-RDI, как кадра, поступившего из доверенного источника при разных значениях времени задержки кадра Eth-AIS. Разработанная модель КА остается чувствительна ко времени $t_{1.2.1}$ задержки кадра Eth-AIS при любых значениях вероятности $P1$.

Таким образом, можно сделать общий вывод: при возникновении условий, позволяющих увеличить вероятность подмены доверенного объекта, в разы уменьшается время реализации КА. Также это позволяет формировать исходные данные по реализации КА типа «Подмена доверенного объекта» в ТрС СЕ для перспективных исследований в области разработки механизмов защиты транспортной сети связи СЕ от целенаправленных воздействий злоумышленника. Анализ полученных результатов показывает, что математическая модель типа «Подмена доверенного объекта» работоспособна, чувствительна к изменению входных данных.

Выводы

ТрС является ключевым элементом телекоммуникационной сети с точки зрения управления и обеспечения безопасности функционирования железнодорожного транспорта. При этом технология СЕ представляет наибольший интерес относительно применения ее принципов в ТрС.

Важным местом относительно воздействия злоумышленника на ТрС является система мониторинга и управления, на которые злоумышленник будет воздействовать сразу несколькими способами. Причем одним из основных видов воздействия является КА типа «Подмена доверенного объекта».

Составляющие каждого цикла функционирования ПТС КА имеют различную длительность, что требует разработки математических моделей КА для их учета при анализе реакции ПТС КА и затрат времени для реализации мер защиты ТрС от КА. Таким образом для разработки модели КА типа «Подмена доверенного объекта» используются профильные модели атак и метод ТПСС.

Разработанные модели КА помогают обоснованно задавать требуемые временные характеристики реализации КА с целью перспективного моделирования механизмов защиты ТрС СЕ, а также позволяют формировать исходные данные по реализации КА типа «Подмена доверенного объекта» в ТрС СЕ для перспективных исследований в области разработки механизмов защиты ТрС СЕ от целенаправленных воздействий злоумышленника.

Учитывая тот факт, что рассмотренный тип компьютерной атаки является лишь одним из основных вариантов воздействия злоумышленника на ТрС, задачи разработки моделей КА представляют дальнейший интерес.

Список используемых источников

1. IEEE 802.1ag Connectivity Fault Management.
2. ITU-T Y.1731 OAM functions and mechanisms for Ethernet based networks.
3. Ануфренко А.В., Волков Д.В., Канаев А.К. Принцип организации узла агрегации мультисервисной сети связи // Актуальные проблемы инфотелекоммуникаций в науке и образовании. IV Международная научно-техническая и научно-методическая конференция: сборник научных статей в 2 томах. 2015. Т. 1. С. 203–206.
4. Степанов С.Н. Основы телетрафика мультисервисных сетей. М.: Эко-Трендз, 2010. 392 с.
5. Цыбаков В.И. Разработка и исследование метода расчета качества обслуживания пользователей широкополосной интегрированной мультисервисной корпоративной сети: дис. канд. техн. наук. М., 2005. 174 с.
6. Боев В.Д., Сыпченко Р.П. Компьютерное моделирование. Элементы теории и практики: учеб. пособие. СПб.: ВАС, 2009. 436 с.
7. Ануфренко А.В., Канаев А.К. Предложения по обоснованию требований к характеристикам оборудования узлов пакетной транспортной сети связи // Т-КОММ: Телекоммуникации и транспорт. 2018. Т. 12. № 2. С. 47–54.
8. Коцыняк М.А., Осадчий А.И., Коцыняк М.М., Лаута О.С., Дементьев В.Е., Васюков Д.Ю. Обеспечение устойчивости информационно-телекоммуникационных сетей в условиях информационного противоборства. СПб: ЛО ЦНИИС, 2014. 126 с.
9. Привалов А.А. Метод топологического преобразования стохастических сетей и его использования для анализа систем связи ВМФ. СПб.: ВМА, 2001. 186 с.
10. Привалов А.А., Карабанов Ю.С., Королев А.И., Кириленко В.О. Разработка структуры программного комплекса моделирования информационного конфликта системы безопасности телекоммуникационного объекта РЖД с подсистемой нарушителя // Интеллектуальные системы на транспорте: материалы V Международной научно-практической конференции «ИнтеллектТранс-2015». 2015. С. 327–332.

* * *

MODEL OF THE ATTACKER'S INFLUENCE ON A FRAGMENT OF TRANSPORT COMMUNICATION NETWORK BASED ON THE CARRIER ETHERNET TECHNOLOGY

A. Anufrenko¹, A. Kanaev², E. Login²

¹Telecommunications Military Academy,
St. Petersburg, 194064, Russian Federation

²PGUPS of the Emperor Alexander I,
St. Petersburg, 190031, Russian Federation

Article info

Article in Russian

For citation: Anufrenko A., Kanaev A., Login E. Model of the Attacker's Influence on a Fragment of a Transport Communication Network Based on the Carrier Ethernet Technology // Proceedings of Telecommunication Universities. 2018. Vol. 4. Iss. 3. PP. 17–25.

Abstract: *The transport communication network plays a key role in ensuring the transfer of traffic between access nodes. At the same time, modern Carrier Ethernet network technology, which implements a number of mechanisms for monitoring and managing the state of the transport communication network fragments, does not have built-in protection mechanisms against the targeted attack of the attacker on the communication network. In this regard, the task associated with the evaluation of the probability-time characteristics of the attacker's attack on the transport communication network, functioning on the basis of Carrier Ethernet technology, is actualized. The article examines the functioning of the Carrier Ethernet transport network in terms of the impact on the computer-style attack "Substitution of the trusted object." A block diagram of the algorithm of the computer attack of this type will be revealed, as well as its mathematical model developed taking into account the profile model of the attack and the method of topological transformation of stochastic networks.*

Keywords: *transport communication network, Carrier Ethernet, computer attack, substitution of the trusted object.*

БИЛИНЕЙНЫЕ МАСШТАБНО-ВРЕМЕННЫЕ РАСПРЕДЕЛЕНИЯ ЭНЕРГИИ АФФИННОГО КЛАССА В ЧАСТОТНО-ВРЕМЕННОМ ПРОСТРАНСТВЕ (продолжение обзора)

С.В. Дворников^{1*}

¹Военная академия связи имени Маршала Советского Союза С.М. Буденного,
Санкт-Петербург, 194064, Российская Федерация

*Адрес для переписки: practicsv@yandex.ru

Информация о статье

УДК 621.391

Язык статьи – русский

Ссылка для цитирования: Дворников С.В. Билинейные масштабнo-временные распределения энергии аффинного класса в частотно-временном пространстве: продолжение обзора // Труды учебных заведений связи. 2018. Т. 4. № 3. С. 26–44.

Аннотация: Представлен обзор теоретических основ синтеза масштабнo-временных распределений аффинного класса. Проанализирована методология вейвлет-преобразований с адаптивным окном анализа в интересах обработки процессов с высокой нестационарностью параметров. Обобщены аналитические подходы синтеза распределений инвариантных к аффинным сдвигам. Показана общность масштабнo-временных и частотно-временных распределений. Приведены аналитические выражения и результаты экспериментов наиболее известных форм распределений аффинного класса.

Ключевые слова: масштабнo-временные распределения, методы обработки процессов с высокой нестационарностью параметров, распределения аффинного класса, вейвлет-преобразования.

Введение

В предшествующих работах по данной тематике [1, 2] достаточно подробно были рассмотрены билинейные частотно-временные распределения (ЧВР), представляющие собой научно-методический аппарат методов гармонического анализа. Основопологающим моментом для ЧВР является функция Вигнера, позволяющая формировать их на своей основе, тем самым объединяя в единый класс распределений Коэна [2–5].

Однако развитие билинейных распределений не ограничивается указанным классом ЧВР, поскольку аффинные финитные функции позволили открыть новую страницу данной научной теории [6]. В связи с этим, в настоящей статье представлен обзор так называемых масштабнo-временных распределений, основанных на кратно масштабных преобразованиях сигнальной энергии, которые в [7] определены как распределения аффинного класса. В общем случае, масштабнo-временные распределения базируются на вейвлет-преобразованиях сигналов. Развитие данного научного направления определили работы таких известных ученых как Bertrand J., Bertrand P., Flandrin P., Baraniuk R., Goncalves P., Auger F и др. [7–12].

Проблемы, определившие переход к масштабнo-временным описаниям энергии

Высокая нестационарность изменения спектрально-временных параметров исследуемых процессов на интервале наблюдения приводит к тому, что оценки на основе ЧВР получаются как бы «размытыми», смазанными [13]. Одним из решений указанной проблемы явилась реализация концепции адаптации длительности окна анализа в зависимости от характера происходящих изменений исследуемого процесса [10]. В частности, в [14, 15] предложен критерий, определяющий сигналы с нестационарным изменением параметров. К таковым следует относить сигналы, у которых коэффициент перекрытия диапазона изменения контролируемых параметров составляет не менее двух на интервале их наблюдения.

Под действие указанного критерия попадают, например, неоднородности оцениваемого спектра переходных процессы при смене частотных позиций и т. д. Интуитивно ясно, что их число будет возрастать с приближением к верхней границе частотного диапазона. Поэтому и разрешающую способность инструмента анализа целесообразно

регулировать в зависимости от наблюдаемой частоты [16].

Действительно, чем выше ее значение в пределах диапазона, тем более широкополосные сигналы могут в нем использоваться, а, следовательно, необходимо увеличивать полосу канальных фильтров с целью недопущения сглаживания высокочастотных составляющих исследуемых процессов.

Приведенные рассуждения позволяют сделать вывод, что анализ быстрых изменений наблюдаемого процесса целесообразно проводить, используя относительно узкое временное окно наблюдения. В тоже время анализ относительно медленных изменений удобнее проводить в пределах широкого временного интервала [14].

Таким образом, необходимо иметь «гибкое» частотно-временное окно, которое автоматически сжимается при высокочастотном изменении параметров сигналов (исследуемых процессор) и расширяется при относительно стационарном их поведении. Указанным требованиям в полной мере удовлетворяет интегральное вейвлет-преобразование (ИВП) [17, 18], методология которого и была положена в основу методов кратномасштабного анализа, у истоков которого стояли Bertrand P. и Flandrin P. [7–10, 12, 19]. Именно они обосновали возможность формирования на основе вейвлет-преобразований сигналов их псевдобилинейных форм масштабно-временных распределений, подробный анализ свойств которых представлен в [6, 14]. В том числе получены удобные компактные формы представления кратковременных фрагментов сигналов для реализации их в алгоритмах обработки [20–26].

Непосредственно сущность аффинной инвариантности, определяющая физическую природу масштабно-временных описаний плотности распределения энергии, была подробно изложена в [10, 17, 27]. Это позволило установить общность аффинных распределений с ЧВР класса Козна [3, 11, 14, 29, 30], а также их практическую применимость [25, 26].

Теоретические основы формирования масштабно-временных распределений на основе вейвлет-преобразований

Рассмотрим ИВП с позиций фильтрации сигналов частотно-зависимым окном, которое определим следующим образом: функция $h \in L^2(\mathbf{R})$, тождественно не равная нулю, называется функцией-окном, если произведение $x \cdot h(x)$ также принадлежит $L^2(\mathbf{R})$, а центр t^* и радиус Δ_h функции-окна h определяются как:

$$t^* = \frac{1}{\|h\|_2^2} \int_{-\infty}^{\infty} x |h(x)|^2 dx \quad (1)$$

и

$$\Delta_h = \frac{1}{\|h\|_2^2} \left\{ \int_{-\infty}^{\infty} (x - t^*)^2 |h(x)|^2 dx \right\}^{1/2}. \quad (2)$$

Согласно введенному определению, ширина функции-окна составит $2\Delta_h$.

Заметим, что выбор вейвлет-функции в качестве функции-окна требует корректного определения базисному вейвлету ψ его образа преобразования Фурье $\hat{\psi}$, а также, соответственно, значений их центров и радиусов: $t^*, \omega^*, \Delta_\psi, \Delta_{\hat{\psi}}$.

Такой подход позволяет четко определить границы временного окна, локализуемые ИВП W_ψ при обработке, в общем случае, аналогового сигнала $z(t)$:

$$(W_\psi z)(b, a) = |a|^{-\frac{1}{2}} \int_{-\infty}^{\infty} z(t) \overline{\psi_n\left(\frac{t-b}{a}\right)} dt, \quad (3)$$

$$[b + at^* - a\Delta_\psi, b + at^* + a\Delta_\psi]. \quad (4)$$

Центр локализации $b + at^*$ и его значимую ширину $2a\Delta_\psi$ называют параметрами временной локализации сигнала (функции) во «временном окне» [14]. Так, на рисунке 1 представлен частотно-временной план разбиения пространства $L^2(\mathbf{Z})$, образованного базисом вейвлетов Гаусса 5-го порядка.

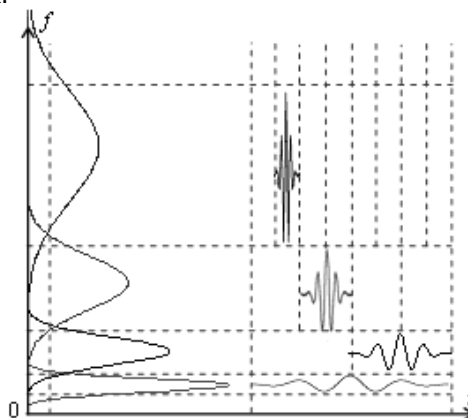


Рис. 1. Фрагмент частотно-временного разбиения пространства в базисе вейвлета Гаусса 5-го порядка

Если положить, что:

$$\hat{v}(\omega) = \hat{\psi}(\omega + \omega^*), \quad (5)$$

то функцию v также можно рассматривать в качестве функции-окна с центром в нуле и радиусом, равным $\Delta_{\hat{\psi}}$. Тогда ИВП (3) примет вид:

$$(W_\psi z)(b, a) = \frac{a}{2\pi\sqrt{|a|}} \int_{-\infty}^{\infty} \hat{z}(\omega) e^{jb\omega} \hat{v}\left(a\left(\omega - \frac{\omega^*}{a}\right)\right) d\omega. \quad (6)$$

Анализ выражения (6) показывает, что его результат $(W_\psi z)(b, a)$ с точностью до множителя $a/2\pi\sqrt{|a|}$ и линейного сдвига по фазе $e^{jb\omega}$, определенного значением параметра b , совпадает с

величиной (3) и дает локализованную информацию о спектре $\hat{z}(\omega)$ сигнала $z(t)$ с «частотным окном»:

$$\left[\frac{\omega^*}{a} - \frac{\Delta\hat{\psi}}{a}, \quad \frac{\omega^*}{a} + \frac{\Delta\hat{\psi}}{a} \right]. \quad (7)$$

Центр локализации ω^*/a и его значимую ширину $2\Delta\hat{\psi}/a$ называют параметрами частотной локализации сигнала (в общем случае функции) в «частотном окне».

Учитывая факт, что выражения (3) и (6) определяют одну и ту же функцию $(W_\psi z)(b, a)$, можно перейти к понятию «частотно-временного окна»:

$$[b + at^* - a\Delta\psi, \quad b + at^* + a\Delta\psi] \times \left[\frac{\omega^*}{a} - \frac{\Delta\hat{\psi}}{a}, \quad \frac{\omega^*}{a} + \frac{\Delta\hat{\psi}}{a} \right], \quad (8)$$

для совместного частотно-временного анализа на основе ИВП. Поскольку с математической точки зрения функция $u(x)$ определена как для положительных, так и отрицательных значений x , то в приложении к практической обработке сигналов базисный (или материнский) вейвлет ψ следует выбирать таким образом, чтобы центр ω^* функции $\hat{\psi}$ был положительным числом, а, с учетом параметра a , значение ω^*/a являлось частотным центром $\left[\frac{\omega^*}{a} - \frac{\Delta\hat{\psi}}{a}, \frac{\omega^*}{a} + \frac{\Delta\hat{\psi}}{a} \right]$ анализируемого диапазона частот.

Тогда отношение частотного центра к ширине диапазона частот равняется:

$$\frac{\omega^*/a}{2\Delta\hat{\psi}/a} = \frac{\omega^*}{2\Delta\hat{\psi}} = Q. \quad (9)$$

Величина (9) называется Q – постоянной частотного анализа [27, 32].

Необходимо отметить важное свойство частотно-временного окна (8). Оно сужается вдоль оси времени при больших значениях частотных центров ω^*/a и расширяется при их малых значениях, в то время как общая площадь частотно-временного окна остается постоянной и равной $4\Delta\psi\Delta\hat{\psi}$ [27] (см. рисунок 2).

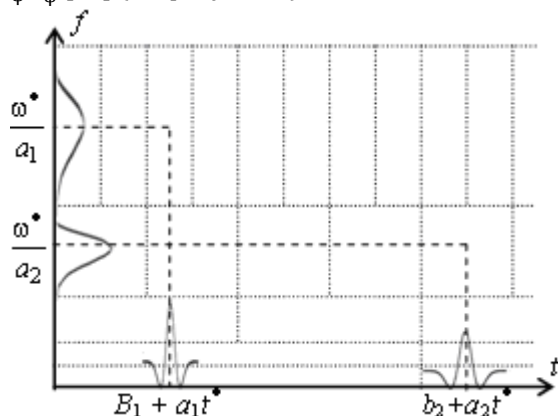


Рис. 2. Фрагменты частотно-временных окон интегрального вейвлет-преобразования

Для доказательства условий, характеризующих частотно-временное разрешение частотно-временного плана ИВП, рассмотрим следующие леммы.

Лемма 1.

Если существует производная u' и $u' \in L^1(\mathbb{R})$, то $\hat{u}'(\omega) = j\omega\hat{u}(\omega)$.

Доказательство.

Для доказательства леммы 1 в ИВП преобразуем правую часть выражения, воспользовавшись формулой интегрирования по частям $\int \vartheta dv = \vartheta v - \int v d\vartheta$:

$$\begin{aligned} \int_{-\infty}^{\infty} u(x) e^{-j\omega x} dx &= \\ &= \left| \begin{array}{l} f(x) = \vartheta \\ e^{-j\omega x} dx = dv \end{array} \right. \begin{array}{l} d\vartheta = u'(x) dx \\ v = -(e^{-j\omega x})/(j\omega) \end{array} \Big|_{-\infty}^{+\infty} = \\ &= -\frac{e^{-j\omega x}}{j\omega} u(x) \Big|_{-\infty}^{+\infty} + \frac{1}{j\omega} \int_{-\infty}^{\infty} e^{-j\omega x} u'(x) dx = \\ &= \frac{1}{j\omega} \int_{-\infty}^{\infty} e^{-j\omega x} u'(x) dx. \end{aligned}$$

При преобразовании учитывалось, что:

$$\int_{-\infty}^{\infty} e^{-j\omega x} dx = -\frac{1}{j\omega} \int_{-\infty}^{\infty} e^{-j\omega x} d(-j\omega x) = -\frac{e^{-j\omega x}}{j\omega}$$

и $u(x) \rightarrow 0$ при $x \rightarrow \pm\infty$. Тогда с учетом преобразования Фурье для $\hat{u}'(\omega) = \int_{-\infty}^{\infty} e^{-j\omega x} u'(x) dx$ окончательно получим $\hat{u}'(\omega) = j\omega\hat{u}(\omega)$.

Лемма 2.

Если функции $u \in L^1(\mathbb{R})$ и $h \in L^1(\mathbb{R})$, то всегда существует функция $g \in L^1(\mathbb{R})$, такая что $g = \langle h, u \rangle$, причем $\hat{g} = \langle \hat{h}, \hat{u} \rangle$.

Доказательство.

$$\hat{g}(\omega) = \int_{-\infty}^{\infty} e^{-j\omega x} \left(\int_{-\infty}^{\infty} u(x-v) h(v) dv \right) dx.$$

Так как произведение $|u(x-v)||h(v)|$ интегрируемо в пространстве $\mathbb{R}^2$, то возможно применение теоремы Фубини [27]. Тогда замена переменных $(x, v) \rightarrow (v = x - v, v)$ дает:

$$\begin{aligned} \hat{g}(\omega) &= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \exp[-j(v+u)\omega] u(v) h(v) dv dv = \\ &= \left(\int_{-\infty}^{\infty} \exp(-jv\omega) u(v) dv \right) \left(\int_{-\infty}^{\infty} \exp(-jv\omega) h(v) dv \right), \end{aligned}$$

что подтверждает правомерность результата $\hat{g} = \langle \hat{h}, \hat{u} \rangle$ [35].

Лемму 2 характеризует условие выполнения свертки в частотной и временной области.

Лемма 3.

Если существуют функции u и h , принадлежащие $L^1(\mathbf{R}) \cap L^2(\mathbf{R})$, то при их формальном равенстве $u = h$ следует, что:

$$\int_{-\infty}^{\infty} |u(x)|^2 dx = \frac{1}{2\pi} \int_{-\infty}^{\infty} |\hat{u}(\omega)|^2 d\omega.$$

Доказательство.

Пусть $g(x) = h^*(-x)$. Тогда, в соответствии с леммой 2 о свертке $\hat{g}(\omega) = \hat{u}(\omega)\hat{h}^*(\omega)$, формула восстановления $u(x) = \int_{-\infty}^{\infty} \hat{u}(\omega) e^{j\omega x} d\omega$ для функции при ее значении $u(0)$ дает значение:

$$\begin{aligned} \int_{-\infty}^{\infty} u(x)h^*(x) dx &= g(0) = \frac{1}{2\pi} \int_{-\infty}^{\infty} \hat{u}(\omega)d\omega = \\ &= \frac{1}{2\pi} \int_{-\infty}^{\infty} \hat{u}(\omega)\hat{h}^*(\omega)d\omega. \end{aligned}$$

Тогда при выполнении равенства $u = h$ имеем:

$$\int_{-\infty}^{\infty} |u(x)|^2 dx = \frac{1}{2\pi} \int_{-\infty}^{\infty} |\hat{u}(\omega)|^2 d\omega. \quad (10)$$

Заметим, что формула (10) согласуется с равенством Планшереля [35].

Рассмотренные леммы дают основание для построения функции с компактной локализацией энергии как во временной области, для u , так и в частотной – для ее преобразования Фурье $\hat{u}$. Регулировать частотно-временные размеры функции, т. е. масштабировать ее возможно при помощи параметра a . Если учесть, что $a < 1$, то тогда для функции $u_a(t) = a^{-1/2}u(t/a)$ ее преобразование Фурье $\hat{u}_a(\omega) = \sqrt{a}\hat{u}(a\omega)$ расширяется в $1/a$ раз, т. е., теряя в частотной локализации, получаем аналогичный выигрыш во временной. Сам частотно-временной разброс концентрации энергии ограничен принципом неопределенности Гейзенберга [36], когда плотность вероятности локализации в момент t будет определяться как $\frac{1}{\|u\|^2}|u(t)|^2$, а для частотной области в момент ω , соответственно – $\frac{1}{2\pi\|u\|^2}|\hat{u}(\omega)|^2$.

Тогда среднее положение по времени может быть найдено как:

$$\nu = \frac{1}{\|u\|^2} \int_{-\infty}^{\infty} t |u(t)|^2 dt, \quad (11)$$

а по частоте:

$$\xi = \frac{1}{2\pi\|u\|^2} \int_{-\infty}^{\infty} \omega |\hat{u}(\omega)|^2 d\omega. \quad (12)$$

Очевидно, что дисперсии разброса во времени и по частоте будут равны:

$$\sigma_t^2 = \frac{1}{\|u\|^2} \int_{-\infty}^{\infty} (t - \nu)^2 |u(t)|^2 dt, \quad (13)$$

$$\sigma_\omega^2 = \frac{1}{2\pi\|u\|^2} \int_{-\infty}^{\infty} (\omega - \xi)^2 |\hat{u}(\omega)|^2 d\omega. \quad (14)$$

Для вейвлет-функции эти значения определяются как $\sigma_t^2 = \Delta_\Psi$, $\sigma_\omega^2 = \Delta_\Phi$.

Теорема 1.

Временная дисперсия σ_t^2 и частотная дисперсия σ_ω^2 функции $u \in L^2(\mathbf{R})$ удовлетворяет неравенству $\sigma_t^2 \sigma_\omega^2 \geq 1/4$. Неравенство переходит в равенство тогда и только тогда, когда существуют $(\nu, \xi, a, b) \in \mathbf{R}^2 \times \mathbf{C}^2$ такие, что:

$$u(t) = a \exp[j\xi t - b(t - \nu)^2].$$

Доказательство.

Предположим, что $\lim_{|t| \rightarrow \infty} \sqrt{t}u(t) = 0$. Тогда, если $\|u\| = 1$, а средняя временная и частотная локализации u соответственно равняются ν и ξ , то среднее временное и частотное положение $\exp(-j\xi t) u(t + \nu)$ будет равно нулю. Следовательно, достаточно доказать выполнение условия, что $\nu = \xi = 0$.

При этом следует отметить, что:

$$\sigma_t^2 \sigma_\omega^2 = \frac{1}{2\pi\|u\|^4} \int_{-\infty}^{\infty} |tu(t)|^2 dt \int_{-\infty}^{\infty} |\omega \hat{u}(\omega)|^2 d\omega.$$

Так как $j\omega \hat{u}(\omega)$ есть преобразование Фурье от $u'(t)$ (согласно лемме 3), то равенство (10), примененное к $j\omega \hat{u}(\omega)$ (ко второму сомножителю правой части равенства), дает:

$$\sigma_t^2 \sigma_\omega^2 = \frac{1}{\|u\|^4} \int_{-\infty}^{\infty} |tu(t)|^2 dt.$$

Применим к полученному выражению неравенство Шварца [35]:

$$\begin{aligned} \sigma_t^2 \sigma_\omega^2 &\geq \frac{1}{\|u\|^4} \left[\int_{-\infty}^{\infty} |tu'(t)|^2 |u^*(t)| dt \right]^2, \\ \sigma_t^2 \sigma_\omega^2 &\geq \frac{1}{\|u\|^4} \left[\int_{-\infty}^{\infty} t/2 [u'(t)u^*(t) + u'^*(t)u(t)] dt \right]^2, \\ \sigma_t^2 \sigma_\omega^2 &\geq \frac{1}{4\|u\|^4} \left[\int_{-\infty}^{\infty} |u(t)|^2 dt \right]^2 = \frac{1}{4}. \end{aligned}$$

Для получения равенства:

$$\sigma_t^2 \sigma_\omega^2 \geq 1/4 \quad (15)$$

неравенство Шварца, примененное к:

$$\sigma_t^2 \sigma_\omega^2 \geq \frac{1}{\|u\|^4} \left[\int_{-\infty}^{\infty} |tu'(t)|^2 |u^*(t)| dt \right]^2,$$

должно перейти в равенство. Это означает, что существует $b \in \mathbb{C}$ такое, что:

$$u'(t) = -2btu(t). \quad (16)$$

Следовательно, существует $a \in \mathbb{C}$ такое, что $u(t) = a \exp(-bt^2)$.

Если $v \neq 0$ и $\xi \neq 0$, то соответствующие сдвиги по времени и по частоте приводят к:

$$u(t) = a \exp[j\xi t - b(t - v)^2]. \quad (17)$$

По своей сути ИВП является процедурой аффинного представления исследуемого процесса в кратномасштабном базисе вейвлет-функций, поскольку каждое окно формируется в результате линейного изменения параметров a и b .

Физический смысл рассмотренных операций заключается в последовательной корреляции исследуемого процесса $z(t)$ с аффинными копиями базисного вейвлета ψ .

Таким образом, результат (3) и (6) $(W_\psi z)(b, a)$ определяет расположение (в терминах $b + at^*$), «размер» (в терминах параметра a) и область изменения $z(t)$ с учетом способностей сжатия (растяжения) и сдвига. Следовательно, в своей совокупности информация, содержащаяся в ИВП $(W_\psi z)(b, a)$, полностью характеризует исследуемый процесс, а значит, на ее основе можно восстановить (репродуцировать) исходное значение $z(t)$.

Однако невозможность обратного репродуцирования говорит о неоднозначности разложения. Примером неоднозначного представления является корреляция и процедура одномерного преобразования Фурье [27].

В общем случае вейвлет-базис $\{\psi_{i,k}\}$ является двойственным базисом разложения, т. е. он позволяет осуществлять обратную репродукцию [37]. Однако, строго придерживаясь математических позиций, возможны четыре ситуации восстановления, определяющие ограничения на область определения на $(W_\psi u)$.

Первое условие: восстановление по $(W_\psi u)(b, a)$, $b, a \in \mathbb{R}$.

Одним из условий данного восстановления является наличие константы:

$$C_\psi = \int_{-\infty}^{\infty} \frac{|\hat{\psi}(\omega)|^2}{|\omega|} d\omega < \infty. \quad (18)$$

Конечность константы (18) сужает класс функций ψ из $L^2(\mathbb{R})$, которые могут использоваться в качестве базисных в ИВП.

Так, если ψ – функция-окно, то она принадлежит пространству $L^1(\mathbb{R})$, т. е.:

$$\int_{-\infty}^{\infty} |\psi(x)| dx < \infty, \quad (19)$$

но в этом случае $\hat{\psi}$ должна быть непрерывной функцией. Докажем это утверждение.

Теорема 2.

Пусть $u \in L^1(\mathbb{R})$. Тогда ее преобразование Фурье $\hat{u}$ обладает следующими свойствами:

- (а) $\hat{u} \in L^\infty(\mathbb{R})$, где $\|\hat{u}\|_\infty \leq \|u\|_1$;
- (б) $\hat{u}$ равномерна и непрерывна на $\mathbb{R}$;
- (в) $u'(\omega) \rightarrow 0$ при $\omega \rightarrow \pm\infty$.

Доказательство.

В условии (а) рассматривается норма значения $\hat{u}$. Учитывая что $L^1(\mathbb{R}) \subset L^\infty(\mathbb{R})$, то любое значение в $L^1(\mathbb{R})$ никогда не превысит его величины в $L^\infty(\mathbb{R})$. Следовательно, при соответствующем нормировании на указанные величины всегда будет выполняться неравенство $\|\hat{u}\|_\infty \leq \|u\|_1$.

Для доказательства (б) возьмем произвольное δ , и рассмотрим:

$$\begin{aligned} & \sup_{\omega} |\hat{u}(\omega + \delta) - \hat{u}(\omega)| = \\ &= \sup_{\omega} \left| \int_{-\infty}^{\infty} e^{-j\omega x} (e^{-j\delta x} - 1) u(x) dx \right| \leq \\ &\leq \int_{-\infty}^{\infty} |\exp(-j\delta x) - 1| \times |u(x)| dx. \end{aligned}$$

Так как:

$|\exp(-j\delta x) - 1| \times |u(x)| \leq 2|u(x)| \in \exp(-j\delta x) - 1 \rightarrow 0$, при $\delta \rightarrow 0$, то по теореме Лебега о переходе к пределу под знаком интеграла имеем, что величина, стоящая справа, стремится к нулю при $\delta \rightarrow 0$.

Утверждение (в) называют леммой Римана–Лебега. При ее доказательстве заметим, что если u' существует и принадлежит $L^1(\mathbb{R})$, то из леммы 2 и утверждения (а) имеем при $\omega \rightarrow \pm\infty$:

$$|\hat{u}(\omega)| = \frac{1}{(\omega)} |\hat{u}'(\omega)| \leq \frac{1}{(\omega)} \|u'\|_1 \rightarrow 0.$$

В общем случае для любого $\varepsilon > 0$ можно найти функцию g такую, что $g, g' \in L^1(\mathbb{R})$ и $\|u - g\|_1 < \varepsilon$. Тогда из утверждения (а) имеем:

$$\begin{aligned} |\hat{u}(\omega)| &\leq |\hat{u}(\omega) - \hat{g}(\omega)| + |\hat{g}(\omega)| \leq \\ &\leq \|u - g\|_1 + |\hat{g}(\omega)| < \varepsilon + |\hat{g}(\omega)|. \end{aligned}$$

Заметим, хотя $\hat{u}(\omega) \rightarrow 0$ при $\omega \rightarrow \pm\infty$ выполняется для любой $u \in L^1(\mathbb{R})$, это не означает, что и $\hat{u}$ обязательно принадлежит $L^1(\mathbb{R})$. Примером является функция Хевисайда:

$$u_H(x) = \begin{cases} 1 & \text{для } x \geq a \\ 0 & \text{для } x < a \end{cases} \quad (20)$$

В $u(x) = e^{-x} u_H(x)$, которая принадлежит $L^1(\mathbb{R})$, в то время как ее преобразование Фурье $\hat{u}(\omega) = \frac{1}{1+j\omega}$ пространству $L^1(\mathbb{R})$ не принадлежит.

Из (18) следует, что функция $\hat{u}$ обращается в нуль в начале координат, другими словами:

$$\int_{-\infty}^{\infty} \Psi(x) dx = 0. \quad (21)$$

Следовательно, базисный вейвлет должен представлять гладкую функцию с высокой временной локализацией, тогда формула реконструкции исходного процесса из ИВП будет иметь вид [14]:

$$u(x) = \frac{1}{C_{\Psi}} \int_{\mathbf{R}^2} \int \left\{ (W_{\Psi} u)(b, a) \right\} \times \left\{ |a|^{-1/2} \Psi\left(\frac{x-b}{a}\right) \right\} \frac{dadb}{a^2}, \quad (22)$$

$$u \in L^2(\mathbf{R}).$$

Здесь $\mathbf{R}^2 = \mathbf{R} \times \mathbf{R}$.

Следует заметить, что в выражении (22) ядро $|a|^{-1/2} \Psi\left(\frac{x-b}{a}\right)$ с точностью до знака комплексного сопряжения используется в ИВП (3) и (6), что говорит о двойственности базисного вейвлета Ψ .

Второе условие: восстановление процесса (сигнала) по его ИВП $(W_{\Psi} u)(b, a)$, $b \in \mathbf{R}$, $a > 0$.

В практических реализациях, когда в качестве анализируемой функции выступает реальный сигнал, область интегрирования при реставрации ограничивается до $\mathbf{R} \times (0, \infty)$. В этом случае для базисных вейвлетов достаточно соответствия условию:

$$\int_0^{\infty} \frac{|\hat{\Psi}(\omega)|^2}{|\omega|} d\omega = \int_0^{\infty} \frac{|\hat{\Psi}(-\omega)|^2}{|\omega|} d\omega = \frac{C_{\Psi}}{2} < \infty. \quad (23)$$

Тогда для любой Ψ , удовлетворяющей (23), формула реконструкции имеет вид:

$$u(x) = \frac{2}{C_{\Psi}} \int_0^{\infty} \left[\int \left\{ (W_{\Psi} u)(b, a) \right\} \times \left\{ |a|^{-1/2} \Psi\left(\frac{x-b}{a}\right) \right\} db \right] \frac{da}{a^2}, \quad (24)$$

$$u \in L^2(\mathbf{R}).$$

Близость выражений (24) и (22) позволяет предположить, что и в этом случае отсутствует свойство единственности двойственного. Однако здесь в большей степени интересным видится выполнение самого условия реконструкции. Докажем правомерность данного утверждения.

Теорема 3.

Пусть $\Psi \in L^2(\mathbf{R})$ – вещественная функция, такая что:

$$C_{\Psi} = \int_{-\infty}^{\infty} \frac{|\hat{\Psi}(\omega)|^2}{|\omega|} d\omega < \infty.$$

Тогда любая функция $u \in L^2(\mathbf{R})$ удовлетворяет равенствам:

$$u(x) = \frac{1}{C_{\Psi}} \int_0^{\infty} \int (W_{\Psi} u)(b, a) \left\{ |a|^{-1/2} \Psi\left(\frac{x-b}{a}\right) \right\} \frac{dadb}{a^2}$$

и

$$\int_{-\infty}^{\infty} |u(x)|^2 dx = \frac{1}{C_{\Psi}} \int_0^{\infty} \int [(W_{\Psi} u)(b, a)]^2 \frac{dadb}{a^2}.$$

Доказательство.

Представим через $u(x)$ – интеграл справа выражения:

$$u(x) = \frac{1}{C_{\Psi}} \int_0^{\infty} \int (W_{\Psi} u)(b, a) \left\{ |a|^{-1/2} \Psi\left(\frac{x-b}{a}\right) \right\} \frac{dadb}{a^2}$$

как интеграл от сверток. Тогда подстановка $(W_{\Psi} u)(b, a) = \langle u, \Psi_a^*(b) \rangle$, где $\Psi_a(b) = a^{-1/2} \Psi(x/a)$, дает:

$$(x) = \frac{1}{C_{\Psi}} \int_0^{\infty} \langle (W_{\Psi} u)(b, a), \Psi_a(x) \rangle \frac{da}{a^2} =$$

$$= \frac{1}{C_{\Psi}} \int_0^{\infty} \langle u, \Psi_a^*(b) \rangle \langle \Psi_a(x) \rangle \frac{da}{a^2}.$$

В то же время преобразование Фурье от $u(x)$ есть:

$$\hat{u}(\omega) = \frac{1}{C_{\Psi}} \int_0^{\infty} \hat{u}(\omega) \sqrt{a} \hat{\Psi}^*(a\omega) \sqrt{a} \hat{\Psi}^*(a\omega) \frac{da}{a^2} =$$

$$= \frac{\hat{u}(\omega)}{C_{\Psi}} \int_0^{\infty} |\hat{\Psi}(a\omega)|^2 \frac{da}{a}.$$

Так как Ψ – вещественная функция, то $|\hat{\Psi}(-\omega)|^2 = |\hat{\Psi}(\omega)|^2$.

Следовательно, сделав замену переменных $\xi = a\omega$, получаем, что:

$$\hat{u}(\omega) = \frac{\hat{u}(\omega)}{C_{\Psi}} \int_0^{\infty} \frac{|\hat{\Psi}(\xi)|^2}{\xi} d\xi = \hat{u}(\omega).$$

Таким образом, если равны преобразования Фурье функций, то, следовательно, равны и их временные образы.

Для доказательства справедливости условия:

$$\int_{-\infty}^{\infty} |u(x)|^2 dx = \frac{1}{C_{\Psi}} \int_0^{\infty} \int [(W_{\Psi} u)(b, a)]^2 \frac{dadb}{a^2},$$

воспользуемся следующим фактом, что преобразование Фурье по b от $(W_{\Psi} u)(b, a)$ есть $\hat{u}(\omega + b) \hat{\Psi}(\omega)$. Тогда в соответствии с леммой 3 формула (10), примененная к правой части, дает:

$$\frac{1}{C_{\Psi}} \int_0^{\infty} \int [(W_{\Psi} u)(b, a)]^2 \frac{dadb}{a^2} =$$

$$= \frac{1}{C_{\Psi}} \int_0^{\infty} \frac{1}{C_{\Psi}} \int |\hat{u}(\omega + b) \hat{\Psi}(\omega)|^2 d\omega db.$$

Из применения теоремы Фубини [35] следует:

$$\frac{1}{C_\psi} \int_0^\infty |\hat{u}(\omega + b)|^2 db = \|u\|^2,$$

отсюда вытекает равенство:

$$\int_{-\infty}^\infty |u(x)|^2 dx = \frac{1}{C_\psi} \int_0^\infty \int_{-\infty}^\infty [(W_\psi u)(b, a)]^2 \times \frac{dad b}{a^2}. \quad (25)$$

Третье условие: восстановление по $(W_\psi u)(b, a)$, $b \in \mathbf{R}$, $a = \frac{1}{2^i}$, $i \in \mathbf{Z}$.

Указанное восстановление рассматривает частотно-временную локализацию с частотными окнами:

$$H_i = [2^i \omega^* - 2^i \Delta_{\hat{\psi}}, \quad 2^i \omega^* + 2^i \Delta_{\hat{\psi}}], \quad (26)$$

при том, что i принимает значения целых чисел, т. е., $i \in \mathbf{Z}$.

Так как ИВП (3) используется для определения временных интервалов $[b + 2^{-i} t^* - 2^{-i} \Delta_\psi, b + 2^{-i} t^* + 2^{-i} \Delta_\psi]$, на которых спектральное содержание сигнала u с частотами из H_i лежит выше некоторого порога $|(W_\psi u)(b, a)|$. В рассматриваемых условиях доступна только часть информации об $W_\psi u$, следовательно, для реализации операции восстановления на базисный вейвлет должны налагаться более жесткие условия:

$$A \leq \sum_{i=-\infty}^\infty |\hat{\psi}(2^{-i} \omega)|^2 \leq B, \quad (27)$$

где A и B – константы, не зависящие от ω , причем $0 < A \leq B < \infty$.

Согласно (27), базисный вейвлет Ψ в этом случае удовлетворяет неравенствам:

$$A \ln 2 \leq \int_0^\infty \frac{|\hat{\psi}(\omega)|^2}{\omega} d\omega, \quad (28)$$

$$\int_0^\infty \frac{|\hat{\psi}(-\omega)|^2}{\omega} d\omega \leq B \ln 2,$$

обозначая, что C_ψ лежит между величинами $A \ln 2$ и $B \ln 2$.

Выполнение условия (27) для базисного вейвлета ψ , говорит о наличии у него двойственного центра ψ^* , преобразование Фурье которого задается формулой:

$$\hat{\psi}^*(\omega) = \frac{\overline{\hat{\psi}(\omega)}}{\sum_{i=-\infty}^\infty |\hat{\psi}(2^{-i} \omega)|}. \quad (29)$$

Теперь формула восстановления, использующая двойственное преобразование (29), определяется следующим образом:

$$u(x) = \sum_{i=-\infty}^\infty \int_{-\infty}^\infty \left\{ 2^{\frac{i}{2}} (W_\psi u)(b, 2^{-i}) \right\} \times \{ 2^i \psi^*[2^i(x-b)] \} db, \quad (30)$$

$$u \in L^2(\mathbf{R}).$$

Базисные функции, удовлетворяющие условию (27), называются двухпараметрическими вейвлетами [32].

Четвертое условие: восстановление по $(W_\psi u)(b, a)$, $b = k2^{-i}$, $a = 2^{-i}$, $k, i \in \mathbf{Z}$.

Данное условие ориентированно на дискретные выборки, когда частотная ось разбивается на диапазоны, использующие степени 2 для масштабного параметра a .

Дальнейшее повышение эффективности вычислительных процедур связано с двухпараметрическими значениями $b = k2^{-i}$ на временной оси при $a = 2^{-i}$, $i \in \mathbf{Z}$ вместо произвольных $b \in \mathbf{R}$.

Для данного случая формула ИВП имеет удобный с точки зрения построения на ее основе вычислительных алгоритмов вид:

$$(W_\psi u)(k2^{-i}, 2^{-i}) = \int_{-\infty}^\infty u(x) \overline{\{ 2^{i/2} \psi(2^i x - b) \}} dx = \langle u, \psi_{i,k} \rangle, \quad (31)$$

где $\psi_{i,k}(x) = 2^{i/2} \psi(2^i x - b)$ при $k, i \in \mathbf{Z}$.

Выполнение указанных условий для ψ позволяет рассматривать базисные вейвлет-функции в качестве R -функций, представляющие базис Рисса в $L^2(\mathbf{R})$ [17], т. е. линейная оболочка $\psi_{i,k}$ при $k, i \in \mathbf{Z}$ достаточно плотна в $L^2(\mathbf{R})$ так, что существуют положительные константы A и B , $0 < A \leq B < \infty$, такие, что выполняется условие:

$$A \|\{c_{i,k}\}\|^2 \leq \left\| \sum_{i=-\infty}^\infty \sum_{k=-\infty}^\infty c_{i,k} \psi_{i,k} \right\|_2^2 \leq B \|\{c_{i,k}\}\|^2, \quad (32)$$

для всех бесконечных суммируемых с квадратом последовательностей $\{c_{i,k}\}$:

$$\|\{c_{i,k}\}\|^2 = \sum_{i=-\infty}^\infty \sum_{k=-\infty}^\infty |c_{i,k}|^2 < \infty. \quad (33)$$

Принадлежность ψ к классу R -функций обусловливает единственность базиса Рисса $\{\psi^{i,k}\}$, а следовательно, единственность разложения любой функции $u \in L^2(\mathbf{R})$ в вейвлет-ряд:

$$u(x) = \sum_{i,k=-\infty}^\infty \langle u, \psi_{i,k} \rangle \psi^{i,k}(x). \quad (34)$$

Для существования ряда (34) должна существовать некоторая функция $\tilde{\psi} \in L^2(\mathbf{R})$, определяющая двойственный базис как:

$$\psi^{i,k} = \tilde{\psi}_{i,k}(x), \quad (35)$$

где $\tilde{\psi}_{i,k}$ определяется $\tilde{\psi}_{i,k}(x) = 2^{i/2} \tilde{\psi}(2^i x - k)$.

Рассмотренный подход позволяет более строго определить вейвлет-функцию следующим образом [38]. Так, R -функция $\psi \in L^2(\mathbf{R})$ называется R -вейвлетом (или вейвлетом), если существует функция $\tilde{\psi} \in L^2(\mathbf{R})$ такая, что $\{\psi_{i,k}\}$ и $\{\tilde{\psi}_{i,k}\}$ являются двойственными базисами в $L^2(\mathbf{R})$. Если ψ – R -вейвлет, то $\tilde{\psi}$ называют двойственным вейвлетом, соответствующим ψ .

Признаком «двойственности» вейвлета является его ортогональность. Однако самая важная особенность двойственного вейвлета состоит в симметричности пары $(\psi, \tilde{\psi})$, т. е. ψ является двойственным для $\tilde{\psi}$, и, как следствие, единственности разложения любой функции $u \in L^2(\mathbf{R})$ в вейвлет-ряд, где каждый коэффициент $c_{i,k}$ есть ИВП u относительно $\tilde{\psi}$ двойственного ψ , вычисленное в точках временной шкалы:

$$(b, a) = (k2^{-i}, 2^{-i}). \quad (36)$$

В соответствии с данным определением каждый вейвлет ψ в порожденном им базисе Рисса $\{\psi_{i,k}\}$ образует так называемое «замыкание» линейной оболочки $\{\psi_{i,k}: k \in \mathbf{Z}\}$ на каждом уровне декомпозиции W_i при $i \in \mathbf{Z}$:

$$W_i = \text{clos}_{L^2(\mathbf{R})}(\psi_{i,k}: k \in \mathbf{Z}). \quad (37)$$

Следовательно, пространство $L^2(\mathbf{R})$ может быть разложено в прямую сумму подпространств W_i :

$$L^2(\mathbf{R}) = \sum_{i \in \mathbf{Z}} W_i, \quad (38)$$

причем для каждого из базисных вейвлетов разложение (23) будет единственным (в (38) берутся прямые суммы).

Очевидно, что при условии ортогональности ψ подпространства W_i также будут ортогональны, и сумма (38) становится ортогональной суммой:

$$L^2(\mathbf{R}) = \bigoplus_{i \in \mathbf{Z}} W_i = \dots \oplus W_{-1} \oplus W_0 \oplus W_1 \oplus \dots, \quad (39)$$

а $W_l \perp W_m$, где $l, m \in \mathbf{Z}$, причем $l \neq m$.

В (39) окружность вокруг знака плюс $\bigoplus$ означает ортогональную сумму. Разложение (39) называют ортогональным разложением пространства $L^2(\mathbf{R})$ [35].

Вместе с тем разложение (39) не использует все ортогональные свойства $\{\psi_{i,k}\}$, а именно то, что для каждого i $\langle \psi_{i,k}, \psi_{i,l} \rangle = \delta_{k,l}$. Это означает, что существует широкий класс вейвлетов, которые могут использоваться для порождения ортогональных порождений $L^2(\mathbf{R})$. В частности, к такому следует отнести полуортогональные вейвлеты, позволяющие конструировать ортогональные базисы.

Вейвлет ψ в $L^2(\mathbf{R})$ называется полуортогональным вейвлетом [27], если порожденный им базис Рисса $\{\psi_{i,k}\}$ удовлетворяет равенству:

$$\begin{aligned} \langle \psi_{i,k}, \psi_{l,m} \rangle &= 0, \\ i &\neq l; \\ i, k, l, m &\in \mathbf{Z}. \end{aligned} \quad (40)$$

Особенностью полуортогональных вейвлетов является возможность порождения разложения $L^2(\mathbf{R})$ в виде прямой суммы ($+$ является знаком прямой суммы) замкнутых подпространств (39), так что для каждого $i \in \mathbf{Z}$ будут существовать замкнутые подпространства вида:

$$V_i = \dots + W_{i-2} + W_{i-1}. \quad (41)$$

Таким образом, обобщая полученные результаты, можно определить следующие важные свойства вейвлет-функций, адаптированные к решению задач синтеза масштабно-временных распределений.

Свойство 1. Локализация базисной функции во временном и спектральном пространстве:

$$\Delta FT = \text{const}. \quad (42)$$

Свойство 2. Наличие нулевого среднего:

$$\int_{-\infty}^{\infty} \psi(t) dt = 0, \quad (43)$$

более жесткие требования – все первые m моментов были равны нулю:

$$\int_{-\infty}^{\infty} t^m \psi(t) dt = 0 \quad (44)$$

(вейвлет-функция, удовлетворяющая условию (44), называется вейвлетом m -го порядка).

Свойство 3. Ограниченность энергии:

$$\int_{-\infty}^{\infty} |\psi(t)|^2 dt < \infty. \quad (45)$$

Оценка локализации и ограниченности представляется следующим образом:

$$|\psi(t)| < \frac{1}{1 + |t|^m} \quad (46)$$

или

$$|\hat{\psi}(\omega)| < \frac{1}{1 + |k - \omega_0|^m}, \quad (47)$$

где ω_0 – доминантная частота вейвлета, при этом k – по возможности большое число.

Данное свойство гарантирует выполнение процедур обратного вейвлет-преобразования для ретрансляции сигнала конечной длительности.

Свойство 4. Сохранение самоподобия для всех вейвлетов, определяющих базис реконструкции. В частности, все вейвлеты одного базиса представ-

ляют масштабные копии друг друга, что допускает использование любого из них в качестве порождающего, т.е. материнского [38].

Формирование кратномасштабных пространств на основе интегрального вейвлет-преобразования

Для практических приложений процедур вейвлет-преобразований к обработке реальных сигналов целесообразно параметр масштабирования a рассматривать, исходя из следующей интерпретации. Поскольку исходный материнский вейвлет локализуется вокруг некоторой ненулевой частоты f_0 при значении параметра масштаба $a = 1$, то процесс масштабирования есть не что иное, как переход к новому значению частоты f_a , вокруг которой теперь уже будет локализоваться базисная вейвлет-функция.

Тогда процедуру ИВП (3) можно представить как совокупность сверток анализируемого сигнала с вейвлетами, базисная частота которых определяется выражением:

$$f_a = a f_0. \quad (48)$$

В результате формируемое ИВП имеет приемлемое частотно-временное разрешение как на крупных, так и на мелких масштабах частоты. В ряде работ [32] процедуру (3) применительно к обработке сигналов называют непрерывным вейвлет-преобразованием (НВП), а ее результат – вейвлет-коэффициентами.

Поскольку в общем случае параметры a и b меняются непрерывно, то множество базисных функций, осуществляющих декомпозицию сигнала, избыточно.

Очевидно, здесь необходимо дискретизацию параметров a и b осуществить таким образом, чтобы при снижении избыточности описания сохранялась общность свойств НВП.

В [40] предложен следующий подход к выбору параметров НВП:

$$a = a_0^i; \quad b = k b_0 a_0^i, \quad i, k \in \mathbb{Z}, \quad a_0 > 1, b_0 \neq 0. \quad (49)$$

Поскольку параметр b_0 в принципе произвольный, то целесообразно положить $b_0 = 1$. Взаимосвязь параметров (49) позволяет сделать вывод, что с увеличением масштаба увеличивается и размер шага сдвига вдоль временной оси.

Рациональность данного подхода объясняется тем, что при анализе с большим масштабом мелкие детали не так уж важны, в то время как при малых масштабах значение малоразмерных деталей значительно возрастает. Тогда в соответствии с четвертым условием восстановления исходный вид базисного вейвлета с дискретными значениями параметров a и b можно представить в следующем виде:

$$\psi_{j,k}(t) = a_0^{-\frac{j}{2}} \psi(a_0^{-j} t - k). \quad (50)$$

Формула (50) является обобщающей по отношению к (31). Следует отметить: условие того, что множество $\{\psi_{i,k}(t)\}_{i,k \in \mathbb{Z}}$ образует ортонормированную систему, т.е. любая функция $u \in L^2(\mathbb{R})$, характеризующая сигнал, может быть представлена рядом в соответствии с (50):

$$u = \sum_{i,k \in \mathbb{Z}} \langle u, \psi_{i,k} \rangle \psi_{i,k}. \quad (51)$$

Аналитическое описание (51) по своим атрибутам близко к разложению в ряд Фурье по экспоненциальным функциям $e_k = e^{jk}$. Однако экспоненциальная функция имеет глобальный носитель, следовательно, все члены разложения Фурье в каждой точке пространства параметра x зависят от значения самой функции u . В противоположность гармоническим функциям, вейвлет имеет компактный носитель, поэтому только члены разложения ряда (51), соответствующие $\psi_{i,k}$ с $i, 2^{-k}$ в окрестности точки x , вносят наибольший вклад в функцию в этой окрестности.

Другой важной особенностью вейвлет-разложения является то, что коэффициенты каждого уровня декомпозиции содержат всю информацию о степени гладкости сигнала.

Дискретизированное преобразование на основе вейвлета (50) в [40] определяют как ряды вейвлетов непрерывного времени по аналогии с терминологией преобразования Фурье.

Представление сигналов на их основе реализуется путем соответствующей дискретизации НВП (3) с учетом (50):

$$W_{i,k} = \int_{-\infty}^{\infty} a_0^{-i/2} \psi(a_0^{-i} t - k) z(t) dt. \quad (52)$$

Репродукция исходного сигнала $z(t)$ из коэффициентов его декомпозиции (52) возможна только в рамках следующих ограничений:

$$A \|z(t)\|^2 \leq \sum_{i \in \mathbb{Z}} \sum_{k \in \mathbb{Z}} |W_{i,k}|^2 \leq B \|z(t)\|^2, \quad (53)$$

где $A > 0$, а $B < \infty$.

Выбор $A = B = 1$ и $a_0 = 2$ обеспечивает полную реконструкцию, поскольку тогда семейство вейвлетов $\psi_{a,b}(t)$ образует ортогональный базис [41]:

$$z(t) = \frac{1}{C_\psi} \sum_{i \in \mathbb{Z}} \sum_{k \in \mathbb{Z}} W_{i,k} a_0^{-i/2} \psi(a_0^{-i} t - k). \quad (54)$$

В общем случае восстановленный сигнал может и не совпадать с исходным, но будет близок к нему с точки зрения обеспечения минимума среднеквадратической ошибки аппроксимации. Рассмотренные условия декомпозиции и реконструкции

полностью вписываются в рамки четвертого условия восстановления ИВП.

Для дискретных параметров a и b , определяемых формулой (49), рациональным видится выбор $a_0 = 2$, а $b_0 = 1$, поскольку в этом случае алгоритмы расчета вейвлет-коэффициентов легко вписываются в концепцию цифровой обработки.

Сетка дискретизации, обеспечиваемая при $a_0 = 2$ и $b_0 = 1$, называется диадической, а формируемое на ее основе вейвлет-преобразование – диадным [42].

Так, на рисунке 3 представлен фрагмент многокомпонентного сигнала и его кратномасштабное представление на основе диадного вейвлет-преобразования (в [32] совместные масштабно-временные распределения на основе вейвлет-преобразований определены как масштабogramмы).

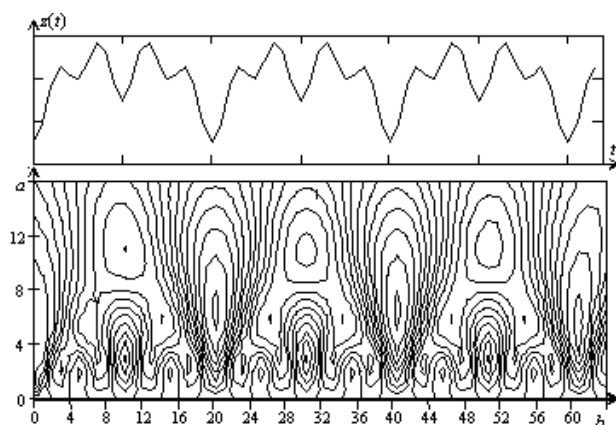


Рис. 3. Многокомпонентный сигнал и его масштабogramма на основе диадного вейвлет-преобразования

Адаптивный характер вейвлет-преобразования, позволяющий в равной степени подчеркнуть, как мелкие особенности анализируемого процесса, так и его крупные детали, позволяет предположить о наличии высоких помехоустойчивых свойств у масштабных представлений. На рисунке 4 представлены фрагменты: исходного тестового сигнала; зашумленного тестового сигнала; масштабogramма (модуль коэффициентов вейвлет-преобразования) зашумленного тестового сигнала.

Анализ результатов, представленных на рисунке 4, показывает, что основные искажения вейвлет-коэффициентов, вызванные воздействием шумов, сосредоточены в высокочастотной области масштабogramмы. Это объясняется низкой стационарностью шумовых компонент. А так как быстрые изменения в частотной области представляются широкополосными процессами, то они как раз и будут сосредотачиваться в части масштабogramмы, соответствующей области с высоким разрешением. Следовательно, если при обратной реконструкции не использовать вейвлет-коэффициенты, пораженные шумом, то восстановленный процесс будет свободен от шумов. Это

позволяет сделать вывод о помехоустойчивости совместных преобразований, построенных на основе вейвлетов.

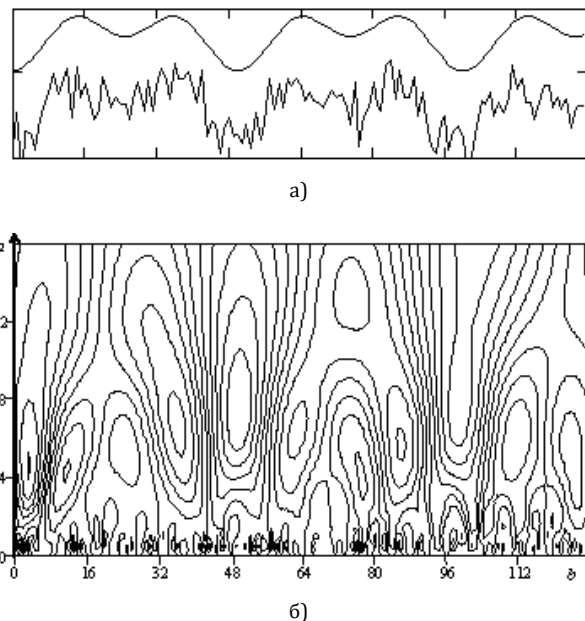


Рис. 4. Графическое представление: а) тестовый многокомпонентный сигнал без шумов и в аддитивных шумах; б) масштабogramма зашумленного тестового сигнала

Следует отметить, что рассмотренный выбор параметров для условия (49) не является единственным.

Так, масштабирование, кратное двум при условии непрерывности временного сдвига, ведет к синтезу вейвлет-фреймов [42].

Вейвлет-фреймы являются промежуточным преобразованием между интегральным и диадным представлением, поскольку непрерывность параметра сдвига ведет к избыточности описания при декомпозиции процессов.

Билинейные формы распределений аффинного класса

Анализ свойств различных форм совместных ЧВР класса Коэна выявил ряд проблем, возникающих при их непосредственном использовании в практических задачах обработки нестационарных процессов [23].

Поиск возможных путей их решений позволил выйти на так называемые гибридные распределения [14]. Данное направление, базирующееся на методологии масштабно-временного анализа, является наиболее продуктивным, поскольку гибридные распределения объединяют в себе помехоустойчивость вейвлет-преобразований и метрологические характеристики ЧВР класса Вигнера.

Успешное решение частных задач на основе гибридных распределений в [14] определяет прагматизм их практического применения.

В свое время Л. Козн установил, что все совместные ЧВР обладают характерным свойством инвариантности частотно-временному сдвигу (см. формулу 40 в [2]). Однако, рассмотрение с этих позиций реализации процесса вейвлет-преобразования позволяет сделать вывод, что в процедурах формирования вейвлет-коэффициентов данное свойство будет трансформировано [14].

Действительно, в классе Козна переход в частную область осуществлялся посредством процедур преобразований Фурье, реализуемых системой однотипных полосовых фильтров, сдвинутых относительно друг друга (рисунок 5а).

При формировании вейвлет-спектра имеем дело с системой фильтров, полосы пропускания которых являются масштабными версиями друг друга (рисунок 5б).

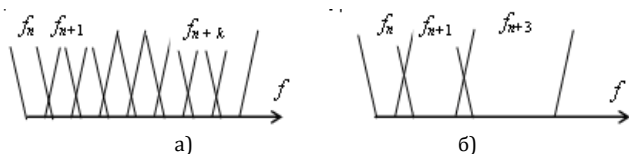


Рис. 5. Система фильтров, обеспечивающая формирование: а) распределений класса Козна; б) вейвлет-преобразований

Очевидно, что частотная инвариантность распределений, синтезированных на основе системы (рисунок 5б), будет определяться законом изменения полосы пропускания образующих ее фильтров.

Учитывая указанные обстоятельства для фильтрации, определяемой положениями кратномасштабного анализа (рисунок 5б), получаемая инвариантность получила название аффинной [10].

Согласно аффинной инвариантности любые масштабные изменения частоты и временные сдвиги, происходящие в сигнале $z(t)$, вызывают аналогичные масштабнo-временные изменения в формируемом распределении $\rho(f, t)$:

$$\begin{array}{ccc} z(t) & \rightarrow & \sqrt{f} z(f(\tau - t)) \\ \downarrow & & \downarrow \\ \rho_z(f, t) & \rightarrow & \varpi_z\left(\frac{v}{f}, f(\tau - t)\right), \end{array} \quad (55)$$

где ϖ_z – масштабнo-временное распределение.

Выражение (55) получено, исходя из условий представления аффинного сдвига в виде: $\frac{1}{\sqrt{a}} z\left(\frac{\tau - t}{a}\right)$, где $a = f_0/f$ – коэффициент масштаба частоты; f_0 – центральная частота, полоса пропускания фильтра (рисунок 5б). Применительно к ИВП – это центральная частота базисного вейвлета [43].

Если положить $f_0 = 1$ Гц, тогда процедуру в ИВП, адаптированную к обработке сигнала, с учетом (48) можно представить в следующем виде:

$$\begin{aligned} W_z(f, t) &= \sqrt{f} \int_{\tau} z(\tau) \psi^*(f(\tau - t)) d\tau = \\ &= 1/\sqrt{f} \int_v \hat{z}(v) \hat{\psi}^*(v/f) \exp(j2\pi f v) dv. \end{aligned} \quad (56)$$

Инвариантность (55) позволила определить аффинный класс масштабнo-временных распределений (МВР) [19]. Если аффинный сдвиг в сигнале по времени рассматривать как:

$$z(t) \rightarrow z_{\vec{a}, \vec{b}}(t) = \frac{1}{\sqrt{|\vec{a}|}} z\left(\frac{t - \vec{b}}{\vec{a}}\right), \quad (57)$$

то его сдвиг в частотной области соответственно можно выразить как:

$$\hat{z}(v) \rightarrow \hat{z}_{\vec{a}, \vec{b}}(v) = \sqrt{|\vec{a}|} \exp(-j2\pi v \vec{b}) \hat{z}(\vec{a} v). \quad (58)$$

Тогда, согласно (57) и (3), а также с учетом (55) билинейное аффинное масштабирование совместного частотно-временного распределения будет определяться следующим образом:

$$\rho(f, t) \rightarrow \varpi_{z, \vec{a}, \vec{b}}\left(\frac{f_0}{f}, t\right) = \varpi_z\left(\frac{t - \vec{b}}{\vec{a}}, \frac{a}{\vec{a}}\right), \quad (59)$$

где $f_0/f = a$ – масштабирующий множитель.

Учитывая возможность синтеза распределений, используя весовую функцию аффинного сдвига $\Pi\left(\frac{t' - t}{a}, af'\right)$, обобщенное МВР представим следующим образом:

$$\varpi_z(a, t; \Pi) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \Pi\left(\frac{t' - t}{a}, af'\right) \times \rho_z(t', f') dt' df'. \quad (60)$$

Масштабнo-временное представление (60) адекватно описывает распределение плотности сигнальной энергии при условии конечности его энергии:

$$\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \Pi(t, v) dt \frac{dv}{|av|} = 1. \quad (61)$$

Таким образом, множество распределений, формируемых в соответствии с (60), определяют аффинный класс представления сигналов. Правомочность указанного вывода подтверждается непротиворечием его по отношению к ЧВР класса Козна.

Действительно, если положить $a = 1$, что не противоречит условиям определения аффинной группы, то выражение (60) трансформируется к виду (40) (см. [2]), при $\Pi(t' - t, f')$ или $\Pi(t, v)$.

Общий анализ (60) показывает, что выбор конкретного типа МВР формально сводится к выбору аффинного ядра $\Pi(t, v)$, которое представляет двумерную сглаживающую функцию с переменной шириной действующей полосы частот.

Другой подход к синтезу обобщенного выражения для описания МВР базируется на обработке узкополосной функции неопределенности (ФН) $\chi(\xi, \tau)$ с использованием обобщенной весовой функции аффинного преобразования $U_A(\xi, \tau)$:

$$\begin{aligned} \varpi_z(a, t; \Pi) &= \\ &= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} U_A\left(\frac{a}{\tau}, a\xi\right) \chi(\xi, \tau) e^{-j2\pi\xi t} d\xi d\tau, \end{aligned} \quad (62)$$

где:

$$U_A(\xi, \tau) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \Pi(v, t) e^{j2\pi(v\tau + \xi t)} dv dt. \quad (63)$$

В [8] и [12] предложен подход определения обобщенного МВР описания сигналов с использованием двухчастотного ядра $\Xi(v, f)$:

$$\begin{aligned} \varpi_z(a, t; \Pi) &= \frac{1}{|a|} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \Xi(v, f) \times \\ &\times \hat{z}\left(\frac{f-v/2}{a}\right) \hat{z}^*\left(\frac{f+v/2}{a}\right) e^{j2\pi vt/a} dv df, \end{aligned} \quad (64)$$

где:

$$\Xi(v, f) = \int_{-\infty}^{\infty} \Pi(t, f) e^{-j2\pi vt} dt. \quad (65)$$

Очевидно, что практическая применимость МВР возможна лишь в том случае, когда они удовлетворяют определенным свойствам, также как ЧВР класса Козна в соответствии с таблицей 1 в [1].

В общем случае МВР обладают рядом положительных свойств с точки зрения их практической применимости в обработке сигналов.

Свойство 1. Сохранение энергии. Совпадение энергии сигнала E_z с энергией МВР, рассредоточенной на масштабно-временной плоскости:

$$E_z = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \varpi_z(a, t; \Pi) dt \frac{da}{a^2}. \quad (66)$$

Свойство 2. Предельные (маргинальные) свойства. Спектральная плотность энергии и мгновенная мощность сигнала могут быть получены из МВР как предельные распределения:

$$\int_{-\infty}^{\infty} \varpi_z(a, t; \Pi) dt = |\hat{z}(v_0/a)|^2, \quad (67)$$

$$\int_{-\infty}^{\infty} \varpi_z(a, t; \Pi) \frac{da}{a^2} = |z(t)|^2. \quad (68)$$

Свойство 3. Сохранение действительных значений. Любое МВР $\varpi_z(a, t; \Pi)$ является действительной функцией, в общем случае, для любых $a \in L^2(\mathbf{R})$ и $t \in L^2(\mathbf{R})$.

Свойство 4. Локализация во времени. При любом аффинном сдвиге сигнала сохраняется его локализация во времени:

$$\hat{z}(f) = \frac{1}{\sqrt{f}} e^{-j2\pi f t_0} \vartheta_H(f) \Rightarrow \quad (69)$$

$$\varpi_z\left(\frac{f_0}{f}, t; \Pi\right) = f \delta(t - t_0) \vartheta_H(f),$$

где: $\vartheta_H(f) = \begin{cases} 0, & f < v_1 \\ (f/v_2 + 1)/2, & v_1 \leq f \leq v_2 \\ 1, & f > v_2 \end{cases}$ – функция Хевисайда [36].

Свойство 5. Адекватность представления. Временная локализация результата скалярного произведения любых двух функций $z(t)$ и $u(t)$ сохраняется и для произведения их МВР:

$$\begin{aligned} \left| \int_{-\infty}^{\infty} z(t) u^*(t) dt \right|^2 &= \\ &= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \varpi_z(a, t; \Pi) \varpi_u^*(a, t; \Pi) dt \frac{da}{a^2}. \end{aligned} \quad (70)$$

Свойство 6. Групповая задержка. Значение групповой задержки для масштабированного сигнала определяется как момент первого порядка его МВР по времени:

$$t_z\left(\frac{f_0}{a}\right) = \frac{\int_{-\infty}^{\infty} t \varpi_z(a, t; \Pi) dt}{\int_{-\infty}^{\infty} \varpi_z(a, t; \Pi) dt}. \quad (71)$$

Свойство 7. Унитарность представления [3]. Для узкополосных сигналов МВР совпадает с ЧВР Вигнера:

$$\varpi_z(a, t; \Pi) = \rho_W\left(\frac{f_0}{a}, t\right). \quad (72)$$

Учитывая, что простейшим совместным распределением сигнала в классе Козна является спектрограмма, то воспользовавшись свойством, аналогичным инвариантности, определяемой выражением (40), см. [2], с учетом аффинного характера вносимой инвариантности, можно получить аналогичную простейшую форму для МВР.

Поскольку аффинная инвариантность присуща вейвлетам, то в качестве весовой функции целесообразно выбрать вейвлет-преобразование вида:

$$\Pi(v, t) = W_\psi. \quad (73)$$

Тогда, в результате применения ядра (73) получим масштабogramму [79]:

$$\begin{aligned} |\varpi_M(a, t; \Pi)|^2 &= \\ &= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \rho_W(\xi, v) W_\psi\left(\frac{v-t}{a}, a\xi\right) dv d\xi. \end{aligned} \quad (74)$$

Адекватность подхода к синтезу распределения (74) позволяет рассматривать масштабogramму как некоторый аналог спектрограммы для аффинного класса.

Сделанные предположения позволяют предположить, что МВР (74) будет удовлетворять свойствам (1) и (3), причем модуль выражения обеспечивает ему наличие положительного результата при любом допустимом значении аргумента.

Для исследования особенностей механизма формирования масштабogramм воспользуемся вейвлетом Морле (вейвлет Гаусса 2-го порядка [28]).

Обозначим для функции-окна вейвлета Морле: $\Delta_{\hat{\psi}}$ – спектральная ширина вейвлета, обеспечивающая сглаживание в частотной области; Δ_{ψ} – действующая ширина вейвлета, обеспечивающая сглаживание во временной области.

В процессе синтеза масштабogramмы согласно (74), значения $\Delta_{\hat{\psi}}$ и Δ_{ψ} будут меняться в зависимости от масштабного множителя [35] и в общем случае можно записать:

$$\Delta_{\hat{\psi}}(f) = f \cdot \frac{\Delta_{\hat{\psi}}}{f_0}; \quad (75)$$

$$\Delta_{\psi}(f) = f_0 \cdot \frac{\Delta_{\psi}}{f}.$$

Результат (75) определяет качество анализа процесса с позиций частотно-временного разрешения.

В отличие от Q -анализа для ЧВР класса Коэна, анализ, определяемый (75), является более гибким инструментом, поскольку его разрешение изменяется в зависимости от диапазона. Следовательно, при одинаковом числе операций, необходимых для синтеза распределений, представления аффинного класса обеспечат более качественную декомпозицию обрабатываемых сигналов в частотно-временном пространстве.

Таким образом, компромисс между разрешением по времени и частоте, определяемый неравенством Гейзенберга, сохраняется и в отношении МВР. Постоянство параметра Q обеспечивается взаимозависимостью $\Delta_{\hat{\psi}}$ и Δ_{ψ} , т. е. улучшение одного показателя происходит за счет ухудшения другого. Продуктивность такого подхода рассмотрим на примере анализа суммы двух тестовых сигналов с гиперболическим законом изменения частоты, ограниченных временным окном.

Так, на рисунке 6 представлен фрагмент масштабogramмы тестового сигнала, синтезированной на основе вейвлета Морле.

Анализ полученных результатов позволяет сделать вывод, что, несмотря на аффинную инвариантность функции-окна, распределение энергии на частотно-временном плане имеет «размытый» характер. В результате сложно определить поло-

жение линии мгновенных частот без априорных сведений о сигнале.

Следует отметить, что применение аналогичных распределений класса Коэна, в частности спектрограммы, также не позволяет получить приемлемых результатов для рассматриваемого случая (рисунок 7).

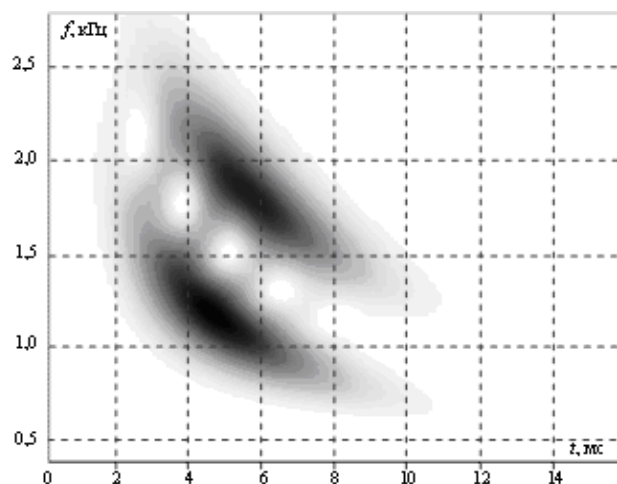


Рис. 6. Масштабogramма тестового сигнала

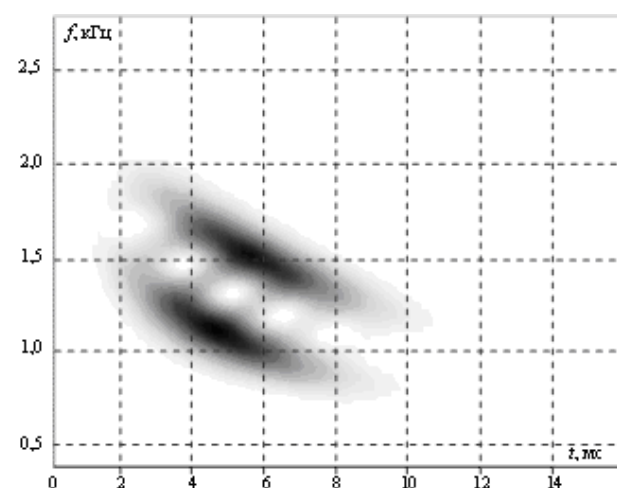


Рис. 7. Спектрограмма тестового сигнала

Более того, результат, полученный на основе спектрограммы, имеет явное смещение относительно истинных частотно-временных координат.

Формальный подход (74), в принципе, может быть обобщен не только на псевдораспределение Вигнера, но и на другие распределения класса Коэна.

С этой целью рассмотрим ядро Коэна с позиций следующей интерпретации [14]:

$$U_A(\xi, \tau) = U_A(\xi\tau) e^{-j2\pi f_0 \tau}, \quad (76)$$

где f_0 – центральная частота исходного порождающего вейвлета.

Тогда с учетом (76) можно установить следующее соответствие между распределениями класса Козна и аффинного класса:

$$\varpi_z(a, t; U_A) = \rho_z\left(\frac{f_0}{a}, t; U_A\right). \quad (77)$$

В соответствии с (77) в аффинном классе могут быть синтезированы распределения, аналогичные распределению Вигнера, Бьорна – Джордона, Рихачека, Чоя – Вильямса и др.

Так, применение раздельного сглаживания распределения по частоте и по времени на основе функции-окна с аффинной инвариантностью позволяет получить аффинное сглаженное псевдораспределение Вигнера (название по аналогии с $\hat{\rho}_w(f, t)$ класса Козна) [8]:

$$\begin{aligned} \hat{\varpi}_w(a, t) = & \frac{1}{a} \int h_0\left(\frac{\tau}{a}\right) \int_{-\infty}^{\infty} g_t\left(\frac{v-t}{a}\right) \times \\ & \times z_a^*\left(v - \frac{\tau}{2}\right) z_a\left(v + \frac{\tau}{2}\right) dv e^{-j2\pi\tau f_0/a} d\tau. \end{aligned} \quad (78)$$

Выражение (78) получено при $f_0 = 1$ Гц.

Аффинное сглаженное псевдораспределение Вигнера удовлетворяет свойствам 1 и 3 при условии, что $g_t \in L^2(\mathbf{R})$, а h_0 удовлетворяет условиям функций Эрмита.

На рисунке 8 представлено распределение тестового сигнала, синтезированное в соответствии с выражением (78).

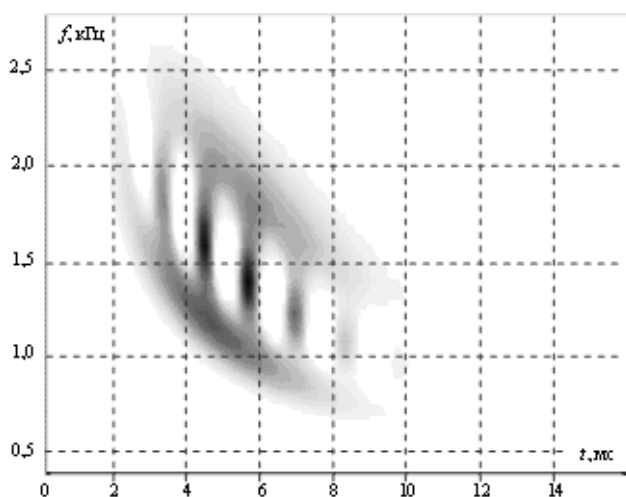


Рис. 8. Аффинное сглаженное псевдораспределение Вигнера тестового сигнала

Отметим, что в отличие от результатов (рисунки 6 и 7) распределение (78) имеет более явно выраженные линии мгновенных частот. Однако ложные всплески энергии, обусловленные межсимвольной интерференцией, не позволяют без априорных данных оценить истинное положение сигнальных компонент.

Следует отметить, что выбор параметров сглаживающих функций позволяет не только управлять разрешением в (78), но и дает возможность осуществлять синтез МВР с заданными свойствами.

В [12] предложено в качестве окон g_t и h_0 использовать различные формы функции Гаусса. Их применение позволяет варьировать результатом (75). В частности, при синтезе масштабограммы задавать значения их действующих полос равным 1, а при синтезе аффинного распределения Вигнера – равным нулю.

В [9] предложен подход к синтезу МВР, базирующийся на выборе соответствующих параметров их двухчастотных ядер $\Xi(v, f)$, определяемых согласно (65). Действительно, если допустить, что:

$$\begin{aligned} \Xi(nu, f) = & G(v)\delta(f - H(v)) - U_A(v, \tau) = \\ = & G(v)\exp(j2\pi H(v)\tau), \end{aligned} \quad (79)$$

где $G(v)$ – произвольная функция (в общем), а $H(v)$ – удовлетворяет условию ее локализации вдоль линии мгновенной частоты $f = H(v)$.

Тогда (64) можно привести к виду:

$$\begin{aligned} \varpi_z(a, t; \Pi) = & \frac{1}{|a|} \int_{-\infty}^{\infty} G(v) \times \\ & \times \hat{z}\left(\frac{H(v)-v/2}{a}\right) \hat{z}^*\left(\frac{H(v)+v/2}{a}\right) e^{j2\pi vt/a} dv. \end{aligned} \quad (80)$$

Анализ аналитического представления (80) показывает, что двухчастотные ядра хорошо локализируются только в том случае, если описывающая их функция изменяется по закону групповой задержки, который для МВР имеет вид экспоненциальный ($t_z(v) = t_0 + cv^{k-1}$), или логарифмический ($t_z(v) = t_0 + c \log v$).

Следовательно, закон изменения двухчастотных ядер можно связать непосредственно с распределением.

С учетом масштабирующего параметра $a = v_0/v$ формальный переход (77) будет иметь следующий вид:

$$\rho_z(v, t; U_A) = \varpi_z\left(\frac{v_0}{v}, t; U_A\right). \quad (81)$$

В соответствии с указанным подходом рассмотрим порядок синтеза МВР.

Если наложить на функции $G(v)$ и $H(v)$ из обобщенного выражения МВР требования соблюдения свойств 4 и 5, в частности:

$$(v) = \frac{v/2}{\sinh(v/2)} \quad \text{и} \quad H(v) = \frac{v}{2} \coth(v/2), \quad (82)$$

то тогда распределение (80) принимает следующий вид:

$$\varpi_B(a, t) = \frac{1}{|a|} \int_{-\infty}^{\infty} \frac{v/2}{\sinh(v/2)} \hat{z} \left(\frac{v \exp(-v/2)}{2a \sinh(v/2)} \right) \times \\ \times \hat{z}^* \left(\frac{v \exp(v/2)}{2a \sinh(v/2)} \right) e^{j2\pi vt/a} dv, \quad (83)$$

где: $\sinh(v/2) = \frac{\exp(v/2) - \exp(-v/2)}{2}$ – гиперболический синус; $\coth(v/2) = \frac{\exp(v/2) + \exp(-v/2)}{\exp(v/2) - \exp(-v/2)}$ – гиперболический котангенс.

В представленном виде распределение (83) аналогично МВР Бертранда [14]. Оно удовлетворяет свойствам 1–7, за исключением предельной временной локализации.

В [8] показано, что только МВР Бертранда из всех распределений, локализуемых посредством двухчастотного ядра, обеспечивает предельную концентрацию энергии сигналов вдоль линии его мгновенных частот, групповая задержка которых изменяется по гиперболическому закону:

$$\hat{z}(v) = \frac{\exp(jU_{ZA}(v))}{\sqrt{v}} \vartheta_H(v). \quad (84)$$

В [12] значение $U_{ZA}(v)$ определяется следующим образом:

$$U_{ZA}(v) = -2\pi \left(vt_0 - a \lg \left(\frac{v}{v_0} \right) \right), \quad (85)$$

$$\varpi_B \left(a = \frac{v_0}{v}, t \right) = v \delta(t - t_z(v)) \vartheta_H(v), \quad (86)$$

где $t_z(v) = -\frac{1}{2\pi} \cdot \frac{dU_{ZA}(v)}{dv}$ – групповая задержка сигнала.

На рисунке 9 представлено МВР Бертранда, синтезируемое в соответствии с выражением (83).

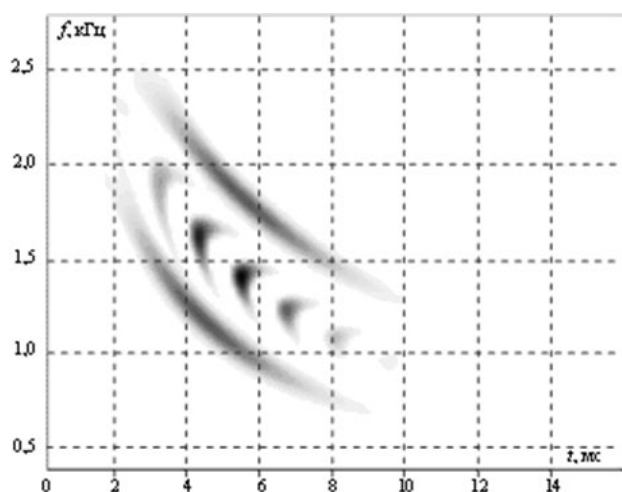


Рис. 9. Масштабно-временное распределение Бертранда тестового сигнала

В общем случае распределение Бертранда в большей степени локализует сигнальную энергию вдоль линий мгновенных частот по сравнению с

масштабограммой. Следовательно, оценки, полученные на его основе, будут более адекватны истинным значениям измеренных параметров.

Вместе с тем следует отметить, что при расчете МВР Бертранда необходимо точно указывать границы частотного диапазона, в пределах которого производятся вычисления. Эти ограничения требуют априорных знаний о частотной области локализации исследуемого процесса (здесь уместней говорить о масштабе частоты).

Ограничения для функций $G(v)$ и $H(v)$, обеспечивают временную локализацию МВР:

$$G(v) = 1 - \left(\frac{v}{4} \right)^2 \text{ и } H(v) = 1 + (v/4)^2, \quad (87)$$

что позволяет получить при подстановке в (80) МВР D -Фландрина [9], которое в некоторой степени представляет собой сглаженную (вдоль оси масштаба и вдоль оси времени) форму МВР Бертранда:

$$\varpi_D(a, t) = \\ = \frac{1}{|a|} \int_{-\infty}^{\infty} (1 - (v/4)^2) \hat{z} \left(\frac{[1 - v/4]^2}{a} \right) \times \\ \times \hat{z}^* \left(\frac{[1 + v/4]^2}{a} \right) e^{j2\pi vt/a} dv. \quad (88)$$

Распределение (88) удовлетворяет свойствам 1–4 и 6–7, и из всех распределений, локализуемых посредством двухчастотного ядра, обеспечивает предельную концентрацию энергии сигналов вдоль линии его мгновенных частот, групповая задержка которых изменяется по закону $1/\sqrt{v}$:

$$\hat{z}(v) = \frac{\exp(jU_{ZA}(v))}{\sqrt{v}} \vartheta_H(v). \quad (89)$$

Здесь значения $U_{ZA}(v)$ определяются следующим образом:

$$U_{ZA}(v) = -2\pi(vt_0 + 2a\sqrt{v}), \quad (90)$$

$$\varpi_D \left(a = \frac{v_0}{v}, t \right) = v \delta(t - t_z(v)) \vartheta_H(v). \quad (91)$$

Частотно-временной план МВР D -Фландрина по своей структуре очень близок к МВР Бертранда, поэтому оценки, полученные на его основе, будут сопоставимы со значениями выражения (83).

Среди возможных билинейных распределений аффинного класса, имеющих простое с точки зрения практической реализуемости ядро, необходимо выделить МВР Ундербергерера.

Синтез МВР Ундербергерера возможен при следующем выборе:

$$G(v) = 1 \text{ и } H(v) = \sqrt{1 + (v/2)^2}. \quad (92)$$

В этом случае получаем активную форму МВР Ундербергера:

$$\varpi_{U_a}(a, t) = \frac{1}{|a|} \times \times \int_0^{\infty} \left(1 + \frac{1}{n^2}\right) \hat{z}\left(\frac{n}{a}\right) \hat{z}^*\left(\frac{1}{na}\right) \exp\left(j2\pi(n-1/n)\frac{t}{a}\right) dn. \quad (93)$$

Распределение тестового сигнала (рисунок 11) вида (93) удовлетворяет свойствам 1–4 и 6–7.

Другой выбор параметров, определяющих характер аффинного сдвига:

$$G(v) = \frac{1}{\sqrt{1+(v/2)^2}} \quad \text{и} \quad H(v) = \sqrt{1+(v/2)^2}, \quad (94)$$

ведет к синтезу так называемой пассивной формы МВР Ундербергера, которая в отличие от формы (93), удовлетворяет свойствам 1–3 и 6–7:

$$\varpi_{U_P}(a, t) = \frac{1}{|a|} \times \times \int_0^{\infty} \frac{1}{n} \hat{z}\left(\frac{n}{a}\right) \hat{z}^*\left(\frac{1}{na}\right) \exp\left(j2\pi(n-1/n)\frac{t}{a}\right) dn. \quad (95)$$

Следует отметить, что главная особенность распределения Ундербергера состоит в том, что оно из всех распределений, локализуемых посредством двухчастотного ядра, обеспечивает предельную концентрацию энергии сигналов вдоль линии его мгновенных частот, групповая задержка которых изменяется по закону $1/v^2$:

$$\hat{z}(v) = \frac{\exp(jU_{Z_A}(v))}{\sqrt{v}} \vartheta_H(v). \quad (96)$$

В [7] значение $U_{Z_A}(v)$ определяется следующим образом:

$$U_{Z_A}(v) = -2\pi \left(vt_0 - \frac{a}{v}\right), \quad (97)$$

$$\varpi_{U_a}\left(a = \frac{v_0}{v}, t\right) = v\delta(t - t_z(v))\vartheta_H(v). \quad (98)$$

Свойства аффинных распределений еще в большей степени становятся предпочтительными для широкополосных сигналов, поскольку для них ЧВР класса Коэна не способны обеспечить адекватность сдвига по групповой задержке во всей полосе анализа.

В первую очередь это относится к сигналам, мгновенная частота которых может значительно изменяться за время анализа, например, в результате эффекта Доплера.

Для класса таких излучений условие Гейзенберга не остается строгим условием, следовательно, на частотно-временной плоскости распределений класса Коэна не будет соблюдаться адекватность их представления. Поскольку для нестационарных сигналов в общем случае закон изменения групповой задержки не является линейным, то, следова-

тельно, применение для их частотной декомпозиции функций Фурье не совсем уместно.

В первую очередь это вызвано тем, что гармонические функции инвариантны к частотным сдвигам сигнала, но не к его аффинному расширению.

Анализ свойств существующих функциональных базисов декомпозиции показал, что рассмотренным условиям соответствует преобразование Меллина. Полученные на его основе представления в общем случае инвариантны расширениям частоты, которые характерны для сигналов с гиперболическим законом изменения несущей.

В [34] предложено следующее аналитическое описание для преобразования Меллина:

$$M_z(\kappa) = \int_0^{\infty} \hat{z}(v) v^{j2\pi\kappa-1} dv. \quad (99)$$

Для исследования свойств инвариантности распределения (99) рассмотрим два сигнала $y(t)$ и $z(t)$ таких, что $\hat{y}(v) = \hat{z}(av)$.

Тогда в соответствии с (99) получим:

$$M_y(\kappa) = a^{-j2\pi\kappa} M_z(\kappa). \quad (100)$$

Выражение (100) однозначно указывает на характер закона инвариантного расширения, определяемого множителем $v^{j2\pi\kappa}$.

С точки зрения понятия базиса разложения распределение Меллина осуществляется совокупностью функций $\{v^{-j2\pi\kappa}\}$, групповая задержка которых определяется как:

$$t_z(v) = \frac{\kappa}{v}. \quad (101)$$

Фактически параметр "κ" управляет функцией групповой задержки и его можно интерпретировать как параметр масштаба изменения скорости модуляции (в данном случае гиперболической). Результаты, представленные на рисунке 10, поясняют физический смысл преобразования Меллина.

Параметр "κ" есть линейная функция времени, поэтому преобразование Меллина (рисунок 10в) можно представлять как относительно "κ", так и относительно "t". Общий вид результата не изменится за исключением шкалы отображения по оси абсцисс.

Для рассмотренного класса сигналов (в общем случае объединяемых условием $\Delta F_z \approx f_0$, где ΔF_z – ширина спектра сигнала; f_0 – центральная частота спектра) их узкополосная ФН не обеспечивает адекватности представления [74]. На рисунке 11 представлена ФН тестового сигнала (по оси ординат – доплеровский сдвиг; по оси абсцисс – величина групповой задержки).

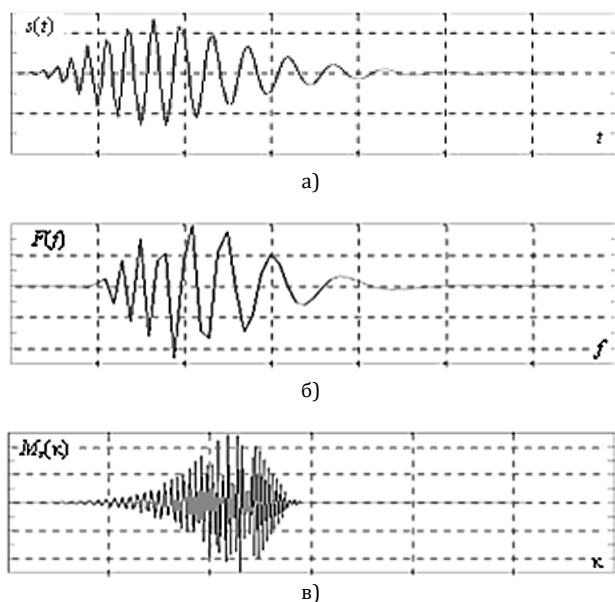


Рис. 10. Доплеровский импульс: а) во временной области; б) в базисе функций Фурье; в) в базисе функций Меллина

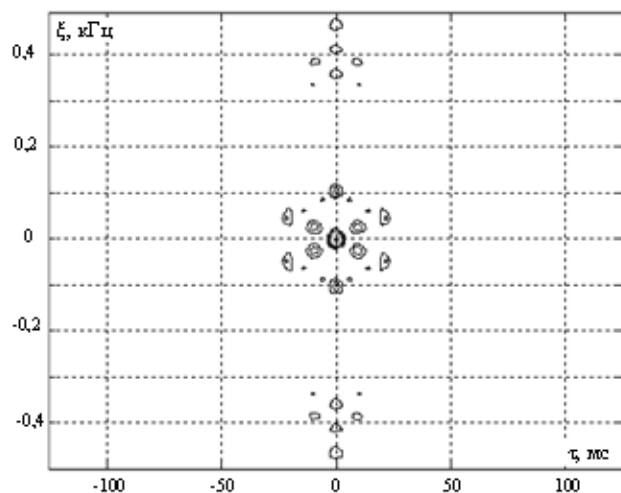


Рис. 11. Функция неопределенности (узкополосная) тестового сигнала

Узкополосная ФН локализует сигнал в центре координатной плоскости и довольно слабо отражает физическую сущность анализируемого процесса. Поэтому для сигналов, имеющих сложный закон изменения мгновенной частоты, в большей степени обосновано применение так называемой широкополосной ФН $\chi_\Delta(a, \tau)$, учитывающей характер нелинейных инвариантных изменений:

$$\begin{aligned} \chi_\Delta(a, \tau) &= a^{-1/2} \int_{-\infty}^{\infty} z(t) z^*(t/a - \tau) dt = \\ &= a^{1/2} \int_{-\infty}^{\infty} \hat{z}(v) \hat{z}^*(av) e^{j2\pi av\tau} dv. \end{aligned} \quad (102)$$

Аналитическое описание (102) фактически соответствует вейвлет-преобразованию аналитического сигнала, где в качестве базисного вейвлета

выступает сам сигнал. Результат обработки тестового сигнала широкополосной ФН представлен на рисунке 12. К сожалению, представление (102), так же, как и остальные МВР аффинного класса, не позволяет полностью исключить результат интерференционного взаимодействия двух составляющих тестового сигнала.

В общем случае формулу (102) можно трактовать как аффинную корреляционную функцию, измеряющую формальное сходство между сигналом его трансформируемой версией. Такой подход удобен с точки зрения синтеза практических алгоритмов расчета широкополосной ФН посредством применения двойного преобразования Меллина.

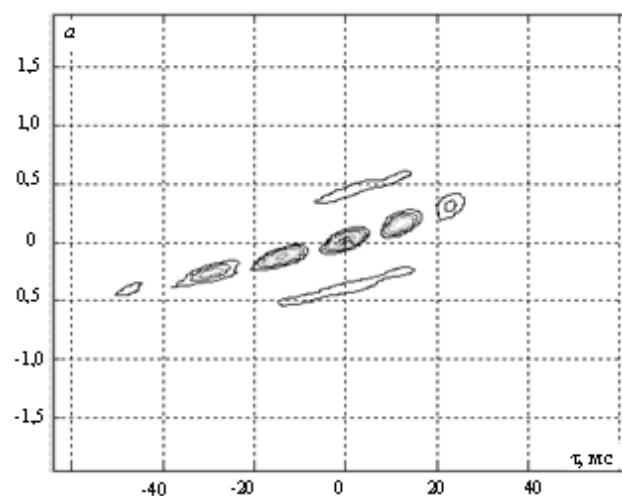


Рис. 12. Функция неопределенности (широкополосная) тестового сигнала

Заключение

Представленные результаты существенно расширяют известные границы частотно-временного анализа, который в настоящем контексте можно интерпретировать как частный случай более общего масштабно-временного анализа.

Предложенные подходы в гораздо большей степени адаптированы к исследованию процессов с нестационарным изменением параметров на интервале обработки и анализа.

Разработанный аналитический аппарат позволяет рассматривать его как инструмент анализа тонкой структуры сигналов. Представленные примеры обработки кратковременных фрагментов сигналов с высоким уровнем нестационарности их параметров

Таким образом, можно заключить, что теория частотно-временного анализа, заложенная Коэнм еще в середине прошлого века, продолжает свое развитие, открывая новые возможности для дальнейшего поиска эффективных процедур обработки сигналов в приложении к практическим задачам всех видов связи и телекоммуникаций [39].

Список используемых источников

1. Дворников С.В. Теоретические основы синтеза билинейных распределений энергии нестационарных процессов в частотно-временном пространстве (обзор) // Труды учебных заведений связи. 2018. Т. 4. № 1. С. 47–60. DOI:10.31854/1813-324X-2018-1-47-60.
2. Дворников С.В. Билинейные распределения с пониженным уровнем интерференционного фона в частотно-временном пространстве: продолжение обзора // Труды учебных заведений связи. 2018. Т. 4. № 2. С. 69–81. DOI:10.31854/1813-324X-2018-2-69-81.
3. Cohen L. Time-Frequency Distribution – a Review // Proceedings of the IEEE. 1989. Vol. 77. Iss. 7. PP. 941–981. DOI:10.1109/5.30749.
4. Claasen T.A.C.M., Meulenbrauker W.F.G. The Wigner Distribution – a Tool for Time-Frequency Signal Analysis. Part 1, 2, 3 // Philips Journal of Research. 1980. Vol. 35. PP. 217–250, 276–300, 372–389.
5. Дворников С.В. Теоретические основы синтеза частотно-временных представлений класса Коэна // Информация и космос. 2008. № 3. С. 16–24.
6. Дворников С.В., Кудрявцев А.М. Теоретические основы частотно-временного анализа кратковременных сигналов: монография. СПб.: ВАС, 2010. 240 с.
7. Bertrand J., Bertrand P. Affine time-frequency distributions in Time-Frequency Signal Analysis – Methods and Applications. Melbourne, Australia: Longman–Cheshire, 1991.
8. Bertrand J., Bertrand P. A class of affine Wigner functions with extended covariance properties // Journal of Mathematical Physics. 1992. Vol. 33. Iss. 7. DOI:10.1063/1.529570.
9. Flandrin P. Time-Frequency / Time-Scale Analysis. San Diego: Academic Press, 1999 (translated by Stöckler from the French editions, Temps-frequency. Paris: Hermes, 1993).
10. Flandrin P., Gonsalves P. Geometry of affine distributions // Proceedings of IEEE-SP International Symposium on Time-Frequency and Time-Scale Analysis, Philadelphia, PA, USA, 1994. PP. 80–83. DOI:10.1109/TFSA.1994.467359.
11. Baraniuk R.G. Covariant Time-Frequency Representations through Unitary Equivalence // IEEE Signal Processing Letters. 1996. Vol. 3. № 3. PP. 79–81. DOI:10.1109/97.481161.
12. Auger F., Flandrin P. Improving the readability of time-frequency and time-scale representation by the reassignment method // IEEE Transactions on Signal Processing. 1995. Vol. 43. Iss. 5. PP. 1068–1089. DOI:10.1109/78.382394.
13. Дворников С.В., Алексеева Т.Е. Распределение Алексеева и его применение в задачах частотно-временной обработки сигналов // Информация и космос. 2006. № 3. С. 9–20.
14. Дворников С.В. Теоретические основы синтеза билинейных распределений. СПб.: ВАС, 2007. 268 с.
15. Алексеев А.А., Кириллов А.Б. Технический анализ сигналов и распознавание радиоизлучений. СПб.: ВАС, 1998. 368 с.
16. Дворников С.В. Проблема поиска сигналов источников информации при радиомониторинге // Мобильные системы. 2007. № 4. С. 33–35.
17. Daubechies I. The wavelet transform, time-frequency localization and signal analysis // IEEE Transactions on Information Theory. 1990. Vol. 36. Iss. 5. PP. 961–1005. DOI:10.1109/18.57199.
18. Meyer Y. Ondelettes et operateurs. Vol. 1. Paris: Hermann, 1990.
19. Rioul O., Flandrin P. Time-Scale Energy Distributions: A General Class Extending Wavelet Transforms // IEEE Transactions on Signal Processing. 1992. Vol. 40. No. 7. PP. 1746–1757.
20. Дворников С.В., Бородин Е.Ю., Маджар Х., Махлуф Ю.Х. Частотно-временное оценивание параметров сигналов на основе функций огибающих плотности распределения их энергии // Информация и космос. 2007. № 4. С. 41–45.
21. Дворников С.В., Яхеев А.Ф. Метод измерения параметров кратковременных сигналов на основе распределения Алексеева // Информация и космос. 2011. № 1. С. 66–74.
22. Дворников С.В., Железняк В.К., Храмов Р.Н., Желнин С.Р., Медведев М.В., Симонов А.Н., Сауков А.М. Метод обнаружения радиоизлучений на основе частотно-временного распределения Алексеева // Научное приборостроение. 2006. Т. 16. № 1. С. 107–115.
23. Дворников С.В., Осадчий А.И., Дворников С.С., Родин Д.В. Демодуляция сигналов на основе обработки их модифицированных распределений // Контроль. Диагностика. 2010. № 10. С. 46–54.
24. Яхеев А.Ф., Дворников С.В. Измерение параметров сигналов на основе оптимизации формы распределения Алексеева // Наукоемкие технологии. 2009. Т. 10. № 1. С. 25–28.
25. Дворников С.В. Демодуляция сигналов на основе обработки их модифицированных частотно-временных распределений // Цифровая обработка сигналов. 2009. № 2. С. 7–11.
26. Дворников С.В., Сауков А.М. Модификация частотно-временных описаний нестационарных процессов на основе показательных и степенных функций // Научное приборостроение. 2004. Т. 14. № 3. С. 76–85.
27. Чуи Ч. Введение в вейвлеты. М.: Мир, 2001. 412 с.
28. Дворников С.В., Дворников С.С., Спирин А.М. Синтез манипулированных сигналов на основе вейвлет-функций. // Информационные технологии. 2013. № 12. С. 52–55.
29. Дворников С.В. Теоретические основы представления сигнала в аналитическом виде функциями его огибающей и полной фазы // Научное приборостроение. 2006. Т. 16. № 4. С. 106–111.
30. Gonsalves P., Baraniuk R.G. A Pseudo-Bertrand Distribution for Time-Scale Analysis // IEEE Signal Processing Letters. 1996. Vol. 3. Iss. 3. PP. 82–84. DOI:10.1109/97.481162.
31. Дворников С.В., Желнин С.Р., Медведев М.В. Метод формирования признаков распознавания сигналов диапазона декаметровых волн по их вейвлет-коэффициентам, рассчитанным на основе лифтинговой схемы // Информация и Космос. 2006. № 2. С. 68–73.
32. Mallat S.G. A theory for multiresolution signal decomposition: the wavelet representation // IEEE Transactions on Pattern Analysis and Machine Intelligence. 1989. Vol. 11. Iss. 7. PP. 674–693. DOI:10.1109/34.192463.

33. Аладинский В.А., Дворников С.В., Сауков А.М., Симонов А.Н. Способ распознавания радиосигналов. Патент на изобретение RUS 2261476 от 26.01.2004.
34. Ovarlez J.-P. La transformation de Mellin et l'analyse des signaux large bande // Journées Thématiques GdR CNRS «Traitement du Signal et Images». 8 et 9 Mars. 1994. INSA Lyon.
35. Малла С. Вейвлеты в обработке сигналов. М.: Мир, 2005. 671 с.
36. Cohen L. Time-Frequency Analysis. Englewood Cliffs, NJ: Prentice-Hall, 1995.
37. Grossmann A., Morlet J. Decomposition of Hardy Functions into Square Integrable Wavelets of Constant Shape // SI-AM Journal on Mathematical Analysis. 2006. Vol. 15. Iss. 4. DOI:10.1137/0515056.
38. Coifman R., Meyer Y., Wickerhauser V. Wavelet Analysis and Signal Processing // In Wavelets and their Applications. Boston: Jones and Bartlett. 1992. PP. 153–178.
39. Дворников С.В. Цифровой синтез спектрально-эффективных сигналов телевидения // Вопросы радиоэлектроники. Серия: Техника телевидения. 2015. № 6. С. 168–173.
40. Cohen A., Daubechies I. Orthonormal Bases of Compactly Supported Wavelets III. Better Frequency Resolution // SIAM Journal on Mathematical Analysis. 2006. Vol. 24. Iss. 2. DOI:10.1137/0524032.
41. Дьяконов В.П. Вейвлеты. От теории к практике. М.: СОЛЮН-Р, 2002. 448 с.
42. Rioul O., Vetterli M. Wavelets and signal processing // IEEE Signal Processing Magazine. 1991. Vol. 8. Iss. 4. PP. 14–38. DOI:10.1109/79.91217.

* * *

BILINEAR SCALE-TEMPORARY DISTRIBUTIONS OF ENERGY OF THE AFFINE CLASS IN THE FREQUENCY-TEMPORARY SPACE (continued review)

S. Dvornikov¹

¹Telecommunications Military Academy,
St. Petersburg, 194064, Russian Federation

Article info

Article in Russian

For citation: Dvornikov S. Bilinear Scale-Temporary Distributions of Energy of the Affine Class in the Frequency-Temporary Space: Continued Review // Proceedings of Telecommunication Universities. 2018. Vol. 4. Iss. 3. PP. 26–44.

Abstract: *The review of theoretical bases of synthesis of scale-time distributions of an affine class is presented. The methodology of wavelet transforms with an adaptive analysis window is analyzed in the interest of processing processes with high parameter non-stationarity. Analytic approaches to synthesis of distributions invariant to affine shifts are generalized. The generality of the scale-time and frequency-time distributions is shown. Analytical expressions and experimental results of the most known forms of affine class distributions are given.*

Keywords: *scale-time distributions, methods for processing processes with high parameter non-stationarity, affine-class distribution, wavelet transforms.*

МОДЕЛЬ РАДИОЛИНИИ СПУТНИКОВОЙ СВЯЗИ В СОСТАВЕ ВЫСОКОСКОРОСТНОЙ СПУТНИКОВОЙ СИСТЕМЫ

В.А. Жиров¹, А.Е. Орлов², А.А. Смирнов²

¹Научно-исследовательский институт радио Санкт-Петербургский филиал (ЛОНИИР),

Санкт-Петербург, 192029, Российская Федерация

²Военная академия связи имени Маршала Советского Союза С.М. Буденного,

Санкт-Петербург, 194064, Российская Федерация

*Адрес для переписки: vorl-2210@mail.ru

Информация о статье

УДК 621.396

Язык статьи – русский

Ссылка для цитирования: Жиров В.А., Орлов А.Е., Смирнов А.А. Модель радиолинии спутниковой связи в составе высокоскоростной спутниковой системы // Труды учебных заведений связи. 2018. Т. 4. № 3. С. 45–53.

Аннотация: В статье предлагается модель радиолинии спутниковой связи в составе высокоскоростной спутниковой системы связи. В отличие от существующих моделей, в которых учитываются тепловые шумы, внутрисистемные и межсистемные помехи, а выполнение требований к готовности по дождевому фактору контролируется после оценки системного запаса, в предлагаемой модели требования к готовности учитываются явно в виде ограничения. Это позволяет ставить и решать задачи оптимизации распределения частотно-энергетического ресурса центральных земных станций и бортовых ретрансляционных комплексов по критерию максимума пропускной способности при выполнении требований к достоверности передачи информации и готовности дуплексного направления спутниковой связи.

Ключевые слова: радиолиния спутниковой связи, частотно-энергетический ресурс, земная станция, бортовой ретрансляционный комплекс, высокоскоростная спутниковая система.

Введение

В настоящее время для предоставления пользователям высокоскоростного доступа к удаленным информационным ресурсам во всем мире создаются и развиваются высокоскоростные спутниковые системы (ВССпС) Ka и Q диапазонов частот [1, 2] с использованием космических аппаратов на геостационарной орбите. Общая архитектура, технические характеристики и ожидаемая пропускная способность таких систем определены в [3].

Развертывание подобных систем для России имеет специфику: обширные территории обслуживания, северное расположение абонентов обуславливают большие потери сигнала, особенно в гидрометеорах. Имеют место значительные различия потерь сигнала в гидрометеорах по территории обслуживания. Ограниченность частотно-энергетического ресурса (ЧЭР) центральных земных станций (ЦЗС) и бортовых ретрансляционных комплексов (БРТК) и необходимость вместе с тем обеспечивать приемлемый уровень готовности спутниковых радиолиний даже в условиях выпа-

дения осадков предопределяют проблему его эффективного распределения.

Стандартные модели спутниковых радиолиний [4] позволяют решать задачи их анализа, но готовность оценивается по результатам оценки системного запаса. Поэтому в задачах распределения ЧЭР необходимо перебирать множество вариантов. Известны модели [5, 6] для нахождения прямого решения, но они имеют серьезные ограничения по условиям применения. В статье предлагается модель спутниковой радиолинии, которая позволяет решать задачи распределения ЧЭР в явном виде без ограничений на условия применения.

Проблемы распределения частотно-энергетического ресурса высокоскоростных спутниковых систем

Особенности ВССпС заключаются в использовании на спутнике многолучевых антенн, обеспечивающих покрытие большой территории узкими лучами с высоким коэффициентом усиления, повторном использовании частот (ПИЧ) в лучах об-

На каждый сетевой сегмент в прямом направлении (ПН) формируется несущая по стандарту DVB-S2 с временным мультиплексированием, которая ретранслируется транспондером одновременно с несущими других сетевых сегментов. Сетевой сегмент поддерживает работу группы АЗС, которые передают свои несущие по принципу многостанционного доступа с частотно-временным разделением через транспондер обратного направления (ОН). Несущие ОН формируются по стандарту DVB-RCS. Задача распределения ЧЭР заключается в делении ограниченной мощности и полосы ЦЗС и БРТК между ГрСС, а внутри группы – между сетевыми сегментами. Принцип распределения ЧЭР представлен на рисунке 1.

Эффективность использования полосы пропускания транспондера в ПН удваивается при использовании ортогональной (двойной) поляризации. Соотношение между полосами частот сетевых сегментов на выходе ЦЗС и транспондера сохраняется. В режиме линейного усиления мощность усилителя мощности транспондера ПН делится между несущими сетевых сегментов аналогично:

где $P_{\text{тр пн}}$ – максимальная мощность усилителя мощности транспондера; P'_{ij} – мощность несущей j -го сетевого сегмента i -ой ГрСС на выходе транспондера.

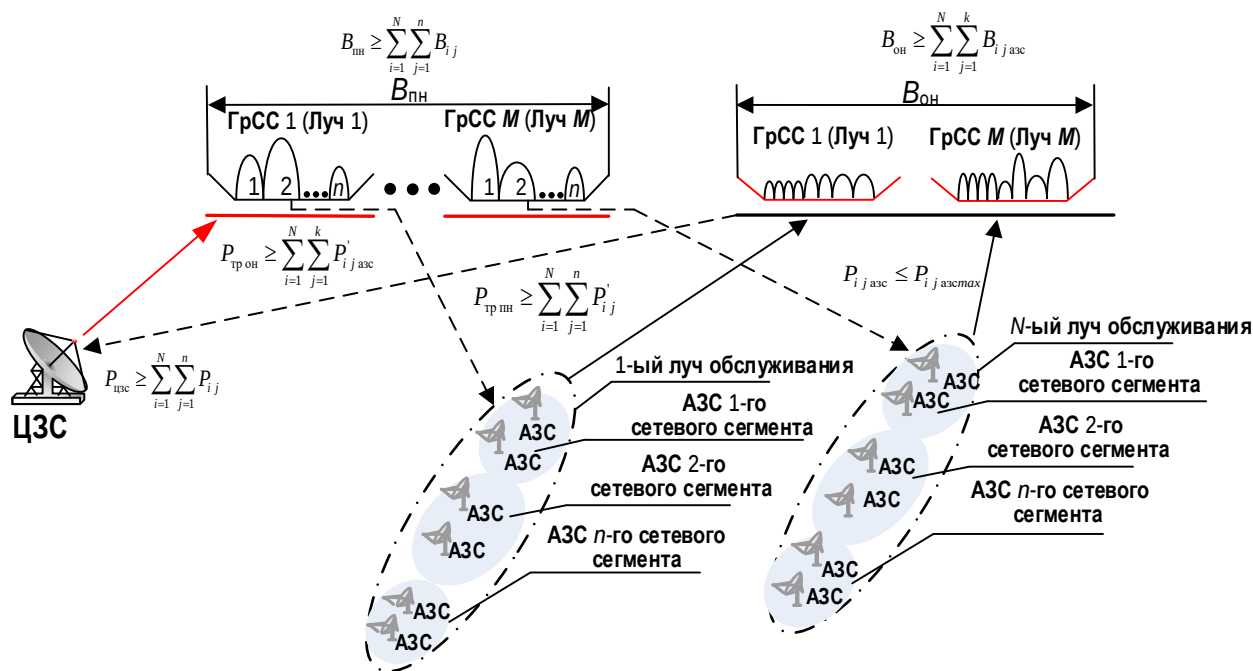


Рис. 1. Распределение ЧЭР ВССпС

Мощность несущей сетевого сегмента на входе и выходе транспондера связаны передаточной характеристикой. В обратном направлении мощность несущей АЗС ограничивается максимальной мощностью передатчика АЗС:

$$P_{i j \text{ азс}} \leq P_{i j \text{ азс max}},$$

где $P_{i j \text{ азс}}$, $P_{i j \text{ азс max}}$ – мощность несущей и максимальная мощность передатчика j -ой АЗС i -го сетевого сегмента, соответственно.

На выходе транспондера ОН мощность усилителя мощности делится между несущими АЗС:

$$P_{\text{тр он}} \geq \sum_{i=1}^N \sum_{j=1}^k P'_{i j \text{ азс}},$$

где $P_{\text{тр он}}$ – максимальная мощность усилителя мощности транспондера ОН; $P'_{i j \text{ азс}}$ – мощность несущей АЗС на выходе транспондера ОН; k – количество АЗС в сетевом сегменте.

Полоса пропускания транспондера ОН делится между сигналами несущих АЗС аналогично:

$$B_{\text{он}} \geq \sum_{i=1}^N \sum_{j=1}^k B_{i j \text{ азс}},$$

где $B_{\text{он}}$ – полоса пропускания транспондера ОН; $B_{i j \text{ азс}}$ – полоса частот, занимаемая сигналом j -ой АЗС i -го сетевого сегмента.

Таким образом, распределение ЧЭР ЦЗС и БРТК – это деление ограниченной мощности и полосы пропускания ЦЗС и БРТК между несущими сетевых сегментов для выполнения требований к пропускной способности каждого сетевого сегмента на выделенном ресурсе (с учетом возможности выбора сигнально-кодовой конструкции (СКК) из имеющегося множества). При этом необходимо учитывать требования к готовности дуплексного направления связи, образуемого любой парой ЦЗС – АЗС. Такая задача является переборной и достаточно трудоемкой.

Обеспечение требований к готовности предполагает учет радиоклиматических условий регионов обслуживания. ВССпС обслуживает обширные территории с существенно разными радиоклиматическими условиями, поэтому одна и та же готовность может обеспечиваться разной мощностью ЦЗС и БРТК. Неготовность будем трактовать в соответствии с [7] как состояние системы передачи продолжительностью не менее 10 с, когда наблюдаются сильно пораженные ошибками секунды, при этом переход в состояние готовности предполагает наличие периода с несильно пораженными секундами продолжительностью более 10 с.

Далее в статье раскрывается модель спутниковой радиолинии в виде функционала для пропускной способности, особенностью которой является явный учет требования к готовности, которая существенно упрощает решение задач распределения ЧЭР.

Модель радиолинии спутниковой связи в составе высокоскоростной спутниковой системы

Известно множество моделей спутниковых радиолиний и сетей спутниковой связи, отличающихся акцентами при постановке задачи. Условно их можно разделить на две группы.

К первой группе относятся модели анализа. Они ориентированы на анализ показателей функционирования линий и сетей спутниковой связи. Модели этой группы достаточно универсальны, но различаются степенью детализации факторов, влияющих на показатели функционирования, а также контролируемыми параметрами. Самыми распространенными в данной группе являются модели, позволяющие оценивать отношение мощности сигнала к мощности шума, а также достоверность в радиолиниях, неготовность по дождю. Затем эти результаты можно использовать для оценки пропускной способности.

Базовый вариант модели данной группы рассматривается в [8]. В [3] предлагается развитие данной модели за счет учета всех составляющих внутрисистемных помех, при этом влияние среды рассматривается в вероятностном смысле, что позволяет ввести в рассмотрение показатели неготовности.

Ко второй группе относятся модели синтеза. Они направлены на получение в явном виде значений технических параметров (земных станций или БРТК) или параметров, характеризующих режимы функционирования. Модели этой группы нацелены на решение задач узкой направленности. Как правило, задачи синтеза удается решить благодаря серьезным ограничениям и допущениям, не позволяющим распространять результаты на другие случаи и условия. Модели данной группы рассматриваются в работах [4, 5].

Отправной точкой для предлагаемой модели служит модель ретрансляции с накоплением шумов и помех (рисунок 2) [4]. Стрелка вниз означает ПН, вверх – ОН.

В ПН информационный сигнал i -го сетевого сегмента с информационной скоростью $R_{bi\downarrow}$, занимаемой полосой частот $df_{ci\downarrow}$ и использованием заданной СКК $S_{i\downarrow}$ поступает на вход усилителя мощности ЦЗС. P_i – мощность несущей сетевого сегмента на выходе усилителя мощности ЦЗС с учетом коэффициента недоиспользования мощности по выходу $OBo_{\text{цзс}}$. Сигнал, принимаемый БРТК, усиливается (с переносом частоты) и излучается в направлении группы АЗС рассматриваемого сетевого сегмента. $EIRP_i$ – значение эффективной изотропной излучаемой мощности на выходе ЦЗС. $G_{t\text{цзс}}$ – коэффициент усиления антенны ЦЗС. G_u – коэффициент усиления приемной антенны транспондера. $(G/T)_u$ и $(G/T)_{\text{азс}}$ – добротность приемной системы транспондера и АЗС, соответственно.

$G_{\text{тр}}$ – коэффициент усиления усилителя мощности транспондера. $ОВо_{\text{тр}\downarrow}$ – коэффициент недоиспользования мощности по выходу усилителя мощности транспондера в ПН. $L_{u\downarrow}, L_{d\downarrow}$ – потери сигнала в свободном пространстве восходящей и нисходящей линий, соответственно. $L_{ou\downarrow}(p_{\text{упн}}), L_{od\downarrow}(p_{\text{дпн}})$ – потери сигнала в гидрометеорах на восходящей и нисходящей линиях, определяемые процентом превышения времени среднего года $p_{\text{упн}}, p_{\text{дпн}}$.

Для ПН учитываются следующие виды помех:

– на ЦЗС: помехи интермодуляции $I_{01\downarrow}$, обусловленные совместным усилением нескольких несущих общим усилителем;

– в БРТК: помехи интермодуляции $I_{02\downarrow}$, обусловленные совместным усилением нескольких несущих общим усилителем мощности транспондера, кросс-поляризационной помехи (КПП) на входе приемной системы транспондера $I_{03\downarrow}$, спектральная плотность шума приемной системы транспондера $N_{0\text{тр}\downarrow}$;

– на АЗС: помехи ПИЧ $I_{04\downarrow}$, спектральная плотность шумов приемной системы АЗС $N_{0\text{азс}\downarrow}$.

В ОН информационный сигнал с информационной скоростью $R_{bi\uparrow}$, занимаемой полосой частот $df_{ci\uparrow}$ и используемой СКК $S_{i\uparrow}$ поступает на вход усилителя мощности АЗС. Сигнал, принимаемый БРТК, усиливается усилителем мощности транспондера и излучается в направлении ЦЗС. $EIRP_{i\text{азс}}$ – эффективная изотропная излучаемая мощность i -ой АЗС, $G_{\text{азс}}$ – коэффициент усиления антенны АЗС, $ОВо_{\text{тр}\uparrow}$ – коэффициент недоиспользования мощности по выходу усилителя мощности транспондера в ОН, $(G/T)_{\text{цзс}}$ – добротность приемной системы ЦЗС, $L_{u\uparrow}, L_{d\uparrow}$ – потери сигнала в свободном пространстве восходящей и нисходящей

линий, соответственно, $L_{ou\uparrow}(p_{\text{уон}}), L_{od\uparrow}(p_{\text{дон}})$ – потери сигнала в гидрометеорах на восходящей и нисходящей линиях, определяемые процентом превышения времени среднего года $p_{\text{уон}}, p_{\text{дон}}$.

Для ОН учтены следующие виды помех:

– в БРТК: помехи интермодуляции $I_{01\uparrow}$, обусловленные совместным усилением нескольких несущих общим усилителем мощности транспондера, помехи ПИЧ $I_{03\uparrow}$, спектральная плотность мощности шумов приемной системы транспондера $N_{0\text{тр}\uparrow}$;

– на ЦЗС: КПП $I_{02\uparrow}$ на входе приемной системы ЦЗС, спектральная плотность шумов приемной системы ЦЗС $N_{0\text{цзс}\uparrow}$.

Общее выражение для отношения мощности сигнала к спектральной плотности мощности шума и помех для спутниковой радиолинии известно [4], мы его представим в виде:

$$Z(\vec{X}) = \left[Z_{N,u}^{-1}(\vec{X}) + Z_{N,d}^{-1}(\vec{X}) + \sum_i Z_i^{-1}(\vec{X}) \right]^{-1}, \quad (1)$$

где $Z_{N,u}(\vec{X})$ – отношение мощности сигнала к спектральной плотности мощности шума на восходящем участке; $Z_{N,d}(\vec{X})$ – отношение мощности сигнала к спектральной плотности мощности шума на нисходящем участке; $Z_i(\vec{X})$ – отношение мощности сигнала к спектральной плотности мощности учитываемых помех (интермодуляции, КПП, ПИЧ, межсистемные и другие виды помех).

Вектор $\vec{X}^T$ включает набор технических параметров оборудования космического и наземного сегментов, набор параметров, определяющих значение потерь сигнала в атмосфере, значения контролируемых параметров (мощность несущей, полоса, занимаемая сигналом).

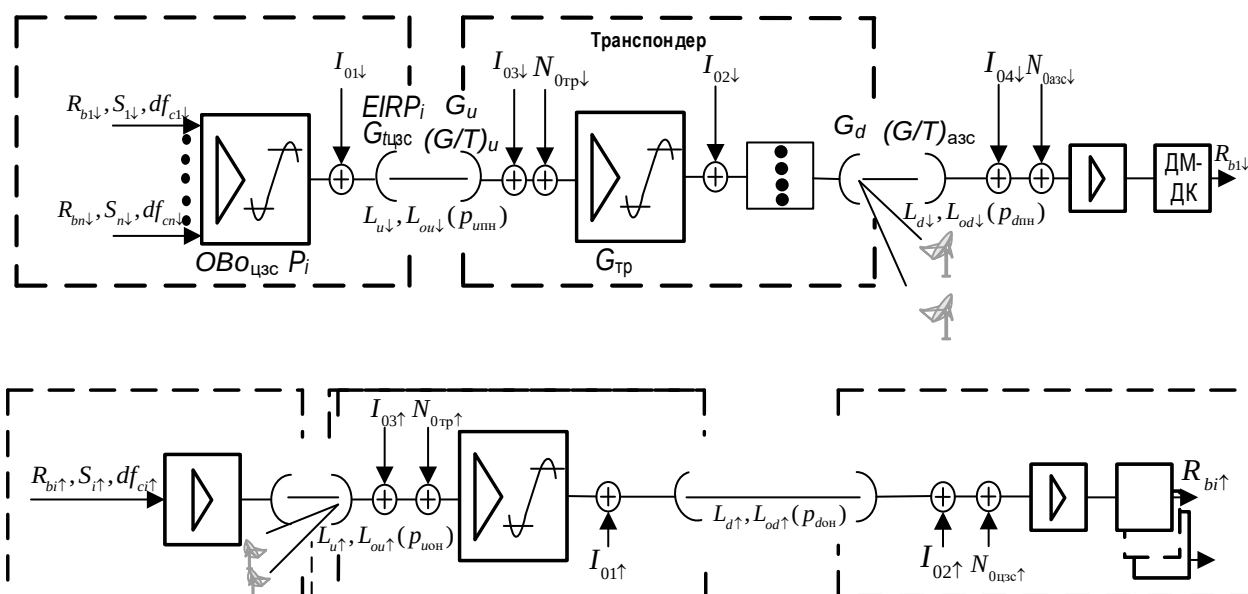


Рис. 2. Модель ретрансляции с накоплением шумов и помех

Формула (1) широко используется для расчета бюджета спутниковой радиолинии: при полностью известных технических параметрах наземной, бортовой аппаратуры и координат размещения станций. Она позволяет оценить ожидаемое отношение мощности сигнала к спектральной плотности шума и помех, а также выбрать СКК и скорость передачи информации. Однако, в задачах распределения ЧЭР использование данной модели неудобно, так как она не связана явным образом с требованиями к готовности и ограничениями по доступному ЧЭР. Для того, чтобы такую связь обеспечить, необходимо провести множество расчетов, чтобы подобрать вариант, удовлетворяющий требованиям.

Оценку пропускной способности симплексного направления связи можно производить с помощью формулы К. Шеннона, представленной в следующем виде [9]:

$$R_b = B_{\text{тр}} \left(1 + \gamma \frac{E_b}{N_0} \right), \quad (2)$$

где E_b/N_0 – отношение энергии сигнала на бит к спектральной плотности шума; γ – частотная эффективность используемой СКК.

Данная формула, определяя границу пропускной способности, будет использована для верификации предлагаемой в статье модели.

Выражение (1) может быть преобразовано в выражение для определения пропускной способности с учетом требований к достоверности информации:

$$R_b = \left[Z_{N,u}^{-1}(\vec{X}) + Z_{N,d}^{-1}(\vec{X}) + \sum_i Z_i^{-1}(\vec{X}) \right]^{-1} / (E_b/N_0)^*, \quad (3)$$

где $(E_b/N_0)^*$ – требуемое отношение энергии сигнала на бит к спектральной плотности шума, соответствующее используемой СКК.

Требование к готовности по дождю учтем следующим образом. К неготовности могут приводить события превышения интенсивности дождя определенных величин на стороне ЦЗС или АЗС (рисунок 3), а также одновременно. Вероятность последнего события обычно невелика, поэтому это событие учитывать не будем.

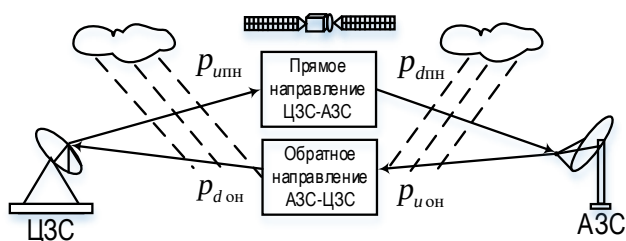


Рис. 3. Моделирование неготовности по дождевому фактору

На рисунке приняты следующие обозначения: $p_{\text{упн}}$ – процентная неготовность ПН по восходящему участку; $p_{\text{дон}}$ – процентная неготовность ОН по

нисходящему участку; $p_{\text{дпн}}$ – процентная неготовность ПН по нисходящему участку; $p_{\text{уон}}$ – процентная неготовность ОН по восходящему участку. Перечисленные процентные неготовности по участкам определяют результирующую неготовность дуплексного направления:

$$T\% = \max(p_{\text{упн}}; p_{\text{дон}}) + \max(p_{\text{дпн}}; p_{\text{уон}}). \quad (4)$$

Определение процентных вероятностей может быть выполнено с использованием моделей прогнозирования потерь сигнала при распространении на трассах «Земля-космос» [10, 11].

Зоны обслуживания разных сетевых сегментов в пределах одного луча (одной группы сетевых сегментов) могут существенно отличаться значениями радиоклиматических параметров. Требования к готовности каждого сетевого сегмента также могут различаться.

Выражение (3) для пропускной способности ПН с учетом выполнения требований к готовности может быть представлено в виде:

$$R_{b,T^*\% \downarrow} = \frac{[Z_{N,p_{\text{упн}\downarrow}}^{-1}(\vec{X}) + Z_{N,p_{\text{дпн}\downarrow}}^{-1}(\vec{X}) + \sum_{i=1}^4 Z_{I,0i\downarrow}^{-1}(\vec{X})]^{-1}}{(E_b/N_0)^*}, \quad (5)$$

где $Z_{N,p_{\text{упн}\downarrow}}(\vec{X})$ – отношение мощности сигнала к спектральной плотности шума с учетом неготовности на восходящем участке; $Z_{N,p_{\text{дпн}\downarrow}}(\vec{X})$ – отношение мощности сигнала к спектральной плотности мощности шума с учетом неготовности на нисходящем участке; $Z_{I,0i\downarrow}(\vec{X})$ – отношение мощности сигнала к спектральной плотности мощности учитываемой помехи; $T^*\%$ – допустимая неготовность.

Использование выражения (5) предполагает алгоритмическую реализацию процедуры поиска значения функционала, соответствующего выполнению условия $T\% = T^*\%$ для дуплексного направления. Эта процедура является базовой в предлагаемой модели. Аналогично для ОН:

$$R_{b,T^*\% \uparrow} = \frac{[Z_{N,p_{\text{уон}\uparrow}}^{-1}(\vec{X}) + Z_{N,p_{\text{дон}\uparrow}}^{-1}(\vec{X}) + \sum_{i=1}^3 Z_{I,0i\uparrow}^{-1}(\vec{X})]^{-1}}{(E_b/N_0)^*}, \quad (6)$$

где $Z_{N,p_{\text{уон}\uparrow}}(\vec{X})$ – отношение мощности сигнала к спектральной плотности мощности шума с учетом неготовности на восходящем участке; $Z_{N,p_{\text{дон}\uparrow}}(\vec{X})$ – отношение мощности сигнала к спектральной плотности мощности шума с учетом неготовности на нисходящем участке; $Z_{I,0i\uparrow}(\vec{X})$ – отношение мощности сигнала к спектральной плотности мощности учитываемой помехи.

Заметим, что выражения (5) и (6) справедливы в пределах выполнения ограничений по доступ-

ному ресурсу. Ограничения по доступному частотному ресурсу для ПН представляется в виде:

$$df_{c\downarrow} = \frac{R_{b,T\%*}(1 + \alpha)}{(\log_2 M \cdot R_c)}, df_{c\downarrow} \leq B_{\downarrow}, \quad (8)$$

для ОН:

$$df_{c\uparrow} = \frac{R_{b,T\%*}(1 + \alpha)}{(\log_2 M \cdot R_c)}, df_{c\uparrow} \leq B_{\uparrow}, \quad (9)$$

где $df_{c\downarrow}$ и $df_{c\uparrow}$ – полосы частот, занимаемые сигналами в прямом и обратном направлении, соответственно; $B_{\downarrow}$ и $B_{\uparrow}$ – значения выделенного частотного ресурса в прямом и обратном направлении, соответственно; α – ролл-офф фактор; R_c – скорость кодирования; M – размерность созвездия сигналов для выбранной СКК.

Ограничения по доступному энергетическому ресурсу для ПН представляется в виде:

$$P_{\downarrow} \in [P_{\min\downarrow}; P_{\max\downarrow}], \quad (10)$$

для ОН:

$$P_{\uparrow} \in [P_{\min\uparrow}; P_{\max\uparrow}], \quad (11)$$

где $P_{\downarrow}$ – мощность несущей сетевого сегмента в ПН; $P_{\min\downarrow}$, $P_{\max\downarrow}$ – максимальное и минимальное значения мощности несущей ПН (в зависимости от энергетического ресурса); $P_{\uparrow}$ – мощность несущей на выходе АЗС; $P_{\min\uparrow}$, $P_{\max\uparrow}$ – минимальное и максимальное значения мощности несущей в ОН (определяется техническими характеристиками АЗС).

Перечень используемых СКК в соответствии со стандартом DVB-S2:

$$S \in \{S_i, \gamma_i, (E_b/N_0)_i^*\}, i = \overline{1, N}, \quad (12)$$

где S_i – вид используемой СКК; γ_i – спектральная эффективность используемой СКК.

Объединяя выражения (5–12) в общую систему уравнений, получим модель радиолинии спутниковой связи в виде функционала для гарантированной пропускной способности прямого и обратного направлений. Эта модель учитывает требования к готовности, достоверности, а также ограничения частотного и энергетического ресурса и связывает два направления показателем неготовности дуплексного направления.

Для оценки влияния помехи интермодуляции необходимо иметь в явном виде взаимосвязь показателей нелинейности и отношения сигнала к спектральной плотности мощности помехи нелинейности. Известны два показателя такого типа: в виде отношения мощности сигнала к мощности продуктов интермодуляции третьего порядка и в виде отношения спектральной плотности мощности сигнала к спектральной плотности мощности продуктов интермодуляции (в англоязычной литературе имеет название Noise Power Ratio). Первый пригоден для оценки помех по соседнему каналу для узкополосных сигналов. В модели рас-

сматривается второй показатель, так как он наилучшим образом соответствует условиям функционирования ВССпС и характеризует уровень продуктов нелинейности по всей полосе полностью загруженного транспондера [11].

Как ранее отмечалось, одной из особенностей ВССпС является ПИЧ на абонентской стороне (как в ПН, так и ОН), что позволяет увеличить пропускную способность. Помехи ПИЧ возникают между лучами с совпадающими частотами. Помехи обусловлены наличием у любого луча боковых лепестков. Гипотетическая модель для спутниковых систем высокоскоростного доступа различных диапазонов волн рассматривается в [3]. В предлагаемой модели ВССпС значение мощности сигнала к спектральной мощности помех от всех лучей с совпадающими частотами оценивается немногим более 90 дБ.

Для оценки влияния помех на значение пропускной способности суммарное значение плотности мощности помех рассматривается относительно спектральной плотности мощности тепловых шумов. В модели принято, что значение плотности всех внутрисистемных помех составляет 50 % относительно плотности мощности тепловых шумов.

При «жестких» требованиях к готовности по дождевому фактору сетевого сегмента ($T\%^* = 0,5 \%$) потери пропускной способности из-за внутрисистемных помех достигают 40 %, при более «мягких» ($T\%^* = 1 \%$) – до 50 %.

Результаты исследований на модели

Возможности модели иллюстрируются на рисунках 4–8, на которых для большей наглядности представлены расчетные значения пропускной способности только для ПН, а вклад ОН обусловлен его влиянием на неготовность дуплексного направления в целом.

Результаты оценки пропускной способности ПН в зависимости от выделенного частотного ресурса в соответствии с моделью (2, 5–12) представлены на рисунке 4.

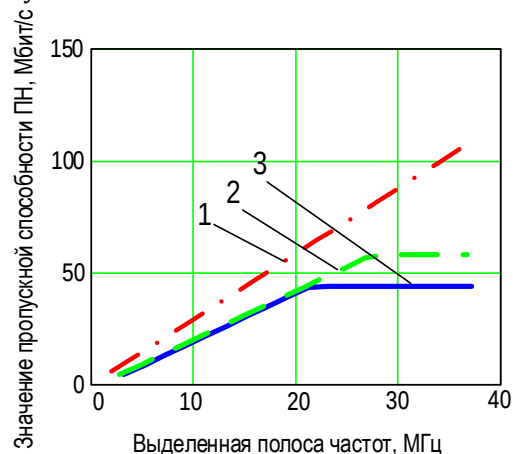


Рис. 4. Зависимость пропускной способности от выделенного частотного ресурса

Линия 1 отражает зависимость в соответствии с моделью (2). Увеличение пропускной способности возрастает при увеличении выделенного частотного ресурса. Линии 2, 3 – в соответствии с моделью (5–12). Линия 2 – при «жестких» требованиях к готовности $T^* = 0,5\%$, линия 3 – при более «мягких» требованиях $T^* = 0,75\%$. Приращение пропускной способности при увеличении доступной полосы ограничивается энергетическим ресурсом. Изменение требований к готовности приводит к изменению значения максимально достижимой пропускной способности. Так «смягчение» требований к готовности для данного примера способствует увеличению пропускной способности на 30 %.

Для проведения исследований на модели далее используются технические характеристики современных БРТК и земных станций: полоса пропускания транспондера ПН для ГрСС – 72 МГц. В ГрСС рассматривается четыре сетевых сегмента, мощность ЦЗС с учетом недоиспользования мощности по выходу в многосигнальном режиме – 50 Вт, соответственно, максимальная мощность на несущую составляет 12 Вт. Для ОН: максимальная мощность несущей на выходе усилителя мощности АЗС – 15 Вт, пропускная способность – 2 Мбит/с, используемая СКК QPSK 5/6. Требование к достоверности передачи информации для прямого и обратного направлений – $P_n = 10E - 7$.

Модель позволяет выявить ограничения по полосе и мощности, что очень важно в задачах распределения ресурса.

Рисунок 5 иллюстрирует ограничение по полосе, когда при увеличении мощности на несущую рост пропускной способности наблюдается до определенного предела, определяемого доступной полосой.

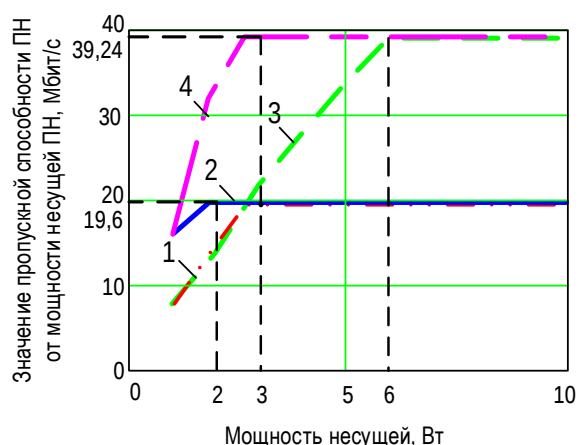


Рис. 5. Ограничение по полосе

Линии 1, 3 – для «жестких» требований к готовности ($T^* = 0,5\%$), линии 2, 4 – для «мягких» ($T^* = 3\%$). Затрат энергетического ресурса для достижения предельной пропускной способности при «мягких» требованиях нужно меньше. Прира-

щение пропускной способности ограничивается выделенным частотным ресурсом. Увеличение выделяемого частотного ресурса в 2 раза дает увеличение пропускной способности также в 2 раза. Так для сценариев 3, 4 предел пропускной способности выше, чем для сценариев 1, 2 в 2 раза.

Рисунок 6 иллюстрирует ограничение по мощности, когда при увеличении полосы на несущую рост пропускной способности наблюдается до определенного предела, определяемого доступной мощностью.

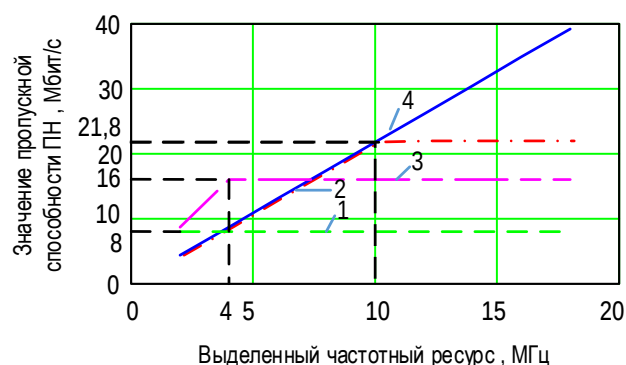


Рис. 6. Ограничение пропускной способности нехваткой энергетического ресурса

Применение энергетически эффективных СКК при «жестких» требованиях к готовности является более эффективным, чем частотно эффективных. Линия 1 при $T^* = 0,5\%$ и использовании 32 APSK 9/10, линия 2 при тех же требованиях к готовности и использовании 8 PSK 3/4. Поведение графика 1 «говорит» о недостаточности энергетического ресурса для приращения пропускной способности. Приращение частотного ресурса после определенного значения к приращению пропускной способности не приводит так как ее предел ограничивается выделенным энергетическим ресурсом. Даже при более «мягких» требованиях к готовности предел пропускной способности при применении частотно эффективных СКК ниже, чем при «жестких» требованиях и применении энергетически эффективных СКК. Для сравнения линия 2 – при $T^* = 0,5\%$ и СКК – 8 PSK 3/4, линия 3 – при $T^* = 3\%$ и СКК – 32 APSK 9/10. Очевидно, что применение СКК меньшего порядка является наиболее эффективным при достаточном частотном ресурсе, «мягких» требованиях к готовности: линия 4 при $T^* = 3\%$ и использовании 8 PSK 3/4.

Особенность предлагаемой модели подчеркивает зависимость пропускной способности от требований к готовности дуплексного направления, представлена на рисунке 7. Эта зависимость также отражает взаимосвязь прямого и обратного направлений.

Самое важное – резкое снижение пропускной способности при жестких требованиях к готовности ($T^* < 0,5\%$).

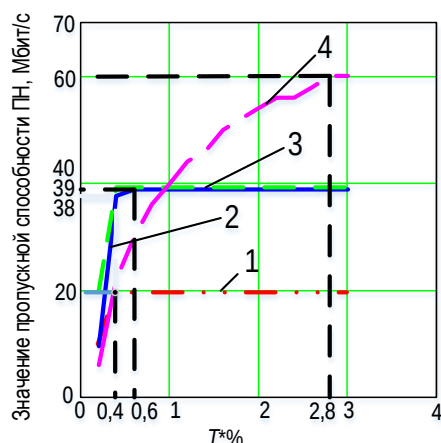


Рис. 7. Зависимость пропускной способности от требований к готовности

При умеренных требованиях к готовности увеличение доступного частотного ресурса в 2 раза способствует увеличению пропускной способности в 2 раза (линии 1, 2). При увеличении мощности несущей в 2 раза способствует достижению максимальной пропускной способности при более «жестких» требованиях к готовности (линии 2, 3). Использование СКК более высокого порядка требует «смягчения» требований к готовности (линии 3, 4).

Взаимосвязь ПН и ОН проявляется и в зависимости пропускной способности прямого направления от требований к пропускной способности обратного направления (рисунок 8).

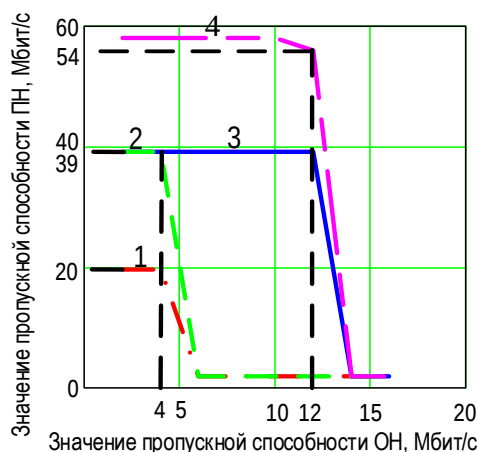


Рис. 8. Связь пропускной способности ПН с пропускной способностью ОН

Применение модели позволяет определить значение пропускной способности ОН, при котором значение пропускной способности ПН поддерживается на необходимом уровне с учетом выполне-

ния требований к готовности, а также «пороговое» значение пропускной способности ОН, при превышении которого значение пропускной способности ПН резко снижается.

Так «смягчение» требования к готовности без изменения выделенной полосы пропускания позволяют увеличить значение пропускной способности ПН: при использовании 8 PSK 3/4 в 2 раза (линии 1, 2) и при использовании 32 APSK 9/10 в 1,5 раза (линии 3, 4). Увеличение выделяемой полосы пропускания при одинаковых требованиях готовности способствует увеличению «порогового» значения пропускной способности ОН. Для сравнения линии 1, 3 при $T\%^* = 0,5\%$ и линии 2, 4 при $T\%^* = 3\%$. При увеличении полосы пропускания в 2 раза способствует увеличению «порогового» значения пропускной способности ОН в 3 раза.

Выводы

1. Предложенная модель радиолинии в составе ВССпС отличается включением в функционал для пропускной способности требования к готовности дуплексного направления спутниковой связи. В данной форме модель позволяет производить оценки не только пропускной способности линий связи и сетевых сегментов, но и решать задачи распределения ЧЭР для достижения целей обеспечения заданной пропускной способности и перераспределения ЧЭР.

2. Разработанная модель связывает варьируемые параметры – мощность несущей, полосу пропускания сетевого сегмента, вид используемой СКК и контролируемые параметры – пропускную способность прямого и обратного направлений спутниковой связи в виде функционала (5–12).

3. Включение в модель ограничивающих условий по допустимой неготовности направления связи предполагает алгоритмизацию расчетов максимальной пропускной способности прямого и обратного направлений связи.

4. Использование ограничивающих условий в явном виде позволяет реализовать наиболее полное использование выделяемого ЧЭР.

5. Результаты, полученные на основе разработанной модели, говорят об адекватности и возможности ее использования при решении задач распределения ЧЭР.

Список используемых источников

1. Inigo P., Vidal O., Roy B., Alberty E., Metzger N., Galinier D., Anzalchi J., Huggins G., Stirland S. Review of terabit/s satellite, the next generation of HTS systems // 7th Advanced Satellite Multimedia Systems Conference and the 13th Signal Processing for Space Communications Workshop (ASMS/SPSC). 2014. PP. 318–322. DOI:10.1109/ASMS-SPSC.2014.6934562.
2. Fenech H., Amos V., Hirsch A., Soumpholphakdy V. VHTS systems: Requirements and evolution // 11th European Conference on Antennas and Propagation (EuCAP). 2017. PP. 2409–2412. DOI:10.23919/EuCAP.2017.7928175.

3. Рекомендации МСЭ-RS.1782 (2007). Возможности для глобального широкополосного доступа в интернет для систем фиксированной спутниковой службы.
4. Tri T.Ha. Digital Satellite Communications (2nd addition). New York: McGraw-Hill, 1990. 370 p.
5. Marshall J., Heissler J. Satcom loading analysis with heterogeneous gain states // Proceeding of MILCOM, Anaheim, CA. 2002. PP. 136–141. DOI:10.1109/MILCOM.2002.1180428.
6. Marshall J., Jo K.Y. Non-homogeneous gain state optimization for transponding satellite communications // IEEE Military Communications Conference (MILCOM–2003), Boston, US. 2003. PP. 517–522. DOI:10.1109/MILCOM.2003.1290156.
7. Рекомендации ITU-T Rec. G.821 (12/2002).
8. Northrop G.M. Aids for the Gross Design of Satellite Communication Systems // Satellite communication systems. 1966.
9. Волков Л.Н., Немировский М.С., Шинаков Ю.С. Системы цифровой радиосвязи. Базовые методы и характеристики: учеб. пособие. М: Эко-Трендз, 2005. 392 с.
10. Рекомендация МСЭ-R P.618-12 (07/2015). Данные о распространении радиоволн и методы прогнозирования, необходимые для проектирования систем связи Земля-космос.
11. Peters R., Woolner P., Ekelman E. Analytic Calculation of Noise Power Robbing, NPR and Polarization Isolation Degradation // 26th International Communications Satellite Systems Conference (ICSSC), San Diego, CA. 2008. DOI:10.2514/6.2008-5509.

* * *

THE MODEL OF THE SATELLITE COMMUNICATION RADIO LINK IN THE HIGH-SPEED SATELLITE SYSTEM

V. Zhironov¹, A. Orlov², A. Smirnov²

¹Radio Research and Development Institute (LONHIR),
St. Petersburg, 192029, Russian Federation

²Telecommunications Military Academy,
St. Petersburg, 194064, Russian Federation

Article info

Article in Russian

For citation: Zhironov V., Orlov A., Smirnov A. The Model of the Satellite Communication Radio Link in the High-Speed Satellite System // Proceedings of Telecommunication Universities. 2018. Vol. 4. Iss. 3. PP. 45–53.

Abstract: *The article is devoted to the model of the satellite communication radio link in the high-speed satellite system. Unlike particular models, which take into account thermal noise, intersystem and adjacent interference, and meeting of readiness requirements with rainfall availability is controlled after the system reserve assessment, in the proposed model, the requirements for availability are explicitly taken into account in the form of a constraint. This makes it possible to set and solve the problems of optimizing the distribution of the frequency-energy resource of central earth stations and on-board retransmission complexes by the criterion of maximum bandwidth while meeting the requirements for the reliability of information transmission and the readiness of the duplex direction of satellite communications.*

Keywords: *satellite radio link, frequency-energy resource, earth station, on-board retransmission complex, high-speed satellite system.*

СТОХАСТИЧЕСКАЯ МОДЕЛЬ МЕТЕОРНОГО РАДИОКАНАЛА

Д.Д. Капралов^{1*}, Д.И. Кирик¹

¹Санкт-Петербургский государственный университет телекоммуникаций им. проф. М.А. Бонч-Бруевича, Санкт-Петербург, 193232, Российская Федерация

*Адрес для переписки: d_i_kirik@mail.ru

Информация о статье

УДК 621.396.24

Язык статьи – русский

Ссылка для цитирования: Капралов Д.Д., Кирик Д.И. Стохастическая модель метеорного радиоканала // Труды учебных заведений связи. 2018. Т. 4. № 3. С. 54–64.

Аннотация: В статье предложена стохастическая модель метеорного радиоканала, отличающаяся от известных тем, что учитывает статистический характер возникновения следов различного типа. По результатам проведенных исследований сформулированы начальные значения параметров квадратурных составляющих сигнала, отраженного от следов различного типа.

Ключевые слова: метеорная радиосвязь, радиоволны, стохастическая модель.

Введение

Развитие современной аппаратуры метеорной радиосвязи (МРС), как правило, характеризуется меньшими габаритами и массой, энергопотреблением и стоимостью по сравнению с аппаратурой других родов радиосвязи. Как показывают теоретические и экспериментальные исследования, МРС может быть с успехом использована для организации сбора данных с удаленных объектов, передачи сигналов метеорологической обстановки и т.д.

Однако метеорная радиосвязь характеризуется рядом существенных недостатков, сдерживающих ее широкое применение. Это, прежде всего, прерывистый характер процесса передачи сообщений в линиях МРС, случайность и широкие пределы изменений значений параметров используемых для связи метеорных следов, сравнительно низкая средняя скорость передачи информации и большое время задержки при передаче сообщений.

Существует большое количество работ, в которых предлагаются различные модели метеорного радиоканала. В частности, в работе [1] отражены от метеорных следов сигналы делятся на два класса – недоуплотненные и переуплотненные, каждый из которых описывается изменением огибающей сигнала в виде детерминированной функции времени, зависящей от нескольких случайных величин (начальное значение огибающей – для недоуплотненного следа, максимальное значение огибающей для переуплотненного следа, средняя длительность интервала времени между следами и средняя длительность следа). Подавляющее

большинство работ [2–5] использует еще более простую модель канала МРС, основанную на представлении сигнала с экспоненциально убывающей огибающей (недоуплотненные следы), при этом в зависимости от решаемой задачи в работах полагаются случайными либо детерминированными параметры огибающей.

В работе [6] разработана модель, учитывающая случайные амплитудные и фазовые изменения сигналов на длительности отражения от метеорного следа, специфические особенности отражений от метеорных следов – эффекты диффузии, рекомбинации, прилипания, однако к недостатку этой модели следует отнести ее непригодность для решения задач анализа эффективности проектируемых систем передачи, так как для проведения моделирования неизвестны распределения начальных значений параметров для следов различных типов.

1. Особенности метеорного распространения радиоволн

Для определения особенностей метеорного распространения радиоволн, необходимо определить принцип происхождения метеорных следов. Результаты исследований позволяют представить следующую картину возникновения метеорных следов [1]. Земная атмосфера непрерывно бомбардируется падающими из космоса метеорными частицами различных масс и скоростей. В сутки на Землю падает $90 \cdot 10^6$ частиц с массой большей $5,5 \cdot 10^{-4}$ гр. Интенсивность потока метеоров зависит от времени суток и от времени года [8]. Мак-

симум приходится на 6 часов утра и на осень, а минимум – на 6 часов вечера и на весну.

Оставшийся после торможения метеора запас кинетической энергии расходуется на разогревание и плавление частиц, а также на испарение и ионизацию атомов частиц. Испарившиеся атомы, обладая большим запасом энергии, производят ионизацию нейтральных молекул воздуха, в основном молекул кислорода. Таким образом, после пролета метеорной частицы остается ионизированный след. Плотность ионизации следа характеризуется величиной линейной электронной плотности, равной количеству образовавшихся свободных электронов на 1 м длины следа. Линейная электронная плотность не остается постоянной вдоль следа и по мере проникновения метеора в атмосферу q плавно возрастает до величины [1]:

$$q_m = 10^{-15} m_m v^3 \cos Z / (\mu_m H), \quad (1)$$

где m_m – первоначальная масса метеора; v – геоцентрическая скорость метеора; Z – угол между осью следа и зенитом; μ_m – атомный вес частиц метеора; H – приведенная высота атмосферы, которая затем уменьшается до нулевого значения. Линейная электронная плотность является функцией высоты над поверхностью Земли. Большая часть пригодных для связи метеорных следов образуется на высоте 85–110 км.

Вылетающие с большой скоростью атомы и молекулы метеорной частицы до момента установления теплового равновесия с окружающей средой успевают удалиться от оси следа на некоторое расстояние, создавая при этом ионизированный столб с эффективным радиусом r_0 , называемым начальным радиусом метеорного следа [7, 8]. Величина начального радиуса является функцией скорости и высоты образования метеорного следа, и по данным [9] его можно аппроксимировать выражением:

$$r_0 = 10^{-2} \sqrt{v} / 6 \cdot \exp(0,05h). \quad (2)$$

Величина начального радиуса метеорного следа определяет начальную объемную электронную плотность следа. Объемная плотность ионизации метеорных следов 10^{12} эл/м³ достаточна для эффективного отражения радиоволн метрового и дециметрового диапазонов. Отражение радиоволн от следа зеркально, т.е. углы между следом и падающим лучом и следом и отраженным лучом равны. Если линейная электронная плотность следа в точке отражения меньше критической q_c , определяемой как [9]:

$$q_c = 7,5 \cdot 10^{13} \exp\left(\frac{16\pi^2 r_0^2}{3\lambda^2 \sec^2 \Phi}\right), \quad (3)$$

где Φ – половина угла между падающим и отраженным лучом; λ – длина волны, рассеивание которой происходит за счет когерентного и независимого

переизлучения всех электронов следа. Следы такого типа называются недоуплотненными.

Сразу же после образования метеорный след начинает расширяться за счет действия амбиполярной диффузии. Она заставляет положительные ионы, находящиеся в ионизированном следе, сдерживать диффузию электронов. В результате этого след будет расширяться медленнее, чем в случае, если бы он состоял из одних лишь электронов. С увеличением диаметра следа рассеянный сигнал претерпевает все более значительные изменения вследствие интерференции полей, возникающих в результате рассеяния от отдельных электронов, что приводит к экспоненциальному спаду амплитуды принимаемого сигнала [9]:

$$U(t) = U_{\max} \exp\left\{-\frac{16\pi^2 D t}{\lambda^2 \sec^2 \Phi}\right\}, \quad (4)$$

где D – коэффициент амбиполярной диффузии, зависящий от высоты точки зеркального отражения, определяемый в соответствии с выражением:

$$D = 10^{(0,067h-5,6)}. \quad (5)$$

Если линейная электронная плотность в точке отражения больше критической, то заметным становится вторичное рассеяние от электрона к электрону, т.е. рассеяние перестает быть независимым и когерентным. За счет рефракции во внешних слоях следа радиоволны не доходят до оси следа на некоторое расстояние. Это расстояние называется критическим радиусом следа. Величина критического радиуса зависит от объемной электронной плотности ионизированных частиц следа. Критический радиус – это расстояние от оси следа, на котором диэлектрическая проницаемость следа становится равной нулю. Величина критического радиуса r_c определяется в виде [9]:

$$r_c = \sqrt{(4Dt + r_0^2) \ln(Q/Q_c)}, \quad (6)$$

где Q_c – критическая объемная электронная плотность, определяемая при условии равенства нулю диэлектрической проницаемости следа на частоте радиосигнала f (кГц).

За счет амбиполярной диффузии радиус следа через время t увеличивается до величины:

$$r_c = \sqrt{(4Dt + r_0^2)}. \quad (7)$$

С увеличением радиуса следа происходит уменьшение объемной электронной плотности следа до величины:

$$Q = \frac{Q_c}{\pi(4Dt + r_0^2)}. \quad (8)$$

В результате на протяжении некоторого времени после пролета метеором зеркальной точки отражения происходит увеличение радиуса следа. Эквивалентная отражающая поверхность следа,

которую можно представить в виде цилиндра радиуса r , возрастает и увеличивается мощность отраженного сигнала. Через интервал времени [9]:

$$t = \frac{\mu_0 e_3^2}{4m} \cdot \frac{q\lambda^2 \sec^2 \Phi}{4\pi^3 D e} \quad (9)$$

отраженная мощность достигает максимального значения [9]:

$$P_M = \sqrt{\frac{\mu_0 e_3^2 \pi q}{4me}} \cdot \frac{1}{32\pi^4} \cdot \frac{1}{P_T \lambda^3 \cos^2 \varphi} \cdot \frac{1}{r_1 r_2 (r_1 + r_2) (1 - \cos^2 \beta \cdot \sin^2 \Phi)} \quad (10)$$

где $\mu_0 = 4 \cdot 10^{-7}$ Гн/м; e_3^2 и m – заряд и масса электрона; P_T – мощность, излучаемая передатчиком; e – основание натурального логарифма; $\cos \varphi$ – множитель поляризации, зависящий от угла φ между векторами поляризации волн, падающих на след и принимаемой антенной; r_1 , r_2 – расстояние зеркального отражения на следе до приемника и передатчика; β – угол между осью следа и плоскостью, проходящей через точки передачи, отражения и приема, называемой плоскостью связи:

$$T_n = 1,13 \cdot \frac{10^{-14} q \lambda^2 \sec^2 \Phi}{16\pi^2 D} \quad (11)$$

Через время диэлектрическая проницаемость следа на оси становится равной нулю, и величина отраженной мощности спадает по экспоненциальному закону. Следы такого типа называются переуплотненными.

Третий тип отражения – отражение от поворотных метеорных следов, возрастание и падение амплитуды сигнала для которого вызвано разворотом метеорного следа атмосферными ветрами и происходит значительно медленнее, чем для отражений от переуплотненных следов.

Метеорные следы всех типов постепенно искривляются и разрываются на отдельные куски атмосферными ветрами и турбулентностью, что ускоряет их разрушение и приводит к появлению дополнительных зеркальных точек. Через некоторое время после формирования ионизированного следа в области первой зоны Френеля появляются дополнительные отражающие точки. Интерференция сигналов, отраженных от этих точек, приводит к замираниям сигналов в точке приема. Обычно замирания начинают наблюдаться лишь спустя 0,2 с после пролета метеором первой зоны Френеля, поэтому кратковременные отражения, соответствующие, в основном недоуплотненным следам, не подвержены замираниям. С течением времени число отражающих точек на следе в результате поворота отдельных его участков увеличивается. К концу отражения объемная электронная плотность быстро уменьшается, что приводит

к быстрому сокращению числа отраженных точек и уменьшению интенсивности замираний.

Замирания наблюдаются при отражении от следов всех типов, а вклад числа отражений с замираниями колеблется от 8 до 40 % для следов первого типа, от 35 до 60 % для следов второго типа и от 60 до 95 % для следов третьего типа.

Установлено, что процентный вклад числа отражений с замираниями является функцией длительности отражения t_c и эффективной длины волны $\lambda_3 = \lambda \cdot \sec \Phi$ и не зависит в первом приближении от минимальной регистрируемой электронной плотности, времени суток и года. Период замираний T_3 является функцией длительности отражения t_c и эффективной длины волны:

$$T_3 = (2,4 \cdot 10^{-2} \cdot \lambda_3 + 0,07) \cdot \lg(t_c) + 1,6 \cdot 10^{-2} \cdot \lambda_3 + 0,27. \quad (12)$$

Как при фиксированном времени, так и при фиксированном замирании распределение близко к Реллеевскому [1].

Вклад каждого типа отражений в общее число отражений зависит от длины трассы, ее ориентации, географических координат и уровня минимальной регистрируемой мощности сигнала. В зависимости от указанных параметров он может составлять следующие значения: процентный вклад числа отражений первого типа – от 50 до 70 %; второго типа – от 10 до 30 %; третьего типа – до 20 %.

Следует заметить, что деление следов на три типа достаточно условное и не исчерпывает всего многообразия возникающих видов отражений. Однако, основываясь на таком представлении процессов отражения и рассеивания, можно перейти к статистическому описанию сигналов на входе демодулятора станции метеорной радиосвязи.

Учитывая случайный характер появления метеорных следов, важным вопросом при изучении метеорного рассеяния радиоволн является исследование численности и длительности метеорных радиоотражений. Большое число работ [1, 9–14] посвящено различным методикам оценки числа отражений, зарегистрированных за единицу времени N и их длительности τ , являющихся случайными величинами. Установлено, что число отражений распределено по закону Пуассона, а длительность зарегистрированных сигналов распределена по экспоненциальному закону для недоуплотненных следов:

$$F(\tau) = \exp\left(-\frac{16\pi^2 D \tau}{\lambda^2 \sec^2 \Phi}\right), \quad (13)$$

где D – коэффициент амбиполярной диффузии.

$$F(\tau) = \tau^{-K}, \quad (14)$$

где K для различных трасс и времени суток изменяется от 0,8 до 1,1 [9].

Метеорные следы, удовлетворяющие условию зеркальности, являются потенциально пригодными для обеспечения связи. Конечная чувствительность приемной аппаратуры позволяет использовать для связи только метеорные следы, создающие в точке приема отраженные сигналы, превышающие некоторый пороговый уровень. Рассмотрим стохастическую модель метеорного радиоканала.

2. Стохастическая модель метеорного радиоканала

При построении модели непрерывного канала метеорной радиосвязи необходимо обеспечить адекватное отражение процессов, происходящих с сигналом в канале, и получить реализуемые аналитические решения задач анализа и взаимодействия систем передач сообщений.

Наиболее широким классом случайных процессов, включающим гауссовские и негауссовские, непрерывные и импульсные процессы с непрерывным и дискретным множеством состояний, являются марковские процессы.

Представим поступающую на вход демодулятора приемника смесь модулированного полезного сигнала $S(t, \vec{\theta}, \vec{\mu})$ и флуктуационного шума $n(t)$ в виде:

$$y(t) = S(t, \vec{\theta}, \vec{\mu}) + n(t), \quad (15)$$

где флуктуационный шум $n(t)$ обладает свойствами:

$$\langle n(t) \rangle = 0, \langle n(t)n(t+\tau) \rangle = 0,5N_0\delta(\tau). \quad (16)$$

При выборе вида сигнала для работы по МРК необходимо учесть и особенности метеорного распространения радиоволн, и степень искажения закона изменения огибающей последовательности сигналов при нелинейном режиме работы усилителя мощности радиопередатчика. В течение времени существования метеорного следа изменяется линейная электронная плотность в точке отражения, что приводит к случайному характеру изменения амплитуды отраженного сигнала. Поэтому для передачи информации по МРК не представляется возможным использование сигналов с амплитудной модуляцией. Таким образом, отсутствие паразитной модуляции сигналов при использовании модуляции с минимальным частотным сдвигом (ММС) позволяет использовать усилитель мощности в режиме класса «С» с отсечкой менее 90° , следовательно, при фиксированной пиковой мощности можно повысить среднюю мощность. Представим полезный ММС-сигнал $S(t, \vec{\theta}, \vec{\mu})$, отраженный от метеорных следов, следующим выражением:

$$\begin{aligned} S(t, \vec{\theta}, \vec{\mu}) = & \sum_{i=1}^{\infty} \sigma(t - t_i) g(t - t_i) \{ \gamma (\mu_H^S \cos(w_1 t + \phi(t, \vec{\theta})) + \\ & + \mu_H^S \sin(w_1 t + \phi(t, \vec{\theta}))) + (1 - \gamma) (\mu_H^S \cos(w_1 t + \phi(t, \vec{\theta})) + \\ & + \mu_H^S \sin(w_1 t + \phi(t, \vec{\theta}))) \}, \end{aligned} \quad (17)$$

где $g(t - t_i)$ – форма огибающей элемента сигнала на i -ом следе; t_i – его случайная задержка на i -ом следе; w_1 – меньшая из угловых частот сигнала; $\mu_H^S(t) - k = \{n, p\}$ и μ_H^S – квадратурные составляющие полезного сигнала, отраженные соответственно от недоуплотненных и переуплотненных следов; γ – индикаторная переменная, характеризующая появление недоуплотненного или переуплотненного следа, принимающая значение 0 или 1 с вероятностями $P_1 = P(\gamma = 1)$, $P_0 = P(\gamma = 0)$; $\sigma(t - t_i)$ – функция Хэвисайда, принимающая значение равно 1 в момент времени t_i появления i -го метеорного следа; $\phi(t, \vec{\theta})$ – функция, описывающая изменение фазы сигнала в зависимости от значения вектора дискретных параметров $\vec{\theta} = \{\Sigma_k, \Theta_k\}$. На длительности T k -ой информационной посылки $t \in [kT, (k+1)T]$ функция $\phi(t, \vec{\theta})$ описывается следующим выражением:

$$\phi(t, \vec{\theta}) = \Sigma_k + \Theta_k \frac{\pi}{T} (t - kT), \quad (18)$$

где Θ_k – дискретный информационный параметр, принимающий значение $\Theta = 1$ или $\Theta = 0$ с равной вероятностью на k -ой информационной посылке; $\Sigma_k = \pi \sum_{q=1}^{k-1} \Theta_q$ – начальная фаза сигнала на k -ом тактовом интервале, равная фазовому набегу за счет манипуляции частоты сигнала на предыдущих $k - 1$ интервалах.

Стохастическая модель сигнала, отраженного от метеорного следа, была предложена в работе [6]. В соответствии с результатами многочисленных наблюдений процессов [1, 9], происходящих при отражении сигналов от метеорных следов различных типов, при описании квадратурных составляющих полагалось, что отражение радиосигналов от недоуплотненных следов характеризуется экспоненциальным спадом напряженности поля в точке приема. Отражение от переуплотненных следов характеризуется тем, что напряженность поля в начале следа растет по закону, близкому к логарифмическому [1], а затем уменьшается в соответствии с законом, близким к экспоненциальному. Отражение от поворотных следов происходит аналогично отражению от переуплотненных, но напряженность поля изменяется значительно медленнее [9].

Процесс формирования поля в месте приема можно представить, как результат рассеяния радиоволн на «горячих» пятнах. В начальный момент времени, сразу после образования метеорного следа, поле в точке приема представляет собой множество когерентных волн, которые с течением времени ввиду расширения и искажения формы следа становятся все более независимыми. Одновременно с расширением следа происходит уменьшение его концентрации за счет амбиполярной диффузии, что приводит к уменьшению интенсивности «горячих» пятен [1]. Вклад отдельных слагаемых в

суммарный процесс очень мал, что вместе с предположением о слабой зависимости отдельных составляемых при рассеянии в канале позволяет использовать центральную предельную теорему и считать коэффициент передачи среды нестационарным процессом с гауссовским распределением. В отличие от процесса, аппроксимирующего рассеяние на неоднородностях ионосферы, процесс, характеризующий рассеяние от метеорных следов, будет нестационарным. Его дисперсия в начальный момент времени, соответствующий моменту образования следа, равна нулю, затем дисперсия вследствие расширения следа увеличивается и, с некоторого момента, дисперсия уменьшается вследствие уменьшения электронной плотности следа.

Таким образом, квадратурные составляющие полезного сигнала, отраженного от метеорного следа $\mu^c(t)$ и $\mu^s(t)$, описываются следующей системой дифференциальных уравнений [6] (нижний индекс $k = \{n, p\}$ в уравнениях опущен, так тип следа определяется значениями постоянных параметров и начальными значениями квадратурных составляющих):

$$\frac{d\mu^{(i)}}{dt} = -\alpha_c \mu^{(i)} - \mu_{cp}^{(i)}(1 + \xi^{(i)}), \quad (19)$$

$$\frac{d\mu_{cp}^{(i)}}{dt} = -\beta \mu_{cp}^{(i)} - D \mu_1^{(i)}, \quad (20)$$

$$\frac{d\mu_1^{(i)}}{dt} = -(\beta + D) \mu_1^{(i)}, \quad (21)$$

$$\frac{d\xi^{(i)}}{dt} = -2\eta \xi^{(i)} - \omega^2 \xi_1^{(i)} + \omega^2 n^{(i)}(t), \quad (22)$$

$$\frac{d\xi_1^{(i)}}{dt} = \xi^{(i)}, \quad (23)$$

где переменным квадратурной составляющей сигнала, отраженного от метеорного следа соответствует индекс $i = c$; η – параметр, характеризующий спектральные свойства случайного процесса $\xi^{(i)}$; индекс $i = s$ является переменной синфазной составляющей; α_c – величина, определяющая ширину спектра флуктуаций квадратурной составляющей: $\alpha_c = 1/\tau_c$, (где τ_c – время корреляции квадратурной составляющей); $n^{(i)}(t)$ – гауссовы шумы со свойствами: $\langle n^{(i)}(t) \rangle = 0$, $\langle n^{(i)}(t)n^{(i)}(t + \tau) \rangle = 0,5\delta^{(i)}N^{(i)}\delta(\tau)$; $\mu_{cp}^{(i)}$ – среднее значение, к которому стремится квадратурная составляющая сигнала по мере разрушения следа за счет амбиполярной диффузии и эффектов рекомбинации и прилипания; D – коэффициент амбиполярной диффузии, значение которого определяется в соответствии с данными; коэффициент β равен сумме коэффициентов прилипания γ_{Π} и рекомбинации γ_p ; $\mu_1^{(i)}$ – величина, характеризующая изменения среднего значения $\mu_{cp}^{(i)}$ квадратурной составляющей сигнала.

Для удовлетворительной аппроксимации сигналов, отраженных от следов всех типов в каче-

стве модели формирующего шума использован квазигармонический случайный процесс – $\xi^{(i)}$, характеристики которого определяются характеристиками замирающих сигналов, в частности, центральная частота процесса ω определяется в соответствии с выражением:

$$\omega = \frac{2\pi}{T_3}, \quad (24)$$

где T_3 – период замираний, вычисляемый в соответствии с (25).

Процесс возникновения метеорных следов, пригодных для связи, может быть охарактеризован экспоненциальным распределением длительности интервалов времени между отраженными сигналами с параметром распределения:

$$\nu = \frac{\zeta}{T_{CP}}, \quad (25)$$

$$T_{CP} = T_{CP0} \left[\frac{G_T G_R P_T}{10R u_p^2} \right]^{-0.6} \left(\frac{f}{37,5^{2.4} g(r)} \right), \quad (26)$$

где $\zeta = p(u_p)$ – характеристика обнаружителя, определяемая как вероятность правильного обнаружения сигналов, пригодных для связи; T_{CP} – средняя длительность интервалов между сигналами; P_T – мощность передатчика; G_T, G_R – коэффициенты усиления передающей и приемной антенн, соответственно; u_p – значение уровня регистрации отраженных сигналов; R – скорость передачи; f – частота передачи, кГц; T_{CP0} – среднее значение интервалов времени между метеорными следами, соответствующее системе со следующими параметрами $G_T = G_R = 10$, $P_T = 200$ Вт, $R = 2000$ Бод, $f = 37500$ кГц, $u_p = u_{p0}$; $g(r)$ – множитель, определяемый в зависимости от протяженности трассы в километрах следующим образом:

$$g(r) = \begin{cases} 1,72, & \text{если } 200 < r < 480 \\ 770/r, & \text{если } 480 < r < 770 \\ 1, & \text{если } 770 < r < 1280 \\ (1 - 0,0006(r - 1280))^{-1}, & \text{если } 1280 < r < 2000 \end{cases} \quad (27)$$

В качестве модели модулирующего процесса $\lambda(t)$ будем использовать случайный процесс, описываемый системой дифференциальных стохастических уравнений [8]:

$$\begin{cases} \frac{d\lambda}{dt} = \lambda_1 \\ \frac{d\lambda_1}{dt} = -2\alpha_{\Pi}\lambda_1 - \nu^2\lambda + \nu^2 n_{\lambda}(t) \end{cases}, \quad (28)$$

где $n_{\lambda}(t)$ – гауссовские шумы, имеющие свойства: $\langle n_{\lambda}(t) \rangle = 0$, $\langle n_{\lambda}(t)n_{\lambda}(t + \tau) \rangle = 0,5N\delta(\tau)$; $N = 4A\alpha_{\Pi}$; α_{Π} – ширина одностороннего энергетического спектра модулирующего сигнала; A – дисперсия процесса $\lambda(t)$ равная средней мощности помехи; ν – основной тон, который примем равным 1 кГц.

Таким образом, данная модель позволяет описать на входе демодулятора станции метеорной

связи аддитивную смесь сигнала, отраженного от следа любого типа, флуктуационного шума. Тип и длительность следа, характер изменения квадратурных составляющих определяются начальными значениями квадратурных составляющих $\mu^{(i)}$, начальными значениями переменных $\mu_{cp}^{(i)}$ и $\mu_1^{(i)}$, а также параметрами, характеризующими скорость изменения квадратурных составляющих и параметров квазигармонического процесса, глубину и длительность замираний сигнала. В работе [6] не определены пределы изменения параметров стохастических дифференциальных уравнений и не определен способ задания начальных значений для квадратурных составляющих и вспомогательных переменных для следов различных типов и длительностей. Для устранения ограничений в применимости указанной модели, разработан алгоритм задания начальных значений квадратурных составляющих для моделирования следов различных типов.

3. Алгоритм задания начальных значений параметров модели сигналов, отраженных от метеорных следов

При проектировании метеорных систем связи используется параметрический анализ, позволяющий оценить влияние на характеристики систем изменения параметров, как самой системы, так и изменение параметров линии. При моделировании радиолинии для определения значения сигнала на входе демодулятора используют уравнение радиолинии, в котором, как правило, априорно известны значения параметров или известны их априорные распределения.

Обычно изменение мощности сигнала на входе демодулятора, отраженного от недоуплотненного следа, описывается известным уравнением [1]:

$$P_{\text{вх}} = \frac{P_{\Pi} G_{\Pi} G_{\Pi\Pi} \lambda^3 q^2 r_e^2 \sin^2 \varphi}{16\pi^2 r_1 r_2 (r_1 + r_2) (1 - \cos^2 \beta \cdot \sin^2 \Phi)} \times \exp \left\{ -\frac{16\pi^2 D t}{\lambda^2 \sec^2 \Phi} \right\}, \quad (29)$$

где r_e – радиус электрона; P_{Π} – мощность, излучаемая передатчиком; G_{Π} , $G_{\Pi\Pi}$ – коэффициенты усиления соответственно передающей и приемной антенн. Максимум мощности принимаемого сигнала от недоуплотненных следов находится вычислением (29) при $t = 0$ [8]:

$$P_{\text{max}} = \frac{P_{\Pi} G_{\Pi} G_{\Pi\Pi} \lambda^3 q^2 r_e^2 \sin^2 \varphi}{16\pi^2 r_1 r_2 (r_1 + r_2) (1 - \cos^2 \beta \cdot \sin^2 \Phi)}. \quad (30)$$

Это значение мощности фактически и является начальным значением мощности сигнала для недоуплотненных следов. В пределах всей диаграммы направленности приемопередающих антенн встречаются следы с различной ориентацией и линейной плотностью. Будем считать, что параметры φ , q , r_1 , r_2 , G_{Π} , $G_{\Pi\Pi}$, β , Φ меняются от следа к

следующему. Линейная электронная плотность q имеет функцию плотности вероятности, описываемую выражением [8]:

$$f(q) = \frac{q_{\min} q_{\max}}{(q_{\max} - q_{\min}) q^2} [\sigma(q - q_{\min}) - \sigma(q - q_{\max})], \quad (31)$$

где $q_{\min}$ и $q_{\max}$ – минимальная и максимальная линейные электронные плотности следов, которые для недоуплотненных следов составляют 10^{12} и 10^{14} электрон/м; $\sigma(x)$ – функция Хэвисайда. Коэффициент ориентации следа β равномерно распределен на интервале $(0, \pi/2)$. Множитель поляризации Φ равномерно распределен на интервале $(\pi/2, \pi)$. Значения параметров r_1 , r_2 , G_{Π} , $G_{\Pi\Pi}$, Φ равномерно распределены на относительно узком интервале вокруг среднего значения [7].

Вычисление функции распределения максимумов уровней недоуплотненных следов как функции от уровня принимаемого сигнала в соответствии с (31) представляет собой вычисление функции распределения с восемью случайными параметрами. Использование метода вспомогательных переменных величин требует расчета матрицы Якоби восьмого порядка, а затем интегрирование по всем восьми вспомогательным переменным для получения функции плотности вероятности, как функции от уровня принимаемого сигнала. Этот процесс не приводит к решению в строгой форме для распределения амплитуд следов.

Однако вместо абсолютных значений мощности принимаемого сигнала воспользуемся логарифмической шкалой в децибелах. Тогда выражение (30) можно записать в следующем виде [7]:

$$P_{dBw} = 10(\lg P + \lg G_{\Pi} + \lg G_{\Pi\Pi} - \lg(r_1 r_2 (r_1 + r_2))) + 30 + 2\lg q + 2\lg \sin \varphi - \lg(1 - \cos^2 \beta \cdot \sin^2 \Phi), \quad (32)$$

где $P = \frac{P_T \lambda^3 r_e^2}{16\pi^2}$ включает в себя все константы линии. Функция плотности вероятности суммы независимых случайных величин вычисляется как свертка плотности вероятности этих независимых величин.

Центральная предельная теорема показывает, что при определенных условиях функция плотности вероятности суммы независимых случайных величин с ростом их числа сводится к гауссовской функции плотности. Известно, что произвольная последовательность случайных величин будет в пределе подчиняться предельной теореме, если сами величины однородно ограничены, т.е. абсолютное значение всех случайных величин конечно. Этому условию удовлетворяют все восемь случайных величин φ , q , r_1 , r_2 , G_{Π} , $G_{\Pi\Pi}$, β , Φ выражения (32).

Таким образом, согласно центральной предельной теореме, функция плотности распределения начальных значений мощности в точке приема сигналов, отраженных от недоуплотненных следов

аппроксимируется гауссовской функцией плотности вероятностей [8]:

$$f(P_{dBw}) = \frac{1}{\sqrt{2\pi}\sigma_P} \cdot \exp\left(-\frac{(P_{dBw} - \bar{P}_{dBw})^2}{2\sigma_P^2}\right), \quad (33)$$

где $\bar{P}_{dBw}$ – среднее значение мощности сигнала на входе приемника; σ_P^2 – дисперсия. В работе [9] показано, что уже при трех случайных величинах из восьми функция распределения мощности сигнала на входе демодулятора достаточно точно аппроксимируется гауссовским распределением. Сравнение теоретической модели (33) с практическими результатами затруднено из-за конечной чувствительности приемника. Регистрируются лишь следы с мощностью превышающей некоторый порог $P_{пор}$. Сигналы, отраженные от остальных следов, не регистрируются.

В таблице 1 представлены значения математического ожидания и дисперсии гауссовского распределения начальных значений мощности сигнала, а также длины выборок N при работе на частотах 45000, 65000, 85000, 104000 кГц.

ТАБЛИЦА 1. Результаты исследования математического ожидания и дисперсии на различных частотах

Месяц	Частота, кГц	$\bar{P}_{dBw}$, дБВт	σ_P^2	N
Февраль	45000	-126,3012	11,14442	14192
Февраль	65000	-123,7344	9,12436	5172
Февраль	85000	-149,0718	15,82285	588
Февраль	104000	-187,1842	43,86346	2860
Март	45000	-131,1799	12,52036	13148
Март	65000	-124,0943	9,203368	5234
Март	85000	-125,2136	9,185085	3012
Март	104000	-140,5984	13,99287	2910
Апрель	45000	-125,3674	11,95014	15525
Апрель	65000	-124,4652	10,32820	7269
Апрель	85000	-126,1734	10,20671	9405
Апрель	104000	-130,6117	11,89659	4383
Май	45000	-122,4071	7,356731	15450
Май	65000	-124,6968	9,294118	9447
Май	85000	-125,2803	9,113829	6227
Май	104000	-125,5103	8,784910	6890
Июнь	45000	-130,3900	10,43953	7985
Июнь	65000	-127,5174	10,48344	11619
Июнь	85	-125,6116	9,389750	9043
Июнь	104	-127,1316	9,626522	9076

Приведенные в таблице 1 данные позволяют задавать начальные значения параметров при моделировании, используя систему дифференциальных уравнений (19–23), квадратурных составляющих сигнала, отраженного от недоуплотненных следов различной амплитуды и длительности.

Используя математическое ожидание $\bar{P}_{dBw}$ и дисперсию σ_P^2 из таблицы 1, генерируется по гауссовскому закону значение мощности на входе приемника P_{dBw} , которое пересчитывается в напряжение огибающей U :

$$U = 10^{0,05 \cdot P_{dBw}}. \quad (34)$$

Значение начального значения одной квадратурной составляющей отраженного сигнала, например, μ_0^C генерируется по равномерному закону в пределах от 0 до U , а значение другой вычисляется по формуле:

$$\mu_0^S = \sqrt{U^2 - (\mu_0^C)^2}. \quad (35)$$

Как ранее отмечалось, что отражение сигнала от недоуплотненного следа характеризуется экспоненциальным спадом амплитуды, поэтому μ_{cr0}^C и μ_{cr0}^S , т.е. начальные значения средних значений квадратурных составляющих, незначительно отличаются от значений самих квадратур. Начальные значения средних квадратурных составляющих μ_{cr0}^C и μ_{cr0}^S генерируются по равномерным законам в пределах $(\mu_0^C, 1,1 \cdot \mu_0^C)$ и $(\mu_0^S, 1,1 \cdot \mu_0^S)$.

Считая, что в уравнениях (20), (21) коэффициенты амбиполярной диффузии, рекомбинации и прилипания не изменяются от следа к следу, т.е. являются постоянными величинами, остаются еще два случайных параметра недоуплотненных следов: $\alpha_c = 1/\tau_c$, μ_{10}^C и μ_{10}^S . Параметр $\alpha_c = 1/\tau_c$, отражающий спектральные свойства процесса, обратно пропорционален средней длительности интервала корреляции. Известно [7], что длительности сигналов, отраженных от недоуплотненных следов, составляют от 0,05 до 0,5 с. Так как среднее время корреляции на порядок меньше длительности сигнала [7, 8], полагаем значение параметра α_c равномерно распределенным в пределах от 0,02 до 0,2 кГц.

Вспомогательные параметры μ_1^C и μ_1^S влияют на скорость уменьшения средних значений соответствующих квадратур с множителем, определяемым амбиполярной диффузией. При разрушении недоуплотненных следов явления рекомбинации и прилипания не успевают проявиться, т.е. следы разрушаются под воздействием амбиполярной диффузии. Значения вспомогательных параметров уменьшаются по экспоненте в соответствии с уравнением (21), поэтому, чем более сопоставимы по величине будут начальные значения параметров μ_{10}^C , μ_{10}^S и начальные значения средних соответствующих квадратур μ_{cr0}^C и μ_{cr0}^S , тем большее влияние на уменьшение средних значений квадратур, а, следовательно, и самих квадратур будет оказывать амбиполярная диффузия. Таким образом, будем считать, что начальные значения вспомогательных параметров μ_{10}^C , μ_{10}^S распределены по

равномерному закону в пределах соответственно $(0,5 \cdot \mu_{cp0}^C, \mu_{cp0}^C)$ и $(0,5 \cdot \mu_{cp0}^S, \mu_{cp0}^S)$.

При задании начальных значений квадратурных составляющих и параметров переуплотненных следов необходимо учитывать особенности возникновения и разрушения этих следов. Учитывая, что в начальный момент проявляются, прежде всего, эффекты вторичного рассеяния и след расширяется, то можно считать начальное значение амплитуды огибающей отраженного сигнала равным пороговому уровню u_p , определяемому чувствительностью приемника. Тогда одна из квадратур, например, μ_0^C распределена по равномерному закону в пределах от 0 до u_p , а вторая вычисляется по формуле:

$$\mu_0^S = \sqrt{u_p^2 - (\mu_0^C)^2}. \quad (36)$$

За счет высокой линейной электронной плотности вторичного рассеяния происходит увеличение отражающей поверхности переуплотненного следа, что, в свою очередь, приводит к увеличению мощности отраженного сигнала до некоторой максимальной величины, после чего происходит почти экспоненциальный спад сигнала в точке приема. Значение максимальной мощности отраженного сигнала определяется согласно (24). Согласно уравнениям (19–23), при достижении максимального значения уровня мощности огибающей отраженного сигнала, описываемая квадратурными составляющими μ^C и μ^S , достигает среднего значения, которое описывается средними значениями квадратур μ_{cp}^C и μ_{cp}^S . Время, через которое амплитуда огибающей отраженного сигнала достигает максимального значения, определяется в соответствии с (9). С другой стороны, сами средние значения μ_{cp}^C и μ_{cp}^S уменьшаются по закону близкому к экспоненциальному.

Таким образом, можно сделать вывод о том, что начальные значения μ_{cp0}^C и μ_{cp0}^S много больше начальных значений квадратурных составляющих. Анализ выражения (24) позволяет сделать вывод о том, что при прочих равных условиях максимальное значение мощности сигнала, отраженного от переуплотненного следа зависит от линейной электронной плотности q : чем больше линейная электронная плотность q , тем больше значение максимальной мощности P_M , и тем, следовательно, больше начальные средние значения квадратурных составляющих μ_{cp0}^C и μ_{cp0}^S . В свою очередь, линейная электронная плотность прямо пропорциональна массе метеора [7]. Известно [7, 8], что массы метеоров распределены по обратностепенному закону:

$$f(m) = \frac{b}{m^s}. \quad (37)$$

Тогда можно считать, что начальное значение среднего значения сигнала U_{cp0} также распределено по обратностепенному закону:

$$f(U_{cp0}) = \frac{b}{(U_{cp0})^s}, \quad (38)$$

где b – постоянная величина, а s – постоянная распределения, которая заключена в пределах от 1,25 до 4,25 [8]. С учетом условия нормировки:

$$C \int_{U_{\min}}^{U_{\max}} \frac{b \cdot dU_{cp0}}{(U_{cp0})^s} = 1, \quad (39)$$

где $U_{\max}$ и $U_{\min}$ – соответственно максимальное и минимальное значения начального среднего значения сигнала, нормирующий множитель C равен:

$$C = \left(\int_{U_{\min}}^{U_{\max}} \frac{b \cdot dU_{cp0}}{(U_{cp0})^s} \right)^{-1}. \quad (40)$$

Минимальное значение $U_{\min}$ не может быть меньше, чем некоторое пороговое значение u_p . Согласно модели [1], описывающей изменение сигнала, отраженного от переуплотненного следа, максимальная амплитуда сравнима с начальным значением амплитуды сигнала, отраженного от недоуплотненного следа. Учитывая, что сигнал, отраженный от переуплотненного следа, достигает максимума через промежуток времени от долей секунды до одной секунды, а среднее значение сигнала U_{cp0} изменяется по закону близкому к экспоненциальному, с вероятностью 0,95 максимальное значение начального среднего значения сигнала, отраженного от переуплотненного следа, не превышает более чем в 10 раз начальное значение сигнала от недоуплотненного следа. Таким образом, можно считать, что $U_{\max} = 10 \times 10^{0,05 \bar{P}_{dBW}}$, где $\bar{P}_{dBW}$ – среднее значение мощности сигнала отраженного от недоуплотненного следа. В работе [8] приводится среднее значение параметра распределения $s = 2,5$.

Таким образом, начальное среднее значение сигнала U_{cp0} распределено по обратностепенному закону в пределах от u_p до $U_{\max}$:

$$f(U_{cp0}) = \frac{(U_{cp0})^{-s}}{\int_{u_p}^{U_{\max}} (U_{cp0})^{-s} dU_{cp0}}. \quad (41)$$

Начальная фаза среднего значения сигнала U_{cp0} имеет случайное значение, поэтому одна из квадратур среднего значения сигнала, например, μ_{cp0}^C распределена по равномерному закону в пределах (u_p, U_{cp0}) , а другая (μ_{cp0}^S) вычисляется:

$$\mu_{cp0}^S = \sqrt{(U_{cp0})^2 - (\mu_{cp0}^C)^2}. \quad (42)$$

Известно, что длительности сигналов, отраженных от переуплотненных следов, составляют от 0,5 до 5 с. Так как среднее время корреляции на порядок меньше длительности сигнала [8, 15], полагаем значение параметра α_c для переуплотненных следов равномерно распределенным в пределах от 0,002 до 0,02 кГц.

Переуплотненные следы разрушаются под воздействием всех трех факторов: амбиполярной диффузии, рекомбинации и прилипания. Значения вспомогательных параметров уменьшаются по экспоненте в соответствии с уравнением (22), поэтому, чем больше будут отличаться по величине начальные значения параметров μ_{10}^c , μ_{10}^s и начальные значения средних соответствующих квадратур μ_{cp0}^c и μ_{cp0}^s , тем большее влияние на уменьшение средних значений квадратур, а, следовательно, и самих квадратур будет оказывать не только амбиполярная диффузия, но и эффекты прилипания и рекомбинации. Таким образом, будем считать, что начальные значения вспомогательных параметров μ_{10}^c , μ_{10}^s распределены по равномерному закону в пределах соответственно $(0,5 \cdot \mu_{cp0}^c)$ и $(0,5 \cdot \mu_{cp0}^s)$.

4. Исследование адекватности разработанной модели метеорного радиоканала

Любая модель, описывающая аналитически физические процессы, происходящие в природе, должна отражать наиболее существенные свойства этих процессов, влияющие на характер изменения случайного процесса во времени. Эксперименты и многочисленные наблюдения показывают, что длительности сигналов, отраженных от недоуплотненных следов, аппроксимируются экспоненциальным распределением. Для проверки адекватности разработанной модели проведем оценку закона распределения длительностей сигналов, отражен-

ных от наиболее часто образующихся недоуплотненных следов.

Для оценки закона распределения проведем моделирование в соответствии с моделью, разработанной в пунктах 2 и 3. Для оценки степени близости закона распределения, полученной выборки к экспоненциальному, будем использовать критерий согласия χ^2 Пирсона.

В таблице 1 приведены значения величины χ^2 , характеризующей степень расхождения теоретического распределения и относительных частот, полученные в результате моделирования недоуплотненных следов при различных объемах выборок N . При моделировании использовалось $r = 20$ интервалов группировки. В таблице 2 приведены значения доверительной границы χ_α^2 при различных значениях уровня значимости.

ТАБЛИЦА 2. Результаты моделирования и теоретического исследования

χ^2	9,41	9,54	13,96	17,24	11,58	7,25	9,31
N	258	268	293	284	277	268	280
χ^2	19,7	13,5	14,09	5,04	21,71	8,11	9,1
N	290	214	288	253	232	267	307
χ^2	9,7	8,5	9,09	7,04	4,71	6,13	8,05
N	264	289	268	295	238	275	262

Приведенные в таблице 2 результаты показывают, что в ряде случаев теоретический закон согласуется лишь с доверительной вероятностью $P\{\chi^2 > \chi_\alpha^2\} = 0,2$ с результатами моделирования (рисунок 1).

В таблице 3 приведены значения доверительной границы χ_α^2 при различных значениях уровня значимости. Из рисунка 2 видно, что теоретический расчет в большинстве случаев согласуется с результатами моделирования.

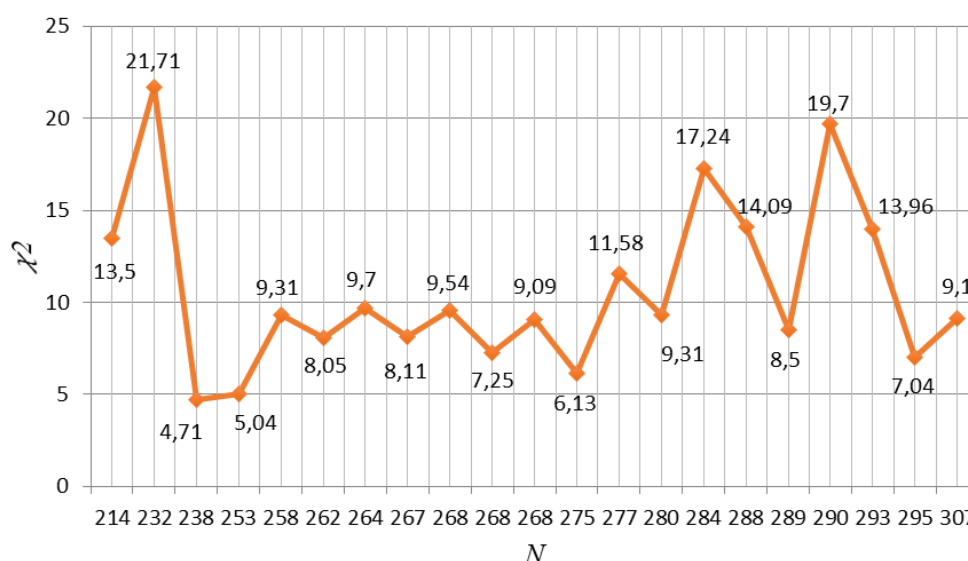


Рис. 1. Оценка степени близости закона распределения полученной выборки к экспоненциальному

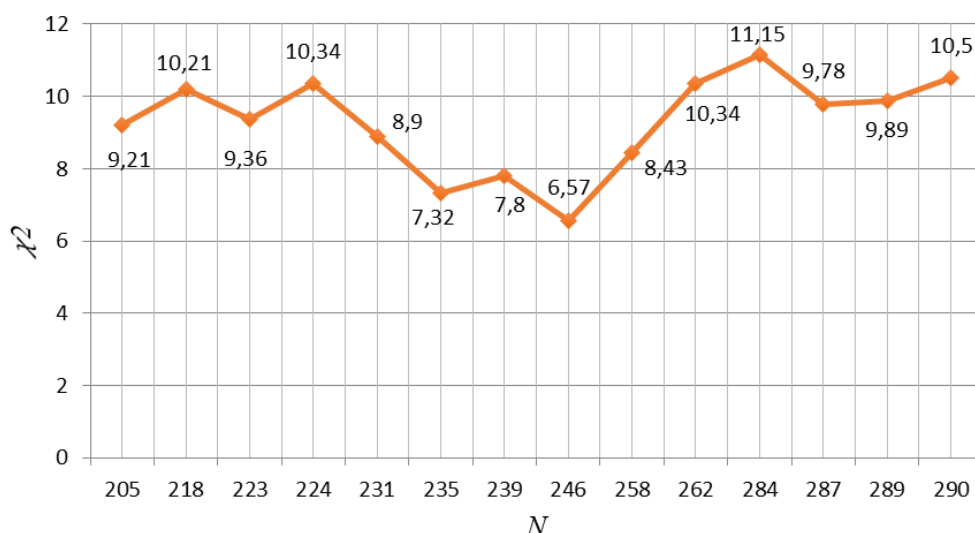


Рис. 2. Сравнение теоретического и статистического распределения

ТАБЛИЦА 3. Значения доверительной границы при разных уровнях значимости

α	0,99	0,98	0,95	0,9	0,8	0,2	0,1
χ^2_{α}	7,015	7,906	9,39	10,87	12,86	22,76	25,99

Однако, если исключить из рассмотрения сигналы малой длительности, меньше некоторого значения $t_{\min}$, которые непригодны на практике для связи из-за своей непродолжительности, то после нормировки теоретического и статистического распределения получаем результаты, приведенные в таблице 4.

ТАБЛИЦА 4. Результаты моделирования и теоретического исследования

χ^2	9,21	10,12	7,32	6,57	9,78	10,34	11,15
N	205	218	235	246	287	262	284
χ^2	10,34	8,43	9,89	10,5	8,9	9,36	7,8
N	224	258	289	290	231	223	239

Из данных, приведенных в таблице 4, видно, что в 92 % случаев теоретическое распределение согласуется с результатами моделирования при уровне значимости $\alpha = 0,9$.

Таким образом, полученные результаты позволяют сделать вывод об адекватности разработанной модели наблюдения за сигналами, отраженными от метеорных следов, сделанными астрономическими лабораториями.

Список используемых источников

1. Метеорная радиосвязь на ультракоротких волнах // Сборник статей под ред. А.Н. Казанцева. М.: Иностранная литература, 1961. 288 с.
2. Metzner J.J. Improved coding strategies for meteor burst communication // IEEE Transactions on Communications. 1990. Vol. 38. Iss. 2. PP. 133–136. DOI:10.1109/26.47842.
3. Pursley M.B., Sandberg S.D. Variable-rate coding for meteor burst communications // IEEE Transactions on Communications. 1989. Vol. 37. Iss. 11. PP. 1105–1112. DOI 10.1109/26.46504.
4. Miller S.L., Milstein L.B. Error correction coding for a meteor burst channel // IEEE Transactions on Communications 1990. Vol. 38. Iss. 9. PP. 1520–1529. DOI 10.1109/26.61393.

Заключение

В работе представлена стохастическая модель метеорного радиоканала, отличающаяся от известных тем, что учитывает статистический характер возникновения следов различного типа, а также распределения начальных значений параметров квадратурных составляющих сигнала, отраженного от различного типа следов.

В модели предложен алгоритм задания начальных значений параметров: начальное значение огибающей сигнала, отраженного от недоуплотненных следов распределено по нормальному закону, а начальное значение огибающей среднего значения сигнала, отраженного от переуплотненного следа распределено по обратностепенному закону.

При проведении оценки адекватности разработанной модели из рассмотрения были исключены сигналы малой длительности, которые непригодны на практике для связи из-за своей непродолжительности. Значение коэффициента χ^2 Пирсона в 92 % случаев при уровне значимости $\alpha = 0,9$, что позволяет сделать вывод об удовлетворительной аппроксимации.

5. Milstein L., Schilling D., Pickholtz R., Sellman J., Davidovici S., Pavelchek A., Schneider A., Eichmann G. Performans of Meteor Burst Communication Channel // IEEE Journal on Selected Areas in Communications. 1987. Vol. 5. Iss. 2. PP. 146–154. DOI:10.1109/JSAC.1987.1146515.
6. Шомников М.А. Стохастическая модель канала метеорной радиосвязи // III Международная научно-техническая конференция. Антенно-фидерные устройства. Системы и средства связи. Воронеж: ВГУ, 1997. Т. 2. С. 126–139.
7. Курганов Р.А. Прогнозирование наклонного рассеивания радиоволн метеорными ионизациями. Казань: КГУ, 1973. 184 с.
8. Фиалко Е.И. Радиолокационные методы наблюдения метеоров. М.: Сов. радио, 1961. 112 с.
9. Аронэ М.Н. Метеорная радиосвязь. М.: Воениздат, 1960. 108 с.
10. Лебедева А.А. Разработка программно-аппаратного комплекса метеорной связи. // Интеллектуальная собственность и современные техника и технологии для развития экономики. Материалы IV республиканской молодежной научно-практической конференции в рамках Всероссийского студенческого форума «Инженерные кадры – будущее инновационной экономики России». 2016. С. 74–76.
11. Рябов И.В., Толмачев С.В., Лебедева А.А. Принципы программно-определяемых радиосистем и их применение в рамках задачи исследования метеорной радиосвязи // Современные наукоемкие технологии. 2016. № 7–1. С. 59–66.
12. Гайдаев А.А., Белькович О.И., Зулиев А.М. Распределение амплитуд метеорных сигналов при наклонном метеорном распространении радиоволн // Известия НЦ СКВШ. 1983. № 4. С. 44–48.
13. Бабаджанов П.Б. Метеоры и их наблюдение. М.: Наука, 1987. 192 с.
14. Greenberg J.M. From Comets to Meteors // Earth, Moon and Planets. 1998. Vol. 82. PP. 313–324. DOI:10.1023/A:1017086832593.

* * *

STOCHASTIC MODEL OF METEOR-BURST RADIO CHANNEL

D. Kapralov¹, D. Kirik¹

¹The Bonch-Bruевич State University of Telecommunications,
St. Petersburg, 193232, Russian Federation

Article info

Article in Russian

For citation: Kapralov D., Kirik D. Stochastic Model of Meteor-Burst Radio Channel // Proceedings of Telecommunication Universities. 2018. Vol. 4. Iss. 3. PP. 54–64.

Abstract: *The article proposes a stochastic model of a meteor-burst communication that differs from the known ones in that it takes into account the statistical nature of the occurrence of traces of various types. Based on the results of the studies, the initial values of the parameters of the quadrature components of the signal reflected from traces of various types.*

Keywords: *meteor-burst communication, radio waves, stochastic model.*

УСОВЕРШЕНСТВОВАННЫЙ ПРОТОКОЛ МАРШРУТИЗАЦИИ EIGRP, ОБЕСПЕЧИВАЮЩИЙ ПОВЫШЕННУЮ УСТОЙЧИВОСТЬ СЕТЕЙ СВЯЗИ

С.И. Макаренко^{1*}

¹ООО «Корпорация «Интел групп»,
Санкт-Петербург, 197372, Российская Федерация
*Адрес для переписки: mak-serg@yandex.ru

Информация о статье

УДК 004.722

Язык статьи – русский

Ссылка для цитирования: Макаренко С.И. Усовершенствованный протокол маршрутизации EIGRP, обеспечивающий повышенную устойчивость сетей связи // Труды учебных заведений связи. 2018. Т. 4. № 3. С. 65–73.

Аннотация: В статье представлен вариант усовершенствованного протокола EIGRP, обеспечивающего повышенную устойчивость сети связи. Повышение устойчивости протокола EIGRP достигается за счет более полного использования им имеющегося топологического ресурса сети, а также за счет способности усовершенствованного протокола одновременно с поиском кратчайших путей формирования дополнительных резервных путей.

Ключевые слова: сеть связи, протокол маршрутизации, устойчивость сети, конвергенция сети, время сходимости, EIGRP.

Введение

Анализ основных тенденций развития телекоммуникационных сетей (ТКС) специального назначения (СН), представленный в работах [1, 2], показал, что их особенностью является функционирование в условиях преднамеренных дестабилизирующих воздействий радиоэлектронного и информационного характера. При этом используемые в ТКС СН протоколы маршрутизации ориентированы на псевдостатическую структуру сети и, как показали теоретические и экспериментальные исследования автора [3, 4], этим протоколам свойственны длительные процессы восстановления связи при изменении топологии сети, централизация принятия решений о перемаршрутизации потоков трафика, недостаточно проработанные механизмы своевременного обнаружения изменений в топологии сети. Основным концептуальным недостатком протоколов маршрутизации является низкая эффективность реакции на изменения топологии сети их математической основы – алгоритмов поиска кратчайших путей.

Подавляющая часть широко используемых протоколов маршрутизации (OSPF, IS-IS, EIGRP, IGRP,

PNNI и др.) основана на алгоритмах Дейкстры и Беллмана-Форда, которые являются «поглощающими» и по своей сути не способны одновременно с поиском кратчайших путей строить в сети множество резервных путей, которые можно использовать в случае изменения топологии. Имеющиеся алгоритмы поиска нескольких путей между узлами (например, алгоритм Йена) являются итерационными и основаны на последовательном поиске все тех же кратчайших путей при удалении отдельных ребер в сети. Однако в таких алгоритмах количество итераций резко возрастает при увеличении размерности сети, что препятствует применению этих алгоритмов в составе протоколов маршрутизации реальных ТКС. Указанные факторы ведут к существенному снижению устойчивости ТКС СН при использовании в них стандартных протоколов маршрутизации.

Одним из направлений устранения вышеуказанных недостатков существующих протоколов маршрутизации и, соответственно, повышения устойчивости ТКС СН является использование имеющегося топологического ресурса в сети и формирование у известных протоколов маршрутизации дополнительной функциональности – способности одно-

временно с поиском кратчайших путей находить дополнительные резервные пути. Данные пути предполагается использовать в случае, если в результате преднамеренных дестабилизирующих воздействий топология сети изменилась и требуется произвести пересчет кратчайших маршрутов, но без прерывания процессов передачи трафика.

Этот подход в теоретическом виде формализован в форме метода обеспечения устойчивости телекоммуникационной сети за счет использования ее топологической избыточности в работе автора [5]. В работе автора [6] на основе данного метода предложена модификация алгоритма поиска кратчайших путей Беллмана-Форда в направлении повышения устойчивости, за счет дополнительного формирования в сети как кратчайших, так и резервных путей. В настоящей статье предлагается рассмотреть применение вышеуказанного метода и модифицированного алгоритма Беллмана-Форда к практике, а именно – к совершенствованию протокола маршрутизации EIGRP (*от англ. Enhanced Interior Gateway Routing Protocol*) в интересах повышения устойчивости ТКС СН, функционирующей на основе данного протокола. Так как протокол является проприетарным, его полная спецификация отсутствует, а при формировании технических предложений были использованы открытые материалы компании-разработчика – Cisco Systems [7].

Особенности протокола EIGRP

Протокол EIGRP является протоколом оценки вектора расстояний (DVA, *от англ. Distance Vector Algorithm*) и основан на алгоритме диффузии обновлений маршрутизации (DUAL, *от англ. Diffusing Update Algorithm*). Он поддерживает бесклассовую междоменную маршрутизацию (CIDR, *от англ. Classless InterDomain Routing*), которая позволяет максимально использовать адресное пространство и маски подсети переменной длины (VLSM, *от англ. Variable-Length Subnet Mask*).

Метрика M отдельного канала связи в протоколе EIGRP вычисляется по формуле:

$$M = \left(k_1 \frac{10^7}{C} + \frac{k_2 \frac{10^7}{C}}{256(1 - \rho)} + k_3 \frac{T_{\text{зад}}}{10} \right) \frac{k_5}{\eta + k_4}, \quad (1)$$

где M – метрика канала связи; C – пропускная способность канала связи в отсутствии нагрузки (бит/с); ρ – нагрузка канала связи, в диапазоне 0..1; η – надежность канала связи, как доля успешно переданных по каналу связи пакетов; $T_{\text{зад}}$ – задержка передачи в канале связи (с); k_1, k_2, k_3, k_4, k_5 – коэффициенты, определяющие значимость соответствующих параметров.

Главным компонентом протокола EIGRP является алгоритм DUAL. Алгоритм DUAL анализирует все маршруты, анонсированные соседними маршрутизаторами, и использует метрику каждого маршрута для их сравнения. Алгоритм DUAL также гарантирует, что маршруты не содержат петель. После соответствующих вычислений алгоритм DUAL заносит маршруты с наименьшей метрикой в таблицу маршрутизации (ТМ). Эти ТМ предоставляют алгоритму DUAL маршрутную информацию для расчета альтернативных маршрутов в случае нарушений в работе ТКС СН. Протокол передает основную маршрутную и топологическую информацию сети соседним маршрутизаторам.

Для повышения эффективности и уменьшения времени конвергенции сети протокол EIGRP рассчитывает несколько маршрутов до узла-получателя. Маршрут, выбираемый в качестве основного для достижения определенного узла, называется первичным. Для каждого конкретного узла может быть сформировано до четырех первичных маршрутов. Они могут иметь разные метрики и рассматриваются как наилучшие свободные от петель маршруты к данному узлу-получателю. Копии первичных маршрутов также помещаются в ТМ. Эти маршруты устанавливаются одновременно с первичными маршрутами по проприетарной технологии и также хранятся в ТМ. Наличие нескольких первичных маршрутов для достижения узла-получателя не является обязательным, но увеличивает надежность маршрутизации в сети.

Маршрутизатор рассматривает устройства на резервных маршрутах как устройства, расположенные ближе к узлу-получателю, чем он сам. Если по каким-либо причинам первичный маршрут не может выполнять свои функции, то алгоритм DUAL инициализирует другой первичный маршрут на основе данных ТМ и формирует новый маршрут к узлу назначения. Если резервный маршрут к узлу назначения отсутствует, то алгоритм DUAL переводит маршрут в активное состояние (active state). Статус маршрута в ТМ может иметь одно из двух значений: активное состояние или пассивное. Пассивный статус имеет маршрут, который устойчив и готов к использованию. Активный статус присваивается временно недоступному маршруту, который в настоящий момент пересчитывается алгоритмом DUAL с использованием математического алгоритма Беллмана-Форда.

Пересчет маршрутов происходит в том случае, когда маршрут становится недоступным, и DUAL не может найти другого альтернативного первичного маршрута. В этом случае маршрутизатор должен запросить дополнительные топологические данные у соседних маршрутизаторов.

Если у соседнего маршрутизатора имеется соответствующий маршрут, то он предоставляет данные об имеющихся у него первичных маршрутах. В этом случае маршрут устанавливается без пересчета. В противном случае, соседний маршрутизатор рассматривает устройства на резервных маршрутах как устройства, расположенные ближе к узлу-получателю, чем он сам.

Если у соседнего маршрутизатора имеется соответствующий маршрут, то он предоставляет данные об имеющихся у него первичных маршрутах. В этом случае маршрут устанавливается без пересчета. В противном случае, соседний маршрутизатор рассматривает устройства на резервных маршрутах как устройства, расположенные ближе к узлу-получателю, чем он сам.

тизатор уведомляет о том, что у него также нет маршрута к требуемому узлу назначения. Если один или более маршрутизаторов, которым был разослан запрос, не отвечает в течение времени, равного длительности «застревания» («stuck in active»), то маршрут (маршруты) переводится в состояние «застревания». В этом случае протокол EIGRP исключает из своей ТМ маршрутизаторы, которые не ответили на запрос, и регистрирует в системном журнале сообщение об ошибке для маршрутов, которые ранее были активными.

Маршрутизаторы на протоколе EIGRP активно устанавливают связи со своими соседними устройствами подобно тому, как это делают маршрутизаторы протокола OSPF. Это отличает EIGRP от маршрутизаторов на обычных дистанционно-векторных протоколах, например, RIP, выполняющих только широковещательную рассылку обновлений маршрутов. Маршрутизаторы EIGRP устанавливают отношения смежности с соседними маршрутизаторами путем рассылки пакетов приветствия Hello. EIGRP предполагает, что до тех пор, пока от известных ему соседних устройств поступают пакеты Hello, они (и их маршруты) остаются действующими (пассивными).

Маршрутизатор протокола EIGRP сохраняет информацию о соседних устройствах в соответствующей таблице, называемой таблицей соседних устройств. В этой таблице записываются номера последних полученных от устройств пакетов протокола EIGRP, а также имеется поле «времени удержания» (Hold Time), в котором записано время получения последнего пакета. Для того, чтобы у соседнего маршрутизатора сохранялся статус пассивного (достижимого и работоспособного), необходимо, чтобы за время удержания от него поступил хотя бы один пакет Hello. Если в течение времени удержания от него не поступило пакетов Hello, то протокол EIGRP рассматривает этот соседний маршрутизатор как неработоспособный и алгоритм DUAL начинает пересчитывать ТМ. По умолчанию, время удержания в три раза больше интервала Hello, однако его можно, при необходимости, переконфигурировать.

Описание временных параметров протокола EIGRP приведено в таблице 1 по данным из работы [7].

Протокол EIGRP независим от используемого сетевого протокола, то есть он не использует протокол TCP/IP для обмена сообщениями маршрутизации. Протокол EIGRP использует свой транспортный протокол для гарантированной доставки информации о маршрутах. Так пакеты обновления ТМ всегда рассылаются методом гарантированной доставки, а пакеты Hello всегда рассылаются ме-

тодом негарантированной доставки и не требуют подтверждения.

ТАБЛИЦА 1. Параметры протокола EIGRP

Временной параметр	Пропускная способность	
	1,544 Мбит/с и менее	более 1,544 Мбит/с
Интервал приветствий (Hello Time)	60 с	5 с
Время удержания маршрута (Hold Time)	180 с	15 с
Время «застревания» маршрута (Stuck in Active)	180 с	

Для того, чтобы обеспечить гарантированную доставку сообщений, EIGRP использует пакеты-подтверждения от принимающего узла для узла-отправителя, представляющие собой пустые пакеты Hello. В отличие от многоадресных пакетов Hello, пакеты-подтверждения являются одноадресными и посылаются конкретному узлу. Подтверждение также может быть осуществлено путем совмещения передачи прямых и обратных пакетов других типов пакетов EIGRP.

Пакеты с обновлениями маршрутов используются в тех случаях, когда маршрутизатор обнаруживает новое соседнее устройство. В этих случаях маршрутизатор EIGRP посылает одноадресные пакеты обновления маршрутов этому новому соседнему устройству для того, чтобы оно могло добавить эту информацию в свою ТМ. Эти пакеты обновления используются также в тех случаях, когда маршрутизатор обнаруживает изменение топологии сети. В этом случае маршрутизатор рассылает многоадресные пакеты обновления всем своим соседним устройствам, предупреждая их об изменении топологии.

Схема функционирования маршрутизатора на базе протокола EIGRP приведен на рисунке 1.

Модифицированный алгоритм Беллмана-Форда

В основе предлагаемого усовершенствования протокола EIGRP лежит использование модифицированного алгоритма поиска кратчайших путей Беллмана-Форда. Данный модифицированный алгоритм был разработан путем применения общетеоретических положений авторского метода [5] к стандартному алгоритму Беллмана-Форда. В результате был разработан новый модифицированный алгоритм Беллмана-Форда, который был ранее опубликован автором в работе [6].

Рассмотрим кратко данный модифицированный алгоритм.

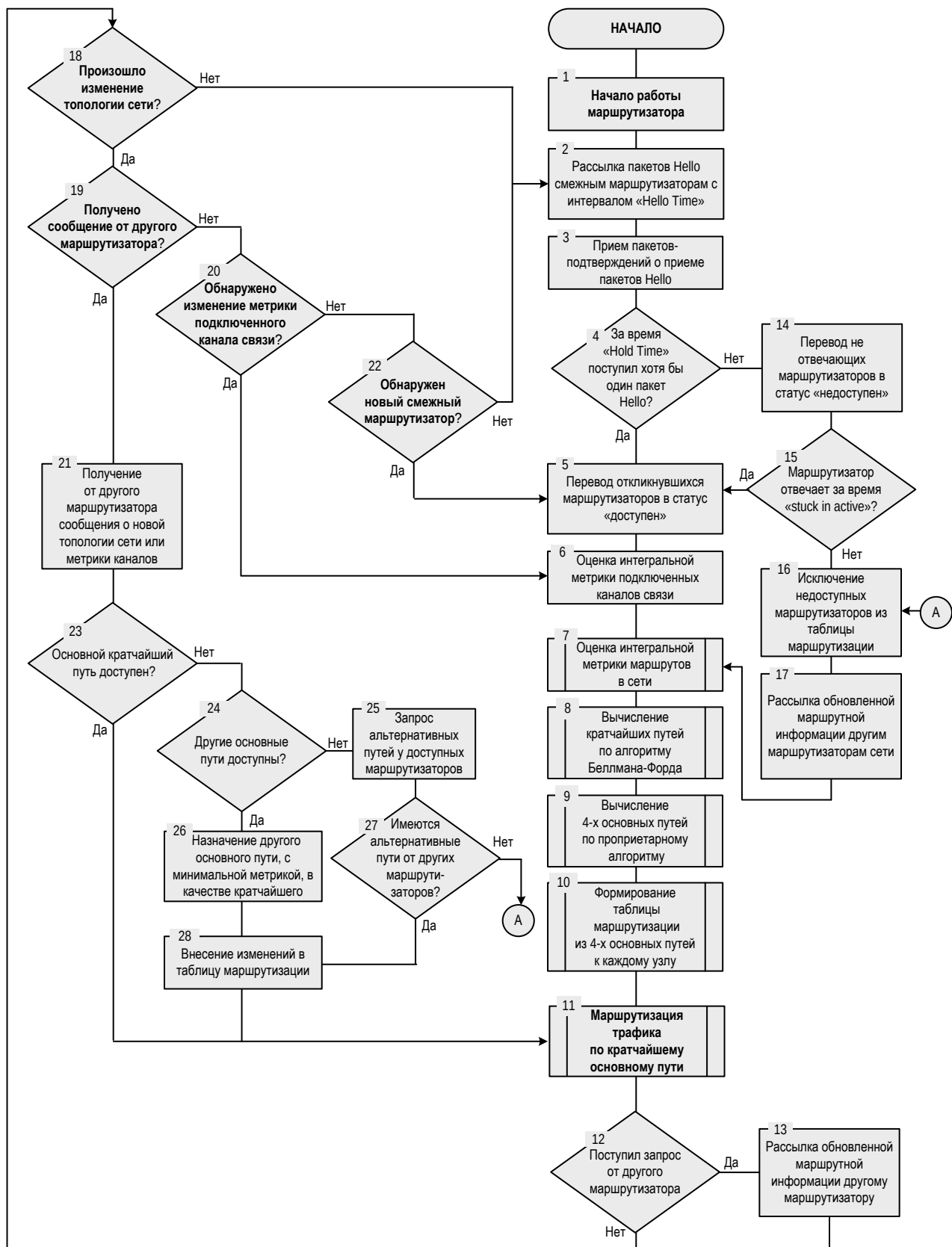


Рис. 1. Схема функционирования маршрутизатора на основе протокола EIGRP

Для формализации модифицированного алгоритма Беллмана-Форда используем следующие обозначения:

$D = \{d_i\}$, $i = 1, \dots, n$ – множество расстояний до помеченных вершин от начальной вершины;

d_i – расстояние от вершины U_1 до вершины U_i , полученное при использовании алгоритма поиска кратчайших путей;

$G(U, V)$ – ориентированный граф, соответствующий сети ТКС СН;

$i = 1, \dots, n$ – переменная, счетчик вершин;

$j = 1, \dots, n$ – переменная, счетчик вершин;

L – множество смежных помеченных вершин, содержащих множество расстояний от начальной вершины до помеченных вершин;

$L = \{l_i\}$, $i = 1, \dots, n$ – множество смежных помеченных вершин;

$l_i = U_j$ – помеченная вершина, через которую достигнута вершина U_i , по окончании работы алгоритма l_i содержит кратчайшие пути к каждой вершине;

n – количество вершин в графе;

P – множество помеченных вершин;

t – переменная, определяющая последнюю помеченную вершину;

U_1 – начальная вершина;

$V(U_i, U_j)$ – вес ребра, соединяющего i -ую и j -ую вершины;

$R = \{r\}$ – множество вершин потенциальных резервных путей (в это множество вносятся достигнутые вершины, смежные рассматриваемой; в дальнейшем, элементы множества используются при нахождении резервных путей);

$C = \{c\}$ – множество весов ребер потенциальных резервных путей (в это множество вносятся веса ребер, исходящих из вершин, вносимых в множество R , и входящих в рассматриваемую вершину);

$Z = \{z\}$ – множество резервных путей в вершину, содержит резервные пути в рассматриваемую вершину, сформированные в результате проведения логических операций над входящими в него элементами и элементами множествами R и L ;

$S = \{s\}$ – множество весов резервных путей к вершине (это множество содержит веса путей из множества Z и используется для ранжировки резервных путей при выводе результатов работы модифицированного алгоритма).

Схема модифицированного алгоритма Беллмана-Форда приведена на рисунках 2а, 2б. К новым элементам модифицированного алгоритма Беллмана-Форда относятся блоки 7, 8, 17–22, 24.

В блоках 7, 8 реализуется формирование элементов множества вершин потенциальных резервных путей R к текущей рассматриваемой вершине. Далее, в блоках 17–22, путем пересечения элементов множества R и L , а также множества Z , осуществляется формирование элементов множества резервных путей в вершину Z . В блоке 24

осуществляется ранжирование резервных путей, находящихся в множестве Z по сумме весов, входящих в их состав ребер, которые в свою очередь находятся в множестве S .

Усовершенствование протокола маршрутизации EIGRP

Протокол EIGRP является проприетарным, при этом алгоритм DUAL в части формирования основных путей не является полностью открытым. В связи с этим совершенствование протокола EIGRP фактически заключается в разработке на его основе нового протокола, в котором исключены «закрытые» проприетарные операции.

В рамках совершенствования протокола EIGRP предлагается реализовать несколько направлений его доработки:

а) необходимо исключить «проприетарные операции» формирования четырех основных путей и заменить их на расчет маршрутов по модифицированному алгоритму поиска кратчайших путей Беллмана-Форда, впервые разработанному в данном исследовании, который одновременно с кратчайшими путями формирует и множество резервных путей;

б) в ТМ вместо четырех основных путей предлагается хранить как кратчайшие, так и непересекающиеся резервные маршруты, ранжированные по уровню увеличения суммарной метрики маршрута;

в) при поступлении от других маршрутизаторов сообщения об изменении топологии ТКС СН или метрики КС, а также в случае самостоятельного выявления подобных фактов маршрутизатор сразу же переходит на резервный маршрут, а уже потом производит оповещение ответственного маршрутизатора, смежных маршрутизаторов, ведет пересчет кратчайших путей и формирование новых маршрутов.

Схема усовершенствованного протокола EIGRP относительно стандартного протокола EIGRP представлена на рисунке 3.

К измененным функциям усовершенствованного протокола EIGRP относятся блоки 6, 9–10 и 14–15 на рисунке 3. К функциям, которые были удалены, относятся блоки 9–10, 23–24, 26 на рисунке 2.

Блок 6 выступает своеобразным «логическим вентилем», который на начальном этапе работы маршрутизатора позволяет сформировать ТМ из кратчайших и резервных путей, а в дальнейшем, при любых изменениях в сети, перенаправляет потоки трафика, на которые влияют произошедшие изменения, на резервные пути (блоки 14, 15). Переход передачи трафика на резервные пути позволяет не прерывать передачу данных при выполнении операций, представленных в блоках 16–20 и 7–10.

Блок 9 отличается тем, что в нем вместо вычисления метрики по стандартному алгоритму Беллмана-Форда реализуется вычисление кратчайших и резервных путей по модифицированному алгоритму Беллмана-Форда.

В блоке 10 реализуется функция добавления в ТМ помимо кратчайших путей еще и резервных путей. Таким образом, маршрут передачи изменяется с резервного пути, введенного в блоке 15, на текущий кратчайший путь.

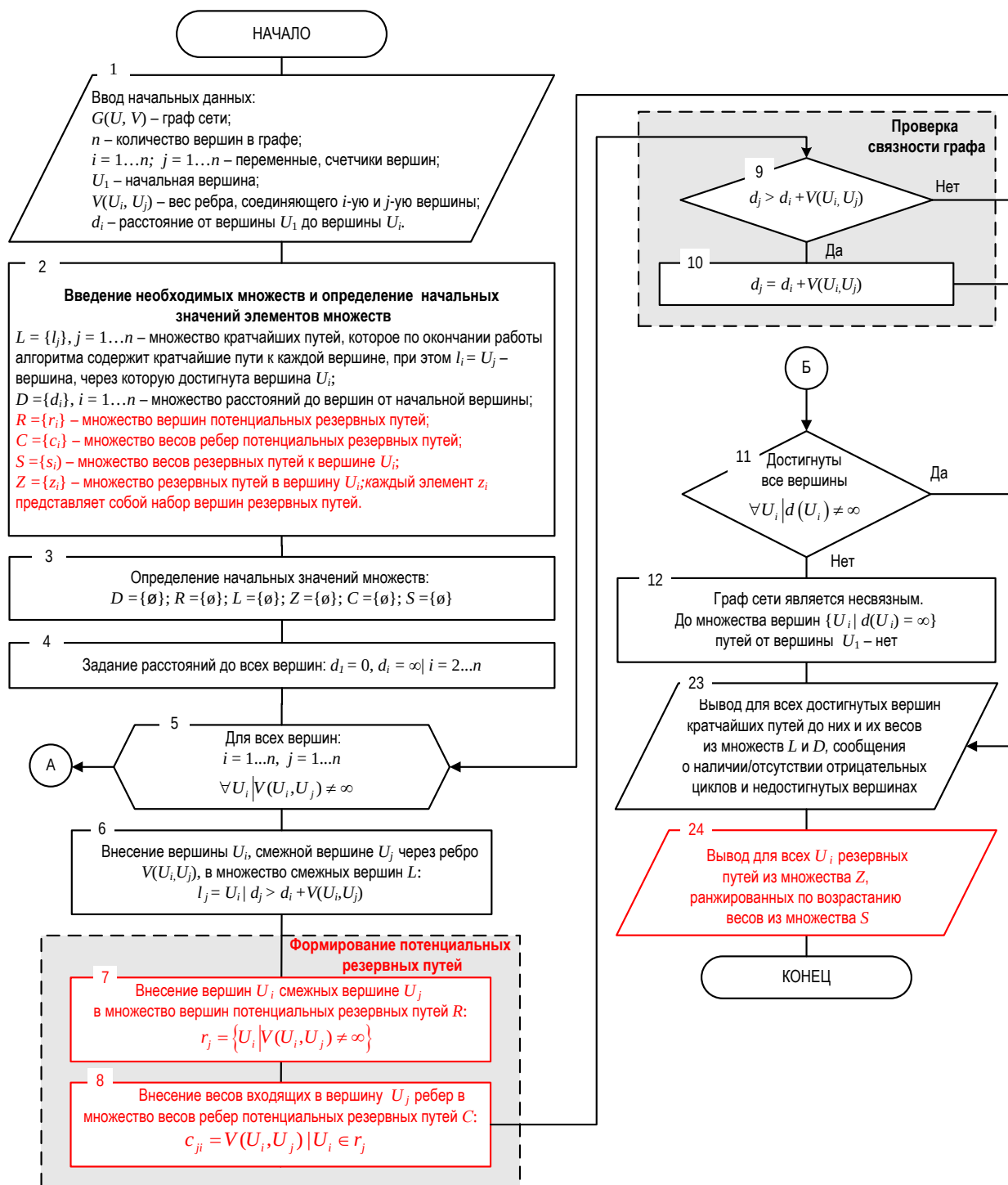


Рис. 2а. Модифицированный алгоритм поиска кратчайших путей Беллмана-Форда [6]

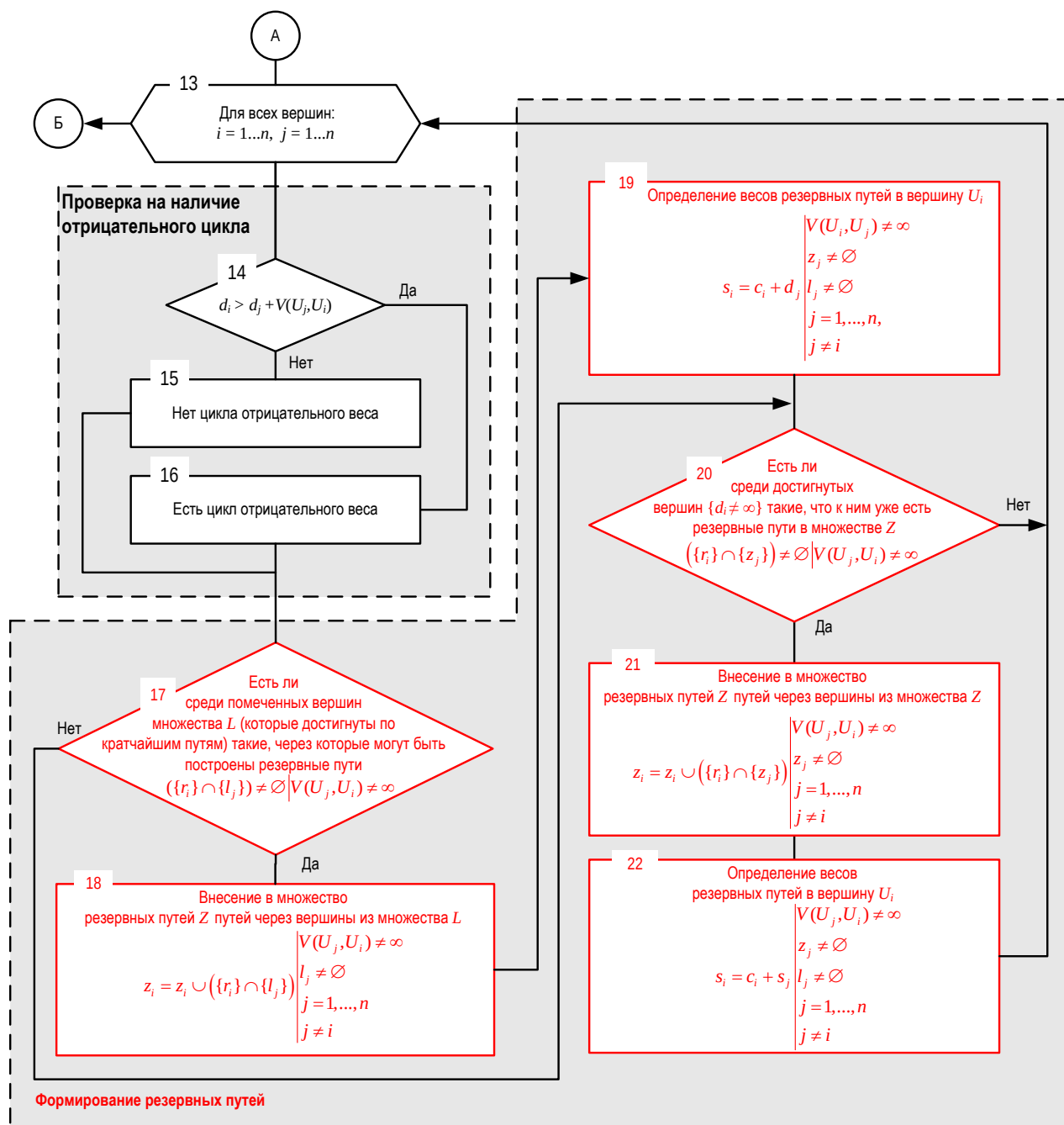


Рис. 26. Модифицированный алгоритм поиска кратчайших путей Беллмана-Форда [6] (продолжение)

Заключение

Технической новизной разработанных предложений по совершенствованию протокола маршрутизации EIGRP, которая отличает данные предложения от уже известных [8–15], является следующее:

1) в составе усовершенствованного протокола EIGRP используется модифицированный алгоритм поиска кратчайших путей Беллмана-Форда [6], который впервые разработан автором и который одновременно с кратчайшими путями позволяет формировать и множество резервных путей;

2) в усовершенствованном протоколе EIGRP изменена последовательность выполнения опера-

ций, а также введены дополнительные операции, выполняемые при изменении топологии сети, что позволяет снизить время восстановления связи при динамическом изменении, как метрики отдельных каналов, так и топологии сети в целом.

Практическая значимость предложений по совершенствованию протокола EIGRP заключается в том, что они применимы к широкому спектру ТКС СН, использующим для маршрутизации трафика протокол EIGRP, и позволяют реализовать повышение устойчивости ТКС СН по двум направлениям – за счет повышения структурной связности сети и за счет снижения длительности переходных режимов маршрутизации.

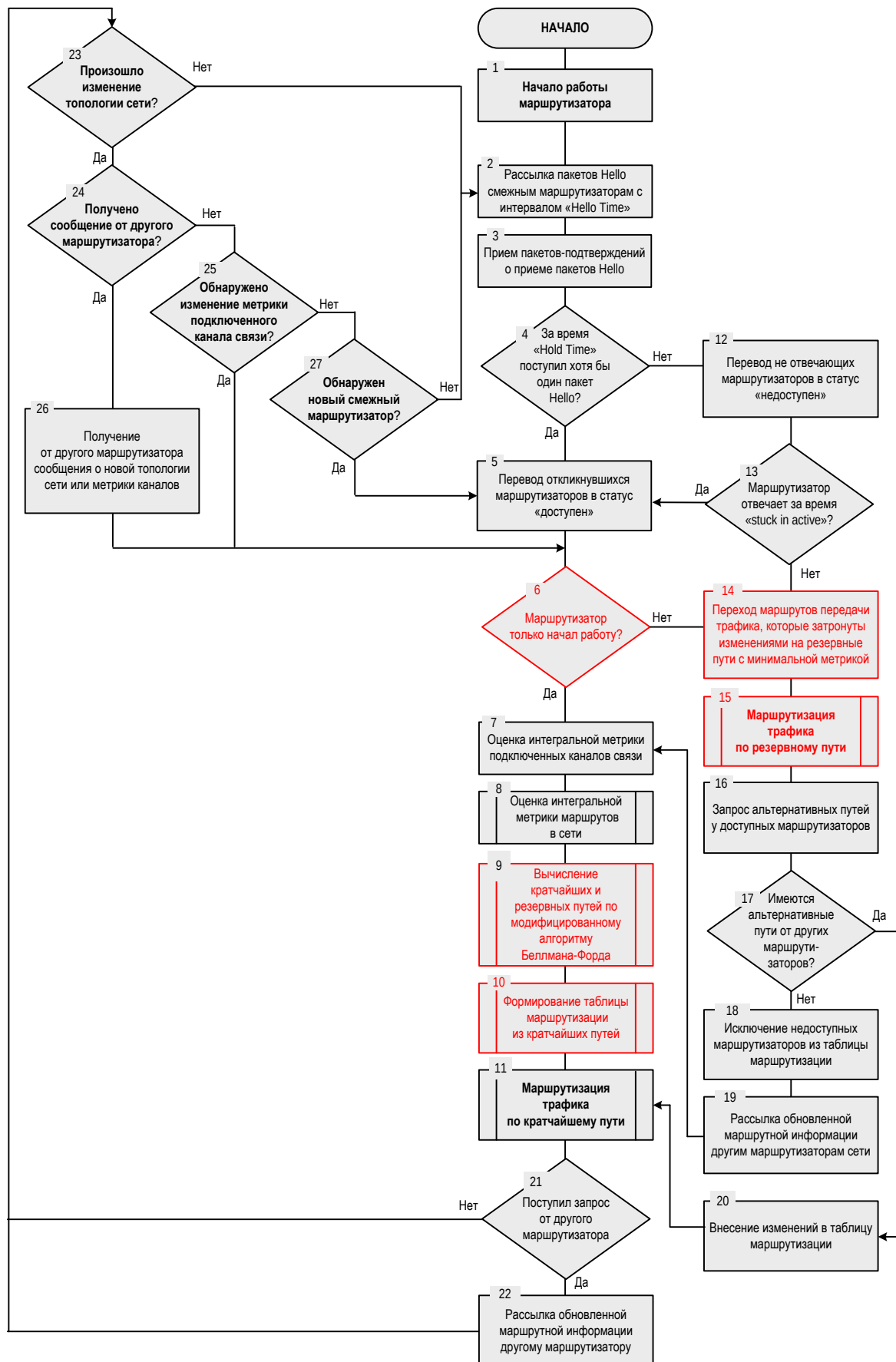


Рис. 3. Схема функционирования маршрутизатора на основе усовершенствованного протокола EIGRP

Список используемых источников

1. Макаренко С.И. Описательная модель сети связи специального назначения // Системы управления, связи и безопасности. 2017. № 2. С. 113–164.
2. Макаренко С.И. Перспективы и проблемные вопросы развития сетей связи специального назначения // Системы управления, связи и безопасности. 2017. № 2. С. 18–68.
3. Макаренко С.И. Время сходимости протоколов маршрутизации при отказах в сети // Системы управления, связи и безопасности. 2015. № 2. С. 45–98.
4. Макаренко С.И., Афанасьев О.В., Баранов И.А., Самофалов Д.В. Экспериментальные исследования реакции сети связи и эффектов перемаршрутизации информационных потоков в условиях динамического изменения сигнально-помеховой обстановки // Журнал радиоэлектроники. 2016. № 4. URL: <http://jre.cplire.ru/jre/apr16/4/text.pdf> (дата обращения: 30.05.2018).
5. Макаренко С.И. Метод обеспечения устойчивости телекоммуникационной сети за счет использования ее топологической избыточности // Системы управления, связи и безопасности. 2018. № 3. С. 14–30.
6. Макаренко С.И., Квасов М.Н. Модифицированный алгоритм Беллмана-Форда с формированием кратчайших и резервных путей и его применение для повышения устойчивости телекоммуникационных систем // Инфокоммуникационные технологии. 2016. Т. 14. № 3. С. 264–274.
7. Программа сетевой академии Cisco. CCNA 1 и 2. М.: Вильямс, 2008. 1168 с.
8. Перепелкин Д.А. Повышение качества функционирования корпоративных сетей на базе протокола EIGRP // Качество. Инновации. Образование. 2012. № 5(84). С. 99–106.
9. Перепелкин Д.А. Алгоритм парных перестановок маршрутов на базе протокола IGRP в корпоративных сетях // Вестник Воронежского государственного технического университета. 2010. Т. 6. № 12. С. 39–43.
10. Перепелкин Д.А. Алгоритм адаптивной ускоренной маршрутизации на базе протокола IGRP при динамическом отказе узлов и линий связи корпоративной сети // Вестник Рязанского государственного радиотехнического университета. 2012. № 42-1. С. 33–38.
11. Михайлов Р.Л. Модели и алгоритмы маршрутизации в транспортной наземно-космической сети связи военного назначения // Системы управления, связи и безопасности. 2015. № 3. С. 52–82.
12. Михайлов Р.Л. Помехозащищенность транспортных сетей связи специального назначения. Череповец: ЧВВИУРЭ, 2016. 128 с.
13. Полукаров Д.Ю. Нечеткая аппроксимация метрики протокола IGRP // Инфокоммуникационные технологии. 2006. Т. 4. № 4. С. 51–54.
14. Красов А.В., Левин М.В., Цветков А.Ю. Метод управления трафиком в гибридной программно-определяемой сети // Информационные технологии и телекоммуникации. 2016. Т. 4. № 2. С. 53–63.
15. Snigurov A., Chakrrian V. Improvement of EIGRP protocol routing algorithm based on information security metrics // 2nd International Scientific-Practical Conference Problems of Infocommunications Science and Technology (PIC S and T 2015), Kharkiv. 2015. PP. 263–265.

* * *

THE IMPROVED EIGRP PROTOCOL FOR HIGH NETWORK STABILITY

S. Makarenko¹

¹Intel Group Corporation ltd,
St. Petersburg, 197372, Russian Federation

Article info

Article in Russian

For citation: Makarenko S. The Improved EIGRP Protocol for High Network Stability // Proceedings of Telecommunication Universities. 2018. Vol. 4. Iss. 3. PP. 65–73.

Abstract: *The improved EIGRP routing protocol is presents in the paper. The improved EIGRP protocol is to use topological network redundancy and create backup paths. This protocol provides high network stability, if the network topology changes rapidly.*

Keywords: *network, routing protocol, network stability, network convergence time, EIGRP.*

ЭРГОДИЧЕСКАЯ МОДЕЛЬ АТОМАРНЫХ КОНСТРУКЦИЙ

Л.М. Макаров¹, А.В. Поздняков², С.В.Протасеня^{1*}

¹Санкт-Петербургский государственный университет телекоммуникаций им. проф. М.А. Бонч-Бруевича, Санкт-Петербург, 193232, Российская Федерация

²Санкт-Петербургский государственный педиатрический медицинский университет, Санкт-Петербург, 194100, Российская Федерация

*Адрес для переписки: saityodabur@yandex.ru

Информация о статье

УДК 519.61

Язык статьи – русский

Ссылка для цитирования: Макаров Л.М., Поздняков А.В., Протасеня С.В. Эргодическая модель атомарных конструкций // Труды учебных заведений связи. 2018. Т. 4. № 3. С. 74–84.

Аннотация: Рассмотрены вопросы естественнонаучного познания законов Природы, актуализирующих проблему позиционирования атомов химических элементов. Используя известные физические концепции и методологию информационного поиска существенных показателей события, формируется эндогенная модель атома. Предлагается набор математических процедур, характеризующих системный характер изменения свойств атомарных конструкций. Реализуется серия вычислительных процедур оценки массы атомов по всему натуральному числовому ряду. Устанавливается наличие несчетного множества атомарных конструкций, обладающих фундаментальным свойством экспоненциального уменьшения времени существования.

Ключевые слова: атом, модель атомарной конструкции, система атомов.

Современная микроэлектроника обладает большим потенциалом развития, который всесторонне обеспечивает физическая химия. Наличие передовых технологий создания перспективных полупроводниковых элементов и технологий точного воспроизведения микроэлектронных средств обеспечивается обширными исследовательскими проектами в области химии. Основой для химических проектов по-прежнему остается современная концепция построения периодической таблицы химических элементов, созданной Д.И. Менделеевым в 1869 г. В 2019 г. мировое научное сообщество отмечает 150-летие создания концепции позиционирования атомарных конструкций химических элементов. Одновременно с этим событием отмечается 100-летие международной организации теоретической и прикладной химии [1]. Эти знаменательные события актуализируют разработки теоретических и практических проектов.

В историческом отношении химия как научная дисциплина создала много специализированных направлений: физическую химию, химию полупроводников, радиационную химию, биохимию и целый ряд других. В этой плеяде научных разработок, в соответствии с современными условиями проведения исследований, все больше внимания

уделяется моделированию, построению моделей развития событий, в том числе и в области химии. Воспроизведение этих проектов на практике требует рассмотрения широкого перечня вопросов, вскрывающих как существо проблемы, так и принципы построения модели.

Выдвижение новых проектов и гипотез по созданию конструктивных образцов материалов требует больших финансовых затрат. С этой точки зрения привлечение информационных технологий на исследования и формирования концептуальных воззрений на Природные процессы позволяет систематизировать накопленный потенциал знаний и априори указать тренд развития событий.

Обычно поиск инновационных проектов проводится с учетом практической значимости предлагаемых концептов. Так, например, в области химии используют в качестве основы суждения группы экспертов, создается оценка величины информационной компоненты проекта, которая безусловно формируется в базисе хороших знаний химии, математики, физики, смежных дисциплин по организации и поддержке научного сегмента промышленного производства, например, микроэлектроники.

Большое количество разработок микроэлектроники основано на фундаментальных знаниях,

полученных в области общей химии и физической химии, с использованием разнообразных средств математики. Физический микромир конструирования изделий микроэлектроники широко оперирует атомарными конструкциями химических элементов [2], в том числе сложными информационными моделями на основе изотопов [3, 4].

Аналогичная ситуация по организации исследовательских работ просматривается в практической медицине – радиоизотопной диагностике. Проекты, ориентированные на пополнение базы знаний о химических – биохимических процессах в живых организмах, требуют систематизации атомарных конструкций химических элементов. Опираясь представлениями о системности атомарных конструкций в терминах математической модели, представляется возможным априори синтезировать новые понятия: адекватности, сопряженности, коррелированности, синхронности, соотносимые с большим количеством биохимических процессов сложной диссипативной системы.

С другой стороны, высокая степень потребности в новых химических материалах, как в области микроэлектроники, так и в области практической медицины, обладающих предсказуемыми свойствами, порождает необходимость систематизации имеющегося потенциала знаний и технологий. Системный подход в области химии формируется на представлении о взаимной связи атомарных конструкций в периодической системе элементов. Это базовый тезис установлен и многократно проверен на практике. Формирование этого тезиса происходило длительное время исключительно в терминах химии. С появлением результатов физических исследований атомарных конструкций базовые положения атомарно молекулярной концепции модифицируются. Физико-химические постулаты формируют квантовую теорию, где постулируется, интуитивно предсказанный Аристотелем, механизм постоянной трансформации атомарной конструкции. Физическая реальность, в философском понимании, постоянно изменяется, хотя этот процесс обладает очень большой продолжительностью. Для каждой атомарной конструкции существует свой период существования. И это соответствует действительности. В ряде случаев такие изменения можно наблюдать на относительно коротком интервале времени. Это феномен распада атомарной конструкции – изотопа, обладающего некоторым обозримым периодом жизни. Обобщенное представление этого процесса формирует представление о существовании процесса трансформации, который может происходить поэтапно. И это действительно наблюдается с некоторыми атомарными конструкциями, например, ураном и торием, которые порождают цепные события. Эти атомарные конструкции с определенным периодом полураспада, в свою очередь, порождают дочерние продукты,

которые также, по истечению определенного времени, распадаются. Установленные явления хорошо известны в квантовой физике и хорошо согласуются законами термодинамики. Стремление любой материальной системы, в том числе и атомарной конструкции, переходить в менее упорядоченное состояние, является предметом рассмотрения в теории динамических систем – эргодических систем. На более высоком уровне организации материи эти закономерности наблюдаются и с планетарными объектами, обладающими также определенным периодом существования в Природе.

Изменение состояния сложных систем изучают давно, и, часто основным инструментом здесь является математическая статистика. Сложные системы, состояние которых меняется, относят к динамическим системам. Отмечено, что для природных динамических систем, к которым можно отнести атомарные конструкции, процесс эволюции сопровождается серией переходов из одного состояния в другое. Формально выделяя трек и точки смены состояний, отмечается близость свойств системы в этих точках. На практике наблюдаемый эффект подвергается детальному изучению с целью установления формализованного описания. Многочисленность природных объектов и явлений, обладающих этим эффектом, сформировали основы теории эргодических систем. В качестве основного инструмента исследования эргодических систем принято рассматривать статистические методы. В ряде случаев это приносит успех исследователю. Некоторые результаты исследований убедительно показывают возможность построения описания эволюции системы, созданного по наблюдению за отдельными элементами. Другие результаты свидетельствуют о возможности получения позитивных знаний об эволюции элементов системы, по результатам наблюдений за всей системой в целом.

Несмотря на немногочисленность таких примеров, отчетливо просматривается общий тезис: система, относимая к категории динамических систем, не является простым набором элементов ее составляющих. Хотя этот тезис в данном виде не демонстрирует возможность построения конструктивных правил априорного развития событий, но допускает существование связи разнесенных по времени событий.

Для атомарных конструкций, которых в настоящее время известно достаточно много, такое суждение актуализирует проблему формализма позиционирования. Другими словами, проблема позиционирования атомарных конструкций может формироваться в терминах и понятиях эргодических систем, с учетом физико-химических свойств.

Действительно, ограничиваясь только понятиями массы (A), заряда (p), количеством нейтронов

(n) и номера (N) атомарной конструкции A_{pn}^N , можно воспроизвести принятую сегодня систему позиционирования химических элементов. Эти представления о принципе позиционировании атомарных конструкций в формате таблицы, почти не изменились за прошедшие 150 лет. Хотя, по фактическому, можно сказать, другому информационному материалу, добытому в физических исследованиях, накопились существенные дополнения к первоначальной концепции.

В анализе имеющегося материала по данной тематической линии можно выделить тезис о ядерных процессах, которые поддерживаются протоном и нейтроном. Это многообразная и сложная тема исследований в ядерной физике. Выделим основные понятия, создаваемые на основе физических терминов: заряда, некоторого количества протонов и нейтронов, находящихся в атоме [5]. В понятиях эргодической системы отдельная атомарная конструкция и набор атомарных конструкций (химических элементов) обладают дуализмом. Другими словами, это две равнозначные и не обладающие общим началом системы. Понятие равнозначности следует трактовать, как способность каждой из систем влиять на эволюцию другой системы. С этой точки зрения понятие общего начала можно трактовать исключительно в физических терминах. Действительно, в качестве первого, наиболее распространенного в Природе, элемента указывается водород. Атом водорода находится на первой позиции периодической системы химических элементов. Атом водорода соотносим со сложной системой, обладающей двумя элементами, определяющими свойство атомарной конструкции. Это ядро и электрон. В последующем ряду трансформации этой конструкции происходит увеличение количества электронов, что повышает сложность атома как системы.

В конструктивном отношении любой атом с массой ($A = p + n$) состоит из нуклонов: положительно заряженных протонов (p) и нейтральных нейтронов (n). Протон и нейтрон относят к элементарным частицам – барионам. Эти частицы очень похожи и обладают почти симметричными свойствами: незначительная разница масс и противоположность зарядов. Протон обладает положительным, нейтрон нулевым зарядом. Сочетание этих свойств на двух частицах в одной системе – атоме, создает прецедент уникальных модификаций ядра.

Количество протонов (p) характеризует величину заряда атома, которая соответствует порядковому номеру (N) атомарной конструкции в периодической таблице химических элементов. Порядковый номер (N) атома представлен натуральным числом. Количество протонов определяет конфигурацию электронной оболочки атома. Вариации по количеству протонов и нейтронов порождают наличие множества модификаций ато-

марной конструкции. Атомные конструкции, обладающие одинаковым числом протонов и разным числом нейтронов, называют изотопами. Нуклиды разных химических элементов, имеющие одинаковое массовое число (A), называются изобарами. Например, изобарами являются ^{40}Ar , ^{40}K , ^{40}Ca . По аналогии названы изотонами нуклиды, имеющие одинаковое количество нейтронов (n), но различающиеся по числу протонов (p) в ядре.

Развивая эти представления, обратим внимание на сложный процесс взаимодействия протонов и нейтронов в ядре атомарной конструкции. Известно, что для каждой атомарной конструкции A_{pn}^N обнаруживается набор изотопов W_p^n . Действительно, $A^N \Rightarrow W_p^n$ [8, 10]. Сегодня известно (обнаружено) более 2000 изотопов [6]. Иллюстрация позиционирования некоторых атомарных конструкций – изотопов W_p^n представлена в таблице 1. Отмечаем, что для любого изотопного набора W_p^n , относительно конкретной атомарной конструкции A_{pn}^N , справедливо условие: $p = \text{const}$ и $n = \text{var}$.

На этом понятии создается общепринятый эталон атомарной массы химического элемента – атомарной конструкции. Фиксация этого понятия проводится на естественнонаучных представлениях о средней массе атома, формируемой на основе разных «природных примесей» и с учетом распространенности атомной конструкции в Природе. Это положение закрепляется международным регламентом, устанавливающим атомарные массы химических элементов. Используя эти рекомендации, создается современная периодическая таблица химических элементов. Принимается концепция: каждая атомарная конструкция обладает уникальной массой, заряд которой соответствует натуральному порядковому числу. Из этого тезиса следует, что все атомарные конструкции, условно, ранжированы по массе и порядковому номеру. Очевидно, что такая концепция построения периодической системы химических элементов имеет существенную доминанту субъективности, так как масса атома определена обобщенно.

Однако есть исключения. Известно, что атомарные конструкции аргона, калия, кобальта, никеля, теллура, йода, тория и протактиния, размещены в периодической таблице с нарушением общего правила возрастающей массы. Этот артефакт нивелируется исключительно физическими понятиями, которые оперируют различиями изотопного состава атомарных конструкций. Другими словами, используя представления о связи $A^N \Rightarrow W_p^n$, декларируется сохранение основного принципа позиционирования атомарных конструкций на основе изотопного набора. Следует признать, что такая трактовка существа проблемы о позиционировании атомарных конструкций имеет высокий уровень фундаментальности понимания процессов Природы.

ТАБЛИЦА 1. Изотопные наборы атомарных конструкций

	1		2		3		4		5		6	
	H		He		Li		Be		B		C	
	<i>p</i>	<i>n</i>	<i>p</i>	<i>n</i>	<i>p</i>	<i>n</i>	<i>p</i>	<i>n</i>	<i>p</i>	<i>n</i>	<i>p</i>	<i>n</i>
Изотопные наборы W_p^n												
0	1	0	2	0	3	0						
1	1	1	2	1	3	1	4	1	5	1		
2	1	2	2	2	3	2	4	2	5	2	6	2
3	1	3	2	3	3	3	4	3	5	3	6	3
4	1	4	2	4	3	4	4	4	5	4	6	4
5	1	5	2	5	3	5	4	5	5	5	6	5
6	1	6	2	6	3	6	4	6	5	6	6	6
7			2	7	3	7	4	7	5	7	6	7
8			2	8	3	8	4	8	5	8	6	8
9					3	9	4	9	5	9	6	9
10					3	10	4	10	5	10	6	10
11							4	11	5	11	6	11
12							4	12	5	12	6	12
13									5	13	6	13
14									5	14	6	14
15									5	15	6	15
16									5	16	6	16
17											6	17

Положим, имеется изотопный набор W_p^n , представленный количеством протонов и количеством нейтронов, образующих нуклид [7, 8].

Введем в рассмотрение метрическое евклидово пространство, для которого определена метрика $L = N(p, n)$ [9]. Представления о таком пространстве позволяют рассматривать пространственные объекты. По существу, современные представления об атомарной конструкции оперируют объемной формой. Для определенности положим, что рассмотрению подлежит сфера. В таком случае проекцией этого объекта на плоскость евклидова пространства является окружность, радиус которой определяется по выражению:

$$R = \frac{\sqrt{p^2 + n^2}}{2}. \quad (1)$$

Принимая во внимание определение метрики L , будем рассматривать вписанный в окружность прямоугольник со сторонами p и n . Избранное построение позволяет установить количественные показатели для набора W_p^n [9, 4], посредством операции вычисления параметра $S_j^N = (p_N * n_j)$, характеризующего площадь прямоугольника со сторонами p и n . Так, например, используя данные из таблицы 1, для атомарной конструкции водорода имеем:

$$H \Rightarrow W_1^6 = (S_{11}, S_{12}, \dots, S_{16}) \Rightarrow (p_{\text{const}-N=1} * n_1^6). \quad (2)$$

По аналогии применим эту процедуру к последующим атомарным конструкциям: $He \Rightarrow W_2^8$; $Li \Rightarrow W_3^{10}$; $Be \Rightarrow W_4^{12}$ и далее.

Формально, многократное использование данной процедуры инициализирует создание матричной структуры метрического пространства.

$$\begin{pmatrix} S_{11}, S_{12}, \dots, S_{1i}, \dots, S_{1N} \\ S_{21}, S_{22}, \dots, S_{2i}, \dots, S_{2N} \\ \dots \dots \dots S_{ji} \dots \dots \dots \\ S_{M1}, S_{M2}, \dots, S_{MN} \end{pmatrix} = F. \quad (3)$$

Метрическое пространство представлено двумя координатными осями: N (номер атомарной конструкции) и n (количественный показатель нейтронов). Заметим, что количественные показатели равны $N = p$. В таком случае, имеем: ось абсцисс – p , а ось ординат – n . Принимая во внимание геометрическую интерпретацию параметров A_{pn}^N , сформируем стек расчетных значений площади прямоугольника, вписанного в окружность (таблица 2).

ТАБЛИЦА 2. Метрическое пространство группы из десяти A_{pn}^N

		H	He	Li	Be	B	C	N	O	F	Ne
	$N=p$	1	2	3	4	5	6	7	8	9	10
e	n	W_p^n									
	0	0	0	0							
5	1	1	2	3	4	5					
6	2	2	4	6	8	10	12				
7	3	3	6	9	12	15	18	21			
8	4	4	8	12	16	20	24	28	32		
9	5	5	10	15	20	25	30	35	40	45	
10	6	6	12	18	24	30	36	42	48	54	60
11	7		14	21	28	35	42	49	56	63	70
12	8		16	24	32	40	48	56	64	72	80
13	9			27	36	45	54	63	72	81	90
	10			30	40	50	60	70	80	90	100
	11				44	55	66	77	88	99	110
	12				48	60	72	84	96	108	120
	13					65	78	91	104	117	130
	14					70	84	98	112	126	140
							90	105	120	135	150
							96	112	128	144	160
								119	136	153	170
								126	144	162	180
									152	171	190
									160	180	200
										189	210
										198	220
											230
											240
		q6	q8	q10	q12						
		6	8	10	12						

Дальнейшее сопоставление результатов вычислений, представленных в таблице 2, абсолютно

Показатель e вычисляется построчно по выражению:

$$e = \frac{S_{i_max}}{S_{i_min}}. \quad (4)$$

На схеме (рисунок 1) представлены изотопы: количество протонов возрастает слева направо; количество нейтронов возрастает сверху вниз [6].

p →	1	2	3																												
n ↓	H	He	Li	4	5																										
0	¹ H	² He	³ Li	Be	B	6																									
1	² D	³ He	⁴ Li	⁵ Be	⁶ B	C	7																								
2	³ T	⁴ He	⁵ Li	⁶ Be	⁷ B	⁸ C	N	8																							
3	⁴ H	⁵ He	⁶ Li	⁷ Be	⁸ B	⁹ C	¹⁰ N	O	9																						
4	⁵ H	⁶ He	⁷ Li	⁸ Be	⁹ B	¹⁰ C	¹¹ N	¹² O	F	10																					
5	⁶ H	⁷ He	⁸ Li	⁹ Be	¹⁰ B	¹¹ C	¹² N	¹³ O	¹⁴ F	Ne	11	12																			
6	⁷ H	⁸ He	⁹ Li	¹⁰ Be	¹¹ B	¹² C	¹³ N	¹⁴ O	¹⁵ F	¹⁶ Ne	Na	Mg	13	14																	
7	⁸ He	¹⁰ Li	¹¹ Be	¹² B	¹³ C	¹⁴ N	¹⁵ O	¹⁶ F	¹⁷ Ne	¹⁸ Na	¹⁹ Mg	Al	Si	15																	
8	¹⁰ He	¹¹ Li	¹² Be	¹³ B	¹⁴ C	¹⁵ N	¹⁶ O	¹⁷ F	¹⁸ Ne	¹⁹ Na	²⁰ Mg	²¹ Al	²² Si	P	16																
9	¹² Li	¹³ Be	¹⁴ B	¹⁵ C	¹⁶ N	¹⁷ O	¹⁸ F	¹⁹ Ne	²⁰ Na	²¹ Mg	²² Al	²³ Si	²⁴ P	S	17																
10	¹³ Li	¹⁴ Be	¹⁵ B	¹⁶ C	¹⁷ N	¹⁸ O	¹⁹ F	²⁰ Ne	²¹ Na	²² Mg	²³ Al	²⁴ Si	²⁵ P	²⁶ S	Cl	18															
11	¹⁵ Be	¹⁶ B	¹⁷ C	¹⁸ N	¹⁹ O	²⁰ F	²¹ Ne	²² Na	²³ Mg	²⁴ Al	²⁵ Si	²⁶ P	²⁷ S	²⁸ Cl	Ar	19															
12	¹⁶ Be	¹⁷ B	¹⁸ C	¹⁹ N	²⁰ O	²¹ F	²² Ne	²³ Na	²⁴ Mg	²⁵ Al	²⁶ Si	²⁷ P	²⁸ S	²⁹ Cl	³⁰ Ar	K	20														
13	¹⁸ B	¹⁹ C	²⁰ N	²¹ O	²² F	²³ Ne	²⁴ Na	²⁵ Mg	²⁶ Al	²⁷ Si	²⁸ P	²⁹ S	³⁰ Cl	³¹ Ar	³² K	Ca	21														
14	¹⁹ B	²⁰ C	²¹ N	²² O	²³ F	²⁴ Ne	²⁵ Na	²⁶ Mg	²⁷ Al	²⁸ Si	²⁹ P	³⁰ S	³¹ Cl	³² Ar	³³ K	³⁴ Ca	Sc	22													
15	²⁰ B	²¹ C	²² N	²³ O	²⁴ F	²⁵ Ne	²⁶ Na	²⁷ Mg	²⁸ Al	²⁹ Si	³⁰ P	³¹ S	³² Cl	³³ Ar	³⁴ K	³⁵ Ca	³⁶ Sc	Ti	23												

Рис. 1. Схема позиционирования атомарных конструкций W_p^n

Анализируя данные таблицы 2, выделим общий элемент S_{ji} матрицы F (3). Отмечаем, что каждое последнее значение в столбце $S_{j_{\max}}^N$ позволяет вычислить показатель q :

$$q_j = \frac{S_j^N}{p_N}, \quad (5)$$

$$q_{p+1} - q_p = h_j = 2 = \text{const.} \quad (6)$$

Представленные показатели q_j и h_j достаточно просто вычисляются на массиве значений матрицы F или по данным таблицы 2.

Выделим базовые положения рассматриваемой эндогенной модели. Отметим, что построение метрического пространства позиционирования атомарных конструкций осуществляется исклю-

чительно на известных физических параметрах W_p^n [10]. Далее, введенная в рассмотрение метрика пространства обеспечивает отображение всех W_p^n , а полученные показатели q_j и h_j обладают системными признаками и представлены действительными числами. Тогда, согласно теореме Кантора [11] множество действительных чисел несчетно, а, следовательно, исключается возможность указать финишную, реально существующую Природную атомарную конструкцию. Количество атомарных конструкций континуально и обладает системной связью, обеспечивающей вычисление количественных показателей.

Проиллюстрируем представленные концептуальные положения на нескольких примерах. Рассмотрим атомарную конструкцию W_p^n – для водорода (таблица 3). Согласно информационному ис-

точнику [10, 6] представим данные и проведем расчет параметра S по выражению:

$$S_j^N = (p_N * n_j). \quad (7)$$

ТАБЛИЦА 3. Атомарная конструкция W_p^n водород

$N = 1$	Водород						
p	1	1	1	1	1	1	1
n	0	1	2	3	4	5	6
S	0	1	2	3	4	5	6

Проведем аналогичный расчет для гелия и лития (таблицы 4, 5).

Таблица 4 Атомарная конструкция W_p^n гелий

$N = 2$	Гелий							
p	2	2	2	2	2	2	2	2
n	0	1	2	3	4	5	6	7
S	0	2	4	6	8	10	12	14

ТАБЛИЦА 5 Атомарная конструкция W_p^n литий

$N = 3$	Литий									
p	3	3	3	3	3	3	3	3	3	3
n	0	1	2	3	4	5	6	7	8	9
S	0	3	6	9	12	15	18	21	24	27

Аналогичные расчеты можно провести для последующих атомарных конструкций. Зафиксируем расчетные показатели S_n^N (рисунок 2).

Отмечаем, что с точки зрения формальной записи, для произвольно выбранной атомарной конструкции однозначно сопоставляется такой набор W_p^n , что любой элемент набора W_p^n характеризуется вполне определенной парой показателей p и n , отождествляемых с координатными осями метрического пространства.

Выделим верхнее и нижнее значение показателя S_{ji}^N . Очевидно, что для любого S^N всегда найдется $S_{\min}^N$ и $S_{\max}^N$. Идентифицируем $S_{\min}^N$ – «верхний бьеф» (top bief) и $S_{\max}^N$ – «нижний бьеф» (lower bief), как границы области существования атомарной конструкции. Тогда выражения для определения границ установленной области представим в виде:

$$H_{\text{topbief}} = N^2 - 4N = H_{\text{tb}} \quad (8)$$

и

$$G_{\text{lowerbief}} = 2N^2 + 4N = G_{\text{lb}}. \quad (9)$$

В таком случае для любой произвольно выбранной атомарной конструкции A_{pn}^N химического элемента можно однозначно указать набор W_p^n , для которого вычисляются показатели H_{tb} и G_{lb} . Этот постулат, базис которого составляют выражения (8) и (9), положен в основу модели метрического пространства атомарных конструкций химических элементов – эргодическую модель. Очевидно, и понятно, что такая процедура на мет-

рическом пространстве с метрикой L однозначно указывает позицию химического элемента. Тогда справедлив тезис о наличии формализованной связи между атомарными конструкциями.

		Количество протонов (p)									
		1	2	3	4	5	6	7	8	9	10
Количество нейтронов (n)	0	H	He	Li	Be	B	C	N	O	F	Ne
	1	0	0	0	0						
	2	1	2	3	4	5					
	3	2	4	6	8	10	12				
	4	3	6	9	12	15	18	21			
	5	4	8	12	16	20	24	28	32		
	6	5	10	15	20	25	30	35	40	45	
	7	6	12	18	24	30	36	42	48	54	60
	8		14	21	28	35	42	49	56	63	70
	9		16	24	32	40	48	56	64	72	80
	10			27	36	45	54	63	72	81	90
	11			30	40	50	60	70	80	90	100
	12				44	55	66	77	88	99	110
	13				48	60	72	84	96	108	120
	14					65	78	91	104	117	130
	15					70	84	98	112	126	140
	16						90	105	120	135	150
	17						96	112	128	144	160
	18							119	136	153	170
	19							126	144	162	180
	20								152	171	190
	21								160	180	200
	22									189	210
	23									198	220
	24										230

Рис. 2. Иллюстрация позиционирования показателей S_n^N для разных атомарных конструкций W_p^n

Проверим этот тезис. Положим, согласно [10] существует атомарная конструкция $N = 11$, для которой установлен изотопный набор W_p^n . Проведем вычисление S_n^{11} по выражению (7). Создадим таблицу 6 расчетных показателей.

Сопоставив адекватность табличных данных и данных, полученных по выражениям модели, определим значение $S_{\min}^N$ по выражению (8) и $S_{\max}^N$ по выражению (9) и выведем, что:

$$H_{\text{tb}} = N^2 - 4N = 11^2 - 4 * 11 = 77, \quad (10)$$

$$G_{\text{lb}} = 2N^2 + 4N = 2 * 11^2 + 4 * 11 = 286. \quad (11)$$

Получили результаты, соответствующие расчетным значениям модели. Установленный принцип формирования суждения о численных граничных значениях области позиционирования атомарной конструкции W_p^n , можно тиражировать на всех значениях натурального ряда чисел. Это означает, что обладая знаниями только о непрерывности и бесконечности натурального ряда чисел, можно определить область допустимых значений S^N , которые являются основой для атомарной конструкции W_p^n . Другими словами, эндогенная модель позволяет предсказать не только размер атомарной конструкции изотопного набора W_p^n – количество изотопов для атома под конкретным позиционным номером, но перечислить все изотопы набора.

ТАБЛИЦА 6. Изотопный набор для атомарной конструкции $N = 11$ (Натрий)

№	P	n	S_n^{11}
1	11	7	77
2	11	8	88
3	11	9	99
4	11	10	110
5	11	11	121
6	11	12	132
7	11	13	143
8	11	14	154
9	11	15	165
10	11	16	176
11	11	17	187
12	11	18	198
13	11	19	209
14	11	20	220
15	11	21	231
16	11	22	242
17	11	23	253
18	11	24	264
19	11	25	275
20	11	26	286

Развивая эти представления, следует принять во внимание ранее установленный факт о бесконечности натурального ряда чисел, которые идентифицируют заряд ядра атома. В таком случае для любого произвольного натурального числа N всегда найдется атомарная конструкция, в которой абсолютно точно и полно представлен полный набор W_p^n . Такое утверждение является следствием наличия функциональной зависимости, представленной выражениями (8) и (9), составляющих основу модели позиционирования атомарных конструкций химических элементов.

Анализируя принцип формирования значений S_j^N , входящих в набор W_p^n , отмечаем, что переход от начального значения H_{tb}^N к финишному G_{lb}^N значению осуществляется по рекурсивной процедуре:

$$n_1 = \frac{H_{tb}^N}{N}, n_2 = \frac{H_{tb}^N + N}{N}, \dots, n_j = \frac{H_{tb}^N + jN}{N}, \quad (12)$$

где $j = \frac{G_{lb}^N}{N}$ характеризует количество нейтронов изотопа в атомарной конструкции W_p^n . Начальный и конечный показатель нейтронов в конструкции определяются по выражению:

$$nH = \frac{H_{tb}^N}{N}, \quad nG = \frac{G_{lb}^N}{N}. \quad (13)$$

Наличие процедуры, определяемой выражением (12), позволяет проводить конструирование атомарных конструкций, принимая во внимание только порядковый номер N атомарной конструкции.

Апробирование указанных положений проведем на основе выражений (6), (7) и (10) применительно к атомарной конструкции под номером 119 – W_p^n , описание которой отсутствует в современных информационных источниках.

Для иллюстративности проведем расчеты по уравнениям модели для нескольких атомарных конструкций, представленных порядковыми номерами: с $N = 119$ по $N = 127$. Поскольку размерность стека значений для каждой атомарной конструкции получается достаточно большой, представим результаты в сокращенном виде в таблице 7 (начальные – часть промежуточных и финишные значения).

ТАБЛИЦА 7. Структура атомарных конструкций по параметру S_j^N

Номер атомарной конструкции									
	119	120	121	122	123	124	125	126	127
S_j									
H_{tb}	13685	13920	14157	14396	14637	14880	15125	15372	15621
	13804	14040	14278	14518	14760	15004	15250	15498	15748
	13923	14160	14399	14640	14883	15128	15375	15624	15875
	14042	14280	14520	14762	15006	15252	15500	15750	16002
	14161	14400	14641	14884	15129	15376	15625	15876	16129
	14280	14520	14762	15006	15252	15500	15750	16002	16256
	14399	14640	14883	15128	15375	15624	15875	16128	16383
.....									
G_{lb}	28798	29280	29766	30256	30750	31248	31750	32256	32766
nH	115	116	117	118	119	120	121	122	123
nG	242	244	246	248	250	252	254	256	258

Полагаем, что для каждого номера атомарной конструкции существует уникальный набор изотопов. Например, для $N = 119$ такой набор формируется на основе данных таблицы 7, посредством процедуры выделения только нейтронной компоненты по выражению (12). Результаты вычислений представлены в таблице 8.

Первый изотоп атомарной конструкции с $N = 119$ обладает $p = 119$ и $n = 115$ (таблица 8). Последующие изотопы этой атомарной конструкции демонстрируют превосходящее над позитронами количество нейтронов.

Так, например, последний в списке изотоп этой конструкции, находящийся под номером 128, обладает количеством нейтронов $n = 15228$. Столь значительное количество нейтронов в атомарной конструкции реализует физическую структуру протон – нейтрон, обладающую большой массой и малым временем жизни.

Отмечено, что с увеличением порядкового номера атома химического элемента повышается масса. Масса атома является одним из основных показателей. Понятие массы атома формируется

на представлении о несхожести атомарных конструкций, образующих группу атомов избранного химического элемента.

ТАБЛИЦА 8. Изотопный набор для атомарной конструкции $N = 119$

W_p^n $p = \text{const} = 119$; $n = \text{var}$									
1	2	3	4	5	6	7	8	9	10
115	234	353	472	591	710	829	948	1067	1186
11	12	13	14	15	16	17	18	19	20
1305	1424	1543	1662	1781	1900	2019	2138	2257	2376
21	22	23	24	25	26	27	28	29	30
2495	2614	2733	2852	2971	3090	3209	3328	3447	3566
31	32	33	34	35	36	37	38	39	40
3685	3804	3923	4042	4161	4280	4399	4518	4637	4756
41	42	43	44	45	46	47	48	49	50
4875	4994	5113	5232	5351	5470	5589	5708	5827	5946
51	52	53	54	55	56	57	58	59	60
6065	6184	6303	6422	6541	6660	6779	6898	7017	7136
61	62	63	64	65	66	67	68	69	70
7255	7374	7493	7612	7731	7850	7969	8088	8207	8326
71	72	73	74	75	76	77	78	79	80
8445	8564	8683	8802	8921	9040	9159	9278	9397	9516
81	82	83	84	85	86	87	88	89	90
9635	9754	9873	9992	10111	10230	10349	10468	10587	10706
91	92	93	94	95	96	97	98	99	100
10825	10944	11063	11182	11301	11420	11539	11658	11777	11896
101	102	103	104	105	106	107	108	109	110
12015	12134	12253	12372	12491	12610	12729	12848	12967	13086
111	112	113	114	115	116	117	118	119	120
13205	13324	13443	13562	13681	13800	13919	14038	14157	14276
121	122	123	124	125	126	127	128		
14395	14514	14633	14752	14871	14990	15109	15228		

Это справедливо, поскольку в Природе не существует абсолютно одинаковых атомарных конструкций. Группа из нескольких конструкций атомов, характеризующая свойства материального объекта, относимого к избранному химическому элементу, соотносится со средним значением массы. Это соответствует действительности, поскольку входящие в группу атомы отличаются по количеству нейтронов. Это так называемый изотопный набор атомарных конструкций, существующий в Природе под установленным номером N . Для такого набора характерно использовать усредненное значение массы. В качестве единицы измерения атомных и молекулярных масс принята $1/12$ часть массы нейтрального атома наиболее распространенного изотопа углерода.

Суммарное количество протонов и нейтронов характеризуется массовым числом. Относительно небольшие различия в массах протона и нейтрона, а

также их количествах в ядре атома, порождает дефект масс, который может иметь как положительный, так и отрицательный знак. Наличие этого природного эффекта в атомарной конструкции укрепляет позиции статистических методов оценки массы атома. Модифицируя эти представления, можно полагать, что дефект масс, в группе атомов можно обосновать нарушением симметричности сферы – гипотетической оболочки атома.

Известно, что наличие набора W_p^n воспроизводит границы области допустимой реализации атомарной конструкции. Численные показатели области вычисляются по выражениям (8) и (9). Оценим размер области допустимой реализации атомарной конструкции. Введем показатель:

$$\lambda = \frac{nG}{nH}. \quad (14)$$

Определим диаметр сферы, с учетом порядкового номера атомарной конструкции, по выражению:

$$D_s = \frac{N}{\lambda} + 2N. \quad (15)$$

Диаметр сферы D_s соотносим с массой атомарной конструкции $A^{N_{pn}}$. $D_s \Rightarrow A_{pn}^N \Rightarrow A$. С учетом известной массы A атома величину ошибки расчета определим по выражение:

$$W\% = ((A - D_s)/A)100. \quad (16)$$

Результаты представлены в таблице 9.

ТАБЛИЦА 9. Расчетные показатели массы атомарных конструкций

№	12	13	14	15	16	17	18
Имя	Магний	Алюминий	Кремний	Фосфор	Сера	Хлор	Аргон
A а.е.м.	24,305	26,98154	28,086	30,97376	32,06	35,453	39,948
H	96	117	140	165	192	221	252
n _H	8	9	10	11	12	13	14
G	336	390	448	510	576	646	720
n _G	28	30	32	34	36	38	40
λ	3,50	3,33	3,20	3,09	3,00	2,92	2,86
μ	3,43	3,90	4,38	4,85	5,33	5,82	6,30
D _s	27,43	29,90	32,38	34,85	37,33	39,82	42,30
W%	12,85	10,82	15,27	12,52	16,45	12,31	5,89

Условные обозначения: A – атомная масса химического элемента; H – «верхний бьеф»; G – «нижний бьеф»; n_H – количество нейтронов на верхнем бьефе; n_G – количество нейтронов на нижнем бьефе; λ – асимметрия сферы; μ – эллиптичность сферы; D_s – показатель расчетной массы атома; $W\%$ – погрешность расчета массы атома.

Процедура расчета оценок массы проведена по установленному ЮПАК регламенту использования численных значений масс известных в настоящее время атомарных конструкций химических элементов [1].

Следует отметить, что относительно широкое распространение в Природе некоторых химических элементов создает метрологические трудности точного определения массы атомов. По мнению ЮПАК это порождает некоторую вариативность значений масс атомов, представленных в многочисленных информационных источниках. Данный факт следует учитывать при создании системы оценок атомарных конструкций.

Следуя этим представлениям, рассмотрим типовой набор атомарных конструкций из средней части известной периодической таблицы химических элементов. В качестве начального элемента этого списка выберем ксенон, обладающий 54 порядковым номером (таблица 10). Можно отметить, что такой выбор первого элемента в списке сделан, но основании известных представлений о инертном газе (ксенон), являющимся типичным представителем этого кластера химических элементов.

ТАБЛИЦА 10. Расчетные показатели массы атомарных конструкций

№	54	55	56	57	58	59	60
Имя	Ксенон	Цезий	Барий	Лантан	Церий	Празеодим	Неодим
A a.e.m.	131,3	132,9054	137,33	138,9	140,12	140,9	144,24
H	2700	2805	2912	3021	3132	3245	3360
n _H	50	51	52	53	54	55	56
G	6048	6270	6496	6726	6960	7198	7440
n _G	112	114	116	118	120	122	124
λ	2,24	2,24	2,23	2,23	2,22	2,22	2,21
μ	24,11	24,61	25,10	25,60	26,10	26,60	27,10
Ds	132,11	134,61	137,10	139,60	142,10	144,60	147,10
W%	0,61	1,28	-0,16	0,51	1,41	2,62	1,98

Для другой группы атомов химических элементов, которые менее распространены в Природе, а возможно известны научному сообществу только благодаря экспериментальным исследованиям, используются четкие методики определения массы. Массы этих элементов обладают малой вариативностью, а, следовательно, ошибки определения массы – малы.

Представляет интерес рассмотреть атомарные конструкции с дальними порядковыми номерами, массы которых не известны, а лишь обсуждаются в научном сообществе (таблица 11). В качестве первого в этом списке химических элементов выберем резерфордий. Этот химический элемент, хотя и обладает малым временем жизни, достаточно полно изучен. Реализуем эти научные устремления расчетами массы атомов химических элементов на основе созданной модели. Проведем необходимые расчеты и представим результаты в таблице 12.

ТАБЛИЦА 11. Расчетные показатели массы атомарных конструкций

№	104	105	106	107	108	109	110
Имя	Резерфордий	Дубний	Сиборгий	Борий	Хассий	Мейтнерий	Дармштадтий
A a.e.m.	261	262	266	267	269	276	277
H	10400	10605	10812	11021	11232	11445	11660
n _H	100	101	102	103	104	105	106
G	22048	22470	22896	23326	23760	24198	24640
n _G	212	214	216	218	220	222	224
λ	2,12	2,12	2,12	2,12	2,12	2,11	2,11
μ	49,06	49,56	50,06	50,56	51,05	51,55	52,05
Ds	257,06	259,56	262,06	264,56	267,05	269,55	272,05
W%	-1,51	-0,93	-1,48	-0,92	-0,72	-2,34	-1,79

ТАБЛИЦА 12. Расчетные показатели массы атомарных конструкций

Имя	Элемент 119	Элемент 120	Элемент 121	Элемент 122	Элемент 123	Элемент 124	Элемент 125
A a.e.m.	вычислено	вычислено	вычислено	вычислено	вычислено	вычислено	вычислено
H	13685	13920	14157	14396	14637	14880	15125
n _H	115	116	117	118	119	120	121
G	28798	29280	29766	30256	30750	31248	31750
n _G	242	244	246	248	250	252	254
λ	2,10	2,10	2,10	2,10	2,10	2,10	2,10
μ	56,55	57,05	57,55	58,05	58,55	59,05	59,55
Ds	294,55	297,05	299,55	302,05	304,55	307,05	309,55
W%	1,8	1,2	1,9	1,82	1,22	1,8	1,15

Эндогенная модель, рассмотренная в работе, создает базис изучения сложных процессов эволюции атомарных конструкций. На базе модели можно рассмотреть атомарные конструкции и с более отдаленными порядковыми номерами. Посредством модели, использующей классические постулаты квантовой теории и методологию информационных технологий анализа физических данных, постулируется наличие бесчисленного множества атомарных конструкций, обладающих массой, вычисляемой по строго регламентированному правилам квантовой механики.

Выделение этих аспектов, в частности акцентирование внимания на показателе массы атома, происходит с учетом прошлых воззрений физической химии, а также вновь созданной модели. На рубеже 150-летнего юбилея открытия периодического закона установлена система Природного позиционирования атомов химических элементов на основе протон – нейтронной пары. Иллюстрацию этого понятия построим на серии декадных оценок массы

атомов химических элементов, полученных посредством эргодической модели (рисунок 3).

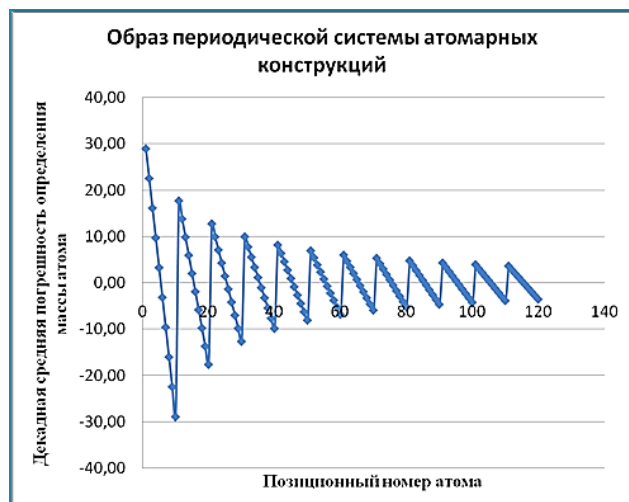


Рис. 3. Декадное отображение расчетной массы атомов

Декадный образ атомарных конструкций химических элементов хорошо иллюстрирует главные тезисы научной работы, в которых указывается на наличие сложных взаимосвязанных процессов, обладающих индивидуальными отличиями, проявляющимися как, в количестве нейтронов в атоме, так и в продолжительности жизни атома. Все это соответствует действительности, которая отчетливо просматривается в экспоненциальной образующей образа периодической системы атомарных конструкций, повсеместно присутствующих в Природе.

Действительно, увеличение атомной массы, закономерно происходящее с возрастанием порядкового номера химического элемента известно давно. Современные накопленные знания позволяют уточнить этот процесс, указав на ведущую роль количественного соотношения протонов и нейтронов в ядре атома. Однако, как известно, значительное увеличение нейтронов в ядре, инициирует процесс создания не только новых атомарных конструкций, но стремительно сокращает время жизни таких материальных тел. Такое суждение является аксиомой современной физики. Следуя этим представлениям, констатируем, что благодаря малому времени жизни таких атомов, оценка подобия увеличивается, а, следовательно, статистическая погрешность определения массы стремится к нулю. Это четко наблюдается как на серии данных модели, так и на графическом образе (рисунок 3) большого набора атомарных конструкций.

Этот феномен, составляющий концептуальную базу рассматриваемой модели, имеет аналог при описании событий регистрации фотонов и электронов, также обладающих высоким уровнем подобия по массе. Лаконичная и понятная физиче-

ская основа созданной модели, позволяет предсказать, априори указать, существование множества атомарных конструкций, обладающих определенной массой, количеством протонов и нейтронов и тем самым обогатить современный социум фундаментальным законом Природы. В работе представлен в качестве примера ограниченный набор атомарных конструкций, с номерами с 119 по 125, описание которых отсутствует в современной литературе. Эта предсказательная парадигма модели представляется новым взглядом на Природные процессы.

Следует признать, что широко обсуждаемые в научной сфере физические проблемы о строении атомарных конструкций с большими порядковыми номерами, действительно актуальны. Признано, что такие конструкции обладают особенностями, в частности малым временем жизни. Теоретическая физика предсказывает наличие этого факта для химических элементов примерно с 90 порядкового номера. Это тезис действительно подтверждается в многочисленных физических экспериментах.

На основании этого можно констатировать, что обнаружение и полное подтверждение этого явления в эксперименте, требует предварительных расчетов и хороших измерительных средств. Проведение предварительных расчетов, предсказывающих особенности строения атомарных конструкций с большими порядковыми номерами целесообразно провести по предложенной модели.

Измерительные процедуры, фиксирующие сам факт существования атомарных конструкций с высокими порядковыми номерами, обладающими уникальным изотопным набором, требует создания высокоточных микроэлектронных сенсорных устройств.

Расширяя эти представления и выделяя изотопные наборы для химических элементов, можно утверждать, что современные представления о биофизических особенностях возрастного изменения живых организмов, значительно обогатятся знаниями, полученными электронными средствами для изотопной диагностики.

Фактически, в настоящее время исключительно микроэлектроника, представляющая основу измерительной платформы современных аппаратных средств, способна установить сам факт физического существования атомарных конструкций с большими порядковыми номерами. Однако даже при наличии таких измерительных средств планирование натурного эксперимента нельзя осуществить в отсутствии модели развертки природных событий.

Список используемых источников

1. Материалы Международного союза теоретической и прикладной химии. URL: <https://iupac.org/what-we-do/periodic-table-of-elements>.
2. Макаров Л.М., Протасеня С.В. Моделирование фрактального образа атома кремния в микроэлектронике // Труды учебных заведений связи. 2018. Т. 4. № 2. С. 91–98. DOI: 10.31854/1813-324x-2018-2-82-90.
3. Вихман Э. Берклеевский курс физики: Квантовая физика. Т. 4. М: Наука, 1986. 392 с.
4. Макаров Л.М. Априорные знания об атомах химических элементов // Проблемы современной науки и образования. 2018. № 5(125). С. 6–15. DOI: 10.20861/2312-8267-2018-49-001.
5. Макаров Л.М. Интеллектуальные системы и поля понятий // Труды учебных заведений связи. 2016. Т. 2. № 3. С. 50–54.
6. Таблица нуклидов. URL: https://ru.wikipedia.org/wiki/%D0%A2%D0%B0%D0%B1%D0%BB%D0%B8%D1%86%D0%B0_%D0%BD%D1%83%D0%BA%D0%BB%D0%B8%D0%B4%D0%BE%D0%B2.
7. Макаров Л.М. Алгоритм позиционирования атомов химических элементов // EUROPEAN RESEARCH: INNOVATION IN SCIENCE, EDUCATION AND TECHNOLOGY. Collection of scientific articles XXXIX International scientific and practical conference. 2018. С. 9–16. DOI: 10.20861/2410-2873-2018-39-003.
8. Макаров Л.М., Поздняков А.В. Формализм позиционирования стабильных изотопов // Проблемы современной науки и образования. 2018. № 6(126). С. 119–124. DOI: 10.20861/2304-2338-2018-126-005.
9. Макаров Л.М. Метрическое пространство атомарных конструкций химических элементов // Наука, техника и образование. 2018. № 8(49). С. 5–9. DOI: 10.20861/2312-8267-2018-49-001.
10. Эмели Д. Элементы. М: Мир, 1993. 255 с.
11. Боголюбов А.Н. Кантор Георг // Математики. Механики. Биографический справочник. Киев: Наукова думка, 1983. 639 с.

* * *

ERGODIC MODEL OF ATOMIC DESIGNS

L. Makarov¹, A. Pozdnyakov², S. Protasenya¹

¹The Bonch-Bruевич State University of Telecommunications,
St. Petersburg, 193232, Russian Federation

²Saint-Petersburg State Pediatric Medical University,
St. Petersburg, 194100, Russian Federation

Article info

Article in Russian

For citation: Makarov L., Pozdnyakov A., Protasenya S. Ergodic Model of Atomic Designs // Proceedings of Telecommunication Universities. 2018. Vol. 4. Iss. 3. PP. 74–84.

Abstract: *The questions of natural science knowledge of the laws of Nature, actualizing the problem of positioning atoms of chemical elements. Using the known physical concepts and methodology of information retrieval of significant indicators of the event, an endogenous model of the atom is formed. A set of mathematical procedures characterizing the systemic nature of changes in the properties of atomic structures is proposed. A series of computational procedures for estimating the mass of atoms over the entire natural numerical series is implemented. The presence of an un-countable set of atomic constructions possessing the fundamental property of exponential reduction of the lifetime is established.*

Keywords: *atom, model of atomic structure, system of atoms.*

COMPARATIVE ANALYSIS OF LOW LATENCY ANONYMOUS COMMUNICATION SYSTEMS

I. Murtala¹, O.A. Tiamiyu^{2*}

¹De-Kharms Telecommunications, Ilorin, Kwara, Nigeria

²University of Ilorin, Ilorin, PMB 1515, Nigeria

*E-mail: ozutiams@yahoo.com

Article info

Article in English

For citation: Murtala I., Tiamiyu O.A. Comparative Analysis of Low Latency Anonymous Communication Systems // Proceedings of Telecommunication Universities. 2018. Vol. 4. Iss. 3. PP. 85–97.

Abstract: *With the rapid development and general acceptance of the Internet globally, concealing one's identity (remaining unknown) for protection and security reasons while communicating over Internet is of great concern. Thus, anonymity has turned into an essential and legitimate aim in numerous applications, including anonymous web surfing, location based services, and E-voting. This study compares existing implementation of various low latency anonymous communication systems, more precisely, onion routing techniques, with respect to user requirements such as performance, reliability and usability to determine their level of efficiency.*

Keywords: *Internet security, anonymity techniques, secure browsing, user experience.*

1. INTRODUCTION

Communication over the Internet has dispersed into the very core fabric of the human relationship. Because of this high diffusion of the Internet adoption, several threats such as identity theft are becoming rampant. This has led to the rise in the usage of anonymity techniques to secure user identity and message content over the Internet.

Anonymity in straightforward terms implies concealing one's identity or in another word remaining unknown while communicating [1–3]. For different reasons, a client may need to remain unknown while communicating over the Internet. Some of the reasons may be institutions worrying that employees might use their systems with a specific end goal to unlawfully get to copyrighted material. Browsing the web also incurs additional security risk as threats such as identity theft may involve spoofing and cloning a client's IP address and using it to launch an attack [4]. Government surveillance also infringes the right of individuals when access to the contents that the government feels may have political, economic or security impact is blocked. In all the aforementioned reasons, a user may choose to utilize an anonymity tool to conceal his identity and thus self-protect from the harms of the Internet and remain unknown to concerned agencies.

The need to protect user's identity propelled the designing of anonymous communication systems [5].

There are several forms of anonymity depending on the intent of the anonymity. Sender anonymity implies that the identity of the originator of a message who communicates over a network stays hidden while

making no consideration to the status of the receiver or the message itself. Receiver anonymity is when the identity of the receiver is unknown to both the sender of the message and any intermediary or third party that may be in between the sender and the receiver. When servers that exist in a communication path are not identifiable as nodes inside the channel, we have a form of node anonymity. This type of anonymity is necessitated where associating with a particular anonymous communication is dangerous. In carrier anonymity, nodes between a sender and a receiver are unidentifiable as carriers for conveying message between them.

Several anonymity tools exist that are based on both high latency anonymity technique and low latency anonymity technique. Different anonymous techniques that exist have different performance and are adopted for different reasons. High Latency Anonymity (HLA) techniques (e.g. Single Mix, and Mix Networks and Mix Cascades) provide stricter encryption but at the cost of network performance, thus, low latency anonymity (LLA) technique (e.g. Single Hop Proxy, DC-Nets and Onion Routing) is preferred. Three major onion routing tools that are widely used for anonymity are Tor, JonDo and I2P [6].

A common belief held by most users of anonymous systems is that anonymity technique though protects their identity on the Internet, limits efficiency and performance of services being sought via them. This believes can be traced to the ideology that there exists server that serves as an encryption tool or in most cases, which fabricate a false address for the host to

use, and as a result, the overhead processes will add to the delay incurred by transmission.

2. LITERATURE REVIEW

Unlike the conventional approach based on only cryptography to securing a communication channel, anonymous communication utilizes both cryptography and necessitates resistance to traffic analysis. Using conventional approach such as encryption will ensure confidentiality, digital certificates, furthermore, Message Authentication Codes (MACs) will ensure authenticity and hash functions will ensure the integrity of message to determine if there has been an unauthorized modification [7]. However, conventional security approach does not protect a session from being analyzed by both an authorized and unauthorized person.

Onion Routing Low Latency Anonymity techniques are more efficient in terms of network performance [8]. Different tools that employ the technique would be analyzed for comparison in this study. However, several onion routing anonymity tools exist and several concerns have been made towards the efficiency of those tools mostly in terms of the delay incurred and the overall throughput of the system. This creates a problem of choice that will best suit user's purpose of anonymity for communication over the Internet. For these reasons and more, this study investigates the performance of existing anonymity tools used in anonymous communication systems to deduce their strengths and weaknesses, and compare them to determine their level of efficiency. The end results of the study will help users and researchers understand performance difference of onion routing based anonymity tools, and will assist them to efficiently choose an anonymity tool that best suits their purpose.

Because of its far reaching use with around 300,000 clients every day [6], existing performance measurements in anonymous systems principally focus on Tor. Snader et al [5] focus on upgrading throughput in Tor and proposed algorithm to accomplish faster execution or higher anonymity. However, importance of latency is as essential as throughput in anonymization networks [9, 10].

Important issues in choosing a proper anonymity tool are considered by researchers occasionally. The ideal system should be dependable, cheap and easy to use. The design of a general model in order to depict the motivations in anonymous networks was expanded. Furthermore, the reason for relaying traffic between nodes using Tor with the aim of improving the overall network performance was described.

Dingledine et al [11] emphasized the usability and network effect in anonymization networks. They argued that usability is important and it will lead to an increase in the user base, and subsequently the achievable anonymity. The time required for sending and receiving message relates to both ease of use and

performance. Notwithstanding, when a client simply surfs the Internet, the client expects a proper performance. The users will most likely not use a service if these expectations are not met. Several studies have attempted to determine the maximum tolerable time for loading a website. Different numbers are estimated for the value of the tolerable time but no consensus was made for the exact value of maximum tolerable time. However, recent research by Kopsel and Wendolsky et al concluded that about four seconds is the maximum tolerable delay for most users [12].

In their work, Stephanie Winkler and Sherali Zeadally (2015) [13], identified, explained and analyzed publicly available online anonymity technologies in terms of their usability and how online anonymity technology can be improved to increase public adoption. They did not compare the existing tools.

Jiejun Kong, Xiaoyan Hong [14] addressed problems for route anonymity by deploying ANODR that ensures that adversaries cannot discover the real identities of local transmitters. The work did not also compare the existing tools. Rather, it was based on "broadcast with trapdoor information", a novel network security concept which includes features of two existing network and security mechanisms, namely "broadcast" and "trapdoor information".

2.1. Onion Routing Tools

TOR – an open source project that provides anonymity service for TCP applications. It is a popular edge network for communicating anonymously over the Internet [11]. TOR is 2 G Onion Router, and it is the most prevalent onion routing technique that is still accessible till date.

TOR (further as Tor) consists of four basic components namely **Client (otherwise referred to as Alice)** that runs a local program called onion proxy to anonymize the client information into Tor; **Server (otherwise referred to as Bob)** which is a TCP application (e.g. web services) that runs on the server and communicates anonymously with Client over the Tor network; **Onion switches** – special proxies that transfer the application information amongst Client and Server (here, Transport Layer Security, TLS) associations are utilized for the overlay link encryption between two onion routers. The application information is stuffed into equal-sized cells of 512 bytes); and lastly, **Directory servers** that hold onion router data such as public keys, legitimate data and directory caches from which onion routers download registry data.

Joining TOR network using onion routing makes it difficult to spy on the nodes in the network being used to communicate. TOR network is freely accessible, and the TOR client application can be downloaded freely as well. Figure 1 shows a node interfacing with the TOR network via the client application. To begin anonymity using TOR, the client associates itself with a catalogue server that assigns the client an IP address that will be used to access the TOR network. After association of the

client to the network occurs the onion routing phase where a server unscrambles one part of the packet sent and gets the IP address to the next onion router. Thereafter, unencrypted traffic is transferred from the client (sender) to the intended server (receiver).

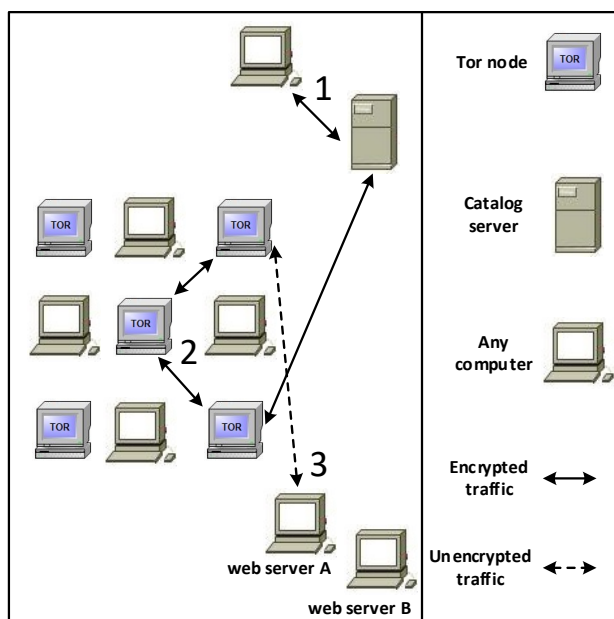


Fig. 1. Sample Tor Network and Tunnel Creation

The anonymization in TOR works with TCP (Transmission Control Protocol) and can be utilized by any application that supports SOCKS (Socket Secure Protocol – a piece of the Internet convention suite as a layer between the application layer and the transport layer). TOR further has novel functionalities including congestion control, forward secrecy, directory servers, configurable exit policies, integrity checking and practical design for location-hidden services through rendezvous point. [12].

Exceptional forward security implies that TOR utilizes an incremental path-building design. The initiator handles session keys used for each hop in the routing. At the point when keys are evacuated, the accompanying hops cannot be decrypted even if the nodes handling the traffic are attacked. The TOR network permits clients at network edges to detect congestion and respond by reducing the data output until the congestion problem is resolved [11]. In contrast to earlier onion routing approaches, TOR network takes an improved view towards sharing data. The network consists of various trusted nodes that act as directory servers and provide signed directories that defines known states and routers [11]. TOR has included the functionality of checking packet integrity with the goal that no node can change the information from one point to another in the network. Clients negotiate rendezvous point in the network to allow them connect to hidden servers and thus provide forward [11].

The term forward security implies that regardless of the possibility that one of the parts in the onion routing is compromised; information about other com-

munication in the network cannot be divulged using that part. TOR utilizes Diffie-Hellman key exchange between each circuit router and the onion proxy for the duration of the lifetime of the circuit so as protect against traffic capturing. Every node in the TOR network has a policy describing hosts and their ports to which it can connect. Some features that are considered unnecessary, such as padding, mixing and traffic shaping are not included in TOR network [9].

TOR enables clients to setup shrouded benefits inside the system. To get a concealed administration running, a client must have a TOR client introduced and should introduce a web server locally on a similar PC. Some arrangement is expected to safeguard that the server does not give away any significant data and remains mysterious. Once the administration is designed, the TOR client will arrange an open key and a private key for the server with the goal that the movement to it can be scrambled [5].

I2P (Imperceptible Internet Project) – I2P is an application that provides anonymity by concealing Metadata. Its objective is to facilitate privacy on communication inside an overlay network over the Internet by ensuring that no useful data can be collected from a packet that will reveal the anonymity of the communicating parties. I2P provides anonymity to all users by routing a message between different routers, thus making it extremely troublesome for an observer who analyses traffic to determine the genuine sender and receiver of packet. I2P encrypts both the content of the message and the identities of the communicating parties and as a result, information cannot be analyzed or leaked. Though ideology in I2P is similar to that of TOR Project [10], however, I2P differs in the terms of the project objectives. While TOR expects to enable anonymous access to the Internet to ensure the clients privacy and access to blocked or confined sites, I2P secures messages inside its system and provides all members access to the messages. Although I2P utilizes proxy services, connection to different networks, as in the Internet or TOR, was not planned [10].

I2P utilizes two diverse encryption plans to secure data (AES256/CBC (symmetric encryption) inside the I2P Tunnels and ElGamal (unbalanced) for every single other message planned for a particular recipient). Figure 2 shows messages encryption processes in I2P.

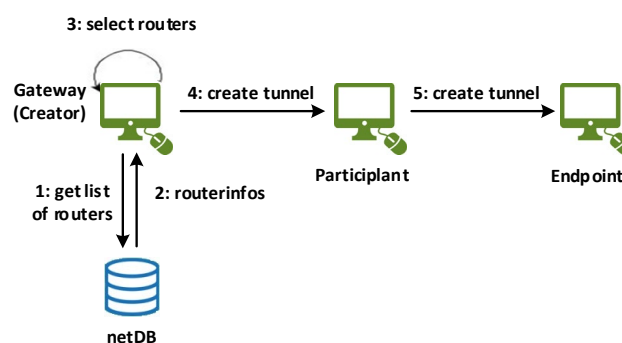


Fig. 2. Messages Encryption Processes in I2P

Each message is encoded with the ElGamal public key of the recipient to protect the message content (including administrative messages that are being exchanged among routers). I2P utilizes a layered encryption with the goal that every router only observes the data it needs.

Normally, the transport instructions revealing to the router where to forward the message. The remaining information in the packet such as further hops, destination or the content of the message is scrambled with AES256/CBC. The sender (outbound gateway) performs this layered encryption by first setting up a session key with every router of the outbound tunnel by utilizing ephemeral Diffie-Hellmann, and then using the key to scramble the message. Furthermore, messages that would later be exchanged in clear content are scrambled using Garlic Encryption with ElGamal to forestall information leakage.

JonDo – an application that spoofs a user's IP address to allow for anonymous Internet access using certain applications e.g. web browsers. It is an open source and free program that offers access to both free and premium services [15, 16]. It encodes information sent over JonDonym (an open source anonymization service also known as JAP (Java Anon Proxy)) network making the plaintext information invisible to neither the ISP (Internet Service Provider) nor JonDonym administrators (with the exception of the last mix in each cascade). JonDo, an Anti-Censorship-Application (ACA) that enables access to generally blocked Internet destinations. Clients may design their JonDo as a sending server with a specific end goal to permit others access to the JonDonym services.

JonDonym depends on the rule of numerous (layered) encryption, dispersion and preparing [17]. This method does not only shield a user's Internet activities from being seen by outsiders but also protects against observation by the individual JonDonym suppliers themselves. The anonymization service of JonDonym is an advancement branch of the ANON Project comprising multiple clients selectable mixed cascades (a cascade comprises of a few independently scrambled mix servers that are operated by non-interrelated associations or private people who all publish their identity). Here, administrators need to abide to strict arrangements, which disallow saving connection information or trading such information with different operators.

Each connection from a client is distinctively encoded for each mix server inside a cascade and exchanged through the cascade to the intended destination such as a website. Therefore, no single mix operator alone can reveal the identity of the communicating parties. Eavesdropper on JonDo connection attempting to sniff packet will only receive garbage data because the connection to each mix is encrypted separately. Likewise, since multiple clients surf the anonymization services at the same time with similar IP address,

all connections to every client must be concealed.

Figure 3 shows the difference in connection to the Internet between using and not using JonDo application.

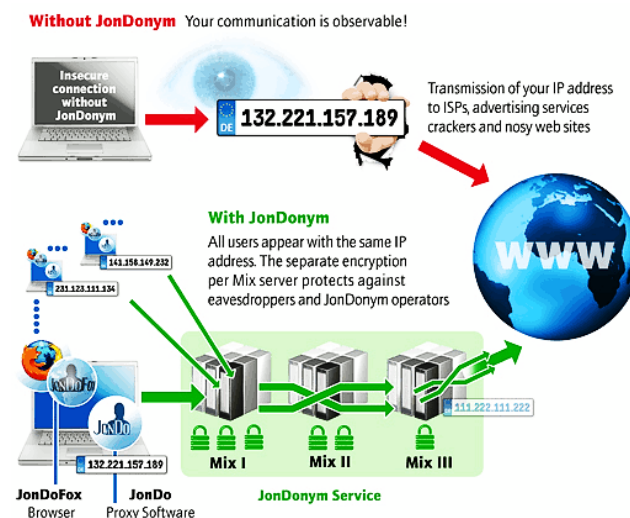


Fig. 3. Channel Creation and Encryption Processes in JonDo

3. METHODOLOGY

This was conducted using GLO (globalcom Nigeria ltd) network through the University of Ilorin as the ISP. The scale of the study was confined to a few tests to allow for timely data collection and analysis. Two servers that are used for the test are amalmabruuk and gnuwget.

Determining the impact of low and high traffic situations on the performance of each tool, systematic procedure of data collection and analysis was carried out in 2 (two) periods, which are on peak period (when the traffic on the network is high or very high) and off peak period (when there is low traffic on the network). The test was performed for 10 consecutive days to have more accurate results to determine the performance efficiency of each tool.

For the test, the followings were required:

- a client system running windows 10 OS 64-bit with hardware configuration of 8 Gb RAM, 2.50 GHz processor and 500 GB hard drive (hard drive size could be less);
- connection to the school network using (i.) above;
- installation of packages for the tools to be tested (they were installed on the client system before the client system was then connected to the network).

Each day, the tests were carried out in 4 phases as follows:

- testing without any anonymization tool running on the client system (this was done to set a benchmark in measuring the anonymity tools);
- testing with JonDo running on the client system;
- testing with I2P running on the client system;
- testing with Tor running on the client system.

3.1. Test 1: RTT, Jitter and Throughput without Anonymization

Public address of the client was first checked using www.cmyip.com. The resulted IP number, 197.211.32.242, along with other information such as the city and the ISP providing the connectivity was displayed as seen in Figure 4.

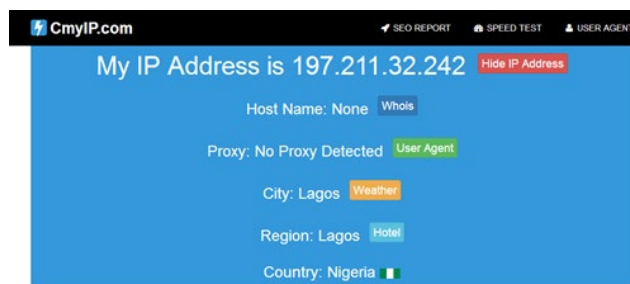


Fig. 4. Public IP Address without Anonymization

The first test was carried out without anonymization on the client system. This requires no system configuration and was done to serve as a baseline for determining the efficiency of the anonymity tools.

PingPlotter (a trace route utility tool that is based on ICMP messages sent between two nodes to determine the availability of a node) was used to ping the

server, amalmabruuk server, while data about Round-Trip Time (RTT) and jitter were recorded as obtained from PingPlotter statistics (Figure 5).

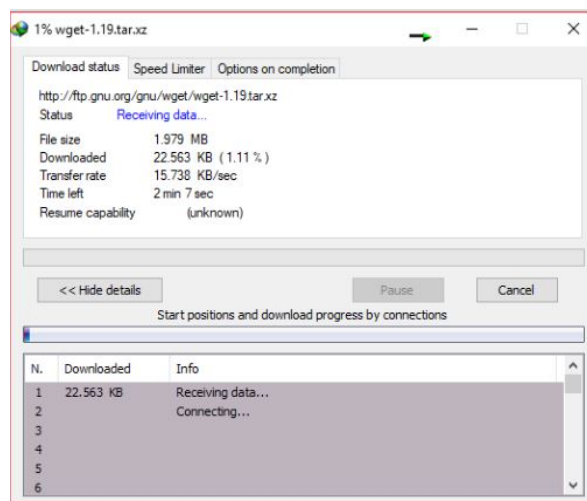


Fig. 5a. Test 1 Result T: Collecting data for throughput

Internet Download Manager (IDM) was also used to collect data for throughput by downloading 1 MB and 2 MB files from the GNU Wget server without anonymization (Figure 6).

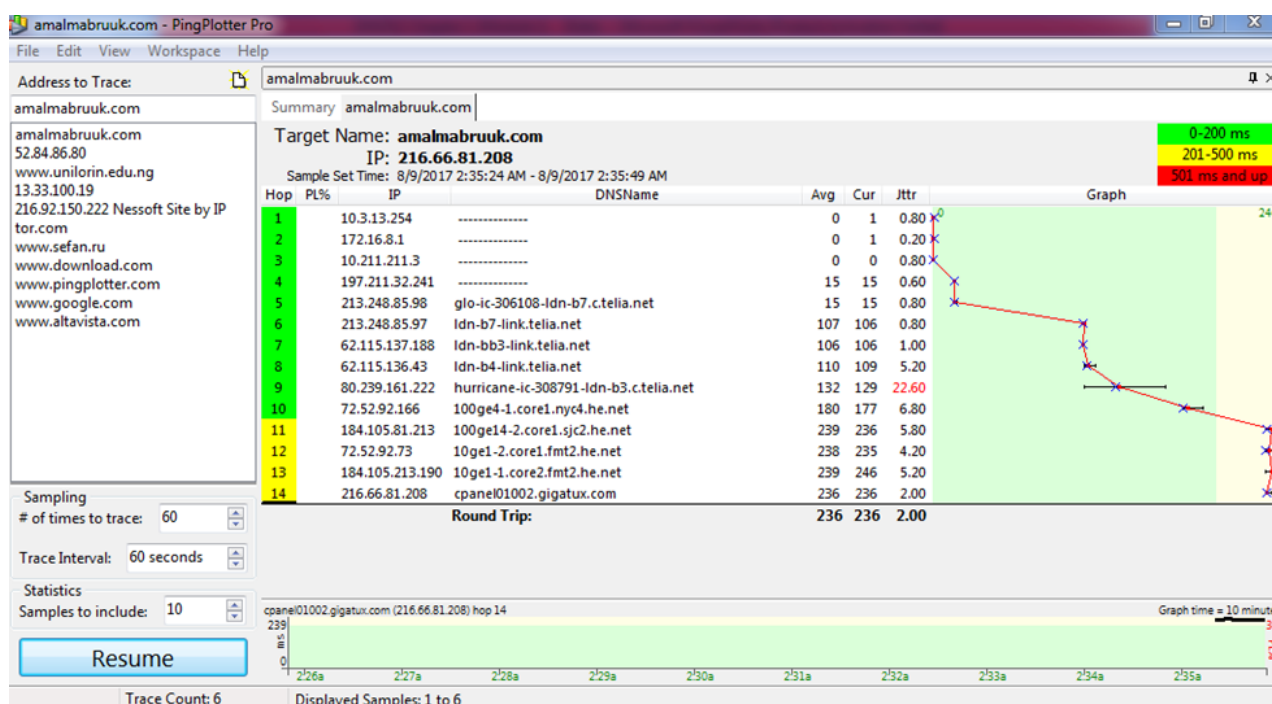


Fig. 5b. Test 1 Result J: Collecting Data for RTT and Jitter without Anonymization

3.2. Test 2: Using JonDo (RTT, Jitter and Throughput)

After installing JonDo on the client host, there is need to perform some configuration before it can be used on a client to provide anonymization. The Internet explorer configuration settings was modified to use JonDo's port 4001. To configure this, the native

browser of the client system (Internet explorer) is launched and the <Tools> button in the menu is pressed. On the <Tools> interface, <Internet options> button is pressed then <connections> then <LAN settings>. In the proxy field after ticking the <use proxy> button, the following proxy address 127.0.0.1 (or localhost) and port number 4001 are entered.

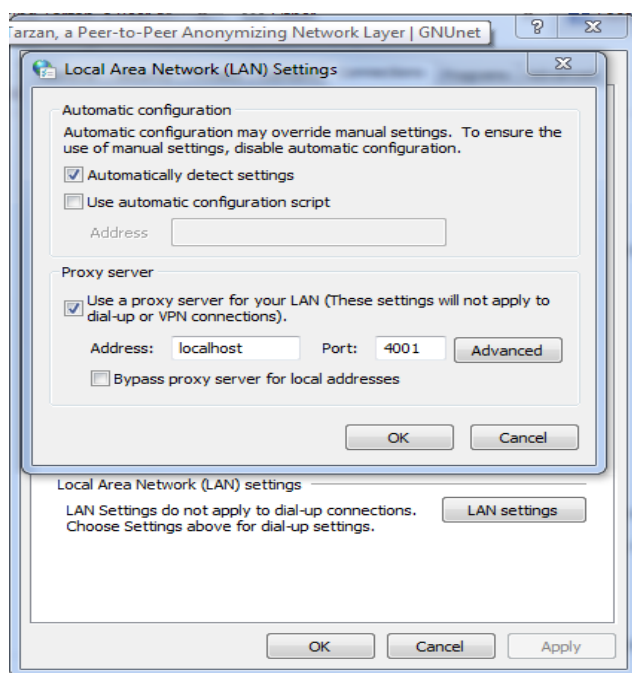


Fig. 6. Setting up JonDo Proxy on Internet Explorer

After installing JonDo on the client host, there is need to perform some configuration before it can be used on a client to provide anonymization. The Internet explorer configuration settings was modified to use JonDo's port 4001. To configure this, the native browser of the client system (Internet explorer) is launched and the <Tools> button in the menu is pressed. On the <Tools> interface, <Internet options> button is pressed then <connections> then <LAN settings>. In the proxy field after ticking the <use proxy> button, the following proxy address 127.0.0.1 (or localhost) and port number 4001 is entered.

After this configuration had been done, the JonDo program was launched from the program list. On the JonDo main interface, the anonymity button was changed from OFF to ON and the connection was tunneled through the JonDo servers. This enabled anonymity for the client station (as seen in Figure 7).

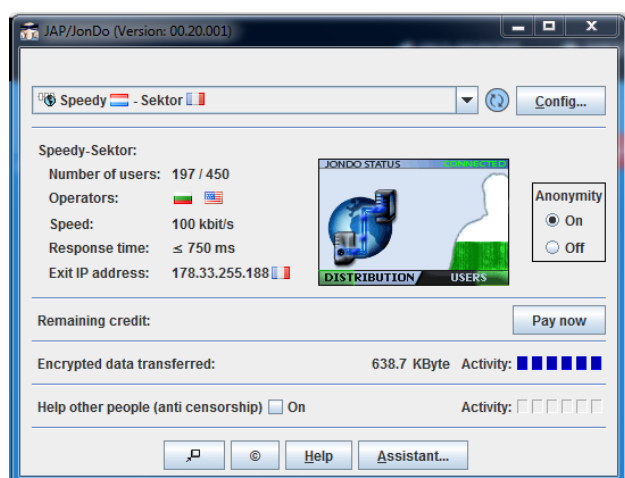


Fig. 7. Enabling Anonymity Using JonDo

After enabling anonymity, the public address of the network is tested again using www.cmyip.com and the result shows a different result from that obtained when no anonymity tool was used (as seen in Figure 8).

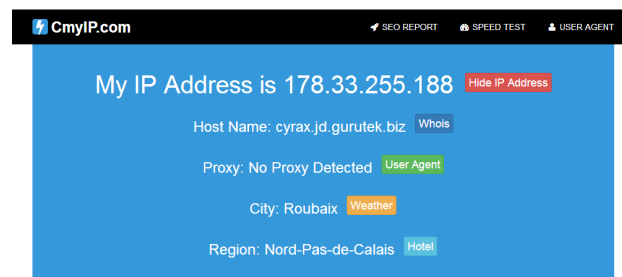


Fig. 8. IP Address as Shown by Cmyip.com after the Anonymization Using JonDo

3.3. Test 3: with I2P (RTT, Jitter and Throughput)

Like JonDo, I2P requires some basic configuration in the Internet explorer. The same procedure as used in JonDo was used to configure I2P. The only difference in the configuration was the proxy port number used in I2P. On the Internet options under connection settings, the proxy address 127.0.0.1 (localhost) and port number is 4444 (Figure 9) was used.

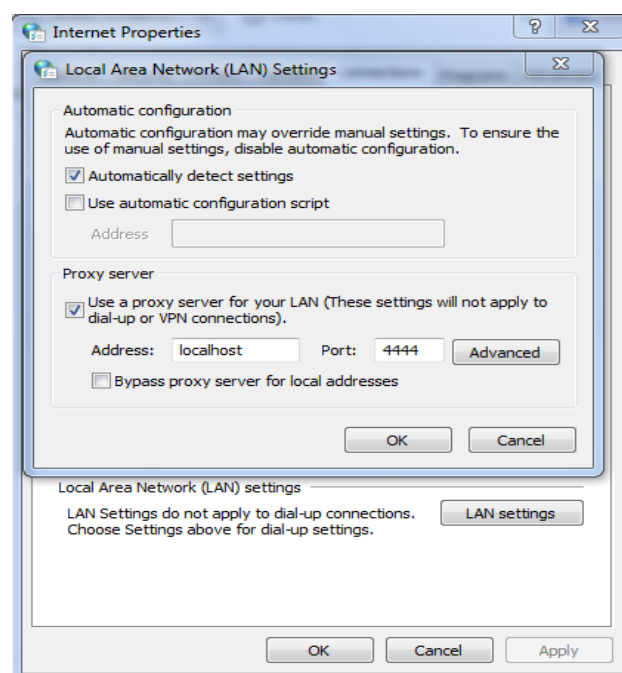


Fig. 9. Configuring I2P Proxy Settings on Internet Connection

After the basic system configuration had been done, the I2P application was launched and it created a tunnel for communication automatically. Using this tunnel, I2P provided anonymity for clients in the overlay network over the Internet.

I2P application shows different information about anonymization such as peers, tunnels and gives more detailed information about client tunnels for shared clients (DSA included) (see Figure 10).

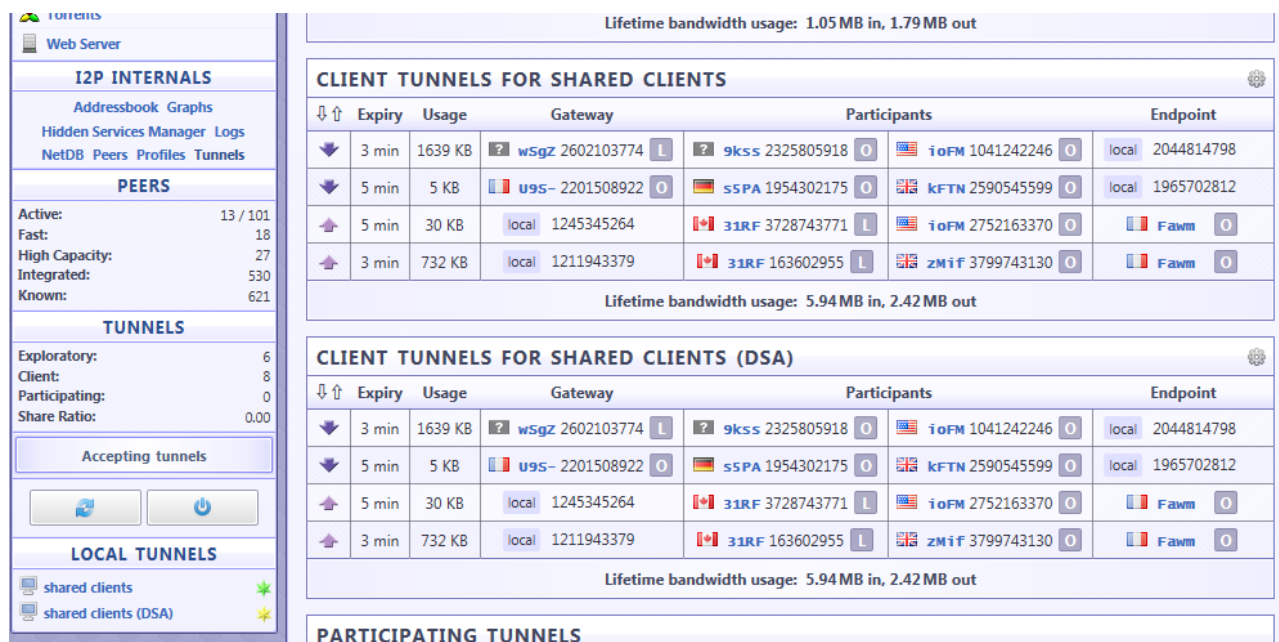


Fig. 10. I2P Interface Displaying the List of Tunnels Used for Anonymization

To confirm if anonymity has occurred, the public address of the client was checked using www.cmyip.com (Figure 11). The check showed that the IP address was different from the case without anonymization.

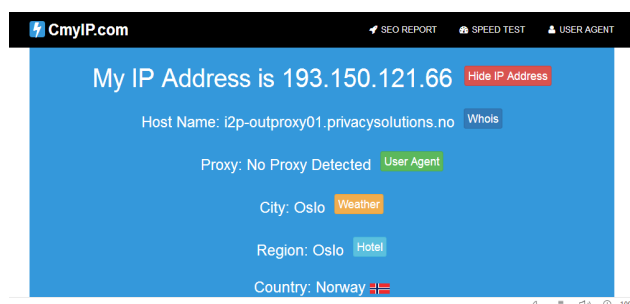


Fig. 11. IP Address as Generated by I2P to Protect Identity of the Source

3.4. Test 4: Tor (RTT, Jitter and Throughput)

Tor is a browser-based anonymity tool and as a result, the entire necessary configuration needed to provide anonymous connection be embedded with the browser. There is no need for additional configuration when using Tor because the browser will automatically connect to the Tor network.

After the browser has been connected to the Tor network, surfing the Internet using Tor browser would create anonymity for the user. Testing if Tor network has successfully protected the identity of the user, the public address was checked using www.cmyip.com (Figure 12).

PingPlotter tool was used to gather data for RTT and jitter for all the cases of the tests (Test 1 to Test 4).

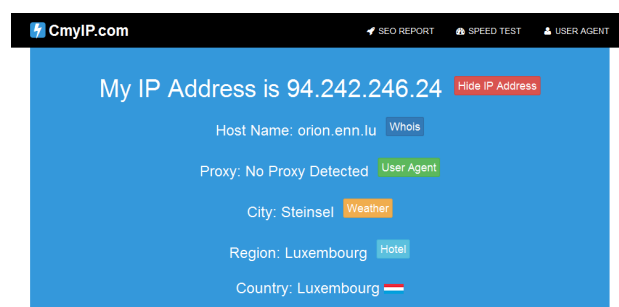


Fig. 12. IP Address when Using Tor Network to Protect Identity of Source

RTT, Jitter and Throughput were analyzed for each of the aforementioned onion routing-based anonymity tools. The tests for RTT and Jitter were analyzed using a trace route utility tool called PingPlotter. PingPlotter was used to capture data of RTT and Jitter. It recorded jitter and packet loss between communications when the fixed-size packet was sent. Each RTT and jitter test was set to 6 traces and the time duration between each trace was set to 20 seconds. This makes each RTT and jitter test last for 120 seconds (2 minutes). Ten (10) RTT and jitter test was conducted and split into five (5) on-peak period and five (5) off-peak periods for each of the tools and without any i.e. without anonymization. This makes the test conducted in a single day to last duration of 80 minutes, 20 minutes for situation without anonymity and 20 minutes each for each of the tools being tested. For the sake of accuracy, the tests and their procedure were repeated for ten (10) days.

While PingPlotter was used for measuring RTT and Jitter, for throughput, IDM is used. IDM is a tool that is used to download files from server, and it showed transfer rate (throughput) of each file being down-

loaded. The throughput test was carried out for 10 consecutive days the same way the RTT and jitter tests were carried out. Two (2) different files of 1 MB and 2 MB were downloaded from a file server <<http://www.gnu.org/software/wget/>>, and the resulting throughput of the download was recorded. The throughput test was first carried out without anonymity on the client system and then for each of the anonymity tools being tested.

4. RESULTS AND ANALYSIS

4.1. Results

4.1.1. Results of RTT, Jitter and Throughput during Peak Period

The table below shows the result obtained from the test when no anonymity tool was used during peak period (when students and staffs were on campus using the Internet as well).

Table 1 shows the result obtained from the test conducted when no anonymity tool was used. As displayed in the table, ten (10) tests were carried out for RTT, Jitter, Throughput (1 MB) and Throughput (2 MB). An average value of 393,1 ms was recorded for Round-Trip Time, Jitter recorded an average value of 61,275, and Throughput (1 MB) recorded an average of 112,2 and Throughput (2 MB) recorded 128,9. Round-Trip Time has a standard deviation of 287,65, Jitter has a standard deviation of 578,73, and Throughput (1 MB) has a standard deviation 23,95 while Throughput (2 MB) has a standard deviation of 40,1.

TABLE 1. Test Results for RTT, Jitter, and Throughput without Anonymity Tool during Peak Period

Tests	RTT (ms)	Jitter (ms)	Throughput for 1 MB file (kB/s)	Throughput for 2 MB file (kB/s)
1	394	77,00	122	137
2	404	99,25	116	127
3	391	41,75	110	124
4	416	95,33	109	116
5	365	43,67	105	127
6	379	49,20	113	126
7	409	73,25	116	132
8	370	37,50	113	137
9	405	32,80	108	131
10	398	63,00	110	132
Average	393,1	61,275	112,2	128,9
STD. Dev.	287,65	578,73	23,95	40,1

Table 2 shows the result obtained using JonDo anonymity tool during peak period.

As seen from the table, the results gotten from the test using JonDo anonymity tool differ from those that were gotten when no anonymity tool was used (comparing tables 1 and 2). 10 test were also carried out in this instance for RTT, Jitter, Throughput (1 MB) and Throughput (2 MB). An average value of 267,7 ms was

recorded for RTT, Jitter recorded an average value of 6,735, and Throughput (1 MB) recorded an average of 40 while Throughput (2 MB) recorded 32,3. RTT had a standard deviation of 36,01, Jitter had a standard deviation of 214,55, while Throughput (1 MB) had a standard deviation of 71,11. Throughput (2 MB) had a standard deviation of 26,01.

TABLE 2. Test Results for RTT, Jitter, and Throughput during Peak Period Using JonDo

Tests	RTT (ms)	Jitter (ms)	Throughput for 1 MB file (kB/s)	Throughput for 2 MB file (kB/s)
1	272	7,40	46	31
2	269	2,25	48	28
3	267	0,00	23	27
4	264	2,60	39	34
5	264	1,75	43	27
6	283	48,00	42	40
7	264	0,60	45	41
8	265	3,00	28	35
9	264	0,75	38	29
10	265	1,00	48	31
Average	267,7	6,735	40	32,3
STD. Dev.	36,01	214,55	71,11	26,01

Table 3 shows the result obtained from the test during peak period using I2P.

TABLE 3. Test Results for RTT, Jitter and Throughput during Peak Period Using I2P

Tests	RTT (ms)	Jitter (ms)	Throughput for 1 MB file (kB/s)	Throughput for 2 MB file (kB/s)
1	229	1,20	10	7
2	241	17,20	13	11
3	232	2,60	11	9
4	231	1,20	16	12
5	233	1,80	15	13
6	231	0,40	17	8
7	235	8,40	16	9
8	232	1,40	20	7
9	233	4,40	23	12
10	232	1,40	23	11
Average	232,9	4	16,4	9,9
STD. Dev.	10,54	26,96	20,48	4,76

The result gotten from the test conducted during peak period using I2P gave an average value of 232,9 ms for RTT while Jitter gave an average value of 4. Throughput (1 MB) gave an average of 16,4 and Throughput (2 MB) gave 9,9. RTT had a standard deviation of 10,54, Jitter 26,96, Throughput (1 MB) had 20,48 while Throughput (2 MB) had 4,76.

Table 4 shows results for Tor anonymity tool during peak period. The result showed a difference when compared to test for both I2P and JonDo. As obvious from the tables, 10 test were also carried out in this

instance for RTT, Jitter, Throughput (1 MB) and Throughput (2 MB). An average value of 312,3 ms was recorded for RTT, Jitter 2,893, Throughput (1 MB) 30,9 ms and Throughput (2 MB) 18,8. RTT had a standard deviation of 761,78, Jitter had 8,476, Throughput (1 MB) had 37,21 while Throughput (2 MB) had 15,95 ms.

TABLE 4. Test Results for RTT, Jitter and Throughput during Peak Period Using Tor

Test	RTT (ms)	Jitter (ms)	Throughput for 1 MB file (kB/s)	Throughput for 2 MB file (kB/s)
1	310	0,45	26	17
2	360	6,17	37	20
3	290	0,90	23	27
4	273	2,30	36	16
5	327	3,69	39	19
6	289	9,20	27	21
7	327	2,00	29	21
8	309	3,63	37	16
9	346	0,27	32	12
10	292	0,32	23	19
Average	312,3	2,893	30,9	18,8
STD. Dev.	761,78	8,476	37,21	15,95

4.1.2. Results of RTT, Jitter and Throughput Off-Peak Period

The test was also carried out during off-peak period when the university campus has less population or users of the network. Table 5 shows the result obtained for cases when anonymity tool was not used.

TABLE 5. Test Results for RTT, Jitter and Throughput during OFF-PEAK Period without Anonymity Tool

Tests	RTT (ms)	Jitter (ms)	Throughput for 1 MB file (kB/s)	Throughput for 2 MB file (kB/s)
1	250	0,40	116	125
2	238	0,20	127	120
3	247	1,20	126	125
4	248	0,80	123	125
5	247	0,40	117	134
6	257	0,67	126	129
7	257	1,00	113	123
8	237	0,60	116	127
9	258	1,20	119	133
10	267	0,40	134	125
Average	250,6	0,687	121,7	126,6
STD. Dev.	86,933	0,1254	43,122	18,711

As seen in the table, an average value of 250,6 ms was recorded for Round-Trip Time, Jitter recorded an average value 0,687 ms, and Throughput (1 MB) recorded an average of 121,7 ms and Throughput (2 MB) recorded 126,6. Round-Trip Time has a standard deviation of 89,933, Jitter has a standard deviation of 0,1254, and Throughput (1 MB) has a standard deviation

43,122 while Throughput (2 MB) has a standard deviation of 18,711 ms.

Using JonDo off peak period, the results in Table 6 were obtained.

TABLE 6. Test Results for RTT, Jitter and Throughput when Using JonDo OFF-PEAK Period

Test	RTT (ms)	Jitter (ms)	Throughput for 1 MB file (kB/s)	Throughput for 2 MB file (kB/s)
1	226	3,40	63	42
2	228	0,80	57	47
3	231	0,00	56	43
4	214	26,40	63	39
5	193	0,70	62	42
6	175	0,60	69	38
7	201	22,3	52	33
8	221	9,20	69	41
9	236	0,07	67	37
10	230	0,00	62	43
Average	215,5	6,347	62	40,5
STD. Dev.	391,83	98,78	31,77	15,6

As seen in the table, the result obtained using JonDo anonymity tool was quite different from the one when anonymity tool was not used (comparing Table 1 and 5). As displayed in the table, 10 tests were also carried out in this instance for RTT, Jitter, Throughput (1 MB) and Throughput (2 MB). An average value of 215,5 ms was recorded for RTT, Jitter recorded an average value 6,347 ms, and Throughput (1 MB) recorded an average of 62 ms while Throughput (2 MB) recorded 40,5. RTT had a standard deviation of 391,83, Jitter had 98,78, and Throughput (1 MB) had 31,77 while Throughput (2 MB) had 15,6 ms.

The same test (off-peak period) was also performed using I2P as the anonymity tool. Table 7 shows the result obtained from the test.

TABLE 7. Test Results for RTT, Jitter and Throughput during OFF-PEAK Period with I2P

Test	RTT (ms)	Jitter (ms)	Throughput for 1 MB file (kB/s)	Throughput for 2 MB file (kB/s)
1	193	2,20	21	15
2	195	0,90	16	21
3	172	0,32	23	17
4	206	0,43	27	16
5	215	0,69	17	12
6	193	2,63	16	17
7	209	11,09	26	15
8	183	7,92	12	19
9	219	9,03	23	23
10	230	0,90	21	12
Average	201,5	3,61	20,2	16,7
STD. Dev.	308,5	16,78	23,28	12,68

As displayed in the table, 10 test was also carried out in this instance for RTT, Jitter, Throughput (1 MB) and Throughput (2 MB). An average value of 201,5 ms was recorded for RTT, Jitter recorded an average value

of 3,61 ms, and Throughput (1 MB) recorded an average of 20,2 ms while Throughput (2 MB) recorded 16,7. RTT had a standard deviation of 308,5, Jitter had 16,78, and Throughput (1 MB) had 23,28 while Throughput (2 MB) had 12,68 ms.

Tor browser was also used as an anonymity tool for protecting user identity, and the test was carried out using Tor browser. The results in Table 8 were obtained when Tor was used as an anonymity tool.

As obvious from Table 8, there is a difference between the results gotten when using other anonymity tools as well as when anonymity tool was not used. Ten (10) tests were also carried out in this instance for RTT, Jitter, Throughput (1 MB) and Throughput (2 MB). An average value of 229,2 ms was recorded for RTT, Jitter recorded an average value of 2,05 ms, and Throughput (1 MB) recorded 41,8 ms while Throughput (2 MB) recorded 24,5. RTT had a standard deviation of 151,73, Jitter had 4,11, and Throughput (1 MB) had 63,733 while Throughput (2 MB) had 12,277 ms.

TABLE 8. Test Results for RTT, Jitter and Throughput during Off-Peak Period with Tor

Test	RTT (ms)	Jitter (ms)	Throughput for 1 MB file (kB/s)	Throughput for 2 MB file (kB/s)
1	250	0,47	46	21
2	243	0,68	57	23
3	229	2,63	33	32
4	232	7,10	36	25
5	243	0,80	38	26
6	222	3,20	47	26
7	216	2,43	49	23
8	220	1,23	37	20
9	217	1,71	32	22
10	220	0,32	43	27
Average	229,2	2,05	41,8	24,5
STD. Dev.	151,73	4,11	63,733	12,277

4.2. Analysis

4.2.1. RTT

Comparing the RTT for the four cases of the test during peak period, it is obvious from Figure 13 that when anonymization tool was not used, there was a high value of RTT (delay). This is due to the presence of proxy server in anonymous systems that downloads data from different sites at constant interval so that it can serve any user the data upon request. I2P proved to be the best anonymity tool when considering delay associated with RTT.

For all cases of the 10 tests, I2P recorded below 250 ms in RTT making it suitable for use especially in real-time communication. It should be noted that I2P is primarily used for overlay networks and as a result, its RTT efficiency can be linked to its design properties. In terms of RTT, with a value below 300 ms, JonDo also performed better than Tor that recorded the worst RTT (308,5 ms) among the tools.

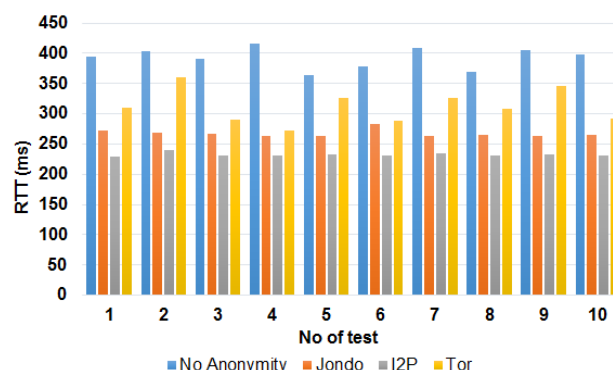


Fig. 13. RTT Comparison for Peak Period when Anonymization Tool Was Used and when It Was Not

However, during off peak period (Figure 14), there is a close similarity between situations when anonymization tool was not used and when it was used. Without anonymity, RTT value was the worst, getting over 250 ms. In terms of RTT, I2P also performed best during off peak period recording value below 200 ms. JonDo and I2P recorded the worst among the anonymity tools. I2P recovered the worst RTT among the tools for most of the tests but at test 9 and 10, JonDo recorded the worst RTT. This fluctuation may be due to the network condition when performing JonDo test for the 9th and 10th trial.

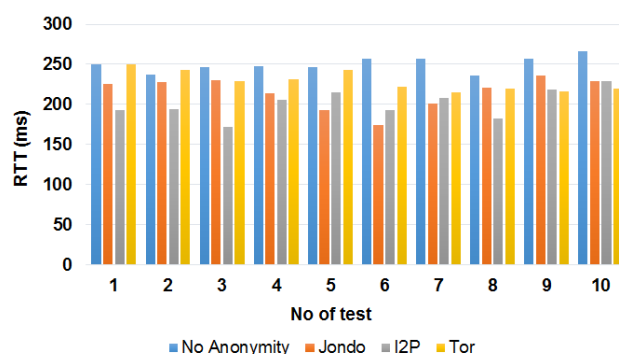


Fig. 14. RTT Comparison for OFF-PEAK Period when Anonymization Tool Was Used and when It Was Not

4.2.2. Jitter

For jitter, the four (4) different cases were also tested (Figure 15). The result shows that jitter during peak period was high when anonymization tool was not used. This is a surprising discovery since jitter was expected to be almost the same for all. The reason for the huge difference between the cases of when anonymity tool was used and when anonymity was not used may be the existence of the proxy server, which acts as a form of buffer between the client and the server. Considering the anonymity tools, there was variation across the level of jitter and the jitter was well below 20 ms except for the JonDo's 6th test, which recorded over 40 ms. This surge can be characterized to be as an effect of the network time between the receptions of packets thus causing delay variation (jitter). In general, Tor was the best tool in terms of jitter

because it recorded almost constant value below 10 ms, and with an average of 2,893.

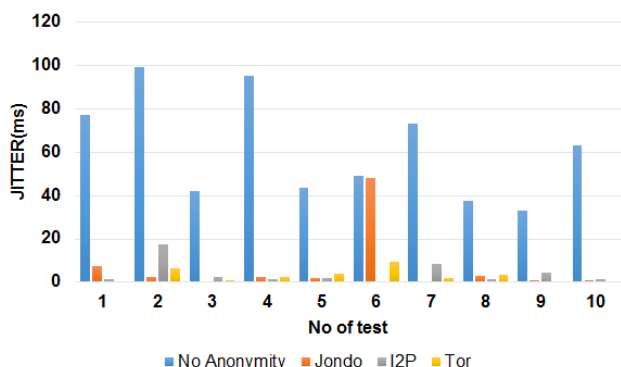


Fig. 15. Jitter Comparison for Peak Period when Anonymization Tool Was Used and when It Was Not

Performing the same test during off-peak period, when anonymization tool was not used, overall jitter performance considering that the jitter values were extremely low 2 ms was better. This shows that in the event of a good network connection, jitter is extremely low for traffic sent across the network. Tor shows the best jitter performance among other tools for all cases except for trials 3, 4 and 5 (Figure 16). JonDo is the worst in terms jitter recording above 25 ms despite the good condition of the network. However, in extreme cases such as during the 3rd and 10th trial, JonDo recorded an efficient jitter value. I2P was average in terms of jitter efficiency but in the 7th trial (7th test), it recorded above 10 ms.

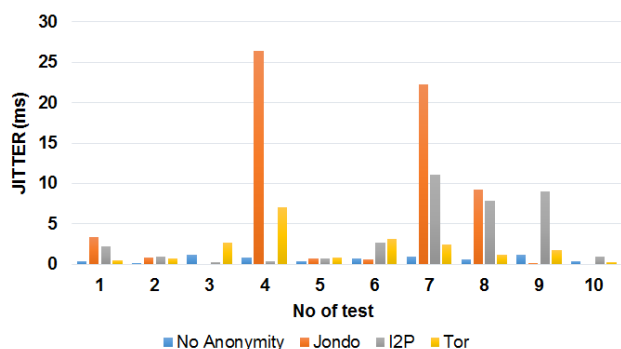


Fig. 16. Jitter Comparison for OFF-PEAK Period when Anonymization Tool Was Used and when It Was Not

4.2.3. Throughput of 1 MB File

To evaluate the efficiency of the tools throughput, two different files were downloaded from the GNU Wget server. A 1 MB file was downloaded from the server during peak period and for all the cases when anonymization tool was used. When anonymization tool was not used, the throughput was the most efficient, close to 140 KBps. Among the anonymity tools, JonDo recorded the best throughput with a value of about 40 KBps (Figure 17). However, I2P was worst in terms of throughput while Tor was average among. I2P recorded below 20 KBps while Tor recorded above 20 KBps for most of the test cases. It ought to be men-

tioned that the throughput when anonymization tool was not used is higher than when anonymity tool was used, this could be because of the time taken for the proxy server to get the intended download file (1 MB in this case) and forward it to the client that requested for it.

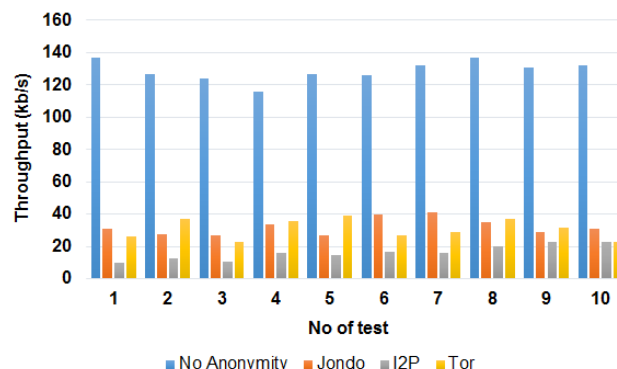


Fig. 17. Throughput (1 MB) Comparison for Peak Period when Anonymization Tool Was Used and when It Was Not

The same test was done during off peak period and the result was the same as cases during peak period. When anonymization tool was not used, throughput was recorded to be close to 140 KBps. This is because of the limitation placed by the university on the transfer rate of files. JonDo improved during off peak period by recording values above 60 KBps compared to during peak period when it recorded about 40 KBps (Figure 18). JonDo still proved to be the best anonymity tool in terms of throughput. Tor also performed better, still average among the anonymity tools. Tor recorded close to 60 KBps during the 2nd test and this is an improvement over the value of about 20 KBps recorded during peak period. I2P is the worst of all the tools in terms of throughput despite improving during off peak period to about 10 KBps.

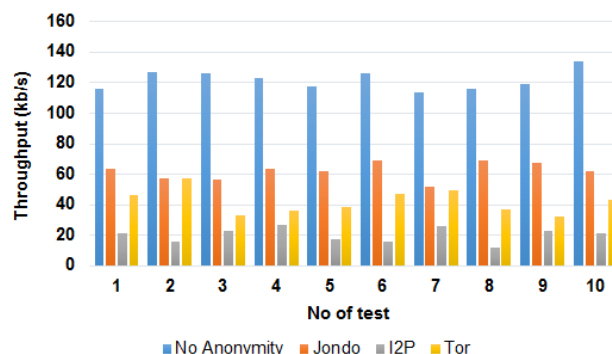


Fig. 18. Throughput (1 MB) Comparison for OFF-PEAK Period when Anonymization Tool Was Used and when It Was Not

4.2.4. Throughput of 2 MB File

To determine if there will be any difference in the performance of the tools when larger files than 1 MB is download, this study considered throughput when 2 MB file was being downloaded. The result however was similar to when 1 MB data was downloaded. When anonymization tool was not used, throughput

was above 120 Kbps (Figure 19), which was the best among all the four cases during peak period. JonDo also performs somehow better during peak period by producing a value above 40 Kbps while Tor produced value close to 40 Kbps. I2P had the worst throughput among all the tools with value close to 20 Kbps.

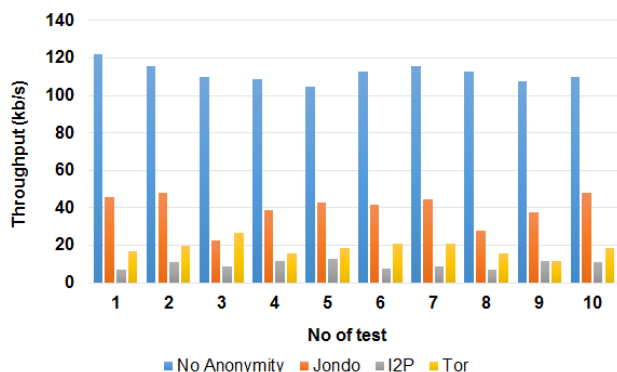


Fig. 19. Throughput (2 MB) Comparison for Peak Period when Anonymization Tool Was Used and when It Was Not

The same test was carried out during off peak period just as it has been with the other cases. Without anonymization, the throughput value improved by some fraction, JonDo by few values. This was because of reduced usage of the Internet connection by people in the campus at the time of the test. I2P recorded the worst among the anonymity tools although also increasing by a few fraction. So also the case with Tor. It could be noted that the changes in the value during off peak period and peak period was not significant as when 1 MB file was used as the test file (Figure 20). Thus, the difference between off peak period and on peak period when 2 MB file is used is inconclusive.

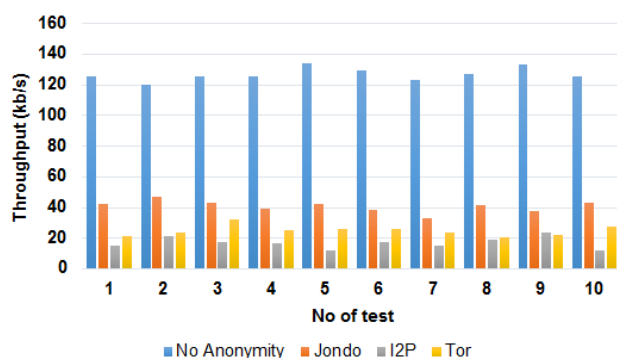


Fig. 20. Throughput (2 MB) Comparison for OFF-PEAK Period when Anonymization Tool Was Used and when It Was Not

References

1. Ling Z., Luo J., Yu W., Fu X. Equal-Sized Cells Mean Equal-Sized Packets in Tor? // IEEE International Conference on Communications (ICC). 2011. DOI:10.1109/icc.2011.5962653.
2. Sun Q., Simon D.R., Wang Y.M., Russell W., Padmanabhan V.N., Qiu L. Statistical Identification of Encrypted Web Browsing Traffic // IEEE Symposium on Security and Privacy (SP). 2002. DOI:10.1109/SECPRI.2002.1004359.
3. Overlier L., Syverson P. Locating hidden servers // IEEE Symposium on Security and Privacy (S&P'06). 2006. DOI:10.1109/SP.2006.24.
4. Wang X., Chen S., Jajodia S. Network Flow Watermarking Attack on Low-Latency Anonymous Communication Systems // IEEE Symposium on Security and Privacy (SP '07). 2007. PP. 116–130. DOI:10.1109/SP.2007.30.
5. Snader R., Borisov N.A Tune-up for Tor: Improving Security and Performance in the Tor Network // 15th Annual Network and Distributed System Security Conference (NDSS'08). 2008.

5. CONCLUSION AND RECOMMENDATION

5.1. Conclusion

Anonymous communication is becoming a widely accepted trend since it provides security for the identity of communication parties. Anonymity tools enhance the privacy of communication over the Internet. This anonymity can be achieved using anonymization tools. Having considered three different anonymity tools that are based on the onion routing approach and the performance of each of the tools, the conclusion is that generally when anonymity tools are used, the throughput of a system or network drops. However, among the onion routing-based anonymity tools and according to the results obtained from the study, JonDo has the highest throughput in terms of the throughput both at peak period and off peak periods. Despite the good performance of JonDo for throughput, the anonymity tool still recorded the highest jitter among anonymity tools tested.

Without the use of anonymization, delay and jitter was the highest for all cases of the test. Tor shows the highest RTT delay among the anonymity tools while JonDo shows the highest jitter. It can thus be concluded that Tor is not the best choice of anonymity tool for real time communication considering its RTT though it provides strong anonymization and is easier to use since it requires no configuration. Furthermore, it has the highest delay.

In terms of throughput, I2P is the worst among though it recorded the best RTT delay among all. It even had better result than the case when the anonymization tool was not used. For this reason, I2P is most suitable anonymity tool for real time communication.

5.2. Recommendation

This study has given an insight to the performance of onion routing-based anonymity tools. And it has explicitly explained which of the tools is suitable for which purpose (real-time communication or otherwise). Thus, the future works could be done developing application(s) anonymous for different purposes taken into consideration the results obtained in this work.

6. Beach A., Gartrell M., Han R. Solutions to Security and Privacy Issues in Mobile Social Networking // International Conference on Computational Science and Engineering. 2009. Vol. 4. PP. 1036–1042. DOI:10.1109/CSE.2009.243.
7. Chen J., Du X., Fang B. An efficient anonymous communication protocol for wireless sensor networks // Wireless Communications and Mobile Computing. 2012. Vol. 12. Iss. 14. PP. 1302–1312. DOI:10.1002/wcm.1205.
8. Jansen R., Syverson P., Hopper N. Throttling Tor Bandwidth Parasites. University of Minnesota TR 11-019, 2011. DOI:10.21236/ada559183.
9. Yu S., Zhou W., Jia W., Hu J. Attacking Anonymous Web Browsing at Local Area Networks Through Browsing Dynamics // The Computer Journal. 2012. Vol. 55. Iss. 4. PP. 410–421. DOI:10.1093/comjnl/bxr065.
10. Wilson M., Bazli B. Forensic analysis of I2P activities // 22nd International Conference on Automation and Computing (ICAC). 2016. PP. 529–534. DOI:10.1109/ICAC.2016.7604974.
11. Dingledine R., Mathewson N., Syverson P. Tor: The Second-Generation Onion Router // 13th USENIX Security Symposium. 2004. Vol. 14. DOI:10.21236/ada465464.
12. Wendolsky R., Herrmann D., Federrath H. Performance Comparison of Low-Latency Anonymisation Services from a User Perspective. Privacy Enhancing Technologies // Lecture Notes in Computer Science. 2007. Vol. 4776. PP. 233–253. DOI:10.1007/978-3-540-75551-7_15.
13. Winkler S., Zeadally S. An analysis of tools for online anonymity // International Journal of Pervasive Computing and Communications. 2015. Vol. 11. Iss. 4. PP. 436–453. DOI:10.1108/IJPC-08-2015-0030.
14. Kong J., Hong X. ANODR: anonymous on demand routing with untraceable routes for mobile ad-hoc networks. // 4th ACM international symposium on Mobile ad hoc networking & computing (MobiHoc '03). 2003. PP. 291–302. DOI:10.1145/778445.778449.
15. Fang Q., Li S., Zhou R. The Application of Onion Routing in Anonymous Communication // 2nd International Conference on MultiMedia and Information Technology (MMIT). 2010. PP. 244–246. DOI:10.1109/MMIT.2010.13.
16. Sitthiworachart J., Joy M. Web-based peer assessment system with an anonymous communication tool // IEEE International Conference on Advanced Learning Technologies. 2004. PP. 918–919. DOI:10.1109/ICALT.2004.1357720.
17. Prakash A., Agrawa D.P., Chen Y. Network Coding Combined with Onion Routing for Anonymous and Secure Communication in a Wireless Mesh Network // International journal of Computer Networks & Communications (IJCNC). 2014. Vol. 6. Iss. 6. PP. 1–14. DOI:10.5121/ijcnc.2014.6601.

* * *

СРАВНИТЕЛЬНЫЙ АНАЛИЗ АНОНИМНЫХ КОММУНИКАЦИОННЫХ СИСТЕМ С НИЗКОЙ ЛАТЕНТНОСТЬЮ

И. Муртала¹, О.А. Тиамийу²

¹De-Kharms Telecommunications, Илорин, Нигерия

²Университет Илорина, Илорин, Нигерия

Информация статье

УДК 004.771

Язык статьи – английский

Ссылка для цитирования: Муртала И., Тиамийу О.А. Сравнительный анализ анонимных коммуникационных систем с низкой латентностью // Труды учебных заведений связи. 2018. Т. 4. № 3. С. 85–97.

Аннотация: *С бурным развитием и тотальным распространением Интернета во всем мире, сокрытие личности его пользователя по соображениям защиты и безопасности является важным проблемным вопросом. Анонимность превратилась в существенную и легальную часть многочисленных приложений, включая безопасный веб-серфинг, услуги на основе определения местоположения и электронное голосование. В статье сравниваются существующие реализации различных анонимных коммуникационных систем с низкой латентностью, точнее, технологии маршрутизации с шифрованием на предмет определения их уровня эффективности по критериям производительности, надежности и удобства использования.*

Ключевые слова: *безопасность в Интернете, методы анонимности, безопасный веб-серфинг, опыт использования.*

МЕТОД КОЛИЧЕСТВЕННОЙ ОЦЕНКИ СЛУХОВОГО ВОСПРИЯТИЯ НЕЛИНЕЙНЫХ ИСКАЖЕНИЙ. ЧАСТЬ 2. ВЫБОР МЕРЫ ТОНАЛЬНОСТИ СИГНАЛА И УЧЕТ ЕЕ ВЛИЯНИЯ НА СУБЪЕКТИВНОЕ ВОСПРИЯТИЕ ИСКАЖЕНИЙ

М.О. Пахомов^{1*}, В.С. Рожнов^{1, 2}

¹Санкт-Петербургский филиал корпорации LG Electronics inc.,
Санкт-Петербург, 191123, Российская Федерация

²Санкт-Петербургский государственный университет телекоммуникаций им. проф. М.А. Бонч-Бруевича,
Санкт-Петербург, 193232, Российская Федерация

*Адрес для переписки: mpakhomov39@gmail.com

Информация о статье

УДК 621.391

Язык статьи – русский

Ссылка для цитирования: Пахомов М.О., Рожнов В.С. Метод количественной оценки слухового восприятия нелинейных искажений. Часть 2. Выбор меры тональности сигнала и учет ее влияния на субъективное восприятие искажений // Труды учебных заведений связи. 2018. Т. 4. № 3. С. 98-105.

Аннотация: Во второй части работы более полно раскрывается тема учета спектральных характеристик тестового сигнала при оценке субъективного восприятия нелинейных искажений. Выбрана наиболее значимая характеристика – тональность и предложена количественная мера этой характеристики – коэффициент тональности. Проведены субъективные тесты прослушивания для определения зависимости слухового восприятия нелинейных искажений от тональной структуры сигнала и спектрального распределения отношения «Шум-Маска».

Ключевые слова: нелинейные искажения, одновременная маскировка, отношение «Шум-Маска», коэффициент тональности, линейная модель множественной регрессии.

Введение

Приступая к описанию задачи, решаемой в представленной работе, прежде всего, следует отметить тот факт, что влияние тональной структуры референсных музыкальных сигналов на слышимость нелинейных искажений был отмечен в середине XX века и подробно описан в работах основоположников психоакустики – Э. Цвикера, Б. Мура и пр. Учет меры предсказуемости спектра сигнала при вычислении допустимого отношения «сигнал-шум» широко используется в современных алгоритмах кодирования музыкальных сигналов, например AAC (от англ. Advanced Audio Coding). Эти исследования помогают более точно вычислять пороги маскировки сигнала нелинейных искажений полезным сигналом, но не обеспечивают необходимой точности для предсказания субъективной оценки в случае, когда искажения заметно превышают порог маскировки. Другими словами, можно ли провести линейную (или нелинейную) зависимость между количеством слышимых шумов и субъективной оценкой. Тесты про-

слушивания показывают, что для набора музыкальных сигналов с разными спектральными характеристиками зависимости между отношением «Шум-Маска» (NMR, от англ. Noise-to-Mask Ratio) и оценкой предпочтения достаточно слабые [6, 7].

В первой части работы был описан способ точной оценки отношения «Шум-Маска» для много-тональных тестовых сигналов. Была выдвинута гипотеза, что влияние слышимых шумов на субъективное предпочтение зависит от частотного диапазона, в котором эти шумы превышают порог маскировки. Данная гипотеза была проверена на одном музыкальном сигнале. В выбранный музыкальный сигнал искусственно вносились нелинейные искажения разных уровней в различные частотные диапазоны и слушателям предлагалось оценить качество звучания с точки зрения предпочтения. Проверка результатов показала различную скорость деградации качества звучания в зависимости от величины «Шум-Маска» для разных частотных диапазонов. Было отмечено, что спектр сигнала различался по своей тональной структуре

в зависимости от частоты. Для объяснения полученных результатов было принято решение провести серию дополнительных тестов прослушивания с набором сигналов, имеющих различную тональную структуру в соответствующих диапазонах частот.

В представленной работе описывается и обосновывается выбор количественной меры тональности сигнала, референсных музыкальных сигналов и способы внесения нелинейных искажений для получения тестовых сигналов. Представлены выходные переменные модели и результаты тестов прослушивания.

1. Выбор меры тональности сигнала

Известно, что тональная структура сигнала играет существенную роль в вычислении порога одновременной маскировки. Рассмотрим известные способы определения тональных и шумоподобных компонент в музыкальном сигнале. Тональными называют те отсчеты быстрого преобразования Фурье, которые обусловлены наличием во входном сигнале продолжительных во времени гармонических составляющих (синусоид). Маскирующий эффект тональных и не тональных компонент различен, и поэтому с целью выявления тональных компонент выполняется специальный анализ спектра.

Спектральный анализ проводится в критических полосах слуха. В первой психоакустической модели (MPEG Layer 1) отыскиваются локальные максимумы энергетического спектра. Если мощность компоненты на определенную величину (7 дБ) превышает соседние локальные минимумы, данная компонента заносится в список тональных.

Во второй психоакустической модели (MPEG Layer 3) не классифицируются компоненты на тональные и шумоподобные. Вместо этого вычисляется индекс тональности, характеризующий близость данной компоненты к тональным или шумоподобным. Индекс тональности характеризует степень предсказуемости спектральной компоненты по значениям, вычисленным на предыдущих кадрах. Пусть $r_\omega = r_\omega(n)$, $\varphi_\omega = \varphi_\omega(n)$ обозначают соответственно амплитуду и фазу компоненты с номером ω в преобразовании Фурье для выборки n . Предсказанные значения амплитуды и фазы вычисляются по формулам (1а), (1б):

$$\hat{r}_\omega = 2r_\omega(t-1) - r_\omega(t-2), \quad (1a)$$

$$\hat{\varphi}_\omega = 2\varphi_\omega(t-1) - \varphi_\omega(t-2). \quad (1b)$$

Мера непредсказуемости определяется следующим образом:

$$c_\omega = \frac{\sqrt{(r_\omega \cos \varphi_\omega - \hat{r}_\omega \cos \hat{\varphi}_\omega)^2 + (r_\omega \sin \varphi_\omega - \hat{r}_\omega \sin \hat{\varphi}_\omega)^2}}{r_\omega + |\hat{r}_\omega|}. \quad (2)$$

Всего вычисляются меры непредсказуемости для 57 поддиапазонов. Для полосы с номером b через $\omega_l(b)$ обозначим нижнюю ее границу, а через $\omega_h(b)$ – верхнюю. Энергия поддиапазона равна:

$$e_b = \sum_{\omega_l(b)}^{\omega_h(b)} r_\omega^2, \quad (3)$$

а взвешенная мера непредсказуемости:

$$c_b = \sum_{\omega_l(b)}^{\omega_h(b)} r_\omega^2 c_\omega. \quad (4)$$

Эта пара величин используется как мера тональности данного поддиапазона.

К числу недостатков данного метода можно отнести высокую вычислительную сложность алгоритмов. Поскольку в будущем планируется предсказывать влияние вносимых искажений на субъективное восприятие музыкальных сигналов при потоковой цифровой обработке (выбор параметров обработки зависит от допустимого значения NMR), решено использовать более простой с точки зрения вычислительных затрат алгоритм, основанный на расчете равномерности спектра в полосе частот.

Количественной мерой тональности предлагается считать коэффициент тональности (КТ) [3] (также известный как коэффициент энтропии Винера), который вычисляется на основе вычисления меры пологости спектра в полосе частот (5):

$$\text{Tonality} = \min \left(\frac{\text{SFM}_{db}}{-60}, 1 \right), \quad (5)$$

где SFM (*от англ.* Spectral Flatness Measure) – мера спектральной пологости: $\text{SFM}_{db} = 10 \cdot \log_{10}(\text{SFM})$, вычисляемая по формуле (6):

$$\text{SFM}(\text{num_band}) = \frac{(\prod_{k \in \text{num_band}} a(k))^{1/K}}{\frac{1}{K} \sum_{k \in \text{num_band}} a(k)}, \quad (6)$$

где $a(k)$ – амплитуды k -ой компоненты спектра.

Коэффициент тональности вычисляется для 4 диапазонов частот от 250 Гц до 8 кГц: базовый (< 500 Гц), низкий средний (500–1000 Гц), высокий средний (1–4 кГц) и высокий (> 4 кГц).

Коэффициент тональности, вычисленный по формуле (5), для белого шума близок к 0, для однотонального сигнала – к единице.

В сверхвысоком диапазоне частот обычно наблюдается понижение энергии компонент с увеличением частоты. Для устранения недостатка данного метода – высокий коэффициент тональности при значительном наклоне огибающей спектра в широком диапазоне, принято решение разбить диапазон сверхвысоких частот на четыре поддиапазона и рассчитывать КТ в этом диапазоне как усредненное значение КТ для поддиапазонов.

2. Сигналы

В предыдущей работе [1] были введены и подробно описаны условные обозначения для акустических сигналов, используемых в исследовании. Здесь приведем только обозначения:

- референсный музыкальный сигнал (МСреф) – неискаженный музыкальный сигнал;
- референсный многотональный сигнал (МТСреф) – многотональный сигнал, синтезированный из референсного музыкального сигнала;
- тестовый музыкальный сигнал (МСтест) – референсный музыкальный сигнал с внесенными нелинейными искажениями;
- тестовый многотональный сигнал (МТСтест) – референсный многотональный сигнал с внесенными нелинейными искажениями;
- очищенный тестовый многотональный сигнал (МТСтест_очищ), имеющий в своем составе только спектральные компоненты тестового многотонального сигнала, содержащиеся в референсном многотональном сигнале;
- сигнал нелинейных искажений (СНИ) – сигнал, содержащий только новые спектральные компоненты тестового многотонального сигнала после его прохождения через звуковой тракт испытуемого устройства, т.е. изначально отсутствующие в референсном многотональном сигнале.

В качестве референсных музыкальных сигналов были отобраны 5 фрагментов музыкальных треков с разными уровнями коэффициента тональности (таблица 1).

ТАБЛИЦА 1. Коэффициенты тональности для МСреф

№ п/п	Фрагмент	Тональность в диапазоне			
		Бас < 400 Гц	Низкие Средние 0,4–1 кГц	Высокие средние 1–4 кГц	Высокие 4–8 кГц
1	Adele, Someone Like You Part 1 (piano)	0,28	0,23	0,2	0,13
2	Adele, Someone Like You Part 2 (voice)	0,23	0,27	0,27	0,8
3	Jamiroquai – Music of the mind	0,25	0,33	0,15	0,09
4	Muse – Follow me	0,28	0,11	0,08	0,06
5	Toto – Simon solo	0,20	0,07	0,06	0,09

Каждому МСреф был создан МТСреф. В данном исследовании мы использовали метод создания многотонального сигнала, основанный на прореживании тональных компонент референсного музыкального сигнала [1]. В основе этого алгоритма создания МТСреф лежит последовательный анализ спектров выборок МСреф, выделение компонент, удовлетворяющих некоторым критериям отбора и обнулением прочих. В случае отсутствия явно выраженных максимумов сигнала в некоторой полосе имеющийся шумовой сигнал заменяется тональной компонентой, расположенной в цен-

тре этой полосы и обладающей такой же энергией. Информация о типе компонент в полосе (тональная или шумоподобная) хранится в специальном файле-паспорте сигнала. Выбор этого типа многотонального сигнала обусловлен тем, что его тональная структура наиболее близка к структуре музыкального сигнала.

После внесения искажений в МСреф и МТСреф получены МСтест и МТСтест для тестов прослушивания и анализа искажений соответственно.

3. Искажения

В процессе исследований в референсные музыкальные и многотональные сигналы вносились нелинейные искажения, как в отдельных узких полосах частот, так и одновременно во всем диапазоне слышимых частот. Уровень сигнала нелинейных искажений был привязан к порогу его маскировки полезным сигналом в каждом диапазоне частот. Для моделирования нелинейных искажений использовались способы, описанные в [1]. Основным алгоритмом создания сигнала нелинейных искажений является жесткое клиппирование исходного музыкального сигнала.

Приведем поэтапное описание процедуры создания сигнала нелинейных искажений.

На первом этапе МСреф обрабатывается фильтром низкой частоты с частотой пропускания 200–400 Гц.

На втором этапе полученный низкочастотный сигнал подвергается процедуре жесткого клиппирования. Уровень ограничителя выбирается с учетом количества искаженных участков сигнала: 1 %, 2 %, 5 % и 10 %. Влияние временного распределения искаженных фрагментов на субъективное предпочтение решено исследовать в заключительной части работы. В данном исследовании используется такой уровень ограничителя, чтобы выше него оказалось 5 % отсчетов низкочастотного сигнала.

На третьем этапе полученный сигнал обрабатывается фильтром высокой частоты с частотой пропускания 200–400 Гц (частота пропускания зависит от верхней частоты исходного низкочастотного сигнала). В результате получается широкополосный сигнал нелинейных искажений.

На четвертом этапе полученный сигнал нелинейных искажений обрабатывается полосовыми фильтрами с полосами пропускания < 600 Гц; 600–1000 Гц, 1–2 кГц, 2–4 кГц, > 4 кГц. В результате получаются сигналы нелинейных искажений в пяти частотных диапазонах.

На пятом этапе происходит корректировка уровней сигналов шумов. Для этого полученные сигналы смешиваются с МТСреф и вычисляются величины NMR. Обычно уровни NMR выше 4 дБ

отмечаются слушателями как критические с точки зрения восприятия.

На заключительном шаге полученные сигналы нелинейных искажений смешиваются с соответствующими МСреф для получения набора МСтест. Полученный набор сигналов используется в тестах прослушивания.

4. Выходные переменные

Для оценки влияния нелинейных искажений на субъективное предпочтение необходимо учитывать эффект одновременной маскировки искажений полезным сигналом и разное восприятие слышимых нелинейных искажений слуховой системы человека в зависимости от характеристик сигнала.

Для учета эффекта маскировки искажений полезным сигналом использовалась упрощенная психоакустическая модель стандарта MPEG-1 ISO/IEC 11172-3.

Анализируемый тестовый многотональный сигнал разбивается на выборки, содержащие 2048 отсчетов сигнала, и для каждой выборки сигнала вычисляется энергетический линейчатый спектр $X(k)$. Для уменьшения краевых эффектов при выполнении БПФ используется окно Ханна $h(n)$. Для полученного энергетического спектра выполняется процедура нормирования, здесь максимальной спектральной компоненте присваивается уровень 96 дБ, а все значения остальных спектральных компонент подвергаются процедуре коррекции относительно этой величины. Для каждой тональной и шумоподобной компоненты (информация о типе компоненты берется из паспорта сигнала) выполняется расчет коэффициентов маскировки и индивидуальных кривых маскировки $M[z(i), z(j)]$, и затем для них вычисляются пороги маскировки $N_{TC}[z(i)]$ и $N_{MC}[z(i)]$. Для аппроксимации порогов маскировки вне критических полос слуха используется выражение [2]:

$$B(\Delta z) = A_0 + 15,81 + \frac{7,5(\Delta z + 0,474) - 17,5\sqrt{1 + (\Delta z + 0,474)^2}}{1}, \quad (7)$$

где $A_0 = 8\min((\Delta z - 0,5)^2 - 2(\Delta z - 0,5), 0)$.

Пример семейства индивидуальных кривых маскировки для различных уровней маскирующего тона приведены на рисунке 1. В завершении, рассчитанные пороги суммируются вместе с абсолютным порогом слышимости для получения глобального порога маскировки $N_{MT}(i)$ в каждой субполосе психоакустического анализа [4–8].

Также следует отметить, что поскольку наблюдаются существенные расхождения между теоретически рассчитанными и измеренными для конкретного слушателя кривыми маскировки, были проведены специальные дополнительные тесты прослушивания для определения индивидуаль-

ных кривых маскировки для выбранных экспертов. Полученные поправки (в основном касающиеся крутизны заднего фланца кривой относительно порога слышимости) использовались для вычисления глобальных порогов маскировки, используемых в работе.

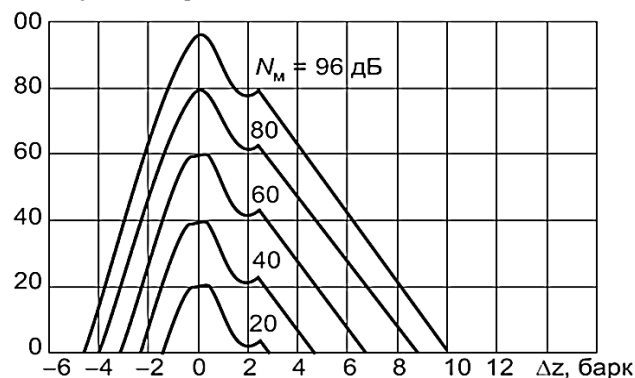


Рис. 1. Аппроксимации кривых маскировки (психоакустическая модель 2 стандарта MPEG ISO/IEC 11172-3)

В качестве количественной характеристики слышимых нелинейных искажений используется часто зависящее отношение «Шум-Маска» [6]:

$$NMR(n) = 10\log_{10} \left(\frac{1}{Z} \sum_{k=0}^{Z-1} \frac{E_{\text{noise}}[k, n]}{M[k, n]} \right), \quad (8)$$

здесь и ниже $E_{\text{noise}}[k, n]$ – энергия k -ой спектральной компоненты сигнала искажений; $M[k, n]$ – уровень порога глобальной маскировки; n – номер выборки.

В первой части работы для связи субъективной оценки с количественной характеристикой шумов через взвешивающую функцию G была введена выходная переменная: воспринимаемый уровень искажений (PNL, от англ. Perceive Noise Level). Было отмечено, что эта функция зависит от частоты:

$$PNL(n) = \sum_{k=0}^{Z-1} NMR(k, n) * G(k). \quad (9)$$

Как было сказано выше, в процессе анализа результатов тестов прослушивания наблюдается зависимость субъективного предпочтения искаженных сигналов не только от уровня нелинейных искажений, но и от тональной структуры сигнала – $T(f)$. Таким образом, выражение (9) можно переписать следующим образом:

$$PNL(n) = \sum_{k=1}^K NMR(k, n) * G(T(k), k). \quad (10)$$

Влияние нелинейных искажений на субъективную оценку качества воспроизводимого испытуемым устройством сигнала определяется средней величиной PNL, полученной для всего сигнала в целом (11):

$$TPNL = \frac{1}{N} \sum_{n=1}^N PNL(n), \quad (11)$$

где N – количество выборок в сигнале; n – порядковый номер выборки.

Основной задачей этой части работы является вычисление взвешивающей функции $G(T(f), f)$.

5. Эксперименты

Тесты прослушивания проводились в звукоизолированной комнате с уровнем фонового шума ~20 dB SPL в диапазоне частот 100 Гц–20 кГц. В качестве экспертов привлекались молодые люди в возрасте от 24 лет до 31 года без дефектов слуха, имеющие аудиометрический порог ниже 20 дБ в полосе частот 125 Гц–8 кГц (8 человек). Для оценки слуховой заметности вносимых искажений использовалась семибалльная шкала, при этом наиболее предпочтительное звучание получало оценку 7 баллов, наименее предпочтительное – 1 балл.

Тестовые музыкальные сигналы через внешнюю 24-х битовую звуковую карту MOTU UltraLite-mk3 подавались на усилитель HiFiMan EF-100, после чего воспроизводились с помощью высококачественных головных телефонов Sennheiser HD 800, имеющих крайне низкий коэффициент нелинейных искажений (< 1 %).

Для определения индивидуальных порогов маскировки для каждого из экспертов были проведены специальные тесты прослушивания, в которых перед экспертами ставилась задача разделить набор тестовых музыкальных сигналов на две группы: в первую группу вошли сигналы, в которых искажения не слышны, во вторую – сигналы со слышимыми искажениями. Разброс индивидуальных порогов маскировки составил в среднем 1 дБ. В представленном исследовании для расчета отношения «Шум-Маска» использовалось среднее значение порога для всех экспертов.

Перед началом исследований экспертам были предъявлены образцы отобранных МСреф (свободные от искажений, с оценкой 7 баллов). Аналогичным образом были предъявлены также образцы сигналов с максимальным количеством внешних нелинейных искажений (как наихудшее звучание, с оценкой 1 балл). Экспертам предлагалось прослушать весь набор записей в любом произвольном порядке для оценки их качества, затем, прослушивая повторно, поставить оценку в соответствии с собственным предпочтением, используя семибалльную шкалу.

Для определения взвешивающей функции $G(T(f), f)$ были проведены две серии субъективно-статистических экспертиз, для которых был подготовлен набор сигналов с известным количеством нелинейных искажений (см. 3. Искажения),

присутствующих, как в отдельных полосах частот, так и в нескольких частотных диапазонах одновременно.

В первой серии тестов прослушивания экспертам предлагалось оценить с точки зрения предпочтительности звучания набор тестовых музыкальных сигналов, имеющих нелинейные искажения только в одной полосе частот. Всего было предложено для прослушивания 5 групп (по количеству МСреф) по 28–33 тестовых фонограммы в каждой. Экспертам требовалось рассортировать их на 7 групп по степени предпочтительности. С помощью этой серии экспериментов удалось определить зависимости субъективной оценки от величины NMR и коэффициента тональности в полосе частот, содержащей нелинейные искажения (рисунок 2).

Во второй серии тестов прослушиваний экспертам предлагалось ранжировать с точки зрения предпочтительности звучания набор тестовых фонограмм, имеющих слышимые нелинейные искажения одновременно в нескольких полосах частот. Для проведения этой серии прослушиваний были также сформированы пять наборов референсных музыкальных сигналов, каждый из которых содержал около 10 образцов. Каждый образец представлял собой фрагмент МСреф, смикшированный с сигналами нелинейных искажений, содержащих различные комбинации уровня искажений в разных диапазонах частот.

В результате этих экспериментов была получена зависимость оценки предпочтения от величины TPNL (11).

6. Результаты

После обработки результатов были рассчитаны зависимости оценки предпочтения от величины NMR и тональности референсного сигнала в полосе частот. Затем полученные зависимости аппроксимированы прямыми вида (12):

$$\text{Score}(\text{band}) = 7 - k_{\text{band}}(\text{Tonality}) * \text{NMR}(\text{band}) + \varepsilon, \quad (12)$$

где $k_{\text{band}}(\text{Tonality})$ – коэффициент регрессии для диапазона частот band и тональности в полосе частот (tonality); ε – погрешность.

Для определения коэффициентов регрессии результаты были сведены в таблицы Excel и построены графики зависимостей и их линейные аппроксимации (см. рисунок 1).

Очевидно, что при дальнейшем увеличении количества слышимых искажений оценка не может снижаться бесконечно, начиная с некоторого уровня зависимости оценок предпочтения от NMR перестают быть линейными. Для определения коэффициентов регрессии мы использовали только линейные участки кривых. В таблице 2 приведены коэффициенты регрессии, полученные для разных сигналов.

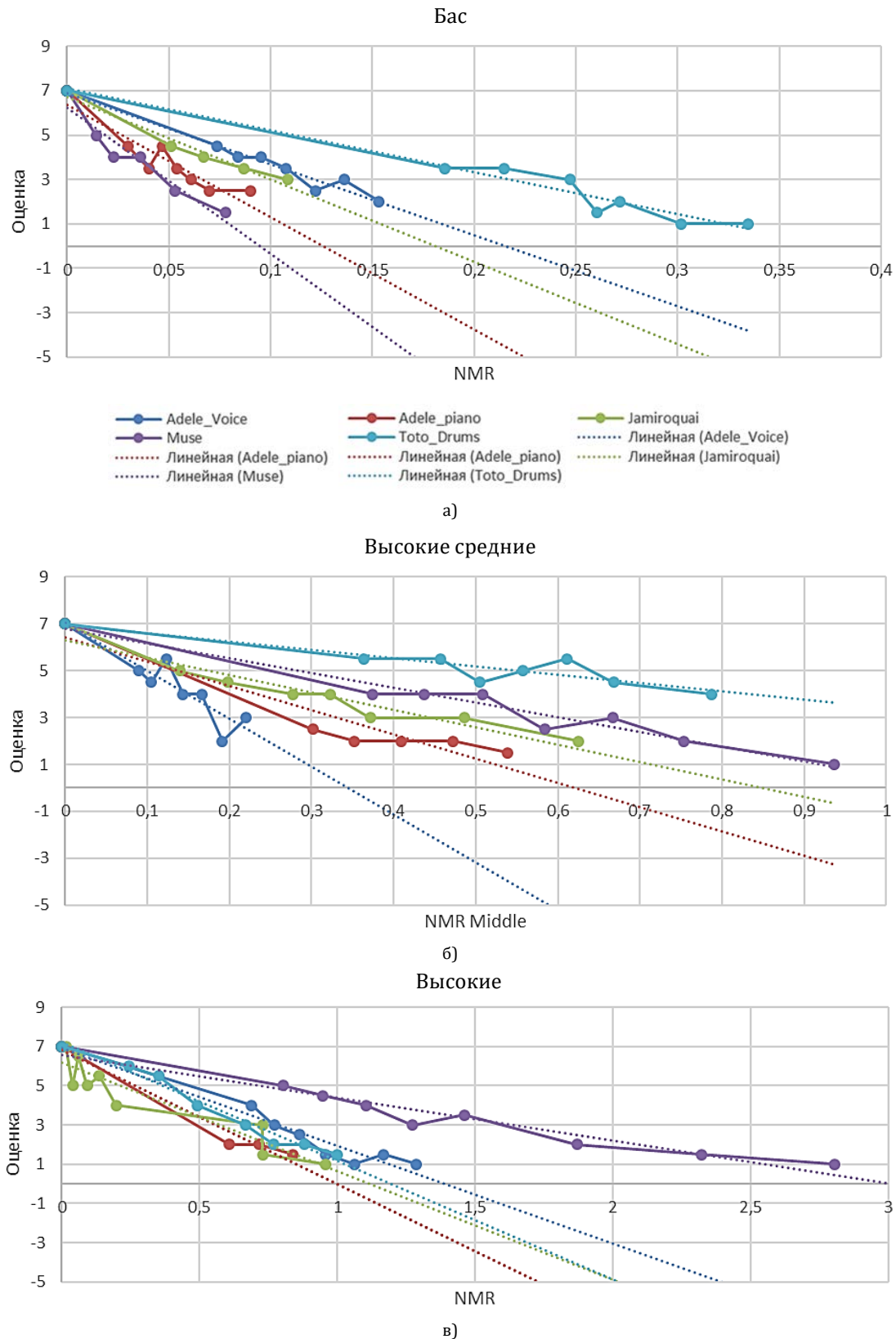


Рис. 2. Кривые изменения предпочтительности звучания от величины отношения «Шум-Маска» при наличии нелинейных искажений в полосе частот: а) 200–400 Гц; б) 1–4 кГц; в) 4–8 кГц

Далее был определен глобальный минимум коэффициентов регрессии и все значения разделены на этот минимум. Полученный набор нормиро-

ванных коэффициентов регрессии задает искомую взвешивающую функцию $G(T(f), f)$ в точках, соот-

ветствующих величинам коэффициентов тональности, полученных в экспериментах (рисунок 3).

Для упрощения расчетов принято решение использовать полиномиальную аппроксимацию усредненных значений коэффициентов регрессии для соответствующих значений тональности – $G(T)$:

$$G(T) = 5349 * T^3 - 2363 * T^2 + 358 * T - 13, \quad (13)$$

где T – рассчитанный коэффициент тональности.

ТАБЛИЦА 2. Коэффициенты регрессии k_{band} (Tonality)

Tonality	k низкие	k средние	k высокие средние	k высокие
0,06	–	–	3,5	2,2
0,08	–	5	6	5
0,09	–	8	6,5	5,4
0,13	–	15	7	6,8
0,15	–	16	–	–
0,2	–	18	–	–
0,23	–	20	–	–
0,25	–	30	–	–
0,27	–	40	–	–
0,28	–	45	–	–
0,3	–	60	–	–
0,33	–	90	–	–

Полученная взвешивающая функция была применена к вычисленным значениям NMR для отдельных выборок тестовых сигналов по формуле (10), значение было усреднено по количеству выборок по формуле (11). Значение субъективной оценки выбиралось по результатам тестов прослушивания, которые приведены на рисунке 3.

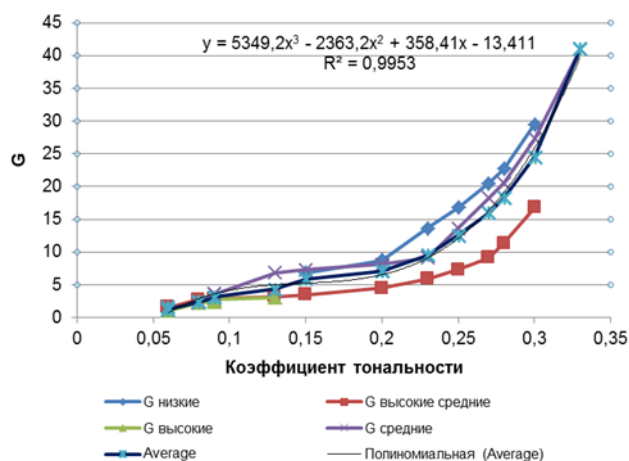


Рис. 3. Зависимости взвешивающей функции от тональности и частоты

Из рисунка 4 видно, что зависимость субъективного предпочтения практически линейно зависит от величины TPNL (корреляция составляет около 0,96 %), в то время как зависимость от величины отношения «Шум-Маска» гораздо менее строгая (корреляция составляет около 27 %).

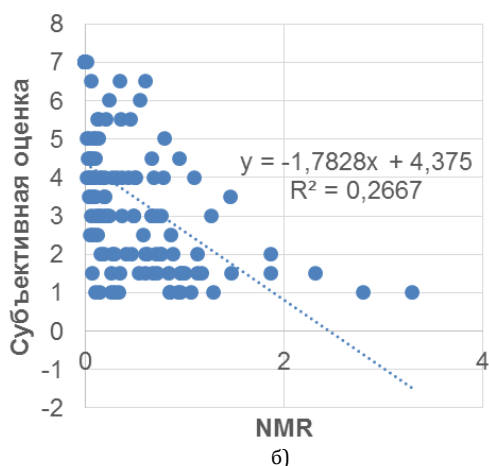
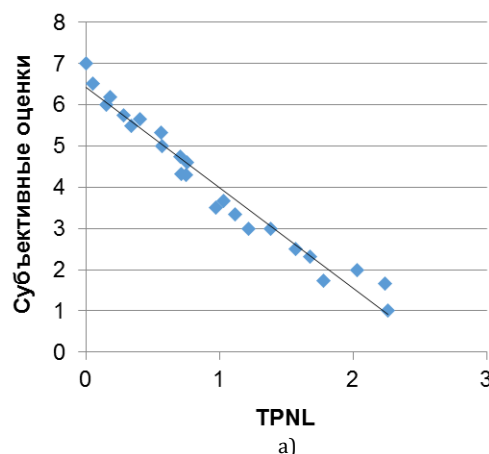


Рис. 4. Зависимость оценки: а) от средней величины воспринимаемых искажений (TPNL) (взвешенных величин NMR); б) от величины NMR

7. Заключение

В работе представлены результаты исследования влияния тональной структуры сигнала на субъективное восприятие музыкальных композиций с заметными нелинейными искажениями. Предложена вычислительно эффективная метрика тональности спектра выборки музыкального сигнала, позволяющая количественно оценивать тональность спектра в режиме обработки потокового аудиосигнала. Показаны зависимости влияния отношения «Шум-Маска» на субъективную оценку от коэффициента тональности. Вычислена взвешивающая функция $G(T)$, позволяющая установить связь между количеством слышимых искажений и субъективным предпочтением. Результаты экспериментов показали высокую точность предсказания субъективной оценки предпочтения сигналов, содержащих широкополосные нелинейные искажения.

Данная работа пока не может претендовать на полноту, так как в ней отсутствуют исследования влияния распределения звуковых искажений во времени и учета пиковых значений звуковых искажений в одном блоке искаженных кадров (параметр, аналогичный выходной переменной мо-

дели RDF в инструменте PEAQ). В заключительной части исследования будут представлены результаты исследования влияния временного распределения искаженных фрагментов на оценку субъек-

тивного предпочтения искаженных музыкальных сигналов, а также влияние сильно искаженных фрагментов.

Список используемых источников

1. Пахомов М.О., Рожнов В.С. Метод количественной оценки слухового восприятия нелинейных искажений. Часть 1. Испытательные сигналы и меры количественной оценки нелинейных искажений // Труды учебных заведений связи. 2018. Т. 4. № 1. С. 75–84. DOI:10.31854/1813-324x-2018-1-75-84.
2. Ковалгин Ю.А., Вологдин Э.И. Цифровое кодирование звуковых сигналов: Учебное пособие. СПб.: Корона Принт, 2004. 240 с.
3. Dubnov S. Generalization of spectral flatness measure for non-Gaussian linear processes // IEEE Signal Processing Letters. 2004. Vol. 11. Iss. 8. PP. 698–701. DOI: 10.1109/LSP.2004.831663.
4. Moore B.C. An Introduction to the Psychology of Hearing. London: Academic Press, 1989.
5. Moore B.C.J. A Model for the Prediction of Thresholds, Loudness, and Partial Loudness // Journal of the Audio Engineering Society. 1997. Vol 45. № 4. PP. 224–240.
6. Tan C.T., Moore B.C.J., Zacharov N., Mattila V.V. Predicting the Perceived Quality of Nonlinearly Distorted Music and Speech Signals // Journal of the Audio Engineering Society. 2004. Vol. 52. № 7/8. PP. 699–711.
7. Temme S., Olive S., Tatarunis S., Welti T., McMullin E. The Correlation between Distortion Audibility and Listener Preference in Headphones // 137th AES Convention, Los Angeles, USA. 2014.
8. Zwicker E., Fastl H. Psychoacoustics: Facts and Models. Berlin, Heidelberg: Springer-Verlag, 1990.

* * *

METHOD FOR QUANTITATIVE EVALUATION OF AUDITORY PERCEPTION OF NONLINEAR DISTORTION. PART 2. METRIC FOR MUSIC SIGNAL TONALITY AND ITS IMPACT ON SUBJECTIVE PERCEPTION OF DISTORTIONS

M. Pakhomov¹, V. Rozhnov^{1, 2}

¹St. Petersburg Branch of the corporation LG Electronics inc.
St. Petersburg, 191123, Russian Federation

²The Bonch-Bruевич State University of Telecommunications,
St. Petersburg, 193232, Russian Federation

Article info

Article in Russian

For citation: Pakhomov M., Rozhnov V. Method for Quantitative Evaluation of Auditory Perception of Nonlinear Distortion. Part 2. Metric for Music Signal Tonality and its Impact on Subjective Perception of Distortions // Proceedings of Telecommunication Universities. 2018. Vol. 4. Iss. 3. PP. 98–105.

Annotation: In the work presented the topic of taking into account the spectral characteristics of the test signal when evaluating the subjective perception of nonlinear distortions is disclosed. The most significant characteristic – tonality is chosen and a quantitative measure of this characteristic is proposed – the tonality coefficient. Subjective listening tests were conducted to determine the dependence of the auditory perception of nonlinear distortions on the tone structure of the signal and the spectral distribution of the noise-mask relationship.

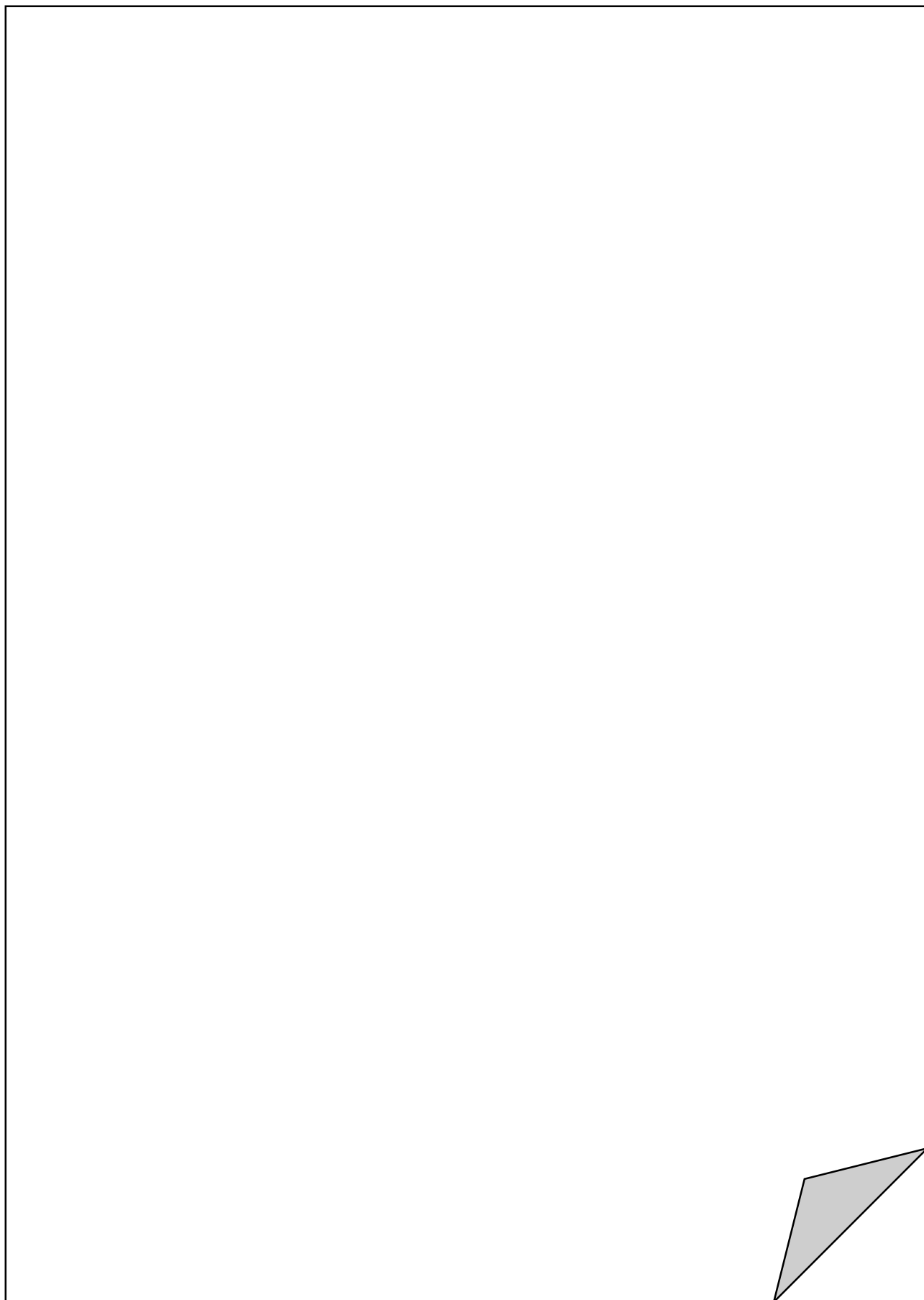
Keywords: multitone signal, nonlinear distortions, simultaneous masking, noise-to-mask ratio, tonality coefficient, linear multiregression model.

СВЕДЕНИЯ ОБ АВТОРАХ

- АНДРЕЕВА Елена Ивановна** кандидат физико-математических наук, доцент, доцент кафедры фотоники и линий связи Санкт-Петербургского государственного университета телекоммуникаций им. проф. М.А. Бонч-Бруевича, AndreevaElenaFils@gmail.com
- АНУФРЕНКО Александр Викторович** научный сотрудник научно-исследовательского центра Военной академии связи им. Маршала Советского Союза С.М. Буденного, leroi88@mail.ru
- БЫЛИНА Мария Сергеевна** кандидат технических наук, доцент, доцент кафедры фотоники и линий связи Санкт-Петербургского государственного университета телекоммуникаций им. проф. М.А. Бонч-Бруевича, BylinaMaria@mail.ru
- ГЛАГОЛЕВ Сергей Федорович** кандидат технических наук, доцент, заведующий кафедрой фотоники и линий связи Санкт-Петербургского государственного университета телекоммуникаций им. проф. М.А. Бонч-Бруевича, GlagolevSF@yandex.ru
- ДВОРНИКОВ Сергей Викторович** доктор технических наук, профессор, профессор кафедры радиосвязи Военной академии связи им. Маршала Советского Союза С.М. Буденного, practicdsv@yandex.ru
- ДОЦЕНКО Сергей Эдуардович** аспирант кафедры фотоники и линий связи Санкт-Петербургского государственного университета телекоммуникаций им. проф. М.А. Бонч-Бруевича, 0472895@gmail.com
- ЖИРОВ Виктор Аркадьевич** доктор технических наук, профессор, начальник научно-исследовательского отдела систем радиосвязи Санкт-Петербургского филиала Научно-исследовательского института радио (ЛОНИИР), vzhspb@loniir.ru
- КАНАЕВ Андрей Константинович** доктор технических наук, профессор, заведующий кафедрой «Электрическая связь» Санкт-Петербургского государственного университета путей сообщения Императора Александра I, kanaevak@mail.ru
- КАПРАЛОВ Дмитрий Дмитриевич** аспирант кафедры конструирования и производства радиоэлектронных средств Санкт-Петербургского государственного университета телекоммуникаций им. проф. М.А. Бонч-Бруевича, dmitry_kapralov@spbgut.ru
- КИРИК Дмитрий Игоревич** кандидат технических наук, доцент, заведующий кафедрой конструирования и производства радиоэлектронных средств, декан факультета радиотехнологий связи Санкт-Петербургского государственного университета телекоммуникаций им. проф. М.А. Бонч-Бруевича, d_i_kirik@mail.ru
- ЛОГИН Элина Валерьевна** ассистент кафедры «Электрическая связь» Санкт-Петербургского государственного университета путей сообщения Императора Александра I, elinabeneta@yandex.ru
- МАКАРЕНКО Сергей Иванович** кандидат технических наук, доцент, заместитель генерального директора по научной работе ООО «Корпорация «Интел групп», mak-serg@yandex.ru
- МАКАРОВ Леонид Михайлович** профессор кафедры конструирования и производства радиоэлектронных средств Санкт-Петербургского государственного университета телекоммуникаций им. проф. М.А. Бонч-Бруевича, elfbio@gmail.com

- МУРТАЛА Ибрахим** сетевой инженер компании «De-Kharms Telecommunications», (г. Илорин, Нигерия), ozutiams@yahoo.com
- ОРЛОВ Алексей Евгеньевич** преподаватель кафедры организации связи Военной академии связи им. Маршала Советского Союза С.М. Буденного, vorl-2210@mail.ru
- ПАХОМОВ Михаил Олегович** кандидат технических наук, senior research engineer Санкт-Петербургского филиала корпорации LG Electronics Inc., mpakhomov39@gmail.com
- ПОЗДНЯКОВ Александр Владимирович** доктор медицинских наук, профессор кафедры медицинской биофизики Санкт-Петербургского государственного педиатрического медицинского университета, pozdneyakovalex@ya.ru
- ПРОТАСЕНЯ Сергей Витальевич** кандидат технических наук, доцент кафедры конструирования и производства радиоэлектронных средств Санкт-Петербургского государственного университета телекоммуникаций им. проф. М.А. Бонч-Бруевича, saitvodabur@yandex.ru
- РОЖНОВ Виктор Сергеевич** аспирант кафедры радиосвязи и вещания Санкт-Петербургского государственного университета телекоммуникаций им. проф. М.А. Бонч-Бруевича, research engineer Санкт-Петербургского филиала корпорации LG Electronics Inc., victor.rozhnov@lge.com
- СМИРНОВ Александр Александрович** кандидат технических наук, докторант кафедры военных систем космической, радиорелейной, тропосферной связи и навигации Военной академии связи им. Маршала Советского Союза С.М. Буденного, midor012@mail.ru
- ТИАМИЙУ Осулале Абдулрахмон** PhD, старший преподаватель кафедры телекоммуникационной науки Университета Илорина (г. Илорин, Нигерия), ozutiams@yahoo.com
- ЧАЙМАРДАНОВ Павел Александрович** аспирант кафедры фотоники и линий связи Санкт-Петербургского государственного университета телекоммуникаций им. проф. М.А. Бонч-Бруевича, pchai@yandex.ru

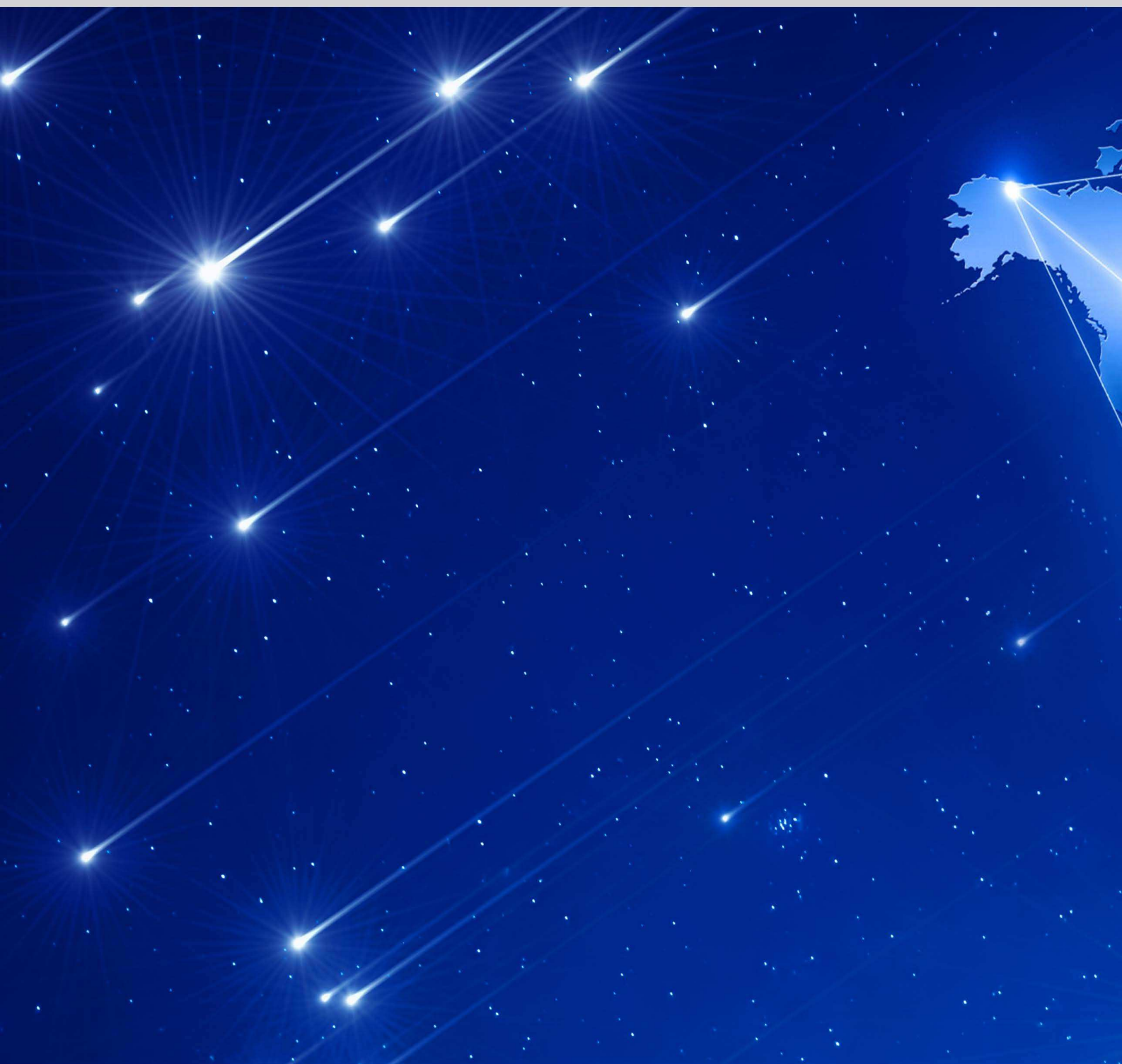
ДЛЯ ЗАМЕТОК



Учредитель и издатель:

Федеральное государственное бюджетное образовательное учреждение
высшего образования "Санкт-Петербургский государственный университет
телекоммуникаций им. проф. М.А. Бонч-Бруевича"

E-mail: tuzs@spbgut.ru Web: tuzs.sut.ru VK: vk.com/spbtuzs



ISSN: 1813-324X

Подписной индекс по каталогу "Межрегионального агентства подписки" – 35664