

НАУЧНЫЙ ЖУРНАЛ

**ТЕЛЕКОММУНИКАЦИИ
И ИНФОРМАЦИОННЫЕ
ТЕХНОЛОГИИ**

№1-2019

(Дата издания: июнь 2019 г.)

РЕДАКЦИОННАЯ КОЛЛЕГИЯ:

Орлов Владимир Георгиевич (*Главный редактор*)

к.т.н., начальник отдела организации научно-исследовательской работы студентов Московского технического университета связи и информатики (МТУСИ), начальник Центра научной работы и технического творчества молодежи МТУСИ, Москва, Россия

Андреев Владимир Александрович

д.т.н., профессор, Поволжский государственный университет телекоммуникаций и информатики, Самара, Россия

Бачевский Сергей Викторович

д.т.н., профессор, ректор Санкт-Петербургского государственного университета телекоммуникаций им. проф. Бонч-Бруевича, Санкт-Петербург, Россия

Зимин Игорь Викторович

Кыргызский государственный технический университет имени И.Раззакова. Институт электроники и телекоммуникаций, Бишкек, Кыргызстан

Ланчиков Павел Николаевич

НП Учебный центр Huawei (Москва), Шеньчжень, Китай

Маркосян Мгер Вардкесович

к.т.н., доцент, Ереванский НИИ средств связи, Ереван, Армения

Прохода Александр Николаевич

к.воен.н., доцент, Балтийский военно-морской институт им. Ф.Ф. Ушакова, Калининград, Россия

Рябко Борис Яковлевич

д.т.н., профессор, Сибирский государственный университет телекоммуникаций и информатики, Новосибирск, Россия

Самойлов Александр Георгиевич

д.т.н., профессор, заместитель директора института информационных технологий и радиоэлектроники Владимирского государственного университета имени Александра Григорьевича и Николая Григорьевича Столетовых (ВлГУ), Владимир, Россия

Рогачев Александр Александрович

д.т.н., в.н.с., Гомельский государственный университет имени Франциска Скорины, Гомель, Республика Беларусь

Суржиков Анатолий Петрович

д.ф.-м.н., профессор, Национальный исследовательский Томский политехнический университет, Томск, Россия

Титов Евгений Вадимович

к.т.н., профессор, Московский технический университет связи и информатики, Москва, Россия

УЧРЕДИТЕЛЬ:

**ОРДЕНА ТРУДОВОГО КРАСНОГО ЗНАМЕНИ ФЕДЕРАЛЬНОЕ ГОСУДАРСТВЕННОЕ
БЮДЖЕТНОЕ ОБРАЗОВАТЕЛЬНОЕ УЧРЕЖДЕНИЕ ВЫСШЕГО ОБРАЗОВАНИЯ
«МОСКОВСКИЙ ТЕХНИЧЕСКИЙ УНИВЕРСИТЕТ СВЯЗИ И ИНФОРМАТИКИ» (МТУСИ)**

РЕДАКЦИОННАЯ ПОДГОТОВКА:

**Отдел организации научно-исследовательской работы студентов
(ОНИРС МТУСИ)**

СОДЕРЖАНИЕ №1-2019

«Цифровые технологии радиосвязи и телерадиовещания»

Сатарова А.И., Варламов В.О., Лобов Е.М.

Проектирование и имитационное моделирование гибридного банка фильтров для обработки широкополосных сигналов низкоскоростными АЦП..... 5

Севериненко А.А., Иванюшкин Р.Ю.

Лабораторный стенд для исследования амплитудно-фазовой конверсии в радиочастотном усилителе мощности на полевом транзисторе..... 11

Платонов О.И., Лобов Е.М.

Исследование гибридных банков фильтров в задачах цифровой обработки сверхширокополосных сигналов.....18

Ионов В.В., Пестряков А.В.

Современное состояние систем космического мониторинга за подвижными объектами на территории России.....24

«Сетевые технологии и системы телекоммуникаций»

Мирошниченко А.В., Шаврин С.С.

Исследование частотно-селективных характеристик авиационных бортовых радиостанций радиотелефонной связи УКВ диапазона.....28

Арсеньева Д.Г., Сызранцев Г.В.

Алгоритмы функционирования маршрутизатора системы сетевого технологического управления при построении таблицы маршрутизации.....34

Кусакина М.С., Нетес В.А.

Проблемы обеспечения надежности программно-конфигурируемых сетей.....39

Коротков А.В., Сенявский А.Л.

Погрешность и неопределенность в метрологии.....44

Денисова М. А., Степанов М.С.

Оценка числа устройств IVR и операторов в контакт-центре с использованием метода декомпозиции.....49

Макарина Д.А., Данилов А.Н.

Анализ технологии мобильной связи пятого поколения.....55

Бельфер Р.А., Глинская Е.В., Кравцов А.В., Орлов В.Г.

Состояние разработки имитатора объединенной сети ПД специального назначения в качестве учебного лабораторного стенда.....61

«Информационные технологии и автоматизация процессов в системах связи»

Авакян С.Т., Шелухин О.И.

Повышение эффективности коллаборативных рекомендательных систем при использовании дополнительных данных.....66

<i>Чистов А.В., Руднев А.Н.</i> Синхронизация в сети как этап обеспечения безопасности данных типа “BIG DATA”.....	77
<i>Зайцев Е.С., Сераджи С.И., Кальфа А.А., Беленькая М.Н.</i> Исследование существующих систем обфускации с целью защиты программных алгоритмов.....	83
<i>Кибедов Р.А., Турута Е.Н.</i> Обеспечение отказоустойчивости ВС на основе статического отказоустойчивого планирования задач.....	89
<i>Пьянкова Т.П., Яшина М.В.</i> Исследование свойств двумерных дискретных преобразований в компьютерном зрении.....	96
<i>Пономарев А.А., Шевелев С.В.</i> Классификация угроз для интернета вещей.....	103
<i>Доткулова А.С., Порунов А.А.</i> Методы и средства разработки интерактивной системы обучения для отслеживания уровня подготовки студентов.....	109
<i>Бушув Э.Ю., Волчков В.П.</i> Разработка алгоритма, реализующего метод спектрального анализа прони в среде матлаб, и исследование чувствительности этого метода к воздействию шумовых помех.....	116
<i>Келимов Э.Р., Гузеев А.В.</i> Применение динамического программирования в выравнивании последовательностей ДНК.....	126
<i>Мирошниченко А.В., Волчков В.П.</i> Применение фреймов Вейля-Гейзенберга в задачах интерполяции и экстраполяции сигналов.....	134
<i>Усачева Д.И., Шишкин М.О., Гадасин Д.В., Гузеев А.В.</i> Применение OLAP-технологий для анализа многомерных данных в контакт-центре.....	142
<i>Артемов М.Д., Воронова Л.И.</i> Разработка подсистемы распознавания простых динамических жестов с помощью 3DCNNLSTM.....	150

**ПРОЕКТИРОВАНИЕ И ИМИТАЦИОННОЕ МОДЕЛИРОВАНИЕ ГИБРИДНОГО БАНКА
ФИЛЬТРОВ ДЛЯ ОБРАБОТКИ ШИРОКОПОЛОСНЫХ СИГНАЛОВ
НИЗКОСКОРОСТНЫМИ АЦП**

Саттарова Анжела Ильдаровна
студентка группы БРР1601 МТУСИ
ang.satt.97@gmail.com

Варламов Владимир Олегович
студент группы БРР1601 МТУСИ
f.vvo@bk.ru

Лобов Евгений Михайлович
МТУСИ, к.т.н., доцент кафедры РТС
lobov@srd.mtuci.ru

Ключевые слова: банки фильтров, аналого-цифровое преобразование, широкополосный сигнал, динамический диапазон.

В настоящее время существует тенденция повышения скорости передачи информации, посредством увеличения полосы используемых частот. Однако, при переходе к широкополосным сигналам при работе в условиях большого уровня помех, требуется обеспечивать широкий динамический диапазон при оцифровке принимаемого сигнала. Для повышения динамического диапазона предлагается использовать гибридный банк фильтров, позволяющий обрабатывать широкополосный сигнал низкоскоростными прецизионными аналого-цифровыми преобразователями. В данной работе были произведены расчет и моделирование гибридного банка фильтров анализа - синтеза для диапазона частот 26,25 МГц-86,25 МГц. При помощи согласованных фильтров оценены искажения линейно частотно модулированного сигнала при прохождении банков фильтров без промежуточной обработки, с проведением понижения и последующего повышения частоты дискретизации промежуточных сигналов с выходов аналоговых фильтров. Также было показано, что в случае приема смеси сигнала и шума, искажения, вносимые системой обработки, несущественно сказываются на отношении сигнал/шум выходного сигнала.

Введение

В цифровой обработке сигналов широко применяются системы многоскоростной обработки, позволяющие при помощи набора фильтров - банка анализа, разделить сигнал на полосы, частота дискретизации которых понижается и на пониженной скорости проводится параллельная обработка фрагментов сигнала. После чего сигнал восстанавливается при помощи банка фильтров синтеза. При этом искажения, вносимые в сигнал фильтрами синтеза и анализа и наложением спектра при децимации, полностью компенсируются. Данные системы получили свое распространение в ситуациях где необходимо производить сложную обработку широкополосных сигналов, например, в системах радиолокации и высокоскоростной связи [1 - 6, 13 - 15]. Необходимость оцифровки широкополосных сигналов с высокой разрядностью и большим динамическим диапазоном привела к появлению гибридных банков фильтров [7 - 9]. В таких системах аналоговый сигнал при помощи фильтров разделяется на полосы, каждая из которых оцифровывается отдельными низкоскоростными высокоточными АЦП. Далее, при необходимости, производится низкоскоростная обработка этих сигналов и при помощи цифрового банка фильтров синтеза восстанавливается широкополосный сигнал. Понижение частоты дискретизации позволяет использовать АЦП большей разрядности. Однако, при этом из-за невозможности идеального согласования банков анализа и синтеза возникают искажения выходного сигнала. Так же

возможна частичная компенсация искажений, создаваемых неравномерностями ФЧХ и ГВЗ аналоговых фильтров, аналогичная методу, используемому в работах [10 - 12].

Моделирование системы обработки

Для моделирования обработки аналогового сигнала был синтезирован банк из восьми полосовых фильтров третьего порядка, покрывающих диапазон от 26,25 МГц до 86,25 МГц. Было проведено схемотехническое моделирование данного банка в пакете “microcap”. Полученные в результате моделирования частотные характеристики были поотсчетно сняты и в дискретном виде перенесены в пакет “matlab”, в котором была произведена дальнейшая обработка. Схема аналогового банка фильтров представлена на Рис. 1. Для симуляции оцифровки сигнала на частоте 15МГц частота дискретизации откликов аналоговых фильтров была понижена в 16 раз.

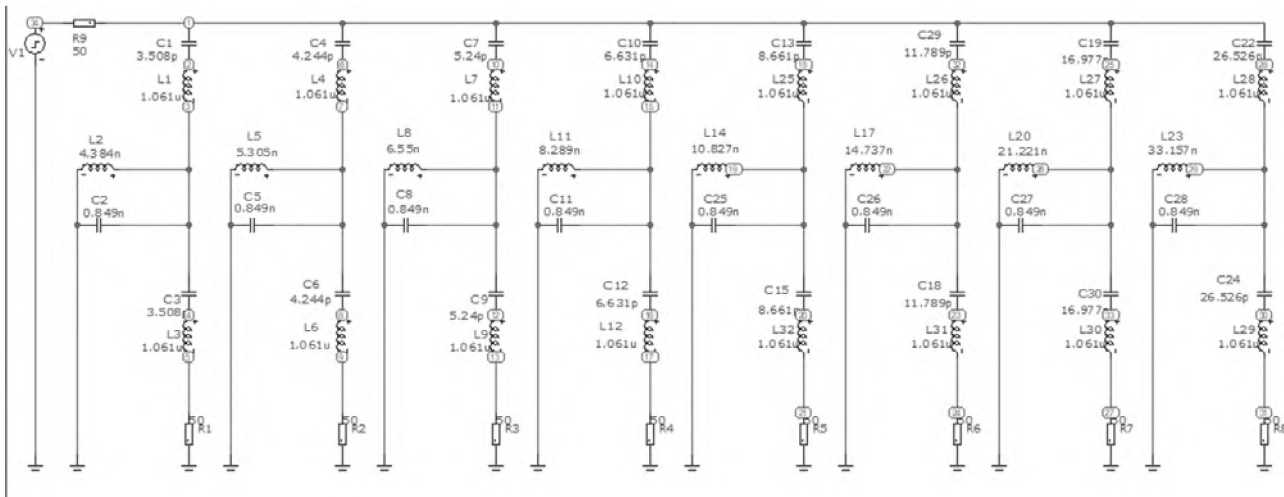


Рис. 1. Принципиальная схема аналогового банка фильтров

Для определения искажений, вносимых системой обработки, был синтезирован тестовый аналитический ЛЧМ сигнал длительностью 200000 отсчетов с частотой дискретизации $f_s=240\text{MHz}$, минимальной частотой 26,25МГц и максимальной 86,25МГц с примесью аддитивного белого гауссовского шума. Таким образом, отношение сигнал шум на входе составило 20 дБ.

$$x(n) = \cos(2\pi * ((f_{\min} + \frac{b}{2} * t) * t)) + i * \sin(2\pi * ((f_{\min} + \frac{b}{2} * t) * t)) \tag{1}$$

где

$$b = (f_{\max} - f_{\min}) / T_s$$

где T_s -длительность сигнала, f_{\min} - минимальная f_{\max} максимальная частота сигнала.

Для оценки вносимых искажений использовались взаимные корреляционные функции исходного (рис. 2) и преобразованного сигналов. Количественной мерой искажений служило отношение квадрата модуля максимума ВКФ к среднему квадрату модуля боковых лепестков минус 1 – R (2), что при уровне шума превышающем уровни боковых лепестков ВКФ позволяет оценивать отношение сигнал-шум обрабатываемого сигнала.

$$R = \frac{|B_{\max}|^2}{\sum_{i=1}^{N-1} |B_i|^2 - |B_{\max}|^2} \tag{2}$$

где N – количество отсчетов отклика СФ, B_i – отклик СФ.

ВКФ рассчитывалась с помощью фильтра, согласованного с исходным ЛЧМ сигналом (СФ).

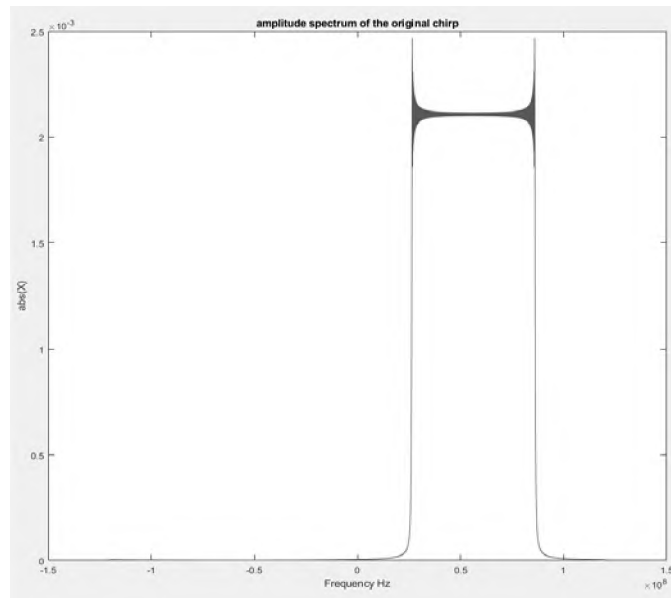


Рис. 2. Амплитудный спектр аналитического ЛЧМ сигнала

Значение R для отклика СФ на используемый ЛЧМ сигнал с примесью шума составило 19.93 дБ, а для сигнала, прошедшего систему обработки, 19.27 дБ (рис. 3), относительно этого значения будут измеряться количественные показатели искажений сигнала.

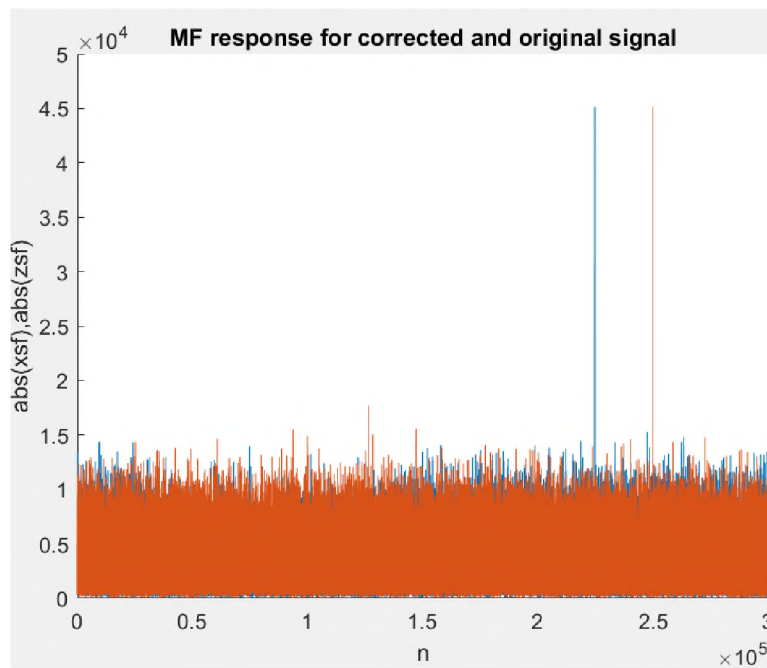


Рис. 3. Реакция согласованного фильтра на ЛЧМ с шумом; синий – оригинальный сигнал, оранжевый – сигнал прошедший систему обработки

Для синтеза сигнала по откликам аналоговых фильтров был рассчитан банк фильтров синтеза посредством экспоненциальной модуляции фильтра прототипа. Импульсная характеристика k-го фильтра синтеза задается выражением (3).

$$h_k(n) = h(n)e^{j\frac{2\pi}{2M}k\left(n - \frac{N-1}{2}\right)} \quad (3)$$

где $n=0-255$; $k=0-31$ - номер фильтра, соответствующий номеру канала. $h(n)$ – импульсная характеристика фильтра прототипа. Из 32 фильтров были взяты 8, у которых центральные частоты

соответствуют положительным центральным частотам аналоговых фильтров. АЧХ фильтров синтеза представлены на рис. 4.

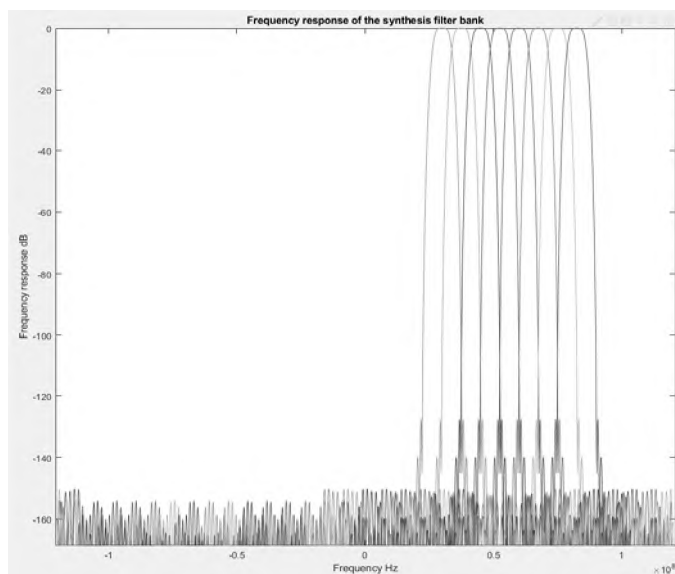


Рис. 4. Частотная характеристика банка фильтров синтеза

Таким образом, передаточную функцию полученной системы можно представить как (рис. 4).

$$T(j\omega)_{\text{aliasing}} + T(j\omega)$$

где $T(z)$ – полезная составляющая передаточной функции (рис. 5), равная

$$\sum_{i=0}^{M-1} H_i(j\omega) * F_i(j\omega)$$

$H_i(j\omega)$ – передаточная функция i -го фильтра анализа, $F_i(j\omega)$ – передаточная функция i -го фильтра синтеза.

Наложенная составляющая $T(j\omega)_{\text{aliasing}}$ (рис. 6) равна:

$$\sum_{i=0}^{M-1} F_i(j\omega) * H_i(j(M\omega - 2\pi m) * f_s) . \text{ Для } m=-M, -M+1, ..M$$

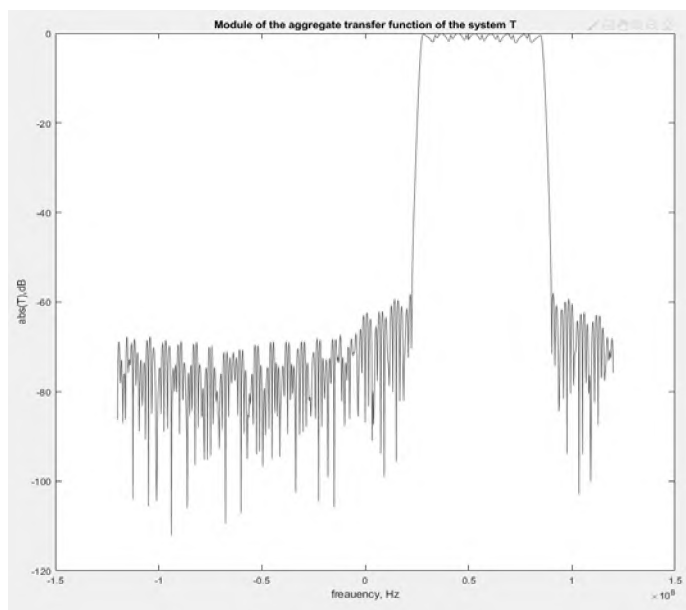


Рис. 5. Модуль полезной составляющей передаточной функции системы

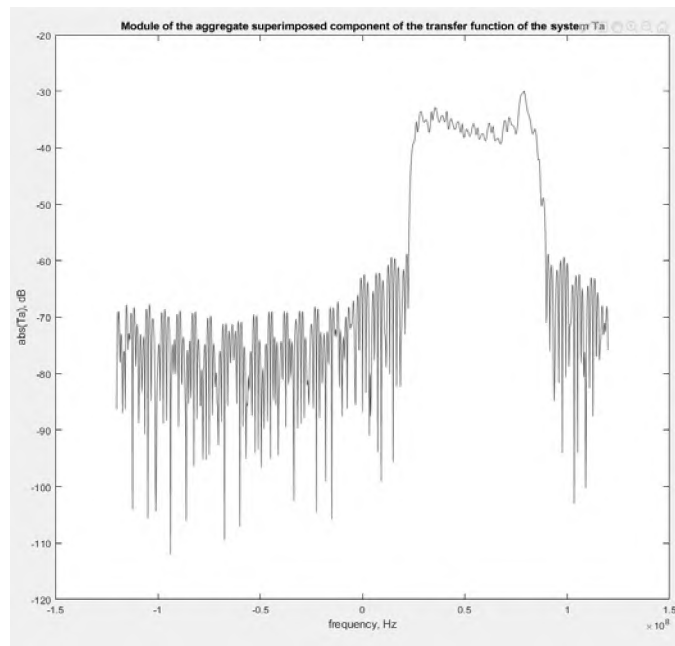


Рис. 6. Модуль наложенной компоненты передаточной функции

Заключение

Было проведено моделирование системы, позволяющей обеспечить низкоскоростную оцифровку широкополосных сигналов, и получены выражения, позволяющие количественно оценивать искажения, вносимые системой. Система подразумевает возможность согласования АЧХ фильтров анализа и синтеза для приближения к идеальному восстановлению. Так же возможна обработка широкополосных сигналов в диапазоне СВЧ посредством двух и более этапного разбиения сигнала на узкополосные каналы.

Литература

1. *Iman Moazzen, Pan Agathoklis.* A General Approach for Filter Bank Design Using Optimization. Department of Electrical and Computer Engineering, University of Victoria. Technical Report. March 2014.
2. *Wu-Sheng Lu,* Use SeDuMi to Solve LP, SDP and SCOP Problems: Remarks and Examples. Dept. of Electrical and Computer Engineering, University of Victoria.
3. *Eero P. Simoncelli and Edward H. Adelson.* Subband Transforms. Chapter 4. MIT Media Laboratory Vision and Modeling Technical Report #137.
4. *Чобану М.К.* Синтез Оптимизированных Многомерных Банков Фильтров. НАУЧНЫЙ ВЕСТНИК МГТУ ГА № 117 2007, сер. Радиофизика и радиотехника. УДК 621.396.
5. *Чобану М.К.* Системы многоскоростной обработки многомерных сигналов. Ч.1. Московский энергетический институт. УДК 621.396.
6. *Чобану М.К.* Синтез основных элементов многомерных многоскоростных систем. Ч.1. Неразделимые матрицы децимации // Сиб. журн. вычисл. матем., 2008, том 11, № 1. С. 95-113.
7. *Tanja Karp, N. J. Fliege.* Modified DFT Filter Banks with Perfect Reconstruction. IEEE Transactions On Circuits And Systems-II: Analog And Digital Signal Processing. Vol. 46. No. 11, November 1999. 1057–7130/99 1999 IEEE.
8. *Anand K. Soman, P. P. Vaidyanathan, Truong Q. Nguyen.* Linear Phase Paraunitary Filter Banks: Theory, Factorizations and Designs // IEEE Transactions On Signal Processing, Vol. 41. No. 12, December 1993. 0018-9464/93 1993 IEEE.

9. *Vaidyanathan P.P.* Multirate Digital Filters, Filter Banks, Polyphase Networks, and Applications: A Tutorial // Proceedings Of The IEEE, Vol. 78, No. 1, January 1990. 00189219/90/0100-0056 1990 IEEE.

10. *Лобов Е.М., Лобова Е.О., Припутин В.С.* Программа, реализующая банк цифровых фильтров синтеза на основе ЕЛТ-преобразований. Свидетельство о регистрации программы для ЭВМ RUS 2017663746 20.10.2017.

11. *Лобов Е.М., Лобова Е.О., Елсуков Б.А.* Устройство компенсации дисперсионных искажений широкополосных сигналов на базе банка цифровых фильтров // DSPA: Вопросы применения цифровой обработки сигналов. 2018. Т. 8. № 3. С. 60-65.

12. *Лобов Е.М., Лобова Е.О., Курочкин А.А.* Вычислительно упрощенная реализация компенсатора дисперсионных искажений на базе банка цифровых фильтров // Системы синхронизации, формирования и обработки сигналов. 2018. Т. 9. № 1. С. 106-110.

13. *Лобов Е.М., Смердова Е.О., Терешонок М.В.* Комплекс лабораторных работ по дисциплине "цифровая обработка сигналов" для студентов направления 11.03.01 и 11.03.02 на базе системы имитационного моделирования "СПЕКТР-2" // Методические вопросы преподавания инфокоммуникаций в высшей школе. 2017. Т. 6. № 2. С. 9-15.

14. *Лобов Е.М., Смердова Е.О.* Оценка импульсной характеристики широкополосного ионосферного канала с помощью метода обратной фильтрации // Т-Comm: Телекоммуникации и транспорт. 2017. Т. 11. № 3. С. 24-27.

15. *Лобов Е.М., Косилов И.С., Кандауров Н.А., Елсуков Б.А.* Методика оценки параметров частотной дисперсии ионосферного канала с помощью широкополосного фазоманипулированного сигнала // Т-Comm: Телекоммуникации и транспорт. 2014. Т. 8. № 9. С. 49-53.

DESIGN AND SIMULATION OF A HYBRID FILTER BANK FOR PROCESSING WIDEBAND SIGNALS WITH LOW-SPEED ADCS

Anjela I. Sattarova

Student of group BRR1601 MTUCI

ang.satt.97@gmail.com

Vladimir O. Varlamov

Student of group BRR1601 MTUCI

f.vvo@bk.ru

Evgeniy M. Lobov

MTUCI, PhD., associate professor of RTS department

lobov@srd.mtuci.ru

Keywords: *filter banks, broadband signals, analog-to-digital conversion, dynamic range.*

Currently, there is a tendency to increase the speed of information transmission by increasing used frequency band. However, when switching to broadband signals and operating under conditions of a high level of noise, a wide dynamic range is required when digitizing the received signal. To increase the dynamic range, it is proposed to use a hybrid filter bank, which allows processing the broadband signal with low-speed precision analog-to-digital converters. In this paper, was calculated and simulated a hybrid analysis-synthesis filter bank for the 26.25 MHz-86.25 MHz frequency range. Using matched filters, we estimated the distortion of a linearly frequency-modulated signal during the passage of filter banks without intermediate processing, after upsampling and downsampling intermediate signals from the outputs of analog filters. To reduce phase shifts and time delays arising during signal processing by synthesis filters, compensation delays and phase shifts were introduced, adjusted to distortions carried by the processing system. It was shown that phase correction allows reducing signal distortions. It was also shown that in the case of receiving a mixture of signal and noise, the distortion introduced by the processing system does not significantly affect the signal-to-noise ratio of the output signal.

ЛАБОРАТОРНЫЙ СТЕНД ДЛЯ ИССЛЕДОВАНИЯ АМПЛИТУДНО-ФАЗОВОЙ КОНВЕРСИИ В РАДИОЧАСТОТНОМ УСИЛИТЕЛЕ МОЩНОСТИ НА ПОЛЕВОМ ТРАНЗИСТОРЕ

Севериненко Александр Андреевич
студент группы МРТ1701 МТУСИ
kuju1ster@yandex.ru

Иванюшкин Роман Юрьевич
МТУСИ, к.т.н., доцент кафедры РОС
rivanyushkin@gmail.com

Ключевые слова: резонансный усилитель мощности, полевой транзистор, амплитудно-фазовая конверсия, фазоамплитудная характеристика, межэлектродные емкости.

Представлена разработка лабораторного стенда на основе резонансного радиочастотного каскада усиления мощности на полевом транзисторе, включая выбор схемы усилителя, подключения измерительных приборов, макетирование, исследование характеристик и сборку лабораторного стенда. Приведены результаты исследования основных энергетических и качественных характеристик разработанного лабораторного стенда.

На сегодняшний день в связи с развитием цифрового теле- и радиовещания, ввиду применения сложных амплитудно-фазомодулированных сигналов, все большую значимость имеет проблема амплитудно-фазовой конверсии – паразитного преобразования амплитудной модуляции в фазовую за счет влияния нелинейных параметрических емкостей усилительных элементов [1 - 6].

Ввиду отсутствия лабораторных стендов, посвященных изучению этого явления [7] принято решение разработать новый лабораторный стенд, позволяющий производить наглядное исследование, как основных характеристик резонансного усилителя мощности, так и явления амплитудно-фазовой конверсии.

При выборе схемы радиочастотного усилителя мощности принято решение использовать резонансный каскад усиления мощности (рис.1) на полевом транзисторе [8, 9], так как при такой схеме максимальная мощность на выходе вырабатывается по первой гармонике, а остальные гармоники значительно подавлены. На выходе усилителя мощности параллельно резистивной нагрузке R_n подключен параллельный колебательный контур $C_{\text{кв}}, L_{\text{кв}}$. Для наблюдения формы импульсов тока установлен резистор $R_{\text{квт}}$. В цепи истока полевого транзистора установлен резистор $R_{\text{квт}}$ для термостабилизации. Питается усилитель от источника питания $U_{\text{пит}}$ через дроссель $L_{\text{д.т}}$. В цепи питания также установлен блокировочный конденсатор $C_{\text{б.т}}$.

На основе показанного на рисунке 1 резонансного каскада усиления мощности была разработана схема лабораторного стенда, представленная на рисунке 2. Для удобства проведения измерений принято решение выбрать относительно невысокую рабочую частоту – 500 кГц. В качестве усилительного элемента при сравнении подходящих, выпускаемых в настоящее время транзисторов выбран полевой транзистор IRF610 за счет небольших межэлектродных емкостей, подходящих рабочих частот, а также удобства организации теплоотвода. Его основные параметры приведены в таблице 1 [10].

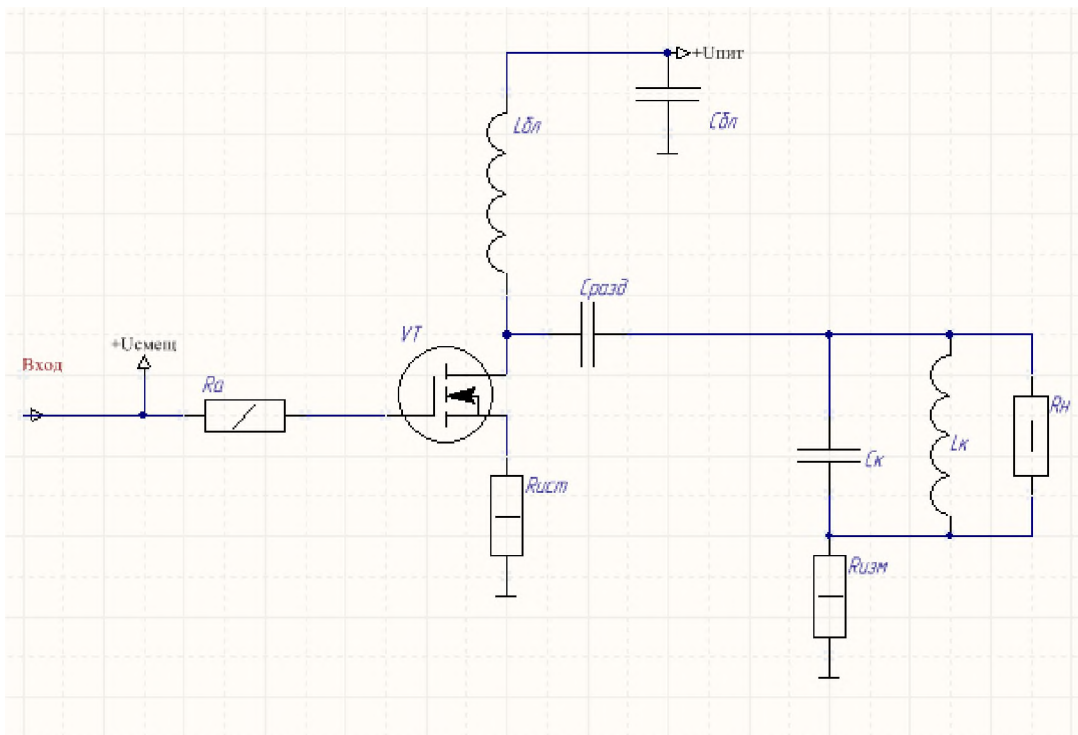


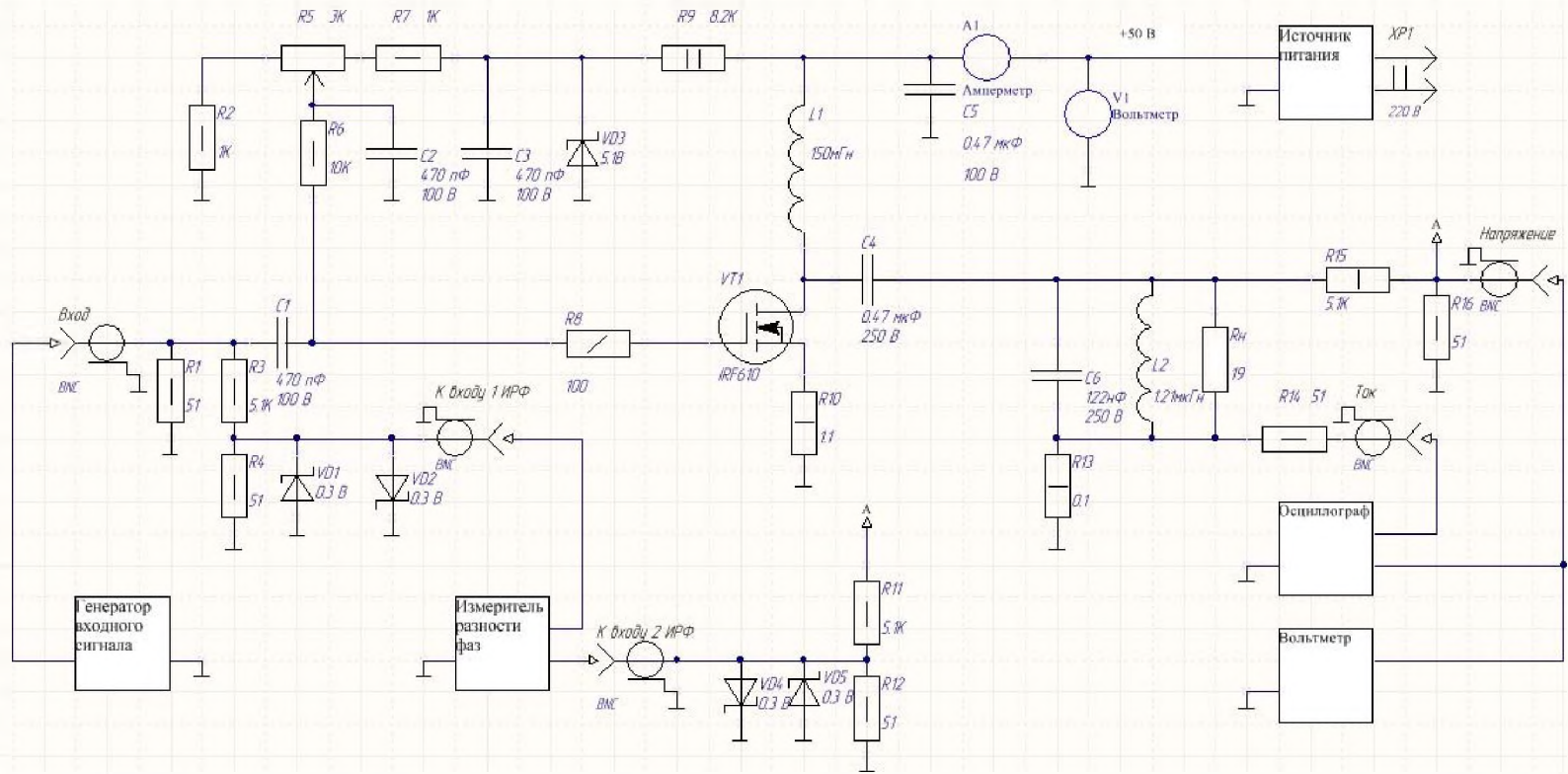
Рис. 1. Принципиальная схема резонансного каскада усиления

Таблица 1

Максимальное напряжение сток-исток, В	Максимальный ток стока, А	Максимальное напряжение затвор-исток, В	Максимальная рассеиваемая мощность, Вт	Тип корпуса	Входная емкость, пФ	Выходная емкость, пФ
200	3,3	20	36	ТО220АВ	140	53

Усилитель мощности работает на колебательный контур L2C6, образующий с резистивной нагрузкой Rн резонансную нагрузку. На выходе усилителя предусмотрены делители для снижения влияния емкостей соединительных кабелей и понижения напряжений на измерительных разъемах. Один из них предназначен для подключения лабораторного вольтметра и осциллографа, другой – для подключения измерителя разности фаз.

На входе также установлен делитель для подключения измерителя разности фаз, так как максимальное рабочее напряжение на его входах составляет 0,3 Вольта. Входной сигнал поступает на усилитель от генератора ГЗ-112, работающего на нагрузку 50 Ом (R1) через разделительный конденсатор C1 и «антипаразитный» резистор R8. Напряжение смещения поступает на затвор полевого транзистора VT1 через цепь стабилизации напряжения R9VD3 и резистивные делители R2, R5-R7. Переменный резистор R5 позволяет корректировать напряжение смещения при настройке стенда. В цепи истока полевого транзистора установлен резистор R10 для термостабилизации, а цепь стока питается от источника питания через дроссель L1, что практичнее резисторной схемы, с точки зрения снижения тепловыделения. Измерительный резистор R10 установлен в цепи колебательного контура и нагрузки для возможности наблюдения формы импульсов выходного тока. Напряжение питания усилителя – 50 Вольт. В цепи источника питания установлены амперметр и вольтметр для контроля потребляемой усилителем мощности.



Далее проведена разработка конструкции и сборка лабораторного стенда. Вид его лицевой панели приведен на рисунке 3, вид со снятой крышкой приведен на рисунке 4.

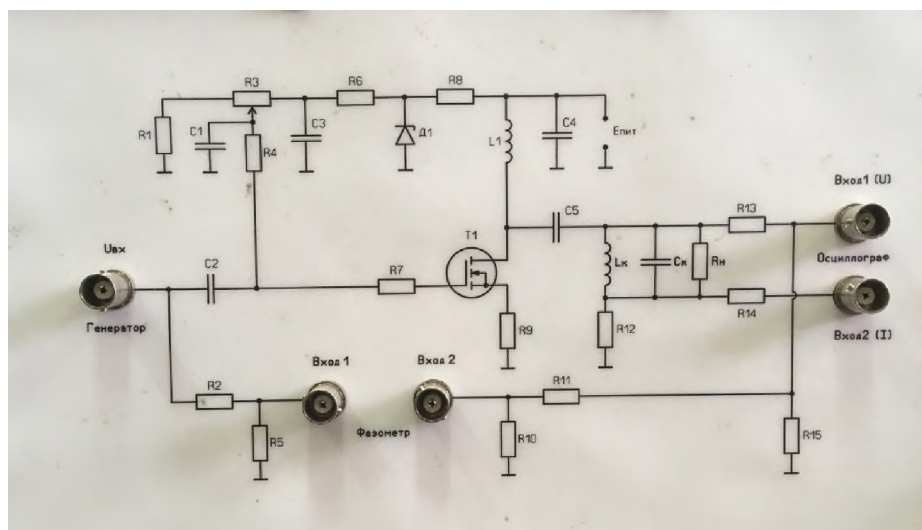


Рис. 3. Вид лицевой панели лабораторного стенда

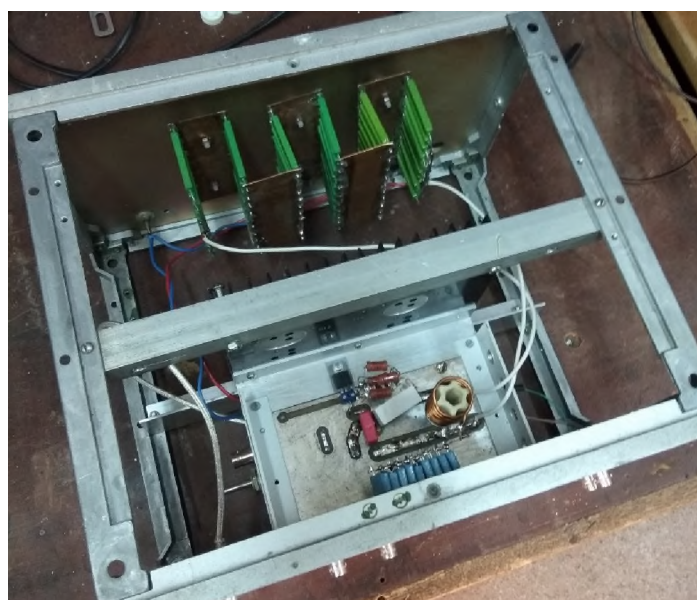


Рис. 4. Вид конструкции лабораторного стенда со снятой крышкой

Проведено исследование основных энергетических и качественных характеристик разработанного лабораторного стенда. График амплитудно-частотной характеристики приведен на рисунке 5, нагрузочные характеристики – на рисунке 6, зависимость КПД от сопротивления нагрузки – на рисунке 7, амплитудная характеристика – на рисунке 8, фазоамплитудная характеристика – на рисунке 9.

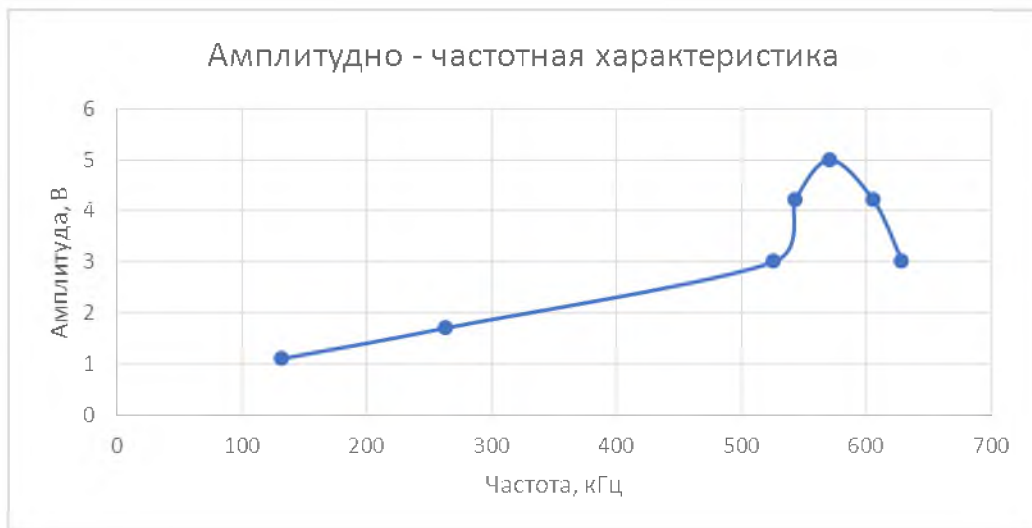


Рис. 5. Амплитудно-частотная характеристика усилителя мощности

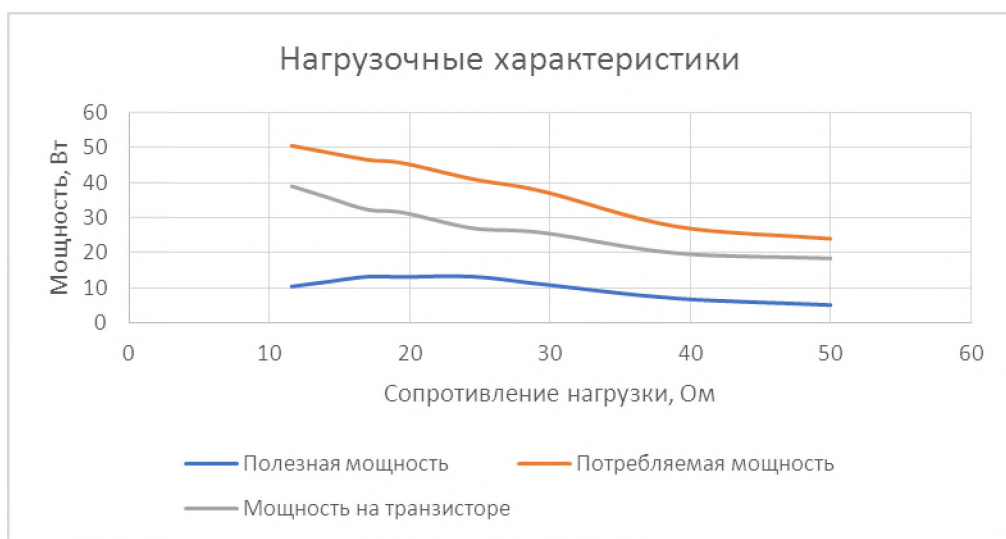


Рис. 6. Нагрузочные характеристики усилителя мощности



Рис. 7. Зависимость КПД усилителя от сопротивления нагрузки

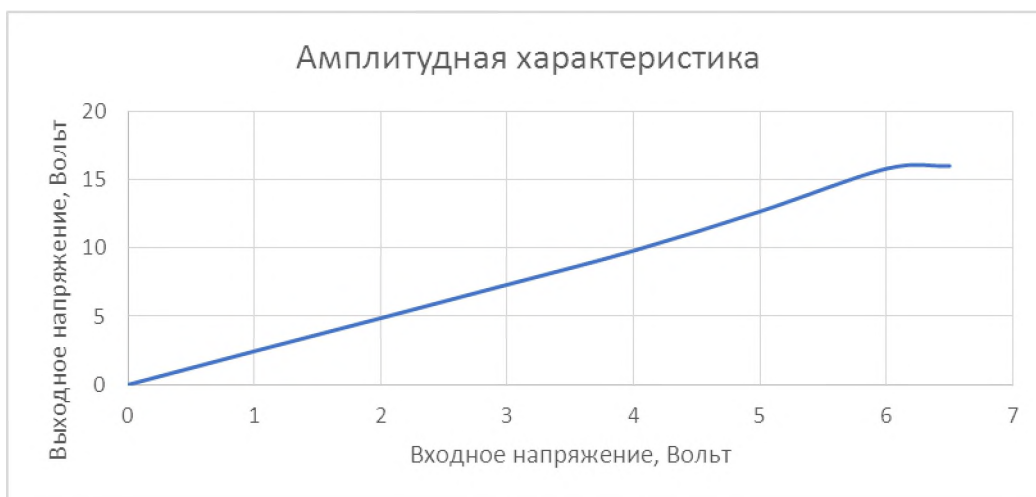


Рис. 8. Амплитудная характеристика усилителя мощности

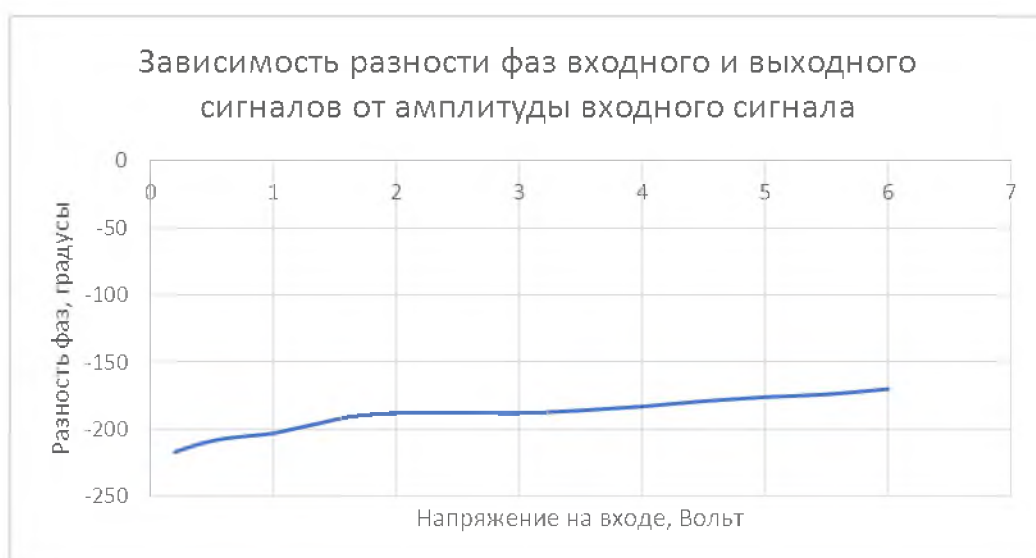


Рис. 9. Фазоамплитудная характеристика усилителя

По результатам исследования характеристик разработанного и изготовленного лабораторного стенда получено:

- резонансная частота – 571 кГц;
- полоса пропускания по уровню -3 дБ равна 63 кГц
- полезная мощность и КПД максимальны при сопротивлении нагрузки, близком к 20 Ом (расчетное – 19 Ом)
 - ограничение по амплитуде наступает при уровне входного сигнала 6 Вольт
 - паразитный набег фазы выходного сигнала за счет амплитудно-фазовой конверсии составляет около 40 градусов (неточности измерений связаны с использованием метода фигур Лиссажу вместо измерителя разности фаз).

Заключение

При анализе полученных результатов можно сделать вывод, что разработанный лабораторный стенд пригоден для постановки новой лабораторной работы, посвященной изучению явления амплитудно-фазовой конверсии.

Литература

1. *Varlamov O.* Research of influence of drm broadcast transmitter nonlinearities onto the output signal parameters // Т-Comm: Телекоммуникации и транспорт. 2014. Т. 8. № 2. С. 59-60.
2. *Иванюшкин Р.Ю., Юрьев О.А.* Способы построения передатчиков цифрового радиовещания диапазона ОВЧ // Системы синхронизации, формирования и обработки сигналов. 2016. Т. 7. № 1. С. 27-29.
3. *Иванюшкин Р.Ю., Юрьев О.А.* Перспективы применения ключевых усилителей мощности классов d и de при построении радиовещательных передатчиков диапазона ОВЧ // Т-Comm: Телекоммуникации и транспорт. 2016. Т. 10. № 5. С. 21-26.
4. *Козырев В.Б., Лаврушиенков В.Г., Леонов В.П., Новиков Г.В., Петяшин Н.Б., Попов И.А., Харитонов А.В., Громорушкин В.Н.* Транзисторные генераторы гармонических колебаний в ключевом режиме. Москва, 1985.
5. *Шмаков Н.Д., Иванюшкин Р.Ю.* Моделирование усилителя с распределенным усилением на полевых транзисторах BLF278 для диапазонов ВЧ и ОВЧ // Т-Comm: Телекоммуникации и транспорт. 2017. Т. 11. № 11. С. 34-39.
6. *Шмаков Н.Д., Иванюшкин Р.Ю.* Исследование твердотельных усилителей бегущей волны диапазона ОВЧ при двух подходах к их расчету // Системы синхронизации, формирования и обработки сигналов. 2017. Т. 8. № 3. С. 88-92.
7. *Варламов О.В., Пестряков А.В., Чугунов И.В.* Исследование усилителей мощности радиосигналов, построенных по методу раздельного усиления составляющих модулированных колебаний, на программно-аппаратном симуляторе // Системы синхронизации, формирования и обработки сигналов. 2014. Т. 5. № 2. С. 69-73.
8. *Хиленко В. И., Малахов Б. М.* Радиопередающие устройства. М.: Радио и связь, 1991. 328 с.
9. *Шахгильдян В.В., Козырев В.Б., Ляховкин А.А. и др.* Радиопередающие устройства; под ред. В.В. Шахгильдяна. 3-е изд., перераб. и доп. М.: Радио и связь, 2003. 560 с.
10. Технические параметры транзистора IRF610/Электронный ресурс, дата обращения 16.02.2019 <https://www.chipdip.ru/product/irf610>.

LABORATORY STAND FOR THE STUDY OF AMPLITUDE-PHASE CONVERSION IN THE RADIOFREQUENCY POWER AMPLIFIER FET

Severinenko Alexander Andreevich

student group MRT1701 MTUCI

kuju1ster@yandex.ru

Ivanyushkin Roman Yur'evich

MTUCI, c.i.s., associate Professor ROS

rivanyushkin@gmail.com

Keywords: *resonant power amplifier, field-effect transistor, amplitude-phase conversion, phase-amplitude characteristic, interelectrode capacitances.*

The development of a laboratory bench based on the resonant RF power amplification stage on the field-effect transistor, including the choice of amplifier circuit, connection of measuring instruments, layout, study of characteristics and Assembly of the laboratory bench. The results of the study of the main energy and quality characteristics of the developed laboratory bench are presented.

ИССЛЕДОВАНИЕ ГИБРИДНЫХ БАНКОВ ФИЛЬТРОВ В ЗАДАЧАХ ЦИФРОВОЙ ОБРАБОТКИ СВЕРХШИРОКОПОЛОСНЫХ СИГНАЛОВ

Платонов Олег Игоревич
студент группы МИТ1801 МТУСИ
lego-9000@yandex.ru

Лобов Евгений Михайлович
МТУСИ, к.т.н., доцент кафедры РТС
lobov@srd.mtuci.ru

Ключевые слова: гибридный банк фильтров, интерполяция, децимация, аналого-цифровой преобразователь, фильтры анализа, фильтры синтеза, сверхширокополосный сигнал, многоскоростная обработка сигнала.

Рассмотрены задача и основные принципы построения гибридных банков фильтров анализа-синтеза сигналов. Приведено сравнение подобных структур с классическими банками цифровых фильтров анализа-синтеза сигналов, показаны преимущества и недостатки гибридных структур при обработке сверхширокополосных сигналов. Описаны основные этапы обработки входного сверхширокополосного сигнала в таких устройствах, а также требования, предъявляемые к характеристикам отдельных элементов, таких как аналоговые фильтры анализа сигналов, аналого-цифровые преобразователи, цифровые фильтры синтеза.

Цифровые многоскоростные банки цифровых фильтров анализа синтеза одним из первых были описаны Крошье в работе [1] для цифровых методов сжатия речи. С течением времени, подобные системы получили широкое распространение, главным образом, для субполосного кодирования звука, изображений и видеоинформации. При разработке основ теории субполосного кодирования исходили из того, чтобы свести к минимуму эффекты наложения и искажения сигналов в системе. Речь шла о расчёте систем узкополосных фильтров, достигающих точного восстановления сигнала на выходе системы, от двухканальных ортогональных банков до многоканальных систем.

Банк фильтров состоит из фильтров анализа и синтеза, образуя систему анализа-синтеза (см. Рисунок 1). При анализе сигнал разделяется в частотной области на M составляющих, в каждой из которых, для уменьшения нагрузки на цифровую систему, понижается частота дискретизации с коэффициентом M при помощи дециматоров для последующей обработки, зависящей от области применения, а при выполнении синтеза составляющие интерполируются с коэффициентом M при помощи интерполяторов и объединяются для получения выходного сигнала.

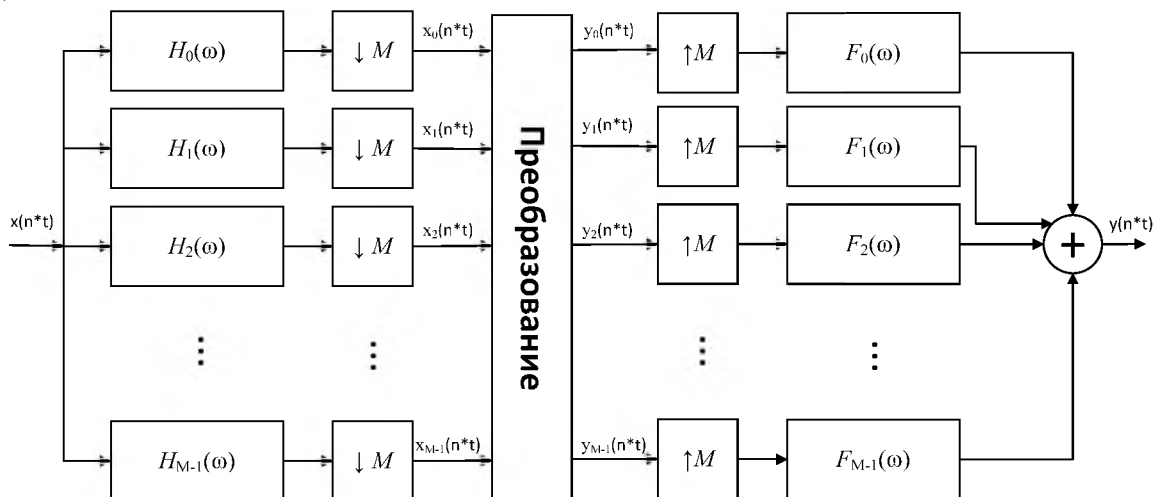


Рис. 1. Банк цифровых фильтров анализа-синтеза сигналов

Многоскоростная фильтрация, предполагающая возможность изменения частоты дискретизации входных цифровых сигналов в ходе их преобразования, является эффективным инструментом уменьшения требуемых вычислительных ресурсов и емкости памяти данных, используемых для реализации микропроцессорных систем.

В настоящее время многоскоростные системы обработки широкополосных сигналов широко применяются в научных исследованиях и технике и во многих сферах они являются единственным эффективным средством решения большого числа прикладных задач. Например, можно назвать такие области применения набора фильтров частотной селекции, как панорамные радиоприемники и анализаторы спектра, полосные вокодеры и системы скрытой связи, трансмультиплексоры и корректоры каналов [5, 6]. Также, банки цифровых фильтров лежат в основе алгоритмов сжатия и кодирования аудиосигналов и видеоизображений стандартов *MPEG-1*, *MPEG-2*, *MPEG-4*.

Гибридные банки фильтров — это такие банки, где фильтрами анализа являются аналоговые фильтры.

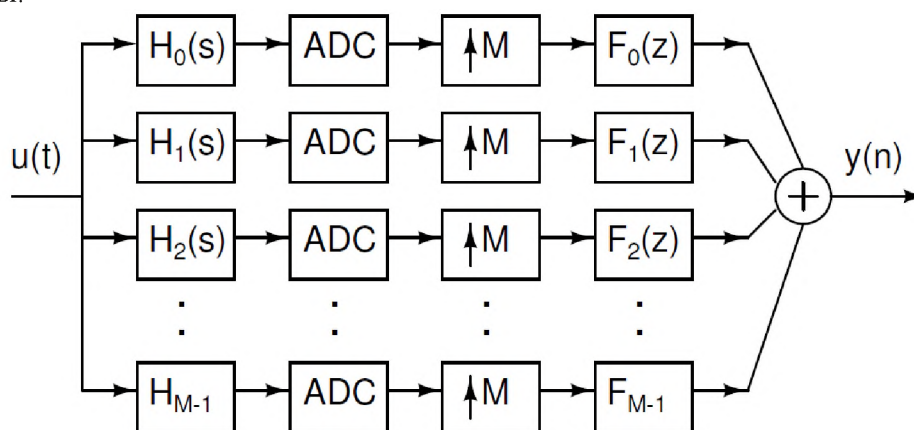


Рис. 2. Гибридный банк фильтров [2]

$H_1(s), H_2(s) \dots H_n(s)$ - аналоговые фильтры анализа

Высокая частота дискретизации и аналого-цифровые преобразователи (АЦП) с высоким разрешением требуются в различных устройствах. Прежде всего высокоскоростное аналого-цифровое преобразование сигнала необходимо в задачах обработки сверхширокополосных сигналов, а также в широкополосных осциллографах и анализаторах спектра. Необходимая частота дискретизации в таких устройствах составляет сотни мегагерц, f разрешение 10-12 бит, что создает определенные трудности при реализации соответствующих АЦП. Данное обстоятельство затрудняет применение классических банков цифровых фильтров анализа-синтеза, показанных в начале настоящей статьи для цифровой обработки сверхширокополосных сигналов. В реальных системах придется жертвовать либо тактовой частотой, либо глубиной дискретизации, что в зависимости от решаемых задач, может быть не приемлемо.

Чтобы решить вышеупомянутые проблемы исследуется подход, основанный на гибридной структуре банка фильтров. Для дискретизации входного сверхширокополосного сигнала в таких структурах вместо одного высокоскоростного АЦП используется множество АЦП с более низкой частотой дискретизации, а входные данные для каждого из этих преобразователей разделены в частотной области с помощью аналоговых фильтров. Обратный синтез же сигнала происходит в цифровой области, и ничем не отличается от синтеза сигналов в классических банках цифровых фильтров.

При проектировании фильтров анализа в таких системах необходимо достичь компромисса. Аналоговые фильтры анализа высокого порядка имеют резкий спад в полосе ослабления, однако высокий порядок фильтра приводит к трудностям в синтезе и конечной реализации, в следствии соответствующего большого числа отводов. Фильтры анализа малого порядка дают в результате АЧХ с медленным спадом, что требует меньшего количества отводов, однако в таком случае возрастает перекрытие АЧХ фильтров в соседних каналах, что увеличивает эффект наложения спектра, затрудняющий дальнейшую обработку сигнала.

Существуют различные методы синтеза необходимых аналоговых фильтров. Можно использовать классические методы синтеза, используя аппроксимации АЧХ Чебышева или Баттерворта. В статье [4] предложен метод, основанный на начальном синтезе цифровых фильтров прототипов анализа, и затем дальнейшем преобразовании их передаточной функции из Z плоскости в передаточную функцию в S плоскости (плоскость параметра операторных функций), получая на выходе передаточные функции аналоговых фильтров. Преимущества данного метода заключаются в упрощении синтеза набора оптимальных фильтров анализа, так как цифровые методы позволяют более гибко варьировать параметрами фильтров, при их синтезе, например, в системе *MatLab*. Однако при данном преобразовании передаточные функции прототипа и конечного аналогового фильтра имеют определенное расхождение, то есть возникает ошибка.

Главное преимущество гибридных банков фильтров – использование в M -раз менее скоростных аналогово-цифровых преобразователей, что позволяет эффективно решить проблему дискретизации сверхширокой полосы частот.

Недостатки гибридных банков фильтров:

- Сложность синтеза оптимальных аналоговых фильтров (большой порядок фильтров), так как структура любого банка фильтров накладывает ограничение на крутизну характеристики фильтров в полосе затухания.
- Фильтры высокого порядка также чувствительны к погрешности компонентов
- Трудности в подавлении эффекта наложения спектра.
- Сложность в создании большого числа каналов в системе, так как каждый аналоговый фильтр требует отдельного синтеза. Кроме того, увеличивается объем необходимой элементной базы.

Параметры АЦП, влияющие на характеристики системы

Ввиду того, что набор АЦП работает одновременно с частотным разнесением, накладываются определенные требования к уровню шумов АЦП в частотной области, так как они напрямую влияют на качество сформированного сигнала при синтезе. Вклады в шум АЦП в частотной области включают в себя гармоническое искажение, тепловой шум, и шум квантования.

Динамические характеристики АЦП определяются с использованием параметров, полученных с помощью спектрального анализа, и измеряются путем выполнения дискретного преобразования Фурье (ДПФ) для выходных отсчетов АЦП

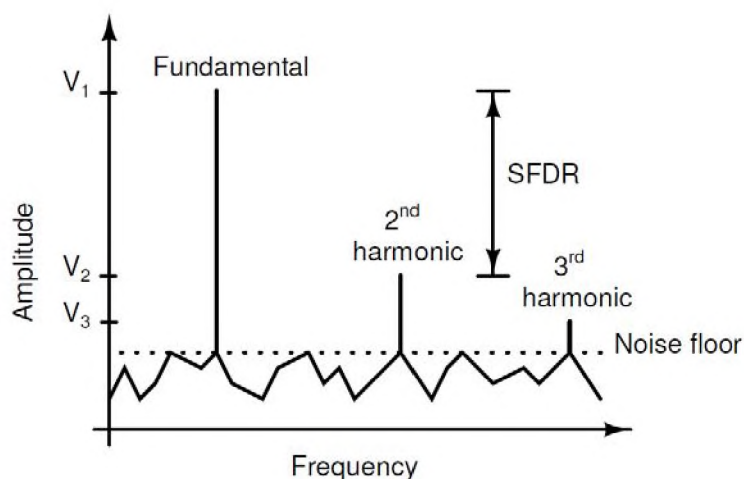


Рис. 3. Анализ работы АЦП в частотной области [2]

Качество работы АЦП можно оценить по динамическому диапазону в частотной области.

Динамический диапазон *SFDR* — это разница между величиной измеряемого сигнала и его самой высоким пиковым импульсом [4]. Как правило, это гармоника измеренного сигнала.

Идеальный АЦП имеет дискретную АЧХ с одной составляющей, равной его частоте дискретизации. Боковые всплески на АЧХ – следствие гармонических искажений, обусловленных нелинейностью компонентов АЦП. Влияния данного вида искажений на выходной сигнал АЦП

можно оценить при помощи параметра «полное гармоническое искажение» (*total harmonic distortion – THD* в зарубежной литературе), и его мощность рассчитывается с использованием формулы [2]:

$$THD = 10 \log \left(\frac{V_2^2 + V_3^2 + \dots + V_n^2}{V_1^2} \right) \quad (1)$$

V_1 – амплитуда полезной составляющей

V_2, V_3, \dots, V_m – амплитуды гармоник

Прочие искажения обусловлены прежде всего тепловыми шумами, шумами квантования и так далее. Их вклад можно оценить, рассчитав величину отношения сигнал / шум плюс искажение (*SNDR*).

Отношение сигнал / шум плюс искажение (*SNDR*) дает описание того, как измеренный сигнал будет соотноситься с шумом и гармониками. *SNDR* может быть вычислен при помощи [2]:

$$SNDR = 10 \log \left(\frac{V_1^2}{V_1^2 + V_1^2 + \dots + V_1^2 + V_u^2} \right) \quad (2)$$

V_u – суммарная мощность шумов

Данные параметры позволяют оценить вклад аналогово-цифровых преобразователей в общий внутренний шум системы гибридного банка фильтров.

Основные этапы преобразования сигнала в гибридном M-канальном банке фильтров

Предположим, что на вход гибридного банка фильтров поступает сигнал

$$X = U(j\omega) \quad (3)$$

Тогда после прохождения аналоговых фильтров анализа в каждом канале имеем:

$$X_k = U(j\omega) * U_k(j\omega) \quad (4)$$

После аналого-цифрового преобразования в каждом канале учитывая, что

$$\omega = \Omega * T_s \quad (5)$$

$$V_k(j\omega) = \sum_{m=-\infty}^{m=+\infty} U(j(\omega - 2\pi m) / T_s) H_k(j(\omega - 2\pi m) / T_s) \quad (6)$$

Затем после этапа интерполяции и фильтрации в каждом канале фильтрами синтеза:

$$W_k(j\omega) = \sum_{m=0}^{m=M-1} F_k(j\omega) U(j(M\omega - 2\pi m) / T_s) H_k(j(M\omega - 2\pi m) / T_s) \quad (7)$$

Наконец, на финальном этапе сигналы со всех поддиапазонов суммируются при помощи сумматора, в результате чего получаем:

$$Y(j\omega) = \sum_{k=0}^{k=M-1} \sum_{m=0}^{m=M-1} F_k(j\omega) U(j(M\omega - 2\pi m) / T_s) H_k(j(M\omega - 2\pi m) / T_s) \quad (8)$$

Выносим входной сигнал, и в результате получаем аналитическое представление сигнала на выходе M-канального гибридного банка фильтров:

$$Y(j\omega) = \sum_{m=0}^{m=M-1} U(j(M\omega - 2\pi m) / T_s) \sum_{k=0}^{k=M-1} H_k(j(M\omega - 2\pi m) / T_s) F_k(j\omega) \quad (9)$$

В идеальном случае выходной сигнал определяется как

$$Y(j\omega) = U(Mj\omega / T_s) \quad (10)$$

что является исходным входным сигналом, дискретизированным с периодом T_s

Компоненты $U(j\omega(M\omega - 2\pi) / T_s, U(j\omega(M\omega - 4\pi) / T_s) \dots U(j\omega(M\omega - (M-1)\pi) / T_s)$ являются внутренними шумами системы, вызванными эффектом наложения спектра.

Для достижения идеальной реконструкции сигнала значения данных компонент должны стремиться к нулю.

Компоненты:

$$T(j\omega) = \sum_{k=0}^{k=M-1} H_k(j(M\omega - 2\pi m)/T_s)F_k(j\omega) \quad (11)$$

также являются функциями наложения спектра. Соответственно, для получения идеальной реконструкции сигнала необходимо решить систему уравнений:

$$\sum_{k=0}^{k=M-1} H_k(j(M\omega - 2\pi m)/T_s)F_k(j\omega) = 0 \quad (12)$$

для k от нуля до M .

Решение данной системы уравнений позволит получить частотную характеристику фильтров синтеза $F_1 \dots F_k$ для предварительно определенного набора фильтров анализа H_k . В таком случае, при синтезе сигнала можно эффективно устранить эффект наложения спектра, возникающий при разделении входного сверхширокополосного сигнала на входе, возникающий вследствие неидеальности аналоговых фильтров анализа в полосе задерживания.

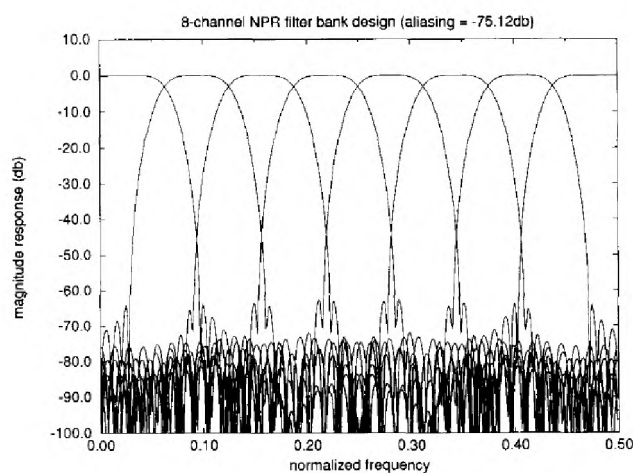


Рис. 4. Демонстрация эффекта наложения спектра [3]

Заключение

В настоящей статье были рассмотрены основные особенности построения гибридных банков фильтров анализа-синтеза сигналов. Можно сделать вывод, что гибридные банки фильтров являются весьма эффективным методом решения проблемы дискретизации сверхширокополосных сигналов. Однако реализация подобных структур требует особого внимания и учета множества параметров при синтезе и подборе всех элементов устройства, так как каждый из них вносит вклад в итоговый внутренний шум системы. Минимизация искажений является главной задачей, при проектировании гибридных банков фильтров.

Литература

1. Croisier A., Esteban D., Galaud C. Perfect channel splitting by use of interpolation/dccimation/tree decomposition techniques // Intern. Confer. Inform. Sci. Svst. Greece, 1976, pp. 443-446.
2. Rajesh Inti. Investigation of hybrid filter bankbased analog-to-digital conversion Department of electrical engineering indian institute of technology, Madras. nov 2007, 113 p.
3. Anand K. Soman, P.P. Vaidyanathan and Q. Nguyen. Linear Phase Paraunitary Filter Banks: Theory, Factorizations and Designs, IEEE transactions on signal processing, Vol. 41, No. 12, December 1993, pp. 3480-3496.
4. Scott Richard Velazquez. Hybrid Filter Banks for analog/digital conversion, Department of Electrical Engineering and Computer Science may 1997, 291 p.

5. Лобов Е.М., Смердова Е.О. Оценка импульсной характеристики широкополосного ионосферного канала с помощью метода обратной фильтрации // Т-Comm: Телекоммуникации и транспорт. 2017. Т. 11. № 3. С. 24-27.

6. Лобов Е.М., Косилов И.С., Кандауров Н.А., Елсуков Б.А. Методика оценки параметров частотной дисперсии ионосферного канала с помощью широкополосного фазоманипулированного сигнала // Т-Comm: Телекоммуникации и транспорт. 2014. Т. 8. № 9. С. 49-53.

RESEARCH OF HYBRID FILTER BANKS IN THE PROBLEMS OF DIGITAL PROCESSING OF ULTRA WIDEBAND SIGNALS

Oleg I. Platonov

Student of group MIT1801, MTUCI

lego-9000@yandex.ru

Evgenii M. Lobov

MTUCI, PhD., associate professor of RTS department

lobov@srd.mtuci.ru

Keywords: *hybrid filter bank, interpolation, decimation, analog-to-digital converter, analysis filters, synthesis filters, ultra wideband signal, multi-speed signal processing.*

Considered the task and the basic principles of constructing hybrid filter banks analysis-synthesis. Present a comparison of such structures with classical banks of digital filter analysis-synthesis, and shown the advantages and disadvantages of hybrid structures in the processing of ultra-wideband signals. Describes the main stages of processing the input ultra-wideband signal in such devices, as well as the requirements for the characteristics of individual elements, such as analog analysis filters, analog-to-digital converters, digital synthesis filters.

СОВРЕМЕННОЕ СОСТОЯНИЕ СИСТЕМ КОСМИЧЕСКОГО МОНИТОРИНГА ЗА ПОДВИЖНЫМИ ОБЪЕКТАМИ НА ТЕРРИТОРИИ РОССИИ

Ионов Владимир Викторович

магистрант МТУСИ

kudecar95@gmail.com

Пестряков Александр Валентинович

МТУСИ д.т.н., профессор кафедры РОС

a.v.pestryakov@mail.ru

Ключевые слова: мониторинг рыболовных судов, радиометки, Argos, Inmarsat, Iridium, Thuraya, Гонец-ДИМ.

Рассмотрена история развития систем космического мониторинга на территории Российской Федерации. Приведены причины отказа от системы «Argos». Описаны доступные на данный момент отечественные и иностранные системы, и проведен их сравнительный анализ.

Во всем мире и, в частности, в СССР началом бурного развития данных систем можно считать 70 – 80 гг. прошлого столетия. Тогда были созданы системы «КОСПАС-SARSAT» (КОСПАС – Космическая Система Поиска Аварийных Судов, SARSAT – Search And Rescue Satellite-Aided Tracking), «Argos» и «Inmarsat». Спустя десятилетие появляются «Iridium» и отечественная система «Гонец». В 1999 г. принимается постановление Правительства РФ №226 от 26 февраля 1999 г., в котором ответственным за мониторинг водных биологических ресурсов назначается Госкомрыболовство России. В этом документе предусматривалось использование систем спутникового мониторинга. В том же году выходит приказ Госкомрыболовства России № 330 от 22 ноября 1999 г. «О Временном положении о спутниковом позиционном контроле иностранных промысловых судов». Этот приказ содержит: общие положения, описывает процедуру регистрации судов в отраслевой системе мониторинга, включает перечень станций систем «Inmarsat» и «Argos», удовлетворяющих требованиям отрасли системы мониторинга, контактные данные и перечень обслуживаемых промысловых районов, форму заявки на тестирование и регистрацию технического средства контроля, форму регистрационной карточки рыболовного судна и акта соответствия оборудования. Это положило начало бурному развитию систем космического мониторинга на территории Российской Федерации.

Следующей значимой вехой развития систем космического мониторинга стал 2007 год. Были проведены работы по проектированию и разработке передающих платформ для экологических исследований, в том числе разработка радиометок проводилась и в МТУСИ (рис. 1). В последующие два года с помощью радиометок были произведены исследования миграции тюленей, белух, амурских тигров, а также ряда других представителей морской и наземной фауны (рис. 2) [1]. В некоторых из проектов принимал участие лично Президент РФ [2,3]. Данные радиометки работали по стандарту системы «Argos», которая до недавнего времени обслуживала рыболовецкий флот РФ. В конце 2015 года вышло постановление Росрыболовства об обязательной замене системы Argos [6] по приказу Министерства сельского хозяйства РФ [4] на другую зарубежную систему «Inmarsat», из-за того что «Argos» не получила разрешение Государственной комиссии по радиочастотам. В начале следующего года Росрыболовство привело ряд причин о незаконности применения французской системы [5]:

- отсутствие согласования Минкомсвязи и ФСБ России;
- противоречие Указу Президента о мерах по обеспечению информационной безопасности РФ при использовании сетей международного обмена;
- нарушение Правил использования спутниковых сетей, находящихся под юрисдикцией иностранных государств;
- нарушение федерального закона «О лицензировании отдельных видов деятельности»;
- нарушение федерального закона «О связи».

Также тогда появились первые упоминания о возможности использования отечественной системы. К концу года было принято окончательное решение о переходе рыболовецкого флота на систему «Inmarsat», и отечественную систему «Гонец-Д1М».



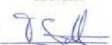

 ARGOS PLATFORM CERTIFICATION 	
Platform Transmitter Terminal A2	
SUBMITTED BY :	MTUCI
ADDRESS :	8a, ul. Aviamotornaya 111024 Moscow Russie
EQUIPMENT TYPE:	MTUCI
OPERATING ENVELOPE	
min. temperature	-20 °C
max. temperature	55 °C
temperature gradient	0,04 °C/min (2,5°C/hour)
CERTIFICATE	
CERTIFICATION : This PTT-A2 :	
is certified for the operating envelope defined above.	
PERFORMANCE : This P.T.T. A2 :	
meets the system performances for location purposes category A	
CNES Argos Project	CLS Argos System Operator
Date : 22/09/2006	Date : 22/09/2006
Signature : 	Signature : 
Page 1	

Рис. 1. Сертификат о соответствии радиобуя международному стандарту

На данный момент на официальном сайте ФГУП «Морсвязьспутник» предоставляются услуги, основанные на технологиях «Inmarsat», «Iridium» и «Thuraya». На официальном сайте спутниковой системы «Гонец-Д1М» также заявлен список абонентских терминалов, подходящих для установки на рыболовецкий флот. Помимо этих систем не прекратилась поддержка и системы «КОСПАС-SARSAT» ввиду ее значимой роли в сохранении жизни людей терпящих кораблекрушение.



Рис. 2. Радиометки для морских млекопитающих: тюлень (слева) и белуха (справа)

Каждая система состоит из трех сегментов:

- космический сегмент, состоящий из рабочих и запасных КА;
- наземный сегмент, включающий в себя технические средства приема, обработки и передачи информации – береговые земные станции, а также средства управления и контроля функционирования сети в целом;
- абонентские спутниковые станции и терминалы: мобильные (морской/речной, железнодорожный, автомобильный, воздушный и др. виды транспорта), носимые и стационарные спутниковые терминалы.

Космический сегмент у систем «Гонец-Д1М» и «Iridium» представлен низкоорбитальными спутниками в количестве - 12 и 66, соответственно. Столь малое число спутников у отечественной системы «Гонец-Д1М» обусловлено направленностью на отечественного пользователя (рис. 3).

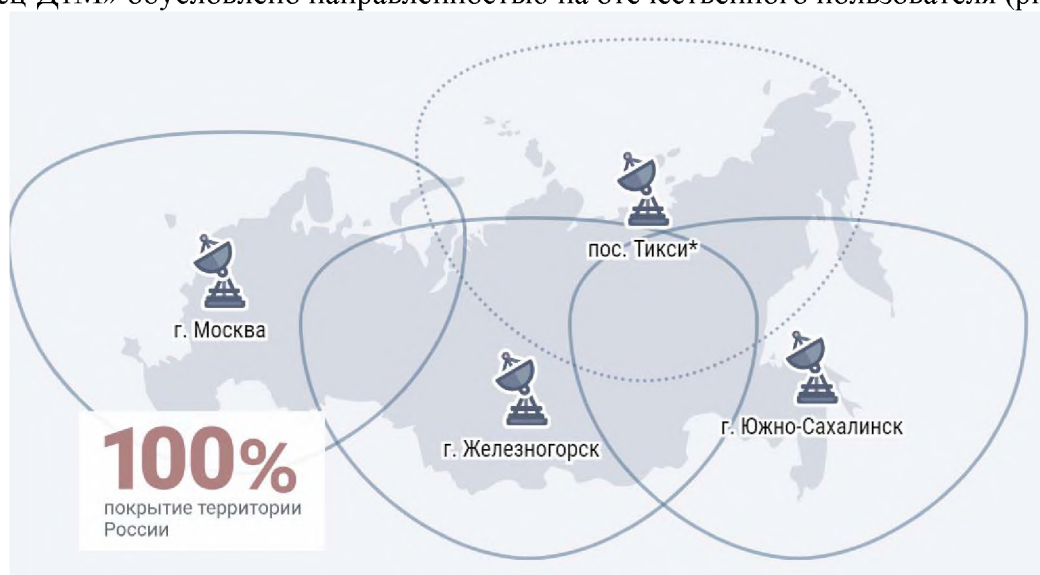


Рис. 3. Зоны радиовидимости каждой региональных станций системы «Гонец-Д1М»



Рис. 4. Зона покрытия ThuraYa

В системе «Iridium» реализован уникальный механизм межспутниковых связей, который используется для передачи сигнала с одного спутника на другой без необходимости ретрансляции этого сигнала на Землю.

Космический сегмент систем Inmarsat и ThuraYa, в отличие от вышеописанных, представлен геостационарными спутниками (в количестве – 14 и 3, соответственно). Использование геостационарной орбиты приводит к неполному покрытию территории РФ. Например, система «ThuraYa» не покрывает не только побережье и водные территории России, но и большую часть Дальнего Востока (рис. 4).

Система «Inmarsat» имеет большее покрытие, однако северные районы также находятся вне зоны покрытия. Эта проблема решается путем комбинирования передающей платформы данной системы и УКВ передатчика системы АИС (Автоматическая идентификационная система).

Исходя из всего вышесказанного, можно сделать вывод, что наиболее подходящей по критерию наибольшей зоны покрытия являются системы «Inmarsat» и «Iridium». Система «Thuraya» удобна для использования рыболовецким флотом только в Охотском и Японском морях. Отечественная система «Гонец-ДИМ» на данный момент находится в стадии доработки. Первоначально орбитальная группировка должна была состоять из двадцати четырех космических аппаратов, а не двенадцати. При двенадцати спутниках на орбите система способна обеспечить передачу данных в режиме реального времени только в северных широтах, и чем ближе к экватору, тем дольше перерывы в передаче сигнала. На экваторе они составляют 15 минут.

Стоимость отечественных абонентских терминалов в 1,5...2 раза меньше чем зарубежных (100...150 тыс. рублей против 200...300 тыс. руб.).

Таким образом, на данный момент на территории Российской Федерации наиболее подходящими системами для мониторинга рыболовецкого флота являются системы «Inmarsat» и «Iridium», которые обеспечивают наибольшее покрытие и качество связи, при соизмеримых тарифах и стоимости на судовые абонентские станции. Однако, для мониторинга за морскими млекопитающими данные системы практически непригодны из-за довольно сложного (и длительного во времени) протокола установления связи.

Литература

1. *Ионов В.В., Пестряков А.В.* Системы космического мониторинга подвижных объектов. XI Международная отраслевая научно-техническая конференция «Технологии информационного общества»: сборник трудов. 2017. М.: Издательский дом «Медиа Паблишер». С. 213-214.
2. Как я ловила белух для Путина // Комсомольская правда, 2009. 4 июля.
3. Рабочая поездка Владимира Путина в уссурийский заповедник [Электронный ресурс] Программа «Амурский тигр» <http://programmes.putin.kremlin.ru/tiger/premier> (Дата обращения 10.05.2019).
4. Проект Приказа Министерства сельского хозяйства РФ "Об утверждении Порядка оснащения судов техническими средствами контроля и их виды" (подготовлен Минсельхозом России 20.02.2016).
5. Замена «Аргоса» на отечественную спутниковую систему [Электронный ресурс] «Маринэк» <https://seacomm.ru/dokumentacija/9707/> (Дата обращения: 10.05.2019).
6. *Козлов А.В., Пестряков А.В.* Развитие спутниковой системы позиционирования и сбора данных ARGOS // Т-Сomm: Телекоммуникации и транспорт. 2012. Т. 6. № 2. С. 36-39.

CURRENT CONDITION OF SPACE MONITORING SYSTEMS FOR MOBILE OBJECTS IN THE TERRITORY OF RUSSIA

*Ionov V.V.,
master MTUCI,*

kudear95@gmail.com

Pestryakov A.V.,

Doctor of Engineering Sciences, MTUCI

a.v.pestryakov@mail.ru

Keywords: monitoring of fishing vessels, beacon, Argos, Inmarsat, Iridium, Thuraya, Gonetz-DIM.

This article describes the history of the development of space monitoring systems in the Russian Federation. The reasons for the rejection of the system «Argos». The currently available domestic and foreign systems are described, and their comparative analysis is carried out.

**ИССЛЕДОВАНИЕ ЧАСТОТНО-СЕЛЕКТИВНЫХ ХАРАКТЕРИСТИК АВИАЦИОННЫХ
БОРТОВЫХ РАДИОСТАНЦИЙ РАДИОТЕЛЕФОННОЙ СВЯЗИ УКВ ДИАПАЗОНА**

Мирошниченко Антон Валерьевич

студент группы М61801 МТУСИ

Mirosh.A.V@yandex.ru

Шаврин Сергей Сергеевич

МТУСИ, д.т.н., профессор кафедры МТС

sss@mtuci.ru

Ключевые слова: амплитудно-частотная характеристика, фильтр, бортовая радиостанция, радиотелефонная связь, VDL-4.

Для решения задачи электромагнитной совместимости линии передачи данных VDL-4 и авиационной радиотелефонной связи рассматривается вопрос частотно-селективных характеристик бортовых авиационных приемников радиотелефонной связи УКВ диапазона. Приводится математическая аппроксимация частотно-селективных свойств приемного тракта в виде БИХ фильтров, для построения математической модели авиационного бортового приемника радиотелефонной связи.

На современном этапе развития авиационной электросвязи наибольший интерес представляет новый стандарт передачи данных VDL-4, который позволяет выполнять множество различных функций: автоматическое зависимое наблюдение-вещание, передача сведений о погоде, осуществление зашифрованной передачи данных и т.д. Однако практическая реализация данной системы вызвала проблемы, которые заключаются в электромагнитной совместимости передающего оборудования VDL-4 с системами связи, уже работающими в общем частотном диапазоне. Наиболее негативный эффект VDL-4 оказал на систему авиационной радиотелефонной связи, работающей в том же частотном диапазоне на соседних каналах.

Для решения задачи электромагнитной совместимости двух систем связи необходимо знать их устройство, в частности, энергетические и частотно-селективные параметры. Если по системе VDL-4 информацию найти не составляет труда, то вопрос об устройстве авиационной радиотелефонной связи или, говоря по-другому, бортовых радиостанций, вызывает затруднение даже у специалистов в данной области. Проблема заключается в том, что авиационная техника «не имеет возраста» и при должном уходе может служить десятилетиями. В настоящее время в одном воздушном пространстве можно встретить как современный лайнер, только что вышедший с завода, так и старенький «кукурузник», возраст которого давно перевалил за полвека.

В данной работе приводится краткий обзор частотно-селективных возможностей авиационных радиостанций радиотелефонной связи в приложении к задачам электромагнитной совместимости.

Радиостанции, выпускаемые до 1945 года

Может возникнуть резонный вопрос: а зачем вообще рассматривать такую старую технику? Дело в том, что самолеты времен второй мировой войны до сих пор принимают участие в различных авиационных шоу и зачастую применяют штатный комплект радиоаппаратуры. Поэтому данная точка отсчета является отправной в представленном исследовании.

Следует напомнить, что под УКВ диапазоном в рамках настоящей статьи рассматривается авиационный диапазон 108 – 137 МГц, причем в системе VDL-4 он разбит более чем на 1000 каналов шириной в 25 кГц каждый.

Если анализировать аутентичные исторические источники [1], можно сделать вывод, что почти все радиостанции работают в диапазоне 2 – 10 МГц, только немецкие радиостанции могут

работать на частотах до 50 МГц. Одной из первых радиостанций УКВ диапазона была американская SCR-522, работающая на частотах 100 – 156 МГц. В дальнейшем радиостанция SCR-522 выпускалась в СССР под обозначением РСИУ-3 (а после модернизации – Р-800).

Все авиационные радиостанции, работающие на частотах от 2 до 40 МГц, обладают схожим строением, поэтому в качестве образца проведенного исследования принят всеволновой приемник типа УС-3. Он представляет собой семиламповый всеволновый супергетеродин.

Сигнал поступает на усилитель высокой частоты, затем попадает на смеситель и первый гетеродин, проходит два каскада усилителей промежуточной частоты, автоматический усилитель громкости, попадает на второй гетеродин, проходит диодный ламповый детектор и через выходной каскад поступает на телефон.

Средняя мощность передатчика могла колебаться в пределах 5-10 Вт, чувствительность приемника достигала 3 мкВ, что позволяло вести передачу на расстояние до 150 км. Все последующие радиостанции, включая современные, имеют схожие энергетические параметры.

Частотная селективность у всех приемников схожа и составляет двукратное (на 6 дБ) ослабление в полосе ± 15 кГц и стократное (на 40 дБ) - в полосе ± 60 кГц. Данные параметры позволяют провести аппроксимацию АЧХ реального фильтра фильтром Баттерворта третьего порядка, с передаточной характеристикой [2,3]:

$$H(j\omega) = \frac{-0.58 \cdot (15000 \cdot 2 \cdot \pi)^3}{(j\omega - 39282 - 68034j)(j\omega - 78552)(j\omega - 39282 + 68034j)} \quad (1)$$

Радиостанции 1945 – 1960 годов

Выше была упомянута американская радиостанция SCR-522, выпускавшаяся с 1944-х годов, однако наибольшую популярность на советском пространстве получил ее аналог РСУИ-4, который устанавливался на самолеты вплоть до конца 60-х годов. Поэтому данную радиостанцию следует отнести к послевоенному периоду.

Важным отличием радиостанции РСУИ-4 от аналогов являлся шаг несущих в 83.33 кГц у РСУИ-4 и 100 кГц у РСУИ-4П вместо стандартных 25 кГц [1]. Это накладывает значительные ограничения при использовании частотного плана с шагом частотной сетки 25 кГц.

Частотная селективность составляет: двукратное (на 6 дБ) ослабление в полосе ± 50 кГц и стократное (на 40 дБ) в полосе ± 210 кГц. Данные параметры позволяют провести аппроксимацию АЧХ реального фильтра, фильтром Баттерворта третьего порядка, с передаточной характеристикой [2]:

$$H(j\omega) = \frac{-0.58 \cdot (50000 \cdot 2 \cdot \pi)^3}{(j\omega - 130940 - 226780j)(j\omega - 261840)(j\omega - 130940 + 226780j)} \quad (2)$$

Радиостанции 1960 – 1990 годов

Данный период времени характерен бурным развитием авиастроения на территории всего земного шара. Именно в этот период бортовые радиостанции получают привычную схему построения и частотно - селективные свойства.

Радиостанции работают в привычной полосе частот 108 – 137 МГц, с шагом в 8.33 или 25 кГц. Прием выполняется по супергетеродинной схеме с однократным преобразованием частоты. В качестве промежуточной частоты может использоваться значение 20 МГц [4].

Важная селективная особенность этих радиостанций заключается в наличии двух выходов радиостанции: одного - на аппаратуру Селкол (аппаратура оповещения пилота о вызове его диспетчером при снятых наушниках), а второго - на телефон пилота. При этом ширина полосы пропускания в тракте приемника составляет 20 кГц, а в качестве модели его частотно-селективных свойств можно применять математический эквивалент кварцевого фильтра ФП2П4-447-01-В – фильтр Баттерворта 10 порядка.

АЧХ такого фильтра при работе в режиме с шагом 8.33 кГц описывается как:

$$H(j\omega) = \frac{0.5791 \cdot (8000 \cdot 2 \cdot \pi)^{10}}{(j\omega - 7439 - 47005j)(j\omega - 21608 - 42399j)(j\omega - 33653 + 33653j)} \cdot \frac{1}{(j\omega - 42399 + 21608j)(j\omega - 47005 + 7439j)(j\omega - 47005 - 7439j)} \cdot \frac{1}{(j\omega - 42399 - 21608j)(j\omega - 33653 + 33653j)(j\omega - 21608 - 42399j)(j\omega - 7439 - 47005j)}$$

АЧХ такого фильтра при работе в режиме с шагом 25 кГц описывается как:

$$H(j\omega) = \frac{0.5791 \cdot (16500 \cdot 2 \cdot \pi)^{10}}{(j\omega - 15350 - 96956j)(j\omega - 44567 - 87456j)(j\omega - 69410 + 69410j)} \cdot \frac{1}{(j\omega - 87456 + 44567j)(j\omega - 96956 + 15350j)(j\omega - 96956 - 15350j)} \cdot \frac{1}{(j\omega - 87456 - 44567j)(j\omega - 69410 + 69410j)(j\omega - 44567 - 87456j)(j\omega - 15350 - 96956j)}$$
 (3)

На выходе приемника к телефонным аппаратам пилотов (наушникам) установлен дополнительный НЧ фильтр со следующими параметрами: затухание 6 дБ на 3000 Гц, и 20 дБ на 3920 Гц. По аналогии можно предположить, что данный фильтр также является фильтром Баттерворта 7-го порядка со следующей АЧХ:

$$H(j\omega) = \frac{0.5793 \cdot (3000 \cdot 2 \cdot \pi)^7}{(j\omega - 3879 + 16999j)(j\omega - 10871 + 13632j)(j\omega - 15707 + 7564j)(j\omega - 17434)} \cdot \frac{1}{(j\omega - 15707 - 7564j)(j\omega - 10871 - 13632j)(j\omega - 3879 - 16999j)}$$
 (4)

На рисунке 1 приведен график АЧХ фильтра, установленного на телефонном выходе приемника (4).

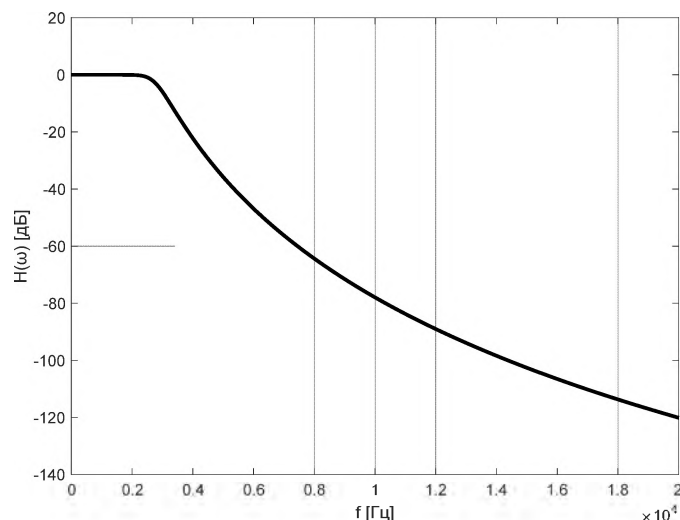


Рис. 1. АЧХ фильтра, установленного на телефонном выходе приемника

Современные радиостанции

На современном этапе развития авиации появилось огромное количество разнообразных радиостанций, преимущество выпускаемых в Китае. По своему строению они аналогичны радиостанции, разобранный в предыдущем разделе, но их частотно селективные свойства были несколько улучшены [5].

Таблица 1. Типовые параметры современного приемника в сравнении с приемником 80 годов.

Приемник	Затухание на 6 дБ, режим 8.33 кГц	Затухание на 60 дБ, режим 8.33 кГц	Затухание на 6 дБ, режим 25 кГц	Затухание на 60 дБ, режим 25 кГц
80-х годов	± 8 кГц	± 18.5 кГц	± 16.5 кГц	± 37 кГц
Современный	± 3.462 кГц	± 8.330 кГц	± 8 кГц	± 17 кГц

На рисунке 2 приведены графики АЧХ фильтров моделирующих приемные тракты радиостанций различных годов.

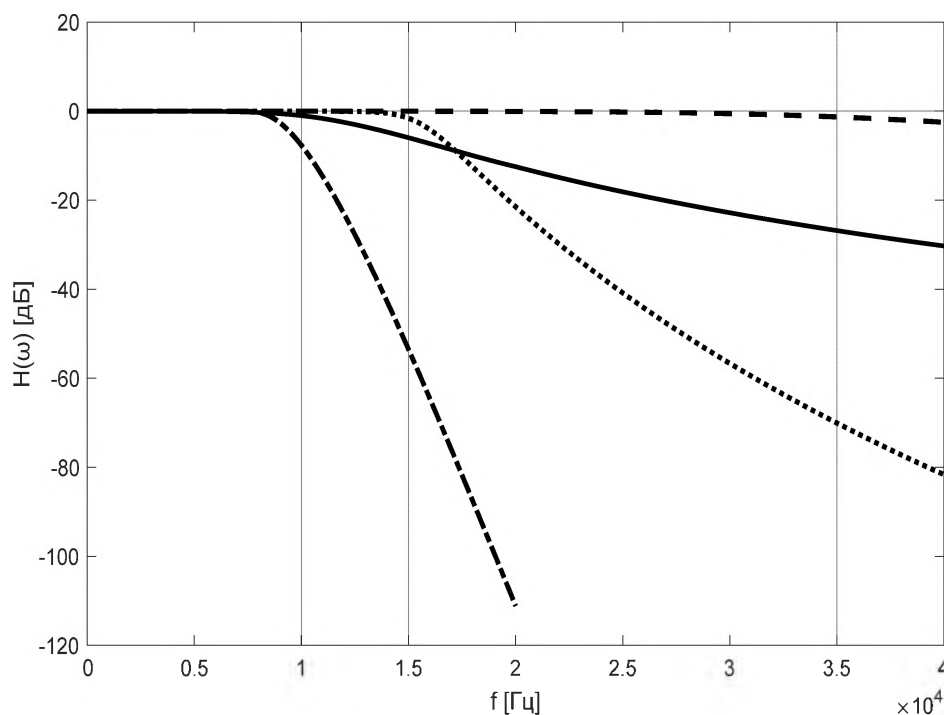


Рис. 2. АЧХ фильтров, моделирующих приемные тракты авиационных радиостанций: сплошная линия – радиостанции до 1945 года, штриховая – радиостанции 1945 – 1960х годов, пунктирная – радиостанции 1960 – 1990х годов, штрихпунктирная – современные радиостанции

Моделирование эффекта проникновения сигнала VDL-4 в тракт голосовой связи

Схема моделирования эффекта проникновения сигнала VDL-4 в тракт голосовой связи представлена на рисунке 3.

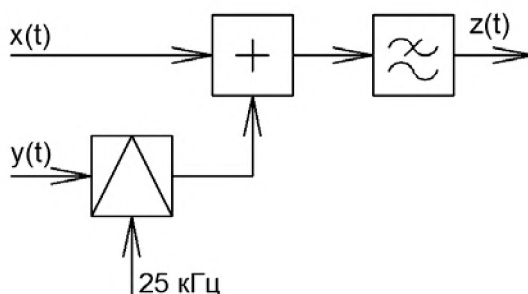


Рис. 3. Схема эксперимента

Она состоит из двух тестовых сигналов: $x(t)$ – записи сигнала речи летчиков и $y(t)$ – записи сигнала передачи данных режима VDL-4 (осциллограммы приведены на рисунке 4). Сигналы VDL-4 представляют из себя пакетную передачу с модуляцией GFSK. Более подробно с

характеристиками сигналов VDL-4 можно ознакомиться в [6 - 10]. Сигнал VDL-4 гетеродином переносится на соседний канал, отстающий на 25 кГц от голосового сигнала. Оба сигнала суммируются и подаются на фильтр с описанной выше АЧХ (в зависимости от приемника, моделирование которого проводится). После этого получается выходной сигнал, схожий с сигналом на выходе наушников пилотов. Результаты моделирования приведены на рисунке 5.

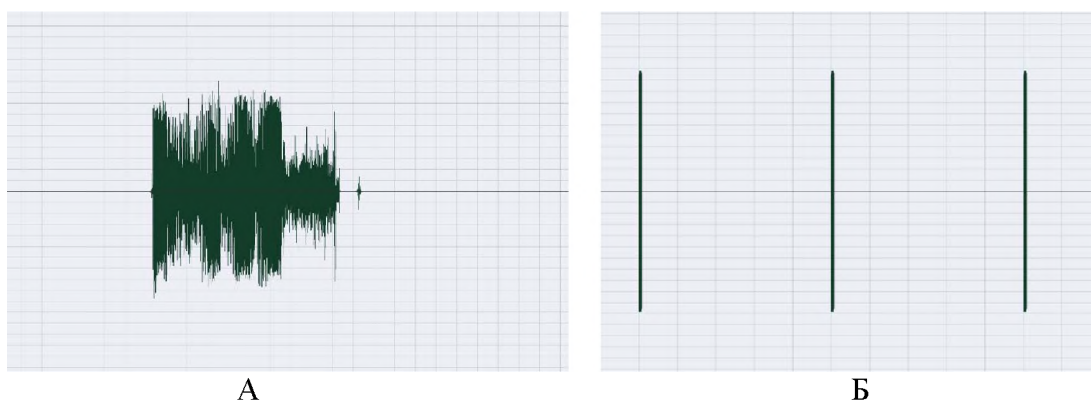


Рис. 4. Исходные сигналы: А – речи, Б – данных режима VDL-4

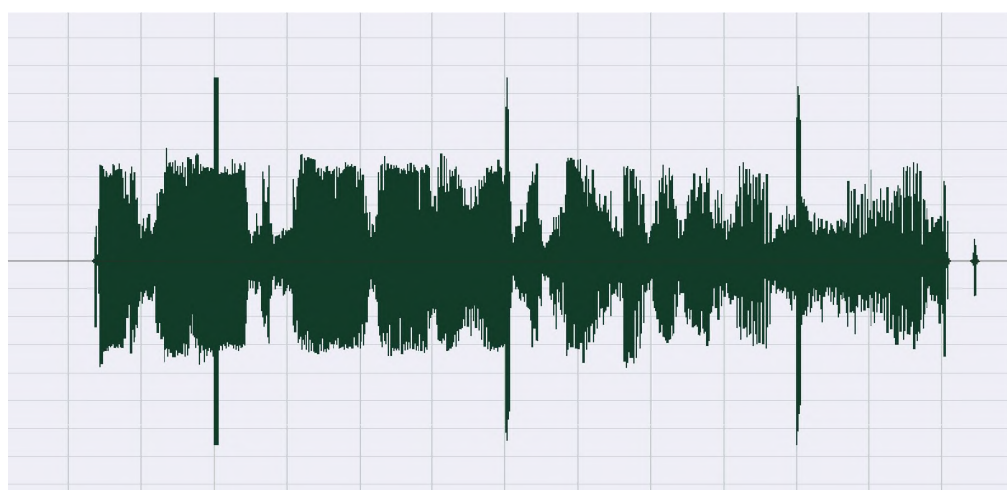


Рис. 5. Результаты моделирования для радиостанции 1960-1990х годов

Результаты моделирования показали неудовлетворительный результат для всех описанных радиостанций. Проникающий в тракт приема сигнал выражается в виде «тикания» отдаленно напоминающего звук часов. Причем для радиостанций с плохой селективной избирательностью (1, 2) в тракт приема попадает весь сигнал данных VDL-4, а в современных радиостанциях (4) только фронты нарастания и спада мощности сигнала VDL-4. На слух это выражается в изменении тональности помехи.

Заключение

1. Для моделирования приемного тракта бортового авиационного приемника можно применять фильтры Баттерворта 3го – 10го порядков, АЧХ которых были приведены в тексте данной статьи.
2. Моделирование показало, что селективных возможностей радиостанций радиотелефонной связи недостаточно. Сигнал передачи данных VDL-4 проникает в голосовой тракт, при работе в соседних каналах (на удалении в 25 кГц).
3. Необходимо применение дополнительных средств защиты, например, компенсаторов помех.

Литература

1. *Иванов В.И.* Справочник по радиоаппаратуре для инженеров и техников связи гражданского воздушного флота. Самолетное радиооборудование. Под. ред. Колодина И.Я. М.: Редакционно-издательский отдел Аэрофлота, 1948. 179 с.
2. *Мирошниченко А.В.* Обобщенный метод вывода канонических форм импульсных характеристик аналоговых фильтров // *Фундаментальные проблемы радиоэлектронного приборостроения*. 2017. Т. 17. № 4. С. 1194-1197.
3. *Мирошниченко А.В., Волчков В.П.* Три канонические формы представления импульсной характеристики аналогового фильтра Чебышева 1-го рода // *Телекоммуникации и информационные технологии*. 2017. № 1. С. 89-94.
4. Радиостанция Баклан. Руководство по технической эксплуатации. ИЖ1.101.017/018 РЭ, 1988 г. 98 с.
5. Радиостанция Орлан. Руководство по технической эксплуатации. ИЖ1.101.039-01.01 РЭ, 2005 г. 201 с.
6. Документ 9816, Руководство по цифровой линии передачи данных (ЛДП) частотного диапазона VHF, перевод с английского ГосНИИАС, часть 1, ИКАО, 2004. 245 с.
7. *Шкодин Д.А., Шаврин С.С.* Перспективы развития голосовой связи в гражданской авиации // *REDS: Телекоммуникационные устройства и системы*. 2016. Т. 6. № 2. С. 280-283.
8. *Клёсова Ю.В., Шаврин С.С.* Компьютерное моделирование метода распознавания речи для организации речевого канала в авиационных системах // *T-Comm: Телекоммуникации и транспорт*. 2016. Т. 10. № 5. С. 54-59.
9. *Григорьев И.Д., Орлов В.Г.* Анализ типов приложений в авиационной системе передачи данных // *Труды Северо-Кавказского филиала Московского технического университета связи и информатики*. 2016. № 1. С. 82-86.
10. *Григорьев И.Д., Орлов В.Г.* Исследование качества связи manet-сети на основе VDL-4 с использованием механизмов канального уровня // *Телекоммуникации и информационные технологии*. 2017. Т. 4. № 1. С. 105-110.

VHF RADIOTELEPHONE AVIATION ON-BOARD RADIO STATIONS FREQUENCY-SELECTIVE CHARACTERISTICS STUDY

Anton V. Miroshnichenko
MTUCI, Student
Mirosh.A.V@yandex.ru
Sergey S. Shavrin
MTUCI, Doctor of Science
[*s.s.shavrin@mtuci.ru*](mailto:s.s.shavrin@mtuci.ru)

Keywords: *amplitude-frequency response, filter, onboard radio station, VDL-4.*

Some on-board aviation radio-telephone receivers of the VHF band frequency characteristics aspects are being studied in order to estimate interband leakage when using VDL mode 4. A mathematical approximation of the frequency-selective properties of the receiving path in the form of IIR filters is proposed as a means to build an aircraft onboard radio telephone receiver mathematical model for electromagnetic compatibility experimental research.

АЛГОРИТМЫ ФУНКЦИОНИРОВАНИЯ МАРШРУТИЗАТОРА СИСТЕМЫ СЕТЕВОГО ТЕХНОЛОГИЧЕСКОГО УПРАВЛЕНИЯ ПРИ ПОСТРОЕНИИ ТАБЛИЦЫ МАРШРУТИЗАЦИИ

Арсеньева Диана Горановна
студентка группы БЗС1501 МТУСИ
ars.dian@yandex.ru

Сызранцев Геннадий Валентинович
МТУСИ, д.в.н., доцент, зав. кафедры СССнН
sysr9959@mail.ru

Ключевые слова: динамическая маршрутизация, система сетевого технологического управления, автоматические сети связи, таблица маршрутизации, алгоритм функционирования.

Рассматриваются варианты функционирования маршрутизатора системы сетевого технологического управления в зависимости от сетевого подключения мультимплексора. Представлены три варианта обработки служебной информации: при получении транзитной информации, информации, адресованной этому мультимплексору, и информации, полученной ошибочно. Сформулированы статические варианты функционирования маршрутизаторов системы сетевого технологического управления.

В настоящее время в условиях высокой интеллектуализации и глобализации промышленности и общественной жизни человечества, высоких темпов их развития имеется острая необходимость в наличии высокодинамичных систем управления, в том числе систем связи, в различных отраслях и направлениях деятельности, в том числе, при чрезвычайных ситуациях, техногенных и природных катастрофах, дестабилизирующих действиях незаконных формирований и др. Такая система связи должна удовлетворять предъявляемым требованиям по мобильности, устойчивости, доступности и управляемости [1]. Для построения высокодинамичных систем (сетей) связи используется телекоммуникационное оборудование, в котором реализована возможность сетевого технологического управления. В ряде научных публикаций [2–4] описаны варианты построения системы сетевого технологического управления, и обосновано, что наиболее рациональным вариантом построения является система на основе маршрутизатора системы технологического управления.

При этом первостепенной является задача маршрутизации информационных пакетов управления от сервера, проходящих через блоки контроля и управления шлюзовых сетевых элементов. Для решения этой задачи маршрутизатор имеет доступ по внутренней шине телекоммуникационного оборудования к любому каналному интервалу линейных интерфейсов Е1.

Маршрутизатор системы сетевого технологического управления – часть оборудования мультимплексора технологии NGPDH и построенной на их основе системы сетевого технологического управления. По этой причине, несмотря на то, что технически он реализован на канальном уровне, он также способен реализовывать функции сетевого уровня, то есть определять пути передачи данных и распознавать IP-адреса сетевых элементов. Данное упрощение также позволяет присваивать маршрутизатору адрес мультимплексора. Маршрутизатор скомутирован с блоком технологического управления оборудованием, что, в свою очередь, обеспечивает поступление данных от управляющей подсистемы (NMS). Это позволяет маршрутизатору заполнять и редактировать таблицу маршрутизации в процессе функционирования.

В зависимости от положения в структуре сети имеется возможность организации функционирования маршрутизатора по трём алгоритмам: при наличии связи с одним сетевым элементом, при наличии связей с двумя сетевыми элементами и при наличии связей с более чем двумя сетевыми элементами. В случае связи с одним или двумя сетевыми элементами для упрощения работы системы сетевого управления и снижения объемов передаваемой служебной

информации используются встроенные программы управления, при которых управляющий элемент системы не рассылает на сетевые элементы пакеты с действующей структурой сети.

Последовательность действий при взаимодействии встроенного маршрутизатора, блока управления и линейного блока мультиплексора технологии NGPDH представлена на рисунке 1. Первоначально определяется количество подключенных линейных элементов и количество IP-адресов. На основании опроса соседних сетевых элементов маршрутизатор принимает решение переводить ли режим работы в автоматический с использованием встроенной программы или же нет.

Если сетевой элемент подключен к более чем двум сетевым элементам и на него поступает транзитная информация, которую необходимо передать одному из подключенных к нему сетевым элементам, используется алгоритм действий, который не является программным. Для этого на сетевой элемент отправляется адресная база всех сетевых элементов, подключенных к данной сети. Когда приходит пакет от системы технологического управления, то маршрутизатор проводит оценку адресной части пакета и определяет сетевой элемент, которому предназначено сообщение. Используя данные таблицы маршрутизации, он отправляет его в сторону нужного сетевого элемента.

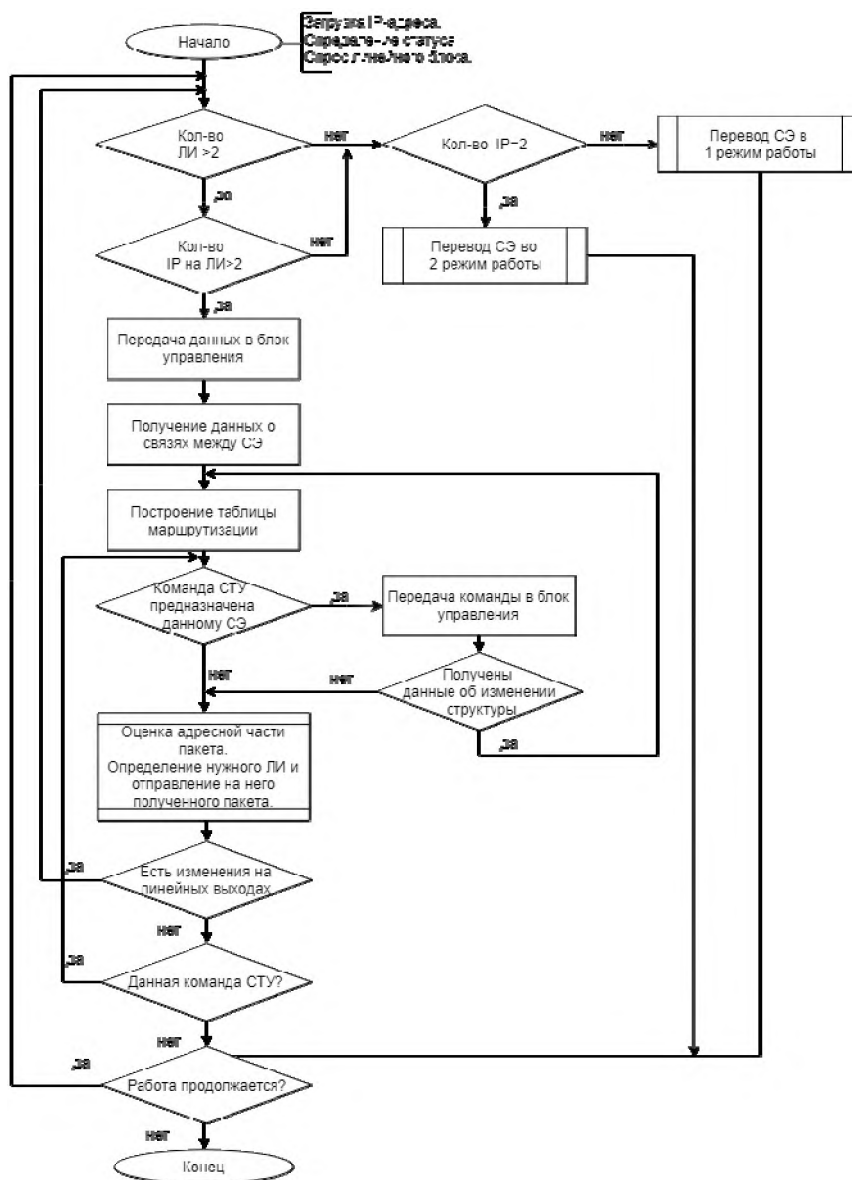


Рис. 1. Последовательность действий при взаимодействии встроенного маршрутизатора, блока управления и линейного блока мультиплексора технологии NGPDH, соответствующая режиму функционирования маршрутизатора № 1.

Первый алгоритм используется редко, так как, как правило, мультиплексоры в аппаратных узлах полевой системе связи подключены к одному или двум сетевым элементам.

Если сетевой элемент подключен к двум сетевым элементам, то используется встроенная программа мультиплексора, последовательность действий которой представлена на рисунке 2. Маршрутизатор проводит оценку линейных интерфейсов для выявления параллельных связей. Если он находит параллельные связи, то элементы ранжируются и группируются по отношению к сетевому элементу, стоящему на «дальнем конце». После чего маршрутизатор строит таблицу маршрутизации с учетом ранжирования и группировки элементов. При поступлении служебного пакета осуществляется оценка принадлежности линейного интерфейса к одной из групп. Затем оценивается адресная часть. Если сообщение адресовано данному сетевому элементу, то оно передается в блок управления, если же нет, то следующему линейному элементу, в соответствии с таблицей маршрутизации. Этот алгоритм назовём первым режимом функционирования маршрутизатора.

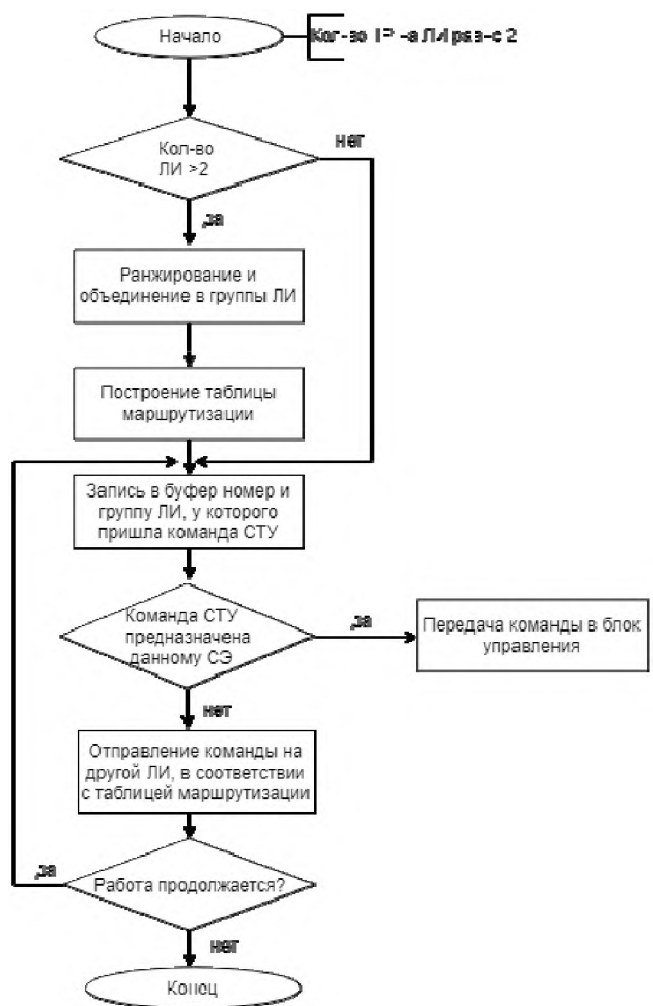


Рис. 2. Последовательность действий маршрутизатора по встроенной программе, соответствующей режиму функционирования маршрутизатора № 2

Если сетевой элемент подключен только к одному сетевому элементу, то используется встроенная программа мультиплексора, последовательность действий которой представлена на рисунке 3. В этом случае маршрутизатор не строит таблицу маршрутизации. При поступлении на сетевой элемент служебного пакета осуществляется оценка адресной части. Если она совпадает с адресом сетевого элемента, то пакет отправляется в блок технологического управления

оборудованием. Если же адреса не совпадают, то пакет удаляется маршрутизатором. Этот алгоритм назовём вторым режимом функционирования маршрутизатора.

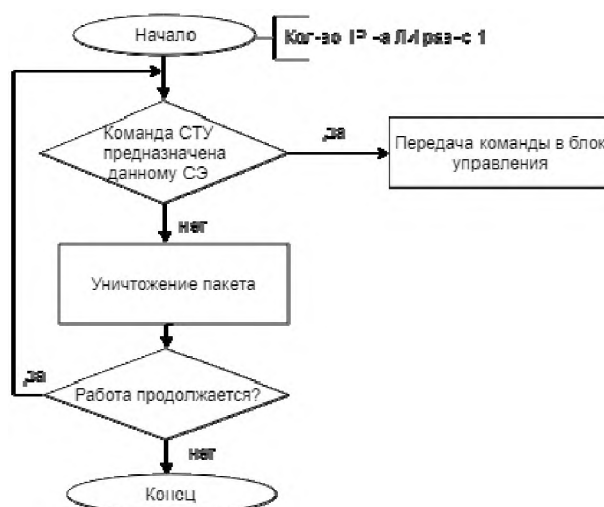


Рис. 3. Последовательность действий маршрутизатора по встроенной программе, соответствующей режиму функционирования маршрутизатора № 3

Заключение

Вышеописанные алгоритмы функционирования маршрутизатора системы сетевого технологического управления позволяют автоматизировать работу при передаче пакетов данных, что, в свою очередь, обеспечивает повышение значений показателей управляемости, доступности, мобильности и устойчивости функционирования полевой системы связи.

Алгоритмы работы выбираются в соответствии со схемой построения сети самим маршрутизатором, что упрощает работу системы сетевого технологического управления и оптимизирует объемы передаваемой служебной информации по каналам управления.

Литература

1. Сызранцев Г.В. Управляемость системы связи // Известия РАРАН. Журнал. М.: Изд. РАРАН. Вып. 2 (72), 2012. С. 81-86.
2. Сызранцев Г.В., Лукин К.И., Иншин Г.В. Модель функционирования автоматической первичной сети связи высокодинамичной системы связи специального назначения // Вопросы оборонной техники. Серия 16. Технические средства противодействия терроризму. М.: НТЦ
3. Сызранцев Г. В., Лукин К. И. Отечественные разработки для построения автоматических первичных сетей связи на технологии NGPDH // Научно-технический журнал «Системы и средства связи, телевидения и радиовещания». Вып. 1, 2. М.: Изд. ООО «ЭКОС». 2012. С. 54-61.
4. Сызранцев Г. В., Лукин К.И., Иншин Г.В., Даниленко А.Н. Модель функционирования автоматической сети связи общего пользования полевой системы связи специального назначения // Вопросы оборонной техники. Серия 16. Технические средства противодействия терроризму. М.: НТЦ «Информтехника» СПб.: Любавич, 2012. Вып. 1-2. С. 85-93.

ALGORITHMS FOR THE FUNCTIONING OF THE ROUTER OF THE NETWORK TECHNOLOGICAL MANAGEMENT SYSTEM WHEN BUILDING THE ROUTING TABLE

Arsenieva Diana Goranovna
Student of group BZS1501 MTUCI
ars.dian@yandex.ru

Syzrantsev Gennadiy Valentinovich
MTUCI, Doctor of Military Sciences,
Associate Professor, Head. Department of SPCS
(Special purpose communication systems)
sysr9959@mail.ru

Keywords: *dynamic routing, network technological management system, automatic communication network, routing table, algorithm of functioning.*

The article discusses the 0 versions for the operation of the network technological management system router depending on the network connection of the multiplexer. Three variants of service information processing are presented: when receiving transit information, information addressed to this multiplexer, and information received erroneously. Static variants of the operation of network technology management routers are formulated.

ПРОБЛЕМЫ ОБЕСПЕЧЕНИЯ НАДЕЖНОСТИ ПРОГРАММНО-КОНФИГУРИРУЕМЫХ СЕТЕЙ

Кусакина Маргарита Сергеевна
студентка группы М61701 МТУСИ
margaritakus@gmail.com

Нетес Виктор Александрович
МТУСИ, д.т.н., профессор кафедры ССисК
v.a.netes@mtuci.ru

Ключевые слова: программно-конфигурируемая сеть, контроллер, обеспечение надежности, аппаратная и структурная надежность, виртуализация, размещение контроллеров, резервирование.

Сформулированы основные задачи обеспечения надежности программно-конфигурируемых сетей (Software Defined Networking, SDN), выделены аппаратный и структурный аспекты надежности. Указана ключевая роль SDN-контроллеров в функционировании сети, рассмотрено влияние их виртуализации на надежность. Предложена модификация задач размещения контроллеров на сети, позволяющая осуществлять оптимизацию с точки зрения надежности. Приведены результаты расчетов надежности для типичного примера, демонстрирующие необходимость сетевого резервирования контроллеров и их связей с коммутаторами.

Одним из активно обсуждаемых и развиваемых в настоящее время направлений в мире телекоммуникаций являются программно-конфигурируемые сети (Software-Defined Networking, SDN). Ряд международных организаций ведут работу в области SDN. В частности, МСЭ-Т посвятил этой тематике серию Рекомендаций Y.33xx. SDN вместе с мобильными сетями пятого поколения (5G), облачными вычислениями, Интернетом вещей и большими данными входит в число направлений, относящихся к так называемым сетям пост-NGN.

Одной из важнейших характеристик сетей связи является надежность. По мере расширения и углубления использования телекоммуникационных систем и сетей растут и требования к их надежности. Поэтому необходимо предпринимать меры, чтобы надежность при использовании новых технических решений, в частности SDN, была бы, по крайней мере, не ниже, а еще лучше выше, чем при традиционном подходе. Важность этой темы обусловлена тем, что SDN должна стать одной из ключевых технологий при построении будущих сетей (Future Networks по терминологии МСЭ-Т), призванных создать инфраструктуру цифровой экономики. В частности, SDN будет одной из базовых технологий для мобильных сетей пятого поколения (5G).

Основная идея SDN – отделение функций передачи трафика от функций управления (как самим трафиком, так и осуществляющими его передачу устройствами) [1]. В традиционных коммутаторах и маршрутизаторах эти процессы реализованы в одном устройстве [6 -9], в то время как в SDN вся логика управления выносится в отдельные устройства, называемые контроллерами, которые способны отслеживать работу сети. Преимуществами SDN являются: возможность упрощения и удешевления сетевого оборудования, сокращение времени на развертывание новых услуг и приложений, снижение трудозатрат на сопровождение сети за счет централизации и автоматизации управления на программном контроллере, повышение эффективности использования ресурсов сети за счет динамического управления и быстрой реакции на увеличение нагрузки.

Как отмечается в [1], в крупномасштабных сетях, надежность является особенно важным фактором, однако характерная особенность логически централизованного управления в SDN состоит в том, что SDN-контроллер имеет тенденцию становиться единой точкой отказа, т.е.

уязвимым местом сети. Поэтому необходимо принятие специальных мер для предотвращения возможности возникновения подобных неблагоприятных ситуаций.

В данной статье анализируются основные аспекты обеспечения надежности SDN, рассматриваются вопросы размещения контроллеров на сети и влияние их виртуализации на надежность, обосновывается необходимость резервирования контроллеров и их связей с управляемыми сетевыми элементами.

Надежность, ее показатели и аспекты

Надежность определяется, как свойство технического объекта сохранять во времени способность выполнять требуемые функции. Она включает в себя целый ряд более простых свойств, из которых центральное место занимает готовность. Основным показателем, количественно характеризующим это свойство, является коэффициент готовности. Он определяется как вероятность того, что рассматриваемый технический объект окажется в работоспособном состоянии в данный момент времени. Именно коэффициент готовности обычно используется при задании требований к надежности сетей связи и их элементов.

В настоящее время общепринятой нормой для коэффициента готовности оборудования связи операторского класса является величина 0,99999 («пять девяток»). Это соответствует среднему времени простоя 5,256 минуты в год. В самом деле, продолжительность года составляет $8760 \text{ ч} = 525600 \text{ мин}$, умножая это число на соответствующую величину коэффициента неготовности $1 - 0,99999 = 10^{-5}$, получим указанное значение.

Традиционно для сетей связи выделяют два аспекта надежности: аппаратный и структурный. Аппаратный рассматривает надежность аппаратуры, отдельных устройств и элементов, входящих в состав узлов и линий сети. Структурный аспект отражает функционирование сети в целом в зависимости от работоспособности или отказов ее элементов (узлов и линий). Рассмотрим эти аспекты применительно к SDN.

Говоря об аппаратной надежности, отметим, что в SDN появляется новый и весьма важный вид аппаратуры – контроллеры. Они выполняют функции управления сетевыми элементами и пакетами данных в сети, поэтому надежность всей сети существенно зависит от надежности контроллеров. Обеспечение надежности контроллеров – типичная задача аппаратной надежности, которой должно быть уделено самое серьезное внимание.

С другой стороны, переход к SDN, с одной стороны, позволяет использовать более простое, а значит и более надежное (при прочих равных условиях) сетевое оборудование. Это обусловлено тем, что его «интеллект» выносится в контроллеры.

К структурному аспекту надежности SDN относятся задачи, связанные с размещением контроллеров и распределением между ними управляемых сетевых элементов, организацией резервирования и взаимодействия между контроллерами. Они будут рассмотрены ниже.

Виртуализация контроллеров

На стыке аппаратного и структурного аспектов надежности лежит выбор способа организации контроллера, для чего возможны два подхода. Во-первых, это может быть автономное оборудование, размещенное в узле связи; во-вторых, возможно виртуализировать функциональность управления коммутаторами, поместив контроллер «в облако». Во втором случае получается симбиоз технологий SDN и NFV (Network Functions Virtualization – виртуализация сетевых функций). При этом SDN-контроллер организуется не в виде специального оборудования, а программно, обычно в центре обработки данных (ЦОД). Такое совместное использование SDN и NFV сулит значительные преимущества и считается весьма перспективным, хотя изначально эти две технологии создавались и развивались независимо друг от друга. Здесь они дополняют друг друга: NFV обеспечивает SDN той платформой, на базе которой работает соответствующее программное обеспечение.

Рассмотрим достоинства и недостатки каждого из способов организации контроллера с точки зрения надежности. Если контроллер реализован традиционным образом и размещен в узле связи, он, как правило, ближе к сетевым элементам, а значит, обеспечивается большая надежность связи с ними, но надежность самого такого контроллера может быть ниже. Если же он реализован в ЦОД, может быть гораздо легче обеспечить его высокую надежность, однако он будет более

удален от сетевых элементов, следствием чего является меньшая надежность связи между контроллером и управляемыми им сетевыми элементами.

Таким образом, здесь необходимы расчеты и сравнительный анализ вариантов для конкретных ситуаций, общая методология которого представлена в [2]. Вообще, подобная ситуация имеет место и для других случаев применения виртуализации и облачных вычислений [2], широко используемых в современных инфокоммуникациях. Общие рекомендации по обеспечению надежности здесь таковы:

- использование ЦОД с высоким уровнем отказоустойчивости инженерной инфраструктуры (не ниже Tier III по классификации Uptime Institute) и высоким уровнем готовности вычислительных ресурсов;
- резервирование ЦОД с их географическим разнесением;
- резервирование каналов передачи данных между клиентским оборудованием (в случае SDN – коммутаторов) и ЦОД.

Размещение контроллеров

Вопросы размещения контроллеров исследовались в целом ряде работ, в частности, в [3, 4] и др. В них рассматривались такие задачи, как нахождение минимального числа контроллеров в сети, выбор мест размещения контроллеров среди узлов-кандидатов, распределение управляемых сетевых элементов между контроллерами, переназначение контроллеров в случае их отказов с минимальным ухудшением качества. При этом учитываются ограничения на задержки передачи пакетов как между контроллером и оборудованием, так и между контроллерами.

При решении задачи размещения контроллеров с учетом задержек передачи пакетов применяются две целевые функции, которые стремятся минимизировать: минисуммная, дающая среднюю (average) задержку, и минимаксная, дающая задержку в наихудшем случае (worst case). Они определяются следующими формулами:

$$L_{avg}(S) = \frac{1}{n} \sum_{v \in V} \min_{s \in S} d(v, s), \tag{1}$$

$$L_{wc}(S) = \max_{v \in V} \min_{s \in S} d(v, s), \tag{2}$$

где S – множество узлов, в которых размещаются контроллеры; V – множество узлов, в которых размещается сетевое оборудование; n – число узлов во множестве V ; d – задержка между рассматриваемыми узлами.

Задержка между парой узлов определяется как сумма задержек во всех линиях, составляющих кратчайший (в смысле задержки) путь между этими узлами:

$$d(v, s) = \sum_{(i,j) \in A} d_{i,j}, \tag{3}$$

где $d_{i,j}$ – задержка в линии между узлами i и j ; A – множество линий, составляющих кратчайший путь между узлами v и s .

В такой постановке задачи размещения контроллеров с целевыми функциями (1) и (2) сводятся к известным в теории графов задачам нахождения медиан и центров графа соответственно. Эти задачи используются во многих приложениях, когда нужно оптимальным образом разместить в сети некоторые пункты обслуживания. В первом случае требуется минимизировать сумму задержек или расстояний от такого пункта до всех обслуживаемых им узлов (минисуммная задача о медианах), во втором – минимизировать задержку или расстояние до наиболее удаленного обслуживаемого узла (минимаксная задача о центрах).

Эти результаты могут быть применены и для выбора вариантов размещения контроллеров, оптимизирующих надежность их связей с узлами сети, в которых находится управляемое ими оборудование. В самом деле, вероятность работоспособности или коэффициент готовности некоторого пути в сети равен произведению вероятностей работоспособности составляющих его линий.

$$p(A) = \prod_{(i,j) \in A} p_{ij} \quad (4)$$

где p_{ij} – вероятность работоспособности линии между узлами i и j . Взяв логарифм от обеих частей (4), получим:

$$\log p(A) = \sum_{(i,j) \in A} \log p_{ij}. \quad (5)$$

Обозначим $c_{ij} = -\log p_{ij}$. Из равенства (5) вытекает, что величины c_{ij} для пути складываются, т.е. для них имеет место аддитивность, аналогичная равенству (3). Поэтому, решая любую задачу, связанную с нахождением кратчайших путей в графе (в частности, задачи размещения медиан или центров) с весовой матрицей дуг $\|c_{ij}\|$, мы получим решение, дающее наиболее надежные пути.

Резервирование контроллеров и связей с ними

Резервирование является одним из важнейших путей обеспечения надежности. Для того чтобы SDN-контроллер не становился единой точкой отказа, в [1] рекомендуется резервировать контроллеры.

Для количественного анализа был рассмотрен пример гипотетической сети на Европейской части России, состоящей из 11 узлов, расположенных в четырех крупных городах. В узлах размещены коммутаторы, которые отвечают за передачу данных от отправителя к получателю. Трафик в городах передается через сети доступа с подключением к двум, трем или четырем узлам в городе.

Рассматривались три варианта размещения и подключения контроллеров в этой сети:

- 1) имеется один контроллер, подключенный к одному из узлов в своем городе (в Москве), резервируются только связи этого узла с узлами в других городах;
- 2) имеется один контроллер, подключенный к двум узлам в своем городе, резервируются связи этих узлов с узлами в других городах;
- 3) имеются два контроллера, расположенные в разных городах (в Москве и Санкт-Петербурге), каждый из них подключен к одному из узлов в своем городе, резервируются связи этих узлов с узлами в других городах.

Для нормальной работы коммутатора, размещенного в узле сети, необходима не только работоспособность его самого, но также работоспособность управляющего им контроллера и наличие связи с ним. Поэтому необходимо стремиться к тому, чтобы вероятность указанного события была бы не меньше, чем коэффициент готовности традиционного оборудования, т.е. 0,99999. Эти вероятности были рассчитаны для наихудшего случая, т.е. для наиболее удаленного от контроллеров узла сети.

Результаты расчетов приведены ниже (Табл. 1). Они показывают, что только в третьем варианте, т.е. при организации сети с двумя контроллерами в разных городах и резервированием путей от каждого узла, можно получить коэффициент готовности, близкий к требуемому значению «пять девяток». Более подробно данный пример описан в [5].

Таблица 1. Результаты расчетов надежности

Вариант размещения и подключения контроллеров	Коэффициент готовности	Коэффициент неготовности	Среднее время простоя, мин/год
1. Один контроллер подключен к одному узлу	0,999965	$3,5 \cdot 10^{-2}$	18,40
2. Один контроллер подключен к двум узлам в одном городе	0,999979	$2,1 \cdot 10^{-2}$	11,04
3. Два контроллера подключены к двум узлам в разных городах	0,999989	$1,1 \cdot 10^{-2}$	5,78

Заключение

Таким образом, обеспечение надежности является важнейшей задачей при создании и эксплуатации программно-конфигурируемых сетей (SDN), поскольку в характерной для них логически централизованной архитектуре управления SDN-контроллер может становиться уязвимым местом сети. При этом необходимо обеспечить как высокую надежность самих контроллеров, так и выбрать рациональные варианты их размещения на сети и организации связи с ними всех узлов. Задачи оптимального размещения контроллеров в сети, ставящие целью максимизировать надежность связей контроллеров со всеми узлами (в среднем или для наихудшего случая), могут быть сведены к известным в теории графов задачам нахождения медиан и центров графа соответственно. Расчеты надежности для различных вариантов размещения и подключения контроллеров на примере типичной сети демонстрируют необходимость сетевого резервирования для контроллеров и их связей с коммутаторами.

Литература

1. ITU-T Recommendation Y.3300 (06/2014). Framework of software-defined networking.
2. *Нетес В.А.* Виртуализация, облачные услуги и надёжность // Вестник связи. 2016. № 8. С. 7–9.
3. *Perrot N., Reynaud T.* Optimal placement of controllers in a resilient SDN Architecture // 12th International Conference on the Design of Reliable Communication Networks (DRCN 2016). Paris, France. March 2016. P. 145–151.
4. *Heller B., Sherwood R., McKeown N.* The controller placement problem // SIGCOM HotSGN. Stanford University, 2012. [Электронный ресурс]. – URL: <http://www.sharecourse.net/sharecourse/upload/quiz/13/21/909838c7e75f081c3ac73dcb0545cf13.pdf>. (Дата обращения: 05.05.19).
5. *Нетес В.А., Кусакина М.С.* Надёжность связи контроллеров с коммутаторами в SDN // Вестник связи. 2018. № 9. С. 10–12.
6. *Нетес В.А.* Мониторинг параметров работы сетей и временная синхронизация // Т-Comm: Телекоммуникации и транспорт. 2014. Т. 8. № 2. С. 36-37.
7. *Сулейманов А.А., Нетес В.А.* Анализ времени подключения к облачной услуге "виртуальный рабочий стол" // Т-Comm: Телекоммуникации и транспорт. 2016. Т. 10. № 7. С. 41-46.
8. *Нетес В.А.* Двусторонние оценки для показателей качества обслуживания с учетом надежности обслуживаемых приборов // Т-Comm: Телекоммуникации и транспорт. 2018. Т. 12. № 8. С. 75-77.
9. *Нетес В.А.* Преподавание теории надежности студентам инфокоммуникационных направлений // Методические вопросы преподавания инфокоммуникаций в высшей школе. 2018. Т. 7. № 2. С. 36-39.

THE PROBLEMS OF RELIABILITY ASSURANCE FOR SOFTWARE DEFINED NETWORKS

Margarita S. Kusakina

Student of group M61701, MTUCI

margaritakus@gmail.com

Victor A. Netes

MTUCI, DSc, professor of CN&SS department

v.a.netes@mtuci.ru

Keywords: *Software Defined Networking, controller, reliability assurance, equipment and system reliability, virtualization, placement of controllers, redundancy.*

The main problems of reliability assurance for Software Defined Networks (SDN) are formulated; equipment and system aspects of reliability are identified. The key role of SDN-controllers in the functioning of the network is specified, the influence of their virtualization on reliability is considered. The modification of the controller placement problem, which allows carrying out optimization from the reliability point of view, is proposed. The results of reliability calculations for a typical example demonstrate the necessity for network redundancy of controllers and their connections with switches.

ПОГРЕШНОСТЬ И НЕОПРЕДЕЛЕННОСТЬ В МЕТРОЛОГИИ

Коротков Александр Валерьевич

студент группы М61801 МТУСИ

kor-avk@ya.ru

Сенявский Александр Леонидович

МТУСИ, к.т.н., профессор кафедры МСиИИ

alsenyavskiy@yandex.ru

Ключевые слова: неопределенность измерений, погрешность измерений, оценка результатов измерений, истинное значение, точность результатов измерений.

В СССР и в России для оценки точности измерений используют анализ погрешностей. В зарубежных странах, последние 20 лет для оценки погрешности измерений используют понятие неопределенность измерений. Такое различие подходов приводит к неоднозначности, например, при сертификации зарубежной продукции. В статье основываясь на существующих нормативных документах рассматриваются особенности этих подходов.

Мир всегда стремится к единству и для этого прикладывает усилия к созданию единого информационного и технического пространства в образовании, технике, науке. Исторически сформировалось так, что в России использовали понятие «погрешность измерений», а за границей применяли понятие «ошибка измерений» (error of measurement, англ. яз.), для оценки достоверности произведенных измерений. Для обеспечения единства в сфере оценки точности результатов измерений было разработано в составе ISO 9000 «Руководство по вычислению неопределенности в измерении», который содержит новую концепцию описания результатов измерений. Это помогло создать новый международный «стандарт», который стал называться неопределенностью измерений. Все технические лаборатории многих стран мира обязаны оценивать точность результатов измерений через понятие неопределенность. В настоящий момент Россия использует обе оценки точности измерений. Данные методы оценки, основываются на главных понятиях классической метрологии, теории ошибок и математической статистики.

Погрешность измерений и оценка результатов измерения

Под понятием погрешность измерений понимают отклонение измеренного значения величины от ее «истинного» значения [3]. Нужно учесть, что такое первичное истинное значение неизвестно.

В процессе измерений обозначают интервал, имеющий в себе это «истинное значение» $X = A \pm \Delta$ (при доверительной вероятности равной 0,95). Интервал от $(A - \Delta)$ до $(A + \Delta)$, как показано на рис. 1, состоит из: истинного значения измеряемой величины (т.е. A) и погрешности измерений.

В соответствии с формой представления погрешности бывают:

- Абсолютная;
- Относительная;
- Приведенная.

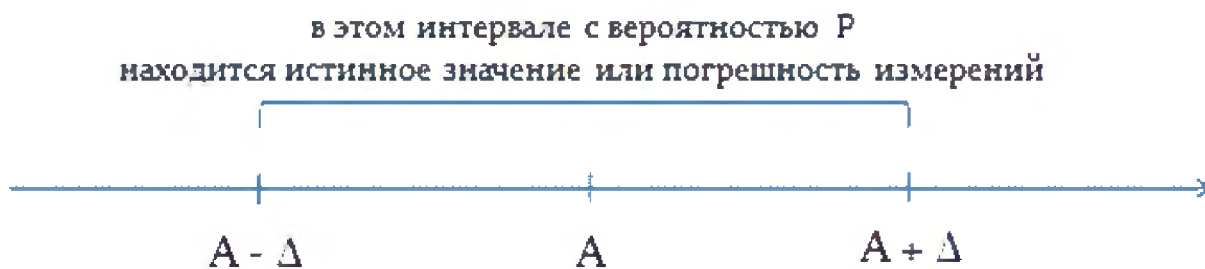


Рис. 1. Погрешность измерений

Неопределенность измерений и оценка результатов измерений

Неопределенность измерения – это параметр, который связан с результатом измерений и характеризуется с разбросом значений, которые могли бы быть логично дописаны измеряемой нами величины. Неопределенность в сравнение с погрешностью не относится к истинному значению, она присвоена именно к измеренному значению величины. Результат измерений описывается выражением: $X = A \pm U$, при вероятности охвата (P) равной 0,95 [1].

Интервал от $(A - U)$ до $(A + U)$, как показано на рис. 2, включает в себя большую часть значений вероятности охвата, которые могли бы быть присвоены к измеряемой величине.

При оценке точности измерений в понятие «неопределенность» полагают, что вычисленная нами величина относится к указанному интервалу значений, при условии если она не выходит за границы данного диапазона значений с учетом выявленной нами неопределенности ($X = A \pm U$).

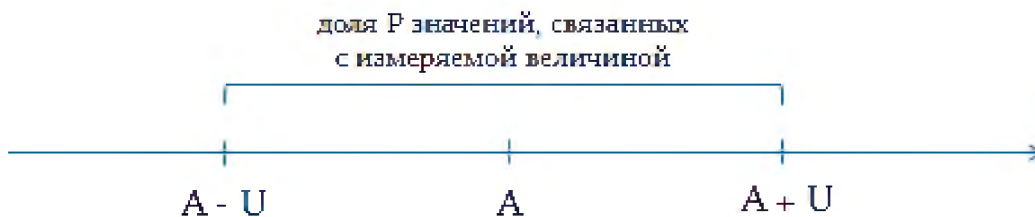


Рис. 2. Неопределенность измерений

Соотношение понятий

В табл.1 проведено соотношение классической теории (погрешности) и теории неопределенности [2], что наглядно показывается их схожесть.

Таблица 1. Соотношение классической теории (погрешности) и теории неопределенности

Классическая теория	Теория неопределенности	Пояснение
Неопределенность результата измерения	Погрешность результата измерения	Наименование получаемого результата
Неопределенность типа А	Случайная погрешность	Входящая в состав погрешности (неопределенности) измерения, которая изменяется случайным способом в ряде полученных измерений и не имеющая систематического характера. Производят оценку статистическими способами, (точнее имеется вероятность обработать и получить большое количество измерений одного значения величины)

Неопределенность типа Б	Систематическая погрешность	Часть неопределенности измерений, которая может быть постоянной или изменяющейся по общеизвестному закону (пример калибровки)
Стандартная неопределенность	Стандартное отклонение погрешности	Неопределенность (погрешность), описываемая стандартным отклонением
Расширенная неопределенность результатов измерений	Доверительные границы	Интервал результата измерений, предполагается, что в него входят почти все или все значения, которые приписаны к данной величине.
Вероятность охвата	Доверительная вероятность	При выборе охвата нужно иметь понятие о типе закона распределения неопределенности.
Коэффициент охвата	Коэффициент распределения погрешности	Коэффициент, который зависит от типа распределения неопределенности.

Различия погрешности от неопределенности измерений

Определим все отличия двух данных понятий и опишем их в табл. 2.

Таблица 2. Различия погрешности от неопределённости

Параметр	Погрешность измерений	Неопределенность измерений
Привязанность	к «истинному» значению (но это значение точно неизвестно)	к измеренному значению
Понятие	относится к определенному измерению, который сделан определенным средством измерения	сомнение в истинности получаемого результата измерения

По таблице видно, что отличий не так много, также они не действуют на подведение окончательных результатов, так как в классическом подходе погрешности измерений также прибавляют результату измерений. Оба понятия дополняют один другого и соединяются в общую стратегию оценки точности результатов измерений.

Оценка результатов измерений при помощи погрешности и неопределенности

Произведем оценку результатов измерений двумя подходами, описанных в прошлых пунктах, а также запишем формулы в табл. 3. Предположим, есть ряд значений, полученных с вольтметра ($N=10$): 100,68; 100,83; 100,79; 100,64; 100,63; 100,94; 100,60; 100,68; 100,76; 100,65.

Таблица 3. Оценка результатов измерений с использованием погрешности и неопределенности измерений

Параметр	Описание результатов измерений с использованием:	
	Погрешности	Неопределенности
Среднее значение	$\bar{X} = \frac{\sum_{i=1}^N X_i}{N}$	$\bar{X} = \frac{\sum_{i=1}^N X_i}{N}$
Результат	100,72	100,72
Разброс результатов измерений	$s_r = +\sqrt{\frac{\sum_{i=1}^N (X_i - \bar{X})^2}{N-1}}$ Стандартное отклонение результатов измерений	Стандартная неопределенность $u(X) = +\sqrt{\frac{\sum_{i=1}^N (X_i - \bar{X})^2}{N-1}}$
Результат	0,034	0,034
Разброс, который характеризует все виды ошибок	Стандартное отклонение суммарной погрешности $s_{\text{сумм}} = +\sqrt{s_r^2 + \frac{\Delta^2}{3}}$	Суммарная стандартная неопределенность $u_c(X) = +\sqrt{u^2(X) + u^2(\delta) + \sum_{j=1}^J c_j^2 u^2(z_j)}$ Член $\sum_{j=1}^J c_j^2 u^2(z_j)$, который обусловлен матричными эффектами, в российской литературе обычно не применяется (считают такие эффекты малыми) [4]
Результат	0,0865	0,0957
Интервал нахождения измеряемой величины	Доверительный интервал $\{\bar{X} \pm k s_{\text{сумм}}\}$	Среднее значение ± расширенная неопределенность $\{\bar{X} \pm U\} = \{\bar{X} \pm k u_c(X)\}$
Результат	100,72 ± 0,173	100,72 ± 0,1914

Расчеты подтверждают, что различия оценок из-за неполного соответствия некоторых алгоритмов расчета и коэффициентов немного разные. Для рядовых измерений будем полагать, что расхождения пренебрежимо малы.

Заключение

Таким образом, погрешность и неопределенность измерений, безусловно разные понятия, но они очень схожи, и на практике используются оба понятия, при этом результаты мало отличаются друг от друга. Наибольшее применение понятия неопределенности находит в теоретической метрологии. В международном сообществе большее предпочтение отдают неопределенности, что было прописано в руководстве в части ISO 9000.

Литература

1. ГОСТ Р 54500.3-2011/Руководство ИСО/МЭК 98-3:2008 Неопределенность измерения. Часть 3. Руководство по выражению неопределенности измерения.
2. ГОСТ Р 54500.3 2011/ Руководство ИСО/МЭК 98-3:2008 Руководство по выражению неопределенности измерения.
3. Погрешность и неопределенность. Сравнительный анализ. <https://scienceforum.ru/2014/article/2014000395>.
4. Оценивание погрешности и неопределенности измерений в лабораторной практике.

ERROR ANALYSIS AND UNCERTAINTY IN METROLOGY

Aleksandr V. Korotkov

Student of group M61801, MTUCI

kor-avk@ya.ru

Aleksandr L. Senyavskiy

MTUCI, PhD., professor of MSiI department

alsenyavskiy@yandex.ru

Keywords: *measurement uncertainty, measurement error, evaluation of measurement results, true value, accuracy of measurement results.*

In the USSR and in Russia, error analysis is used to assess the accuracy of measurements. In foreign countries in last 20 years the concept of measurement uncertainty is used to estimate measurement errors. Such a difference in approaches leads to ambiguity, for example, in the certification of foreign products. The article based on existing regulatory documents discusses the features of these approaches.

ОЦЕНКА ЧИСЛА УСТРОЙСТВ IVR И ОПЕРАТОРОВ В КОНТАКТ-ЦЕНТРЕ С ИСПОЛЬЗОВАНИЕМ МЕТОДА ДЕКОМПОЗИЦИИ

Денисова Мария Алексеевна
студентка группы БСС1501, МТУСИ

marija.masha1801@yandex.ru

Степанов Михаил Сергеевич

к.т.н., доцент кафедры ССuСК, МТУСИ

mihstep@yandex.ru

Ключевые слова: контакт-центр, устройства IVR, операторы, модель Эрланга, поток вызовов, среднее время обслуживания, потери.

Представлена математическая модель контакт-центра с автоответчиками IVR и операторами. Описана схема данной модели, которая разделена в соответствии с методом декомпозиции. Рассматривается модель Эрланга с потерями. Приводится расчет оптимального количества устройств IVR с использованием формулы Эрланга при различных значениях среднего времени обслуживания.

В настоящее время контакт-центры являются неотъемлемой частью организации взаимодействия между клиентом и компанией. Для этого используются самые разные каналы связи: голосовая, веб-чаты, электронная почта, социальные сети. Поступающие по данным каналам запросы обрабатываются операторами. При возникновении ситуации, когда все линии заняты, т.е. достигается максимальная емкость, клиенты помещаются в очередь. Это приводит к длительному времени ожидания, а клиент отключается, не дожидаясь ответа.

Для решения проблемы занятости линий вводится система интерактивного речевого взаимодействия (IVR - англ. Interactive Voice Response), основанная на заранее записанных голосовых сообщениях, которые перенаправляют абонента в нужный раздел [1]. Поэтому необходимо оценить количество устройств IVR в контакт-центре при различных параметрах.

В данной статье объектом исследования является математическая модель контакт-центра с автоответчиками IVR и операторами, представленная на рис.1. Данная модель делится на блоки, каждый из которых рассчитывается отдельно. В нашем случае мы делим модель в соответствии с этапами обслуживания – на IVR и операторов. Поток вызовов, поступающий на IVR, подчиняется пуассоновскому закону с интенсивностью λ . Для расчета количества IVR используется модель Эрланга с потерями, математическое описание которой приводится в следующих разделах статьи.

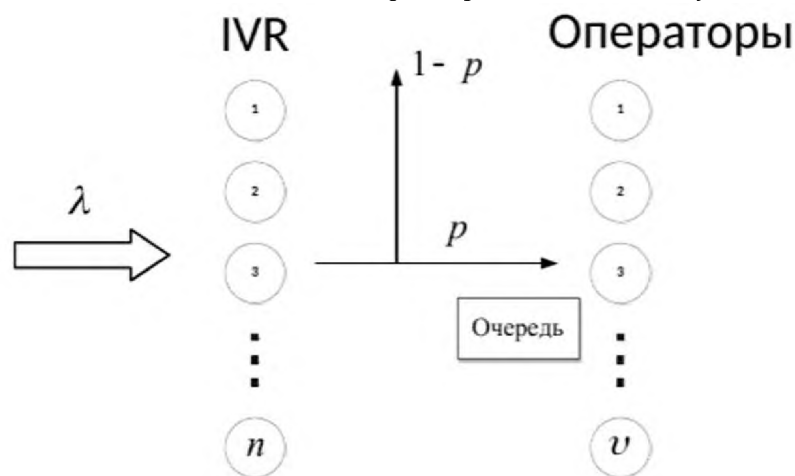


Рис. 1. Математическая модель контакт-центра.

Математическое описание модели Эрланга.

Сначала рассмотрим более простую модель: предположим, что имеется только один поток вызовов, интенсивности λ .

Данную модель создал датский учёный А.К. Эрланг для оценки требуемого числа линий связи в зависимости от количества звонков, средней продолжительности разговора и среднем времени ожидания ответа на звонок [2]. В модели Эрланга учитываются потерянные вызовы от потенциальных клиентов, вызванные недостаточностью числа операторов контакт-центра.

В основе модели Эрланга лежит формула расчета нагрузки для телекоммуникационной системы, которая учитывает поступление случайных сигналов и постановку их в очереди ожидания. Для моделирования случайного процесса поступления звонков используется распределение Пуассона.

Обозначим через $i(t)$ число заявок, находящихся на обслуживании в момент времени t . Изменение состояний модели задаётся случайным процессом $r(t)=i(t)$, определённым на бесконечном пространстве состояний $S=\{i, i=0, 1, \dots, v\}$.

В соответствии с реализованной процедурой построения модели интервалы времени между событиями имеют экспоненциальное распределение и не зависят друг от друга, так же, как и вероятности перехода из состояния в состояние. То есть случайный процесс $r(t)$ - марковский.

В подтверждение этого приведем определение: случайный процесс $r(t)$, описывающий изменение состояний некой физической системы в зависимости от значения параметра времени t , называется марковским, если вероятность любого будущего состояния системы зависит только от ее состояния в настоящий момент и не зависит от того, каким образом система попала в это состояние. Другими словами, состояние модели в рассматриваемый момент времени содержит всю информацию об изменении ее состояний в прошлом, необходимую для описания будущей эволюции модели.

Время пребывания случайного процесса $r(t)$ в состоянии (i) из пространства состояний имеет экспоненциальное распределение с параметром $\lambda + \mu$. После этого с вероятностью совершается переход $r(t)$ в состояние $(i+1)$, а с дополнительной вероятностью в состояние $(i-1)$. Диаграмма переходов $r(t)$, иллюстрирующая динамику изменения состояний рассматриваемой модели Эрланга показана на рисунке 2.

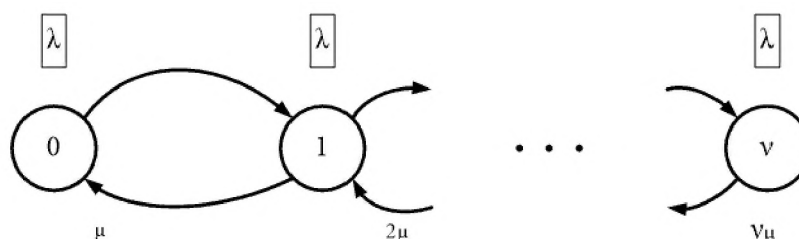


Рис. 2. Динамика изменения состояний (диаграмма переходов) модели Эрланга - модели контакт-центра без деления вызовов на потоки с разным приоритетом.

Обозначим через $p_i(t) = p\{r(t) = i\}$ вероятность нахождения процесса $r(t)$ в состоянии (i) в момент времени t . Через $p_{i,j}(t) = p\{r(t) = j \mid r(0) = i\}$ обозначим условную вероятность перехода $r(t)$ за время t из состояния (i) в состояние (j) . При исследовании марковского процесса будем предполагать, что условные вероятности $p\{r(t+s) = j \mid r(s) = i\}$ не зависят от s , т.е. анализируемый процесс является однородным.

Предположим, что рассматриваемый марковский случайный процесс $r(t)$ обладает предельными вероятностями, т.е. независимо от j выполняются соотношения

$$\lim_{t \rightarrow \infty} p_{j,i}(t) = p(i),$$

где $p(i)$ - распределение вероятностей, т.е. сумма $p(i)$ равна единице и каждое из них неотрицательно.

Система уравнений равновесия, связывающая значения ненормированных вероятностей $P(i)$ для модели Эрланга имеет следующий вид:

$$P(0)\lambda = P(1)\mu;$$

$$P(i)(\lambda + i\mu) = P(i-1)\lambda + P(i+1)(i+1)\mu, \quad i=1, \dots, v-1; \quad (1)$$

$$P(v)v\mu = P(v-1)\lambda.$$

Для $P(i)$ в этой модели выполняется условие нормировки:

$$\sum_{i=0}^v P(i) = 1.$$

Формула Эрланга и доля потерянных заявок.

Рекуррентное соотношение, связывающее значения ненормированных вероятностей $P(i)$ для модели Эрланга имеет вид

$$P(i)\lambda = P(i+1)(i+1)\mu, \quad i=0, 1, \dots, v-1. \quad (2)$$

Используя это соотношение, значение всех ненормированных вероятностей $P(i)$, $i=0, 1, \dots, v$, можно выразить через любую одну из рассматриваемых вероятностей, например, через $P(0)$:

$$P(i) = P(0) \left(\frac{\lambda}{\mu} \right)^i \frac{1}{i!}, \quad i=0, 1, \dots, v. \quad (3)$$

Введем обозначение

$$a = \frac{\lambda}{\mu}. \quad (4)$$

Эта величина является интенсивностью потока вызовов, поступивших за время обслуживания одного вызова, и измеряется в эрлангах (Эрл), т.е. в среднем числе потенциальных вызовов.

С учетом (3) из (4) и условия равенства единице суммы всех вероятностей, можно получить, что нормированное значение $p(0)$ может быть вычислено по формуле:

$$p(0) = \frac{1}{\sum_{i=0}^v \frac{a^i}{i!}}. \quad (5)$$

Из (3) и (5) получаем выражение для $p(i)$:

$$p(i) = \frac{\frac{a^i}{i!}}{\sum_{j=0}^v \frac{a^j}{j!}}, \quad i=0, 1, \dots, v. \quad (6)$$

Основным показателем качества обслуживания поступающих вызовов является доля времени занятости всех операторов. Эта характеристика называется вероятностью потерь по времени и обозначается буквой π . Из определения ясно, что выполняется равенство

$$\pi = p(v). \quad (7)$$

Для модели Эрланга вероятность потерь вызовов π обозначается $E(v, a)$. Из формул (6) и (7) получаем формулу Эрланга для вычисления $E(v, a)$:

$$E(v, a) = \frac{a^v}{\sum_{j=0}^v \frac{a^j}{j!}}. \quad (8)$$

Примеры графиков зависимости $E(v, a)$ от a при различных фиксированных значениях v представлены на рисунке 3.

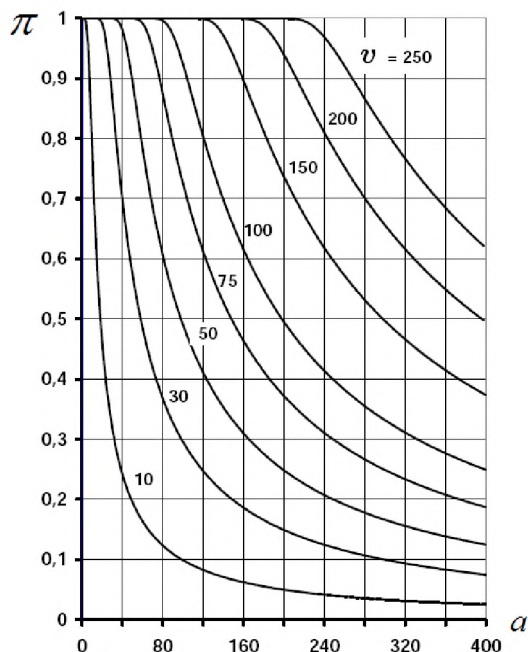


Рис. 3. Зависимости $E(v, a)$ от a при различных фиксированных значениях v .

Формула (8) позволяет подобрать число операторов v , необходимое для того, чтобы обеспечить долю потерянных вызовов не выше некоторого заданного порога. Подбор оптимального числа операторов может осуществляться как при помощи непосредственных расчетов методом перебора, так и графически (по построенному графику) [3-6].

Расчет числа устройств IVR.

Для примера возьмем следующие входные параметры: $\lambda = 1$ (один вызов в секунду). Далее будем менять среднее время обслуживание, тем самым проследим характер изменения функции, показывающей зависимость числа устройств IVR от доли потерянных заявок.

По данным различных контакт-центров можно сделать вывод, что обычно

$1/\mu = 45$ с - среднее время обслуживания. Начинаем расчет методом перебора, каждый раз увеличивая значение r (число устройств IVR). Получаем следующие значения:

Таблица 1. Результаты расчета числа устройств IVR

Число устройств IVR	Величина потерь
1	0.978261
2	0.956542
3	0.934846
...	
49	0.063555
50	0.054104
51	0.045564

Так как предельно допустимая доля потерянных вызовов 5.00%, то можно сделать вывод, что данное требование Π меньше 0.05 выполняется при количестве операторов IVR равном 51.

Далее возьмем среднее время обслуживания равным 40 с ($1/\mu = 40$ с). После аналогичных расчётов получим:

Таблица 2. Результаты расчета числа устройств IVR

Число устройств IVR	Величина потерь
1	0.97561
2	0.951249
3	0.926918
...	
44	0.064597
45	0.054301
46	0.04509

Требование Π меньше 0.05 выполняется при количестве операторов IVR равном 46. А также проведем расчет для величины $1/\mu = 50$ с:

Таблица 3. Результаты расчета числа устройств IVR

Число устройств IVR	Величина потерь
1	0.980392
2	0.960799
3	0.941223
...	
54	0.062482
55	0.053749
56	0.045792

Здесь требование Π меньше 0.05 выполняется при количестве операторов IVR равном 56. Построим график для полученных результатов.

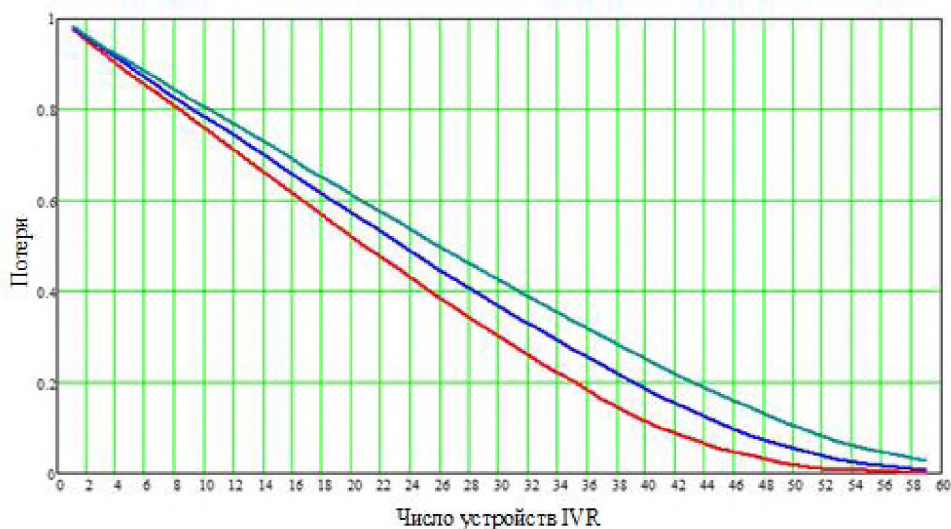


Рис. 4. Зависимость числа устройств IVR от доли потерянных заявок.

Из полученного графика можно увидеть то, как уменьшается доля потерь вызовов с ростом количества операторов при трех параметрах среднего времени обслуживания ($1/\mu = 40$ с – красная кривая, $1/\mu = 45$ с – синяя кривая, $1/\mu = 50$ с – зеленая кривая). Также можно сделать вывод, что при неизменном параметре $\lambda = 1$ (один вызов в секунду) потери в одноименных точках тем больше, чем больше среднее время обслуживания.

Заключение

Таким образом, в работе проанализированы процессы обработки и передачи данных в контакт-центрах, приведено описание математической модели Эрланга и произведен расчет оптимального числа устройств IVR при различных значениях среднего времени обслуживания.

Литература

1. Гольдштейн Б.С., Фрейнкман В.А. Call-центры и компьютерная телефония. СПб.: БХВ – Санкт-Петербург, 2002. 372 с.
2. Степанов С.Н. Теория телетрафика: концепции, модели, приложения // Серия "Теория и практика инфокоммуникаций". М.: Горячая линия – Телеком, 2015. 868 с.
3. Степанов С.Н., Степанов М.С. Алгоритмы оценки показателей пропускной способности обобщенной модели контакт-центра // Автоматика и Телемеханика. 2016. № 7. С. 86-102.
4. Степанов М.С. Планирование числа операторов и линий доступа в современных контакт-центрах // Т-Сomm: Телекоммуникации и транспорт. 2014. Том 8. №7. 89-91.
5. Степанов М.С. Определение и свойства входных параметров обобщенной модели контакт-центра // Т-Сomm: Телекоммуникации и транспорт. 2015. Том 9. №7. С. 25-30.
6. Дудина В.А., Журко А.М., Степанов М.С. Модель контакт-центра с учетом навыков операторов и нетерпеливости абонентов // Т-Сomm: Телекоммуникации и транспорт. 2017. Т. 11. № 12. С. 43-48.

ASSESSMENT OF THE NUMBER OF IVR AND OPERATORS IN THE CONTACT CENTER USING THE DECOMPOSITION METHOD

Maria A. Denisova

Student of group BSS1501, MTUCI

marija.masha1801@yandex.ru

Mikhail S. Stepanov

Ph.D., associate professor, NaCS department, MTUCI

mihstep@yandex.ru

Keywords: *contact-center, IVR, operators, Erlang model, call flow, average service time, losses.*

Here is presented a mathematical model of the contact-center with IVR and operators. The scheme of this model, which is divided according to the decomposition method, is described. The Erlang model with losses is considered. The calculation of the optimal number of IVR using the Erlang formula for different values of the average service time is given.

АНАЛИЗ ТЕХНОЛОГИИ МОБИЛЬНОЙ СВЯЗИ ПЯТОГО ПОКОЛЕНИЯ

Макарина Диана Александровна,

студент группы ИМ61801, МТУСИ

MakarinaDA@yandex.ru

Данилов Алексей Николаевич

МТУСИ, к.т.н., доцент кафедры ССисК

alexnicdaniilov@yandex.ru

Ключевые слова: мобильная связь, Интернет вещей, технология, протокол, сотовая сеть, сервис.

Производится анализ технологии мобильной связи пятого поколения. Выделяются и описываются предпосылки появления, характерные особенности основных технологий реализации и перспективы применения сетей мобильной связи пятого поколения. Дается сравнение мобильных технологий третьего, четвертого и пятого поколения. Показано использование технологии мобильной связи пятого поколения для современной концепции Интернета вещей.

С момента появления и до сегодняшнего дня сети мобильной связи прошли большой путь развития. Появились новые типы пользовательских устройств – смартфоны и планшеты. Возможности, которые открывают мобильные технологии сегодня, уже давно вышли за рамки голосовых услуг, создавая новые способы общения, обмена данными и бизнес-модели.

Технологии продолжают свое развитие в направлении к более высокой производительности и всё большему числу возможностей. В дополнение к существующим технологиям радиодоступа появляются новые технологии, позволяющие решать те задачи, которые невозможно решить с помощью протоколов 3G/4G. Прозрачная интеграция существующих и новых технологий будет способствовать повышению качества пользовательского опыта и появлению целого ряда новых услуг.

В ближайшей перспективе будут внедряться технологии мобильной связи пятого поколения 5G, т.е. набор органически интегрированных технологий радиодоступа [1].

Технологии LTE/LTE Advanced - это эволюционный шаг в развитии технологий мобильной связи наряду с протоколами HSPA и Wi-Fi. Даже технология GSM будет играть важную роль, оставаясь доминирующей технологией во многих уголках земного шара и после 2020 года. Ввиду этого, речь идет не о замене существующих технологий на 5G, а, скорее, об их развитии в дополнении с новыми технологиями радиодоступа, предназначенными для конкретных сценариев и определенных целей.

В настоящее время активно развивается Интернет вещей (IoT), использующий мобильные сети пятого поколения 5G (см. рис. 1). Стационарные и мобильные устройства-автоматы связываются в единое целое, став частью новой промышленной и экономической революции. Современные устройства доступа являются сильным катализатором для развития телекоммуникационной отрасли. Особенно это касается высоких скоростей, виртуализации и облачных сервисов.

При появлении пятого поколения сетей технологии беспроводной связи (предположительно в 2020 году) ожидается, что сеть будет в состоянии обрабатывать примерно в 1000 раз больше мобильных данных, чем современные сотовые сети.

Инновационный цикл генерации нового поколения мобильной связи

Появление технологии 5G должно служить преодолению тех вызовов, которые сейчас наблюдаются на мобильном рынке, а именно:

1. Взрывной рост мобильного трафика передачи данных, как абонентского, так и служебного.
2. Переход от соединений, сосредоточенных на задачах обеспечения связи в цепочке «человек - человек» (H2H) или «человек - машина (сервер)» (H2M) на соединения, нацеленные на обеспечение связи в цепочке «машина - машина» (M2M).
3. Необходимость снижения капитальных затрат на развертывание сетей нового поколения по сравнению с инвестициями в сети предыдущего поколения и минимизации операционных затрат для повышения их экономической эффективности.

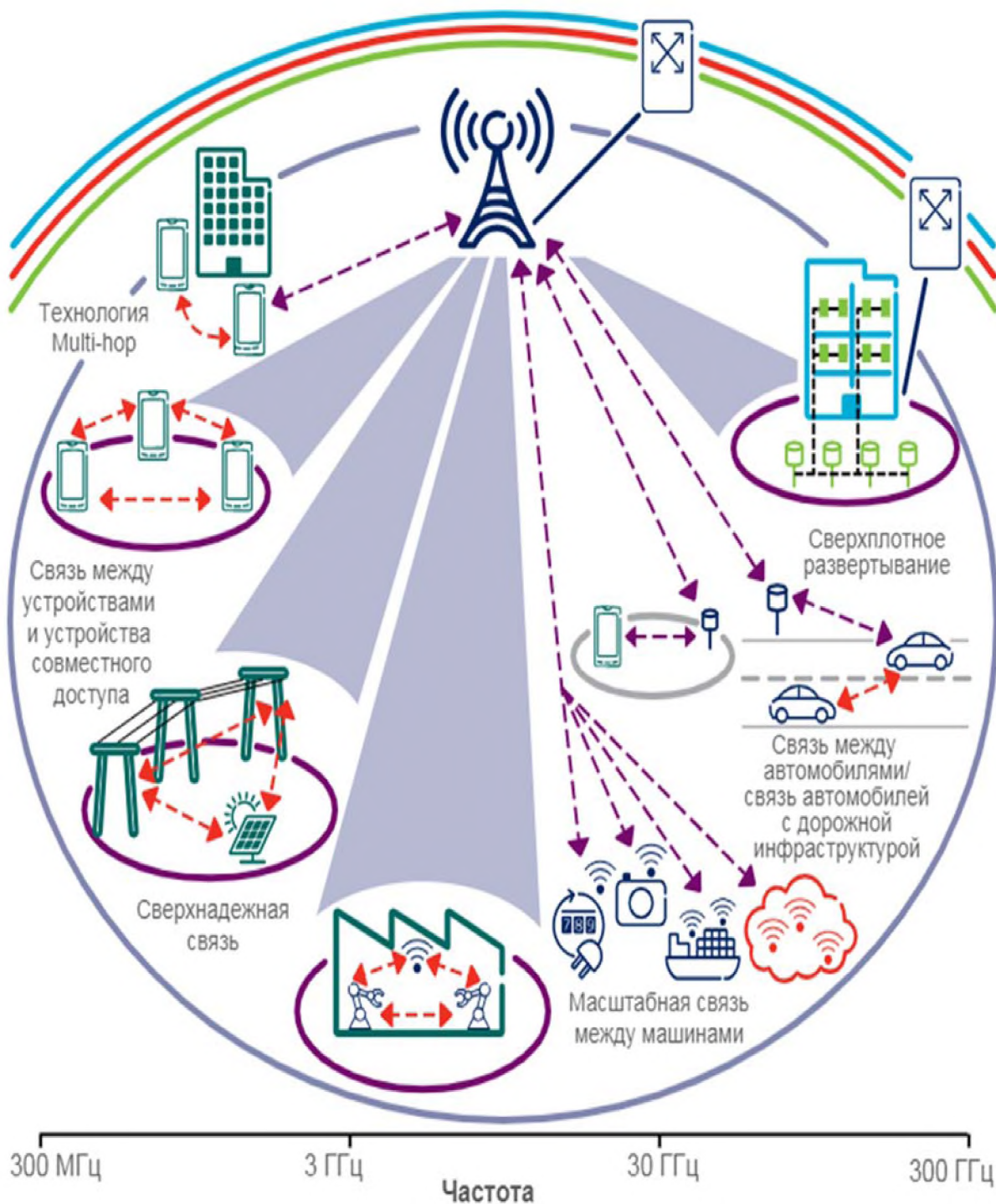


Рис. 1. Использование технологии 5G для Интернета вещей

Технические требования к сетям мобильной связи 5G

1. Рост скорости передачи данных в 10-100 раз в расчёте на абонента - до 10 Гбит/с (UL) и до 5 Гбит/с (DL).
2. Рост потребляемого трафика в 1000 раз - до 500 Гбит на пользователя в месяц.
3. Увеличение количества подключаемых абонентских устройств к соте в 10-100 раз.
4. Десятикратное увеличение времени автономной работы абонентских устройств с небольшим энергопотреблением, таких как сенсоры M2M.
5. Сокращение времени задержки в цепочке E2E с 5 мс до 1 мс и менее.
6. Снижение стоимости эксплуатации и энергопотребления сетей 5G по сравнению с 4G.

Для того, чтобы успешно справиться с задачами, описанными выше требуется разработка новых, более совершенных и эффективных решений беспроводного доступа.

Усовершенствованные версии существующих технологий радиодоступа будут дополнены новыми, нацеленными на конкретные сценарии и способы применения. Таким образом, 5G можно рассматривать как результат развития существующих технологий радиодоступа. Разработанная в итоге система, включающая в себя различные технологии радиодоступа 2-го, 3-го и 4-го поколений, полностью интегрированная и используемая как единое целое, может называться системой 5-го поколения (5G), так как она позволяет вывести качество обслуживания пользователей и производительность сети на тот уровень, который недоступен системам 4-го поколения. Протоколы технологии 5G позволят пользователям полностью раскрыть потенциал технологий мобильной связи и новых сервисов, а оператору связи - получить дополнительные источники дохода. При этом для концепции 5G важен не столько технологический аспект, сколько улучшение пользовательского опыта посредством интеграции современных технологий [2].

Поддержка большого количества маломощных M2M устройств

Мероприятия, направленные на расширение существующих технологий (главным образом LTE/LTE Advanced) в целях поддержки большого количества подключенных M2M устройств уже разработаны в рамках стандартов 3GPP. Однако стандарты LTE/LTE Advanced не смогут удовлетворять высочайшим требованиям некоторых приложений, поэтому необходимо внедрять альтернативные технологии, например, для поддержки устройств с крайне высокими ограничениями по энергопотреблению. Такие технологии будут полностью интегрированы в сотовые технологии пятого поколения, что позволит обеспечить пользователей возможностью непрерывного доступа. К примеру, капиллярные сети, использующие альтернативные технологии, будут соединяться с остальным миром посредством сотовых сетей.

Применение проксимальной связи

В ситуациях, когда пользователи находятся близко друг от друга, и, в особенности, когда информация специфична для конкретного места использования (безопасность движения, государственная безопасность и охрана общественного порядка (NSPS) и общие службы ближнего действия, когда пользователь взаимодействует и обменивается информацией с непосредственным окружением), имеет смысл организация обмена данными непосредственно между устройствами по протоколу «устройство-устройство» (D2D) нежели посредством инфраструктуры сети. Под управлением сетевого протокола D2D обеспечивается локальным службам надежность класса оператора мобильной связи, т. к. сеть сможет управлять трафиком D2D в лицензированном диапазоне. В настоящее время (в рамках стандартов LTE/LTE Advanced) уже делаются первые шаги к интеграции D2D в сетевые коммуникационные технологии. Наряду с этим, D2D сможет послужить важным компонентом для приложений NSPS, поскольку позволяет использовать локальную связь даже в случае повреждения сетевой инфраструктуры.

Обеспечение сверхнадежной связи

В промышленности и социальных службах, работе электронного здравоохранения и в управлении умными городами современные радиосети не всегда могут удовлетворять требованиям к надежности. В некоторых случаях, связанных с использованием средств связи для

интеллектуальных энергосистем или безопасности движения требования ко времени задержки, например, могут быть выше, чем могут предложить современные системы, т.е. не более нескольких миллисекунд при сквозной передаче данных. Надежность в значительной степени зависит от архитектуры и конфигурации сети, а также достаточного количества ресурсов для обработки пиковых нагрузок. В отношении инфраструктурного оборудования нового поколения необходимо будет дифференцировать различные виды трафика на уровне сети и в первую очередь обрабатывать критически важные. Значительная сложность заключается в сочетании сверхмалых значений времени задержки и высокой надежности. Для этого потребуется пересмотреть взаимодействие элементов современных систем мобильного широкополосного доступа, которые в основном нацелены на пропускную способность, покрытие сети и скорость передачи данных [3 - 11]. В целях оптимизации сетей и обеспечения малых значений времени задержки следует изменить соотношение между конструкцией управляющего канала, кодированием, адаптивной модуляцией и управлением радиоресурсами. Там, где необходимо крайне малое время задержки, например, одна миллисекунда или ниже, необходимо внедрять новую технологию мобильной связи пятого поколения, позволяющую обеспечить более короткие временные интервалы передачи данных.

Использование новых сигнально-кодовых конструкций в сетях 5G

Одним из условий будущего развития 5G будет повышение спектральной эффективности передаваемых сигналов за счет применения новых сигнально кодовых конструкций на основе неортогональных сигналов и FTN-сигналов отличных от OFDM -сигналов, используемых в сетях 4G (см. рис.2).

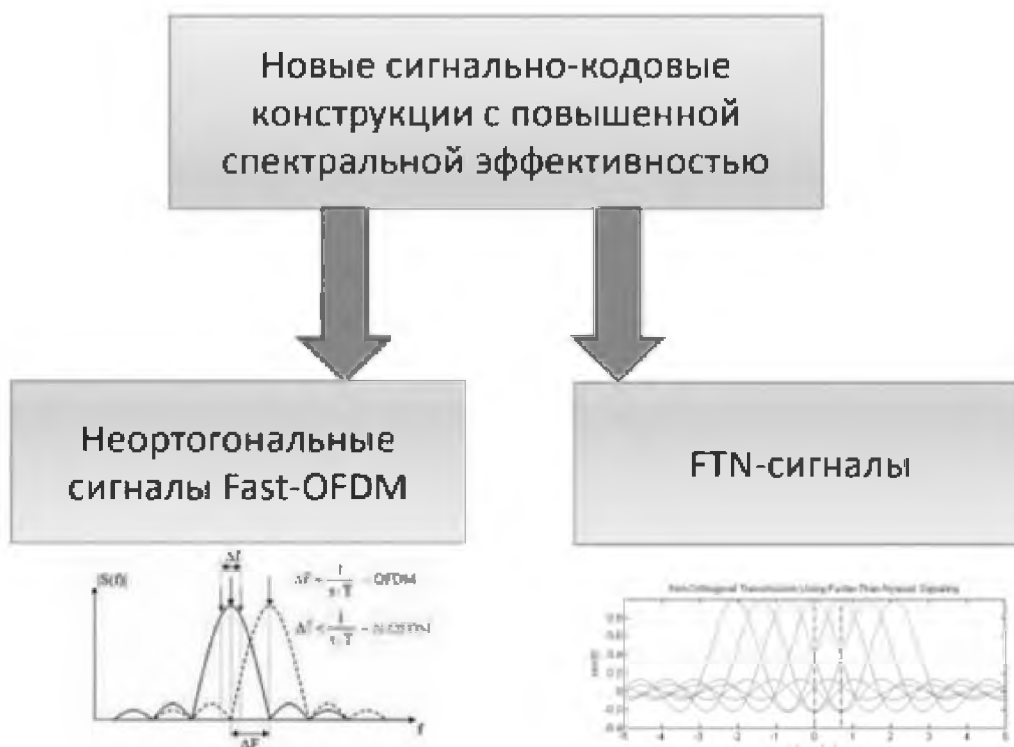


Рис. 2. Сигнально кодовые конструкции на основе неортогональных сигналов и FTN-сигналов

Заключение

1. Реализация проекта «5G к 2020 году» сделает возможным неограниченный доступ к информации.

2. Для реализации перспективной концепции мобильной связи пятого поколения необходимо будет решить проблему возрастающего числа подключающихся к сети устройств.

3. Разнообразие подключенных устройств, приложений и задач позволяет предположить, что 5G станет результатом интеграции существующих технологий радиодоступа, как усовершенствованных LTE/LTE Advanced и HSPA, так и более специализированных, направленных на конкретные цели.

4. Недавно учрежденный консорциум METIS (Mobile and wireless communications Enablers for the Twenty-twenty Information Society) нацелен на разработку фундаментальных концепций 5G и определения дальнейших путей развития мобильной связи.

5. Можно предположить, что концепция системы 5G, разработанная, например, в рамках проекта METIS, будет предложена к стандартизации уже через несколько лет и в таком случае постепенный переход к сетям 5G может начаться уже в 2020 г.

Литература

1. Бакулин М.Г., Крейнделин В.Б., Панкратов Д.Ю. Технологии в системах радиосвязи на пути к 5G. М.: Горячая линия – Телеком, 2018. 280 с.

2. Тухвинский В.О. [Электронный ресурс] // «Перспективы развития сетей 5G: технологии и особенности использования спектра». <http://www.myshared.ru/slide/975111>.

3. Петров Максим [Электронный ресурс] // «Технологии 5G откроют перспективы для развития многих отраслей бизнеса» <https://iot.ru/gadzhety/tekhnologii-5g-otkroyut-perspektivy-dlya-razvitiya-mnogikh-otrasley-biznesa>.

4. Данилов А.Н., Максимов С.П., Крылов Р.Х. Исследование процессов обмена сигнальными сообщениями протокола SIP в сетях NGN // Т-Comm: Телекоммуникации и транспорт. 2014. Т. 8. № 8. С. 32-36.

5. Степанов С.Н., Степанов М.С., Маликова Е.Е., Цогбадрах А., Ндайикунда Ж. Построение и анализ обобщенной модели разделения ресурса для LTE технологий с функциональностью NB-IoT // Т-Comm: Телекоммуникации и транспорт. 2018. Т. 12. № 12. С. 71-77.

6. Мазуркевич Д.О., Орлов В.Г. Эволюция систем безопасности сетей сотовой связи разных поколений // Т-Comm: Телекоммуникации и транспорт. 2011. Т. 5. № 1. С. 38-40.

7. Корионов И.П., Орлов В.Г. Пользовательские аспекты безопасности в сетях LTE // Телекоммуникации и информационные технологии. 2017. Т. 4. № 2. С. 16-21.

8. Пушкарев А.В., Орлов В.Г. Эволюция технических средств формирования и доставки ТВЧ на мобильные терминалы пользователей // Т-Comm: Телекоммуникации и транспорт. 2015. Т. 9. № 1. С. 11-16.

9. Тухвинский В.О. Пятый элемент мобильного мира: итоги MWC-17 // Т-Comm: Телекоммуникации и транспорт. 2017. Т. 11. № 3. С. 4-11.

10. Tikhvinskiy V., Bochechka G. Concept and QOS requirements in 5G networks // Т-Comm: Телекоммуникации и транспорт. 2014. Т. 8. № 9. С. 99-101.

11. Тухвинский В.О. 5G World summit - 2014: курс прежний – от 4G к 5G // Т-Comm: Телекоммуникации и транспорт. 2014. Т. 8. № 7. С. 95-96.

ANALYSIS OF FIFTH GENERATION MOBILE COMMUNICATION TECHNOLOGY

Diana A. Makarina

student of group IM61801, MTUCI

MakarinaDA@yandex.ru

Alexey N. Danilov

MTUCI, PhD., associate professor of CN and SS department

alexnicdanilov@yandex.ru

Keywords: *mobile communications, Internet of things, technology, protocol, cellular network, service*

Fifth generation mobile communication technology is being analyzed. Prerequisites of occurrence, characteristic features of the main implementation technologies and prospects for the use of fifth-generation mobile communication networks are identified and described. Comparison of third, fourth and fifth generation mobile technologies is given. The use of fifth generation mobile communication technology for the modern concept of the Internet of Things is shown.

СОСТОЯНИЕ РАЗРАБОТКИ ИМИТАТОРА ОБЪЕДИНЕННОЙ СЕТИ ПД СПЕЦИАЛЬНОГО НАЗНАЧЕНИЯ В КАЧЕСТВЕ УЧЕБНОГО ЛАБОРАТОРНОГО СТЕНДА

Р.А. Бельфер

к.т.н., доцент кафедры ИБ МГТУ им. Н.Э. Баумана

a.belfer @ yandex.ru,

Е.В., Глинская

ст. преподаватель кафедры ИБ МГТУ им. Н.Э.Баумана

glinskaya-iu8@rambler.ru

А.В. Кравцов

старший научный сотрудник НИИЦ (Москва) ЦНИИ ВВКО

skyak78@gmail.com

В.Г. Орлов

к.т.н. доцент МГУСИ,

Ключевые слова: имитатор объединенной сети, частная сеть, информационная безопасность, сеть передачи данных, маршрутизация.

Рассматривается состояние разработки учебного лабораторного стенда имитатора единой сети передачи данных категории специального назначения, включающей одну или несколько сетей передачи данных разных государственных ведомств. Конфигурация имитатора содержит оконечные пункты, шесть узлов коммутации, центр эксплуатации сети и единый удостоверяющий центр. Приводится краткое описание предлагаемых алгоритмов некоторых функций сети.

Введение

В отличие от материалов, содержащихся в работе [1], в настоящей статье приводится состояние разработки имитатора не для одной сети ПД специального назначения для одного ведомства из указанных в ФЗ «О связи» “для нужд органов государственной власти, нужд обороны страны, безопасности государства и обеспечения правопорядка” [2], а объединенной сети ПД. Под объединенной сетью понимается созданная единая сеть, включающая одну или несколько изолированных (частных) сетей ПД специального назначения для разных ведомств из указанных в категории сетей специального назначения. Создание такой объединенной сети было предложено исходя из экономических соображений сотрудником Научно-исследовательского испытательного центра ЦНИИ Войск ВКО МО РФ А.В. Кравцовым. Для объединенной сети, кроме сохранения функций изолированных сетей поставлена задача обеспечения возможности установления соединения между определенными оконечными пунктами разных частных сетей. Такое соединение будем называть смешанным.

В отношении изолированных частных сетей объединенная сеть является аналогичной виртуальной частной сети зарубежных производителей VPRN (Virtual Private Routed Network), построенной на базе многопротокольной коммутации по меткам MPLS (Multiprotocol Label Switching). Однако, как показано в [3], в обеспечении показателя информационной безопасности VPRN имеется множество незащищенных угроз. Кроме того, технология VPRN не предусматривает необходимых требований по другим важным для сетей ПД категории специального назначения показателям качества обслуживания, таким как: надежность, задержка сообщений, вероятность доставки сообщения и др. В то же время VPRN является одной из используемых в РФ технологией виртуальных частных сетей. В настоящей статье изложено состояние разработки учебного лабораторного стенда (УЛС) имитатора объединенной сети ПД специального назначения и приведены предложения по дальнейшему развитию этих работ.

Конфигурация имитатора объединенной сети ПД

На рис.1 представлена схема гипотетической конфигурации имитатора сети ПД, используемой для выполнения студентами научно-практические работ по имитации объединенной сети ПД специального назначения.

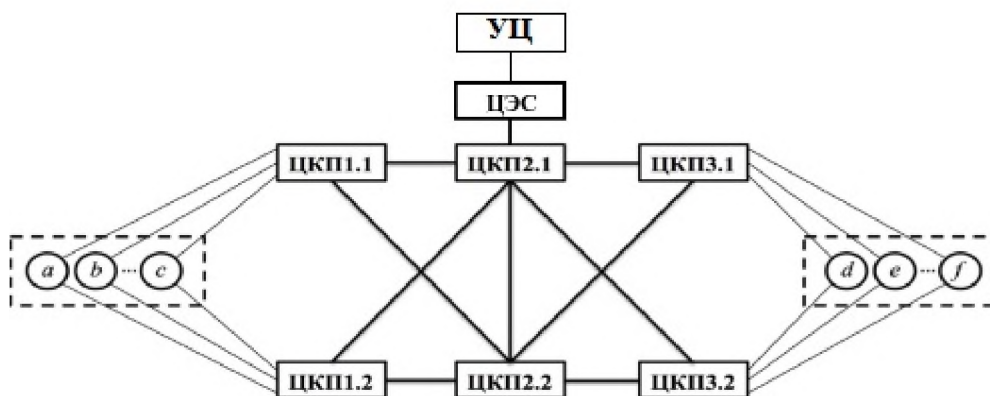


Рис. 1. Конфигурация имитатора объединенной сети ПД с центром эксплуатации сети и единым удостоверяющим центром.

Конфигурация имитатора сети ПД в УЛС предусматривает создание пучка маршрутов из четырех путей маршрутизации одного соединения между окончными пунктами. Каждый путь маршрутизации состоит из трех центров коммутации пакетов (ЦКП). К двум граничным ЦКП абонентского доступа в каждом пути маршрутизации (ЦКП 1.1 - адрес 11, ЦКП 3.1 - адрес 31 и ЦКП 1.2 - адрес 12, ЦКП 3.2 - адрес 32) подключены окончные пункты, а ЦКП 2.1 - адрес 21 и ЦКП 2.2 - адрес 22 являются транзитными (или транспортными). Доступ окончных пунктов (a, b, \dots, c и d, e, \dots, f) к граничным центрам коммутации пакетов удаленный. Оконечные пункты коммутируемых виртуальных каналов обмениваются данными одновременно по четырем путям маршрутизации. Примем, что окончные пункты a (ОПа) и f (ОП f) принадлежат частной сети 1 (ЧС1); окончные пункты c (ОП c) и e (ОП e) принадлежат частной сети 2 (ЧС2), окончные пункты b (ОП b) и d (ОП d) принадлежат частной сети 3 (ЧС3). Присвоим окончным пунктам ЧС1 физические адреса от 1 до 999, ЧС2 - от 1001 до 1999, ЧС3 - от 2001 до 2999. Присвоим физический адрес 101 для ОПа, 601 - для ОП f , 1101 - для ОП c , 1601 - для ОП e , 2101 - для ОП b , 2601 - для ОП d . Номер частной сети обозначим Z .

Алгоритм аутентификации в имитаторе объединенной сети

В системе ОКС№7 сети связи ISDN используется стандартизированный алгоритм взаимной аутентификации (ETSI ETS 300 841) на базе инфраструктуры открытых ключей РКІ (Public Key Infrastructure). Этот механизм используется в работе [1]. Для такого механизма ИБ характерны большие задержки на установление соединения, что не имеет значение для сети связи общего пользования ISDN. Однако для частных сетей ПД, входящих в объединенную сеть ПД категории специального назначения, характерны высокие требования по этой характеристике. Аналогичная задача имеет место в сетях ПД модернизированных энергосистем smart grid [5]. Для исключения указанного недостатка предлагается использовать принцип механизма взаимной аутентификации устройств сети ПД, который изложен в работе [6] для сетей ПД усовершенствованной инфраструктуры счета АМІ (Advanced Metering Infrastructure) модернизированной энергосистемы smart grid. Взаимная аутентификация здесь построена на использовании двух механизмов – РКІ и механизма аутентификации на основе идентификатора устройств. Управление в АМІ взаимной аутентификацией бытового устройства и интеллектуального счетчика осуществляется удаленно из центра управления сетью и подключенного к нему единого удостоверяющего центра (УЦ) [6].

В имитаторе предлагается применить алгоритм взаимной аутентификации смежных узлов сети ПД категории специального назначения так же с использованием двух механизмов – РКІ и

механизма аутентификации на основе идентификатора устройств. В конфигурации имитатора объединенной сети (рис. 1) приведен единый удостоверяющий центр (УЦ), который предоставляет центру эксплуатации сети (ЦЭС) каталог сертификатов всех устройств сети ПД. Удостоверяющий центр УЦ (CA, Certificate Authority) аутентифицирует ЦЭС и предоставляет ему функцию взаимной аутентификации всех смежных узлов имитатора сети ПД. Сертификат ЦЭС выдан этим УЦ, а сертификаты устройств имитатора сети могут быть выданы этим удостоверяющим центром. Приведем последовательность операций по выполнению взаимной аутентификации между всеми смежными узлами на примере одной из пар смежных узлов имитатора сети ПД:

- a) создание и подключение к центру эксплуатации сети единого удостоверяющего центра УЦ имитатора сети ПД. В УС содержится каталог из нескольких сертификатов каждого устройства сети;
- b) УЦ аутентифицирует ЦЭС и пересылает ему указанный выше каталог сертификатов;
- c) аутентификация каждого устройства этой пары смежных устройств на основе их идентификаторов и закрытого ключа сертификата предыдущей взаимной аутентификации;
- d) передача из ЦЭС в успешно аутентифицированные устройства новых сертификатов;
- e) проведение взаимной аутентификацию пары смежных устройств на основе, использования механизма PKI.

Алгоритм формирования канальных ключей между смежными ЦКП

В конкретной паре смежных ЦКП число канальных ключей равно числу частных сетей, используемых в этой паре.

Приведем пример установления канальных ключей для трех частных сетей на примере участка транспортной сети ЦКП 2.1-ЦКП 1.1. По команде с ЦУС в одном из этих ЦКП производится генерация ключа.

Обозначения: R - сгенерированный ключ в ЦКП 2.1.

При использовании между ЦКП 2.1-ЦКП 1.1 путей маршрутизации во всех трех частных сетях (ЧС1, ЧС2, ЧС3) в ЦКП 2.1 создаются канальные ключи для каждой частной сети: KS_{ijk} , где S – номер частной сети, i и j – физические адреса смежных ЦКП, k – номер устанавливаемого КВК. Номер устанавливаемого коммутируемого виртуального канала (КВК) используется для того, что бы канальные ключи были разовыми для каждого соединения. Это относится к тем из частных сетей ПД, требования к которым ИБ более высокие по сравнению с другими ЧС. Для каждой из остальных частных сетей канальные ключи между смежными ЦКП изменяются периодически с помощью генерирования нового R.

Например, $K_{1_{11213}} = \text{hash}(R||1||1\text{ИД}||2\text{ИД}||3)$ для ЧС1 между физическими адресами смежных ЦКП - 11 и 21, номер КВК – 3); $K_{2_{11212}} = \text{hash}(R||2||1\text{ИД}||2\text{ИД}||2)$ для ЧС2, номер КВК – 2; $K_{3_{11213}} = \text{hash}(R||3||1\text{ИД}||2\text{ИД}||3)$ для ЧС3, номер КВК – 3. Все эти ключи передается в ЦКП 1.1 зашифрованными с использованием открытого ключа ЦКП 1. В ЦКП 1.1 производится их дешифрация закрытым ключом ЦКП 1.1.

Архитектура очередей таблиц маршрутизации объединенной сети ПД категории специального назначения

Во всех используемых ЦКП создается строка маршрутизации КВК определенной частной сети. При установлении КВК для каждой используемой в ЦКП частной сети создается строка маршрутизации по логическим адресам. При установлении через ЦКП нескольких КВК, относящихся к определенной одной частной сети, все строки маршрутизации формируют таблицу маршрутизации этой частной сети. Таким образом, в ЦКП несколько таблиц маршрутизации, причем, каждая таблица маршрутизации соответствует определенной частной сети.

В работе [7] предложена для использования архитектура таблицы маршрутизации по логическим адресам на основе очередей строк маршрутизации. Для ЦКП 1.1 принятой

конфигурации в качестве примера приводятся две таблицы маршрутизации объединенной сети ПД специального назначения из двух частных сетей ЧС1 и ЧС3. В начале настоящей статьи приведен принятый диапазон логических адресов для этих частных сетей. Предложенная архитектура таблиц маршрутизации на основе очередей строк маршрутизации позволит реализовать сложное программное обеспечение имитатора сети ПД для дальнейших научно-практических работ по созданию имитатора объединенной сети ПД, включающей частные сети специального назначения нескольких ведомств. О необходимости создания не менее 90% отечественного ПО для использования в госструктурах отмечалось на совещании 10 января 2019 г. премьер-министром РФ Д.А. Медведевым [8].

Маршрутизация в объединенном имитаторе сети ПД специального назначения между оконечными пунктами разных частных сетей

В работе [9] приводится алгоритм маршрутизации смешанного соединения (КВК) между принадлежащим ЧС1 оконечным пунктом ОПа и принадлежащим ЧС3 оконечным пунктом ОПд. Обозначим такое соединение СС13.

Для решения поставленной задачи используются приведенная на рис.1 конфигурация имитатора сети, физические адреса устройств имитатора сети ПД, принадлежность оконечных устройств ЧС1 и ЧС3, а также диапазоны логических номеров [7] очереди свободных номеров ЧС1 и ЧС3 в ЦКП. На основании этих данных необходимо сформировать строки таблиц маршрутизации по логическим адресам во всех ЦКП четырех путей маршрутизации КВК частных сетей ЧС1 и ЧС3. Для ЦКП 1.1. строки таблиц маршрутизации могут использоваться из работы [7].

Из приведённого примера следует, что создание смешанного соединения возможно при условии, если в каждом из путей маршрутизации встречается, хотя бы один ЦКП, который выполняет функции обеих частных сетей (в примере - ЧС1 и ЧС3).

Планируемые работы по развитию функциональных возможностей имитатор сети ПД специального назначения

Дальнейшее совершенствование имитатора объединенной сети ПД в части аппаратной составляющей, алгоритмов функционирования и программного обеспечения связано с следующих задач:

- установка/сброс соединения коммутируемого виртуального канала (КВК) объединенной сети, состоящей из двух частных сетей;
- передача пакета данных по установленному КВК частных сетей;
- установка соединения КВК между оконечными пунктами, принадлежащими разным частным сетям.

Следует отметить сложность программного обеспечения этих функций.

Заключение

Приведённые в статье результаты работ по созданию имитатора сети ПД специального назначения свидетельствуют, что его использование в качестве учебного лабораторного стенда позволит обеспечить подготовку специалистов способных решать задачи создания экономически эффективной отечественной сети ПД.

Литература

1. Басараб М.А., Бельфер Р.А., Кравцов А.В. Состояние разработки имитатора сети ПД специального назначения в качестве учебного лабораторного стенда (без учета выполнения функций объединенной сети). Настоящий выпуск.
2. Федеральный закон Российской Федерации «О связи» № 126-ФЗ от 07.07.2003 (редакция от 27.12.2018 г.).

3. *Michael H. Behringer, Monique J. Morrow.* MPLS VPN Security Cisco Press, 2005, p. 312.
4. *Басараб М.А., Бельфер Р.А., Кравцов А.В., Никулина Т.П.* Алгоритм аутентификации и формирования разовых ключей в имитаторе лабораторного стенда объединенной сети ПД специального назначения, Первая миля №2, 2019. С.62-68.
5. *Басараб М.А., Бельфер Р.А., Глинская Е.В., Якушева Н.* Требования к инфраструктуре сетей связи в составе Smart Grid, Первая миля №4. 2018. С. 78-84
6. Sangji Lee; Jinsuk Bong; Sunhee Shin; Yongtae Shin. A security mechanism of Smart Grid AMI network through smart devicemutual authentication. The International Conference on Information Networking 2014 (ICOIN2014). 2014. pp 592-595.
7. *Басараб М.А., Бельфер Р.А., Кравцов А.В.* Архитектура очередей таблиц маршрутизации объединенной сети ПД категории специального назначения // Электросвязь. № 2. 2019. С.33-37.
8. *Медведев Д.* создание в РФ отечественного ПО не означает самоизоляции страны. – URL: <http://www.finmarket.ru/news/4920471>. Российская Газета от 11 января 2019 года.
9. *Басараб М.А., Бельфер Р.А., Кравцов А.В.* Маршрутизация в объединенном имитаторе сети ПД специального назначения между оконечными пунктами разных частных сетей // Электросвязь. № 3. 2019. С. 51-53.

THE STATE OF DEVELOPMENT OF THE SIMULATOR OF THE INTEGRATED DATA NETWORK OF A SPECIAL PURPOSE AS AN EDUCATIONAL LABORATORY BENCH

R. A. Belfer

candidate of technical Sciences

docent of Information Security, MSTU

a.belfer @yandex.ru

E. V. Glinskaya

senior lecturer of Information Security, MSTU

glinskaya-iu8@rambler.ru

A. V. Kravtsov

senior researcher RaTC

skyak78@gmail.com

V. G. Orlov

MTUCI, PhD., associate professor of MTUCI

Keywords: united network simulator, private network, information security, data transmission network, routing.

The state of development of an educational laboratory bench for a simulator of a single special-purpose data transmission network category, including one or several data networks of various government departments, is considered. The simulator configuration includes end points, six switching nodes, a network operation center, and a single certification center. A brief description of the proposed algorithms for some network functions is given.

ПОВЫШЕНИЕ ЭФФЕКТИВНОСТИ КОЛЛАБОРАТИВНЫХ РЕКОМЕНДАТЕЛЬНЫХ СИСТЕМ ПРИ ИСПОЛЬЗОВАНИИ ДОПОЛНИТЕЛЬНЫХ ДАННЫХ

Авакян Самвел Тигранович
студент группы М111701(73) МТУСИ
avsam@rambler.ru

Шелухин Олег Иванович
д.т.н., профессор кафедры ИБ, МТУСИ
sheluhin@mail.ru

Ключевые слова: машинное обучение, дополнительная информация, кластеризация, коэффициенты сходства и отталкивания, пользовательские признаки, нечисловые характеристики.

Рассмотрен один из способов достижения прироста эффективности работы коллаборативной рекомендательной системы (РС) путём добавления в качестве исходных данных информации о пользователях. Подобное решение представляет собой *гибридную* коллаборативную РС. Проведен сравнительный анализ алгоритмов кластеризации данных и гибридации коллаборативных РС. Построены гистограммы коэффициентов точности и полноты рассмотренных алгоритмов. Показано, что результат работы гибридных РС практически не зависит от выбора алгоритма кластеризации.

Постановка задачи

Построение РС является задачей интеллектуального анализа данных (data mining).

Различают РС на основе содержания (content-based) и на основе рейтингов (collaborative filtering – CF). В ряде случаев используются гибридные алгоритмы, обладающие достоинствами обоих вышеупомянутых видов рекомендательных систем [1].

В зависимости от обрабатываемых данных, коллаборативные рекомендательные системы делятся на две категории:

- *Item-based*. С целью поиска подходящего объекта, рекомендуемого пользователю, сравниваются рейтинги оценённых им объектов. В результате строится таблица item-item, в строках и столбцах которой указаны все объекты. В ячейках на пересечении содержатся коэффициенты сходства между ними, на основе которых формируются рекомендации;

- *User-based*. Для поиска подходящего объекта, рекомендуемого пользователю, сравниваются наборы рейтингов. Для пользователей, обладающих высоким коэффициентом сходства, формируются рекомендации из объектов, оценённых одним пользователем и одновременно отсутствующих в наборе оценок другого.

Целью статьи является анализ эффективности гибридных коллаборативных РС при использовании различных алгоритмов кластеризации при наличии дополнительной информации о пользователях.

Методы кластеризации пользовательских признаков

В качестве исходных данных для коллаборативной фильтрации могут выступать, в частности, матрицы «пользователь-элемент», в которых содержатся оценки пользователей для элементов, количество посещений страниц (прослушиваний аудиозаписей), просмотров видео, частота запросов и т.п.

Добавление информации о пользователях в виде таблицы «пользователь-признак» позволяет снизить влияние разреженности таблицы «пользователь-элемент». Разреженность, т.е. низкая степень заполнения исходной рейтинговой таблицы, свойственна dataset'ам всех крупных компаний, фиксирующих подобную информацию для формирования рекомендаций. Данная проблема возникает, когда количество элементов в системе так велико, что большинство пользователей физически не способны оценить их все за ограниченный промежуток времени. В результате из-за недостатка исходной информации задача выявления сходства по предпочтениям усложняется. Используя базу данных о предпочтениях пользователей, алгоритмы коллаборативной фильтрации помогают предугадать темы или продукты, которые с наибольшей вероятностью интересны пользователю [2].

Пусть сценарий CF имеет следующую структуру: имеется список из m пользователей $\{u_1, u_2, \dots, u_j, \dots, u_m\}$ и список из n элементов $\{i_1, i_2, \dots, i_k, \dots, i_n\}$, и каждый пользователь u_j имеет список оцененных им элементов I_{uj} . В качестве рейтингов могут быть использованы как явные показатели, (например, в масштабе 1-5), так и неявные (количество посещений, история покупок и т.д.) [3]. На основе этих данных строится матрица «пользователь-элемент», в которой по строкам располагаются пользователи U , а по столбцам – элементы I . На пересечении строки и столбца указывается рейтинг, выставленный пользователем u_j для элемента i_k . Данная матрица всегда имеет незаполненные ячейки. Элементы, в ячейке которых не указана оценка, будут представлены в качестве рекомендации пользователю u_j , если эти элементы получили хорошую оценку других пользователей, имеющих схожие показатели с пользователем u_j .

Признаки пользователей чаще всего представлены в виде слов и словосочетаний из сгенерированного заранее или введённых пользователями вручную набора вариантов. Пример формата матрицы «пользователь-признак» представлен в таблице 1. Наиболее очевидными являются два способа представления информации такого формата в пригодный для использования вид: применение метода нечисловой кластеризации и преобразование признаков пользователей в набор числовых значений, сопоставленных с каждым признаком.

Таблица 1. Пример заполнения матрицы «пользователь-признак»

Пользователь	Признак		
	<i>пол</i>	<i>возраст</i>	<i>спорт</i>
Андрей	мужской	18	теннис
Пётр	мужской	25	футбол
Евгения	женский	19	плавание
Анастасия	женский	21	теннис

От выбора способа представления зависит структура учёта признаков пользователей.

Методы нечисловой кластеризации. Под *кластеризацией* будем понимать процесс объединения наиболее схожих между собой сущностей из выборки в группы, называемые *кластерами*. В этом случае рекомендации для пользователей, не имеющих достаточного количества ненулевых оценок в таблице «пользователь-элемент», будут формироваться из сущностей, высоко оценённых группой других пользователей из кластера, в который входит данный пользователь. Данный метод, помимо гибридации за счёт добавления информации о пользователях, дополняется *алгоритмом на основе модели кластеризации* [5]. Процесс составления рекомендаций с помощью коллаборативной РС будет осуществляться для менее объемного набора наиболее схожих между собой пользователей в каждом кластере отдельно.

Достоинством данного подхода является увеличение точности рекомендаций при корректном распределении пользователей на кластеры, а также повышение скорости формирования рекомендаций для новых пользователей. Недостатком метода кластеризации нечисловых данных является проблема выбора коэффициента, влияющего на степень сходства между собой сущностей в кластере.

Среди таких алгоритмов наиболее известен алгоритм CLOPE [7], основанный на частоте попадания каждого из вариантов всех признаков в кластер. Регулировка размеров кластеров и,

соответственно, их количества осуществляется за счёт изменения *коэффициента отталкивания* r . При его увеличении, количество кластеров увеличивается, а степень сходства между пользователями в нём снижается, и наоборот.

Для набора объектов $X = \{x_1, x_2, \dots, x_n\}$, каждый из которых обладает вектором признаков $\{i_1, i_2, \dots, i_m\}$, разбиение на множество кластеров $C = \{C_1, C_2, \dots, C_k\}$ формируется таким образом, что $C_i \neq \emptyset$; $C_i \cap C_j = \emptyset$. Для кластера C_i определяется количество $W(C_i)$ входящих в него уникальных признаков i и среднее значение $H(C_i)$ их повторений.

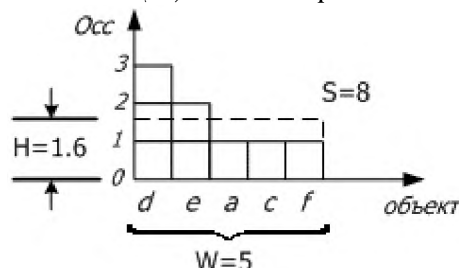


Рис. 1. Иллюстрация гистограмма кластера

Результаты оценки размера кластеров, полученных в результате обработки указанных данных, приведены в таблице 2.

Таблица 2. Размеры кластеров, полученных для данных UCI с помощью алгоритма CLOPE

Коэффициент r	Количество кластеров	Размеры кластеров
1.6	4	23, 31, 13, 34
1.8	5	29, 11, 20, 18, 23
1.9	7	23, 13, 20, 11, 5, 17, 12
2	7	18, 13, 20, 18, 15, 9, 8
2.1	7	18, 13, 20, 22, 15, 5, 8
2.2	8	16, 13, 20, 14, 4, 6, 15, 13
2.3	10	17, 13, 20, 6, 16, 8, 6, 4, 6, 5
2.4	12	13, 11, 20, 5, 10, 7, 6, 6, 12, 6, 3, 2
2.5	13	18, 11, 20, 3, 6, 6, 4, 17, 3, 1, 8, 3, 1
2.6	13	10, 11, 20, 3, 6, 6, 4, 13, 4, 1, 8, 3, 12
2.7	13	11, 11, 20, 4, 4, 6, 4, 12, 3, 4, 4, 6, 12
2.8	14	14, 9, 20, 4, 6, 6, 4, 10, 11, 4, 4, 4, 4, 1

Оценку качества кластеризации алгоритма CLOPE осуществляют по следующим критериям [15], характеризующим «коэффициент сходства»:

- *Индекс оптимального соответствия* I_{OC} , определяющий насколько разбиение близко к оптимальному. Перед его оценкой необходимо найти соответствие между составленными кластерами и эталонными. С этой целью выявляют, в каком кластере находится наибольшее количество элементов из каждого эталонного кластера. Найденное значение совпадений по каждому из кластеров будет соответствовать значению индекса оптимального соответствия.

$$I_{OC} = \frac{\sum_{i=1}^k \min(n_i, c_i)}{k}, \quad (1)$$

где k – эталонное количество кластеров, n_i – размер i -го кластера при эталонном разбиении, c_i – количество элементов i -го эталонного кластера в текущем i -ом кластере;

- Индекс Жаккара [16], представляющий собой отношение числа совпадающих объектов C кластера к сумме числа объектов в кластерах A и B , не вошедших ни в один из кластеров.

$$I_{JA} = \frac{C}{A+B} \quad (2)$$

Для оценки эффективности метода гибридизации проведём кластерный анализ набора данных из репозитория UCI Machine Learning [14]. Для определения индекса оптимального соответствия выбор dataset'a основывался на необходимости наличия эталонного разбиения на кластеры. Данные представлены в булевом формате, в котором каждая единица указывает на наличие конкретного признака. Такое преобразование называется Dummy-кодированием [19].

Результаты оценки коэффициентов, характеризующих качества кластеризации, для базы данных UCI представлены на рисунке 2. Видно, что индекс оптимального соответствия показывает более оптимистичные результаты, чем индекс Жаккара.

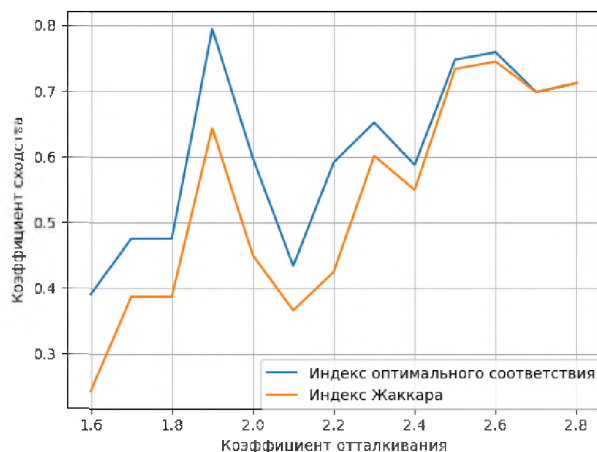


Рис. 2. График зависимости коэффициентов сходимости алгоритма CLOPE от коэффициента отталкивания γ

Преобразование признаков пользователей в набор числовых значений. При переводе нечисловых исходных признаков в числовые становится возможным использования единого алгоритма фильтрации данных для обеих таблиц «пользователь-объект» и «пользователь-признак». Кроме того, при подобном представлении для каждого пользователя формируется однозначно определенный ненулевой набор параметров.

Для того, чтобы предоставить пользователям возможность ручного ввода вариантов каждого из признаков, необходимо провести классификацию текста [10]. Реализовать подобную классификацию возможно с помощью специальных библиотек различных языков программирования (к примеру, scikit-learn для python [8]) или отдельных технических средств, таких как word2vec [9] от компании Google. Полученные результаты позволяют составить набор обобщенных вариаций для наиболее схожих признаков, введенных вручную. В этом случае аналогично матрице «пользователь-элемент» составляется матрица «пользователь-признак», в которой столбцы соответствуют полученному набору вариаций. Если пользователь обладает тем или иным признаком, в матрице на пересечении будет указано значение, отличное от нуля. Полученный набор данных можно обрабатывать методами коллаборативной фильтрации, подходящими для работы с первоначальной матрицей «пользователь-элемент».

Данный подход обладает следующими характерными особенностями:

- Возможность применения всех числовых методов для нахождения коэффициента сходимости между пользователями;
- Возможность для пользователей вводить произвольные признаки вручную, что повышает вариативность;
- Масштабируемость подхода ограничена лишь функционалом выбранного алгоритма коллаборативной фильтрации.

К недостаткам подхода относится то, что при высокой уникальности каждого из признаков, размерность полученной матрицы и, соответственно, её разреженность возрастает. В этом случае возникает необходимость применения методов снижения размерности [11].

Помимо описанной концепции, существуют другие способы преобразования конечного набора признаков в числовой формат [12]:

Dummy-кодирование – заключающееся в преобразовании изначального набора признаков в двоичный формат;

Произвольная перенумерация – редко применяющийся метод кодирования, заключающийся в присвоении каждому признаку случайного числа [13];

Перекодирование частотой встречаемости (КЧВ). Каждая пара вариаций признаков a_1 и a_2 объекта x_i перекодируется в соответствии с частотой совместной встречаемости.

Аппроксимация с помощью взвешенных разреженных разложений матриц. Под «взвешенностью» понимается, что на степень разреженности данных влияют пропущенные значения, а не нулевые. Благодаря этому становится возможным восстановление даже неизвестных значений меток. Сама перекодировка заключается в использовании латентных векторов значений. В этом случае значение a_k соответствует признаку t_1 , значение b_l - признаку t_2 . Найденные значения будут представлены, соответственно, строкой матрицы G и столбцом матрицы H [19]. В результате каждый признак будет перекодирован новыми $2r$ значениями, благодаря чему становится возможным более глубокое исследование закономерностей в данных.

Рассмотрим кластеризацию данных [14] на основе dummy-кодирования. С этой целью проанализируем наиболее распространённые алгоритмы кластерного анализа, применимые в рамках поставленной задачи.

Наиболее известным и одновременно вычислительно простым является **алгоритм k-средних** (k-means) [19]. Разбиение набора $X = \{x_1, x_2, \dots, x_n\}$, $x_i \in R^d$, $i = 1, \dots, n$ на k кластеров $S = \{S_1, S_2, \dots, S_k\}$ осуществляется путём нахождения таких значений центров масс $\mu = \{\mu_1, \mu_2, \dots, \mu_k\}$ кластеров, при которых минимизируется сумма квадратов расстояний $\rho(x, \mu_i)^2$ от каждой точки кластера до его центра:

$$\arg \min_{\mu} \sum_{i=1}^k \sum_{x \in S_i} \rho(x, \mu_i)^2 \quad (8)$$

На каждой итерации рассчитывается расстояние от каждого объекта x до каждого из центров μ , и наиболее близкие к ним объекты формируют кластеры S . Затем пересчитываются значения новых центров масс для каждого кластера:

$$\mu_i = \frac{1}{S_i} \sum_{x_j \in S_i} x_j \quad (9)$$

Условием остановки алгоритма считается отсутствие изменений состава кластеров и, соответственно, центров масс.

Иерархическая кластеризация представляет собой методику кластерного анализа, при которой кластеры вложены друг в друга. Можно предположить, что изначально все объекты входят в один кластер и далее он делится на более мелкие, или же наоборот, обозначить каждый объект в качестве отдельного кластера и далее объединять схожие в более крупные кластеры. В зависимости от выбора между этими двумя подходами различают, соответственно, **агломеративную и дивизионную иерархическую кластеризацию**. Расстояние в таких методах определяется по формуле Ланса-Уильямса [21]:

$$R(U \cup V, S) = \alpha_U R(U, S) + \alpha_V R(V, S) + \beta R(U, V) + \gamma |R(U, S) - R(V, S)| \quad (10)$$

Где U, V, S – точки; $R(U, S)$ – расстояние между двумя точками U и S ; α, β, γ – коэффициенты.

При определенных значениях коэффициентов уравнение приобретает вид аналогичный уравнениям других известных способов расчета расстояний между кластерами. К примеру, расстояние ближайшего соседа (минимальное из расстояний между объектами из двух кластеров):

$$R^0(W, S) = \min_{w \in W, s \in S} \rho(w, s) \quad (11)$$

$$a_U = a_V = \frac{1}{2}; \beta = 0; \gamma = -\frac{1}{2}$$

Или расстояние между центрами кластеров:

$$R^2(W, S) = \rho^2 \left(\sum_{w \in W} \frac{w}{|W|}, \sum_{s \in S} \frac{s}{|S|} \right), \quad (12)$$

$$a_U = \frac{|U|}{|W|}; a_V = \frac{|V|}{|W|}; \beta = -a_U a_V; \gamma = 0$$

Ещё одним распространённым алгоритмом кластерного анализа является плотностный алгоритм пространственной кластеризации с присутствием шума **DBSCAN** [18] (Density-Based Spatial Clustering of Applications with Noise). Алгоритм оперирует плотностью данных, которая характеризуется радиусом окрестности ϵ и количеством соседей m . Задав симметричную функцию расстояния $\rho(x, y)$, выделяется область $E(x)$, для которой $\forall y: \rho(x, y) \leq \epsilon_x$. Если объект x имеет в пределах своего радиуса m объектов, то такой объект называется корневым или ядерным:

$$|E(x)| \geq m \quad (13)$$

Объект p непосредственно плотно-достижим из объекта q , если $p \in E(q)$ и q – корневой объект, таким образом p войдет в тот же кластер, что и q .

Численная оценка эффективности гибридных РС

Произведём расчёт сформулированных выше коэффициентов сходства для наборов кластеров с эталонным разбиением. Результирующие оценки работы алгоритмов на dataset'ax UCI [14] представлены в виде гистограмм на рисунке 3. Видно, что четыре рассмотренных алгоритма кластеризации демонстрируют близкие результаты.

Существует широкий спектр различных вариантов гибридизации рекомендательных систем, различающихся как по способу обработки исходных данных, так и по вычислительной сложности. В таблице 3 представлены распространенные методы гибридизации. [17]

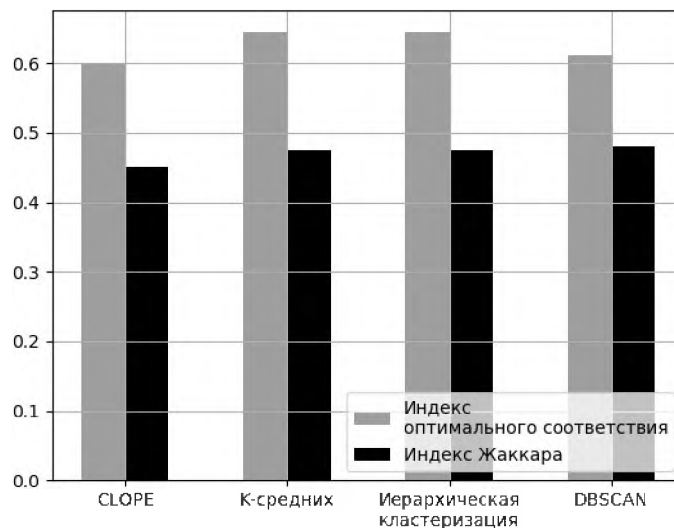


Рис. 3. Гистограмма коэффициентов сходства работы алгоритмов кластерного анализа для данных [14]

Таблица 3. Методы гибридизации рекомендательных систем

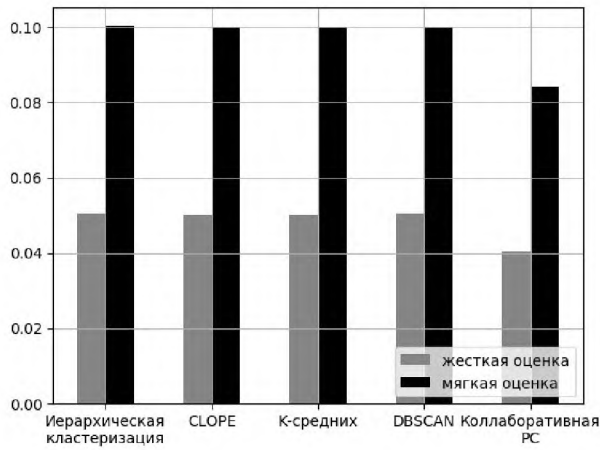
Метод	Краткое описание
Weighted	Комбинирование результатов нескольких независимых друг от друга РС
Switching	Выбор одного из результатов работы РС относительно некоторого порога
Mixed	Список рекомендаций формируется частично несколькими РС
Cascade	Поэтапное применение нескольких РС

Для оценки качества рекомендаций рассматривался dataset, уже содержащий в себе рейтинговые и признаковые таблицы – movielens-100k [6]. В наборе содержится 100 тысяч оценок для 1682 объектов, выставленных 943 пользователями. Признаками пользователей является информация о возрасте, поле и профессии. Для оценки качества рекомендаций рейтинговые таблицы разбивают на обучающую и тестовую выборки. Результаты кластеризации представлены в виде гистограмм коэффициентов точности и полноты [22] сформированных с их помощью рекомендаций (рис. 4).

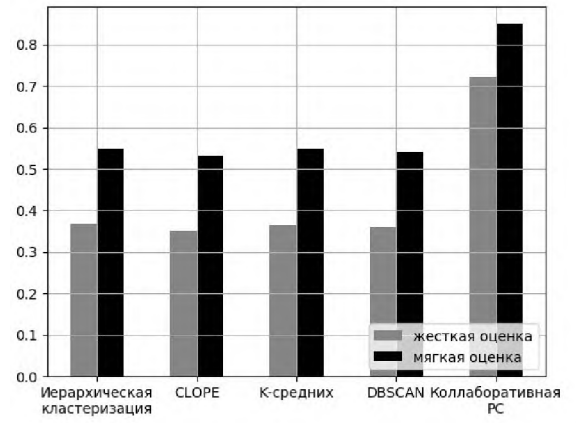
В алгоритмах на базе методов *weighted*, *switched* и *mixed* рекомендации строились на основе оценок других пользователей на базе одного и того же кластера. На этом этапе для каждого из рассмотренных выше алгоритмов кластеризации: *CLOPE*, *k-means*, *agglomerative* и *DBSCAN* формировались четыре матрицы сходства. Выбор между рекомендациями, полученными с помощью коллаборативной фильтрации и кластеризации с помощью алгоритма *weighted*, производился на основе наличия объекта рекомендации в обеих сформированных выборках. Поскольку часть объектов при расхождении оценок более чем на единицу приводит к пропуску текущего объекта в конечной рекомендательной выборке, это снижает показатель «полнота» (recall), что и видно из рисунка 4-б. Тем не менее рекомендации, сгенерированные данным алгоритмом, обладают большей точностью (precision) по сравнению с обыкновенной коллаборативной РС (рис.4-а). В алгоритме *mixed* осуществлялось смешение результатов коллаборативной фильтрации и кластерного анализа в соотношении 1 к 4. В результате 80% рекомендаций были получены на основе результатов обработки рейтинговой таблицы, а 20% - на основе пользовательской. Результаты соответствующих этому алгоритму представлены на рис. 4-в и 4-г.

Алгоритм *switched* предполагает переход между алгоритмами построения рекомендаций в гибридной РС на основе некоторого правила, выбор которого зависит от полноты исходных данных и определяется спецификой поставленной задачи. Таким правилом может быть отсутствие оценок по конкретному объекту у пользователей, схожесть с которыми по величине индекса Жаккара не превышала 0,2. Результаты, соответствующие этому алгоритму, представлены на рис. 4-д и 4-е.

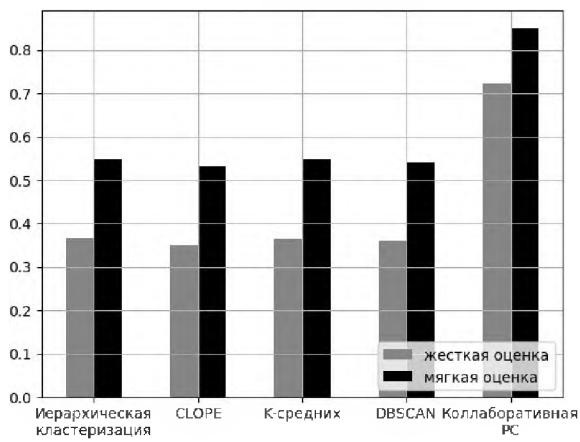
В отличие от предыдущих алгоритм *cascade* не подразумевает независимого составления рекомендаций частями гибридной системы. Все входящие в него алгоритмы применяются последовательно. На первом этапе набор пользователей делится на кластеры по признакам. На втором этапе на основе рейтингов в рамках каждого кластера рассчитывается коэффициент сходства. Результаты, соответствующие этому алгоритму, представлены на рис. 4-ж и 4-з.



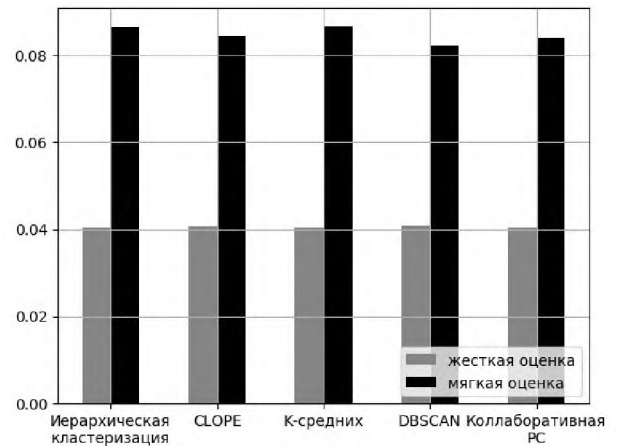
(a)



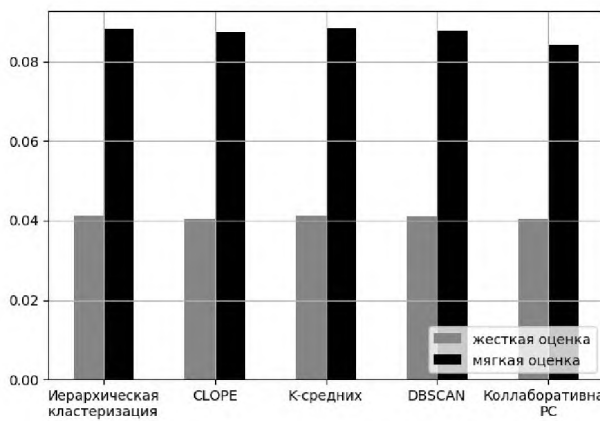
(г)



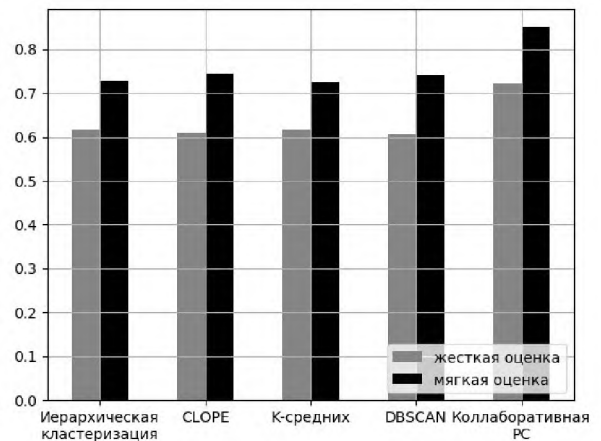
(б)



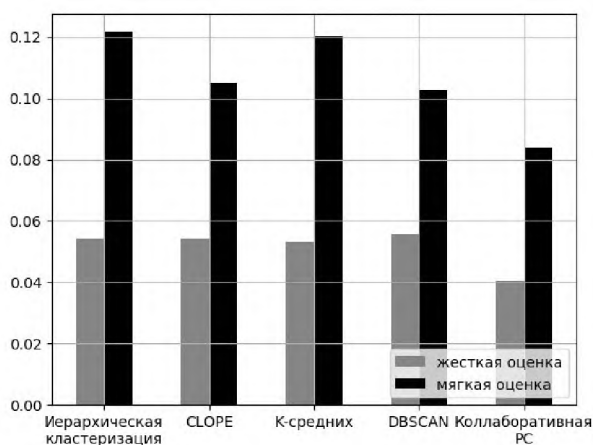
(д)



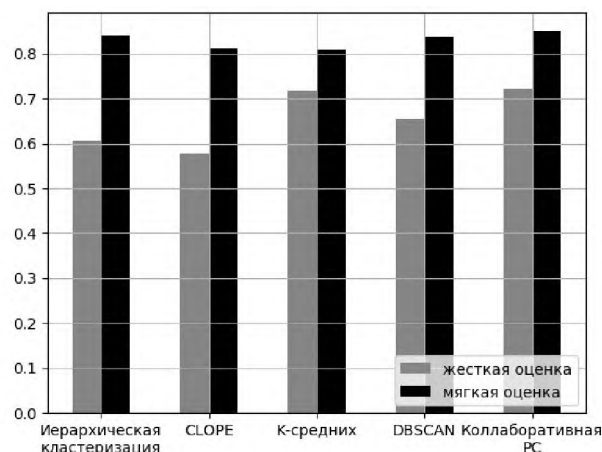
(в)



(е)



(ж)



(з)

Рис. 4. Гистограммы коэффициентов точности и полноты гибридных алгоритмов коллаборативной фильтрации (а - точность *weighted*, б – полнота *weighted*, в – точность *mixed*, г - полнота *mixed*, д – точность *switched*, е – полнота *switched*, ж – точность *cascade*, з – полнота *cascade*)

Сравнение гистограмм, приведенных на рисунках 4-а ...4-з показывает, что наибольшей точностью (precision) обладает метод гибридации РС типа *cascade* (рис.4-ж). Незначительно уступает ему метод *weighted*. Характеристики методов гибридации *mixed* и *switched* наихудшие.

Заключение

При относительно небольшом объеме исходных данных гибридная, подразумевающая обработку двух или более независимых наборов данных для формирования одного конечного набора рекомендаций в параллельном режиме, может привести к увеличению затрат ресурсов без заметного прироста эффективности. Вместе с тем, методы гибридации, предполагающие последовательную обработку двух или более наборов исходных данных, могут существенно повысить точность (precision) рекомендаций, сохранив при этом их полноту (recall).

Результаты работы гибридных РС практически не зависят от выбора алгоритма кластеризации.

Учет дополнительных признаков пользователей в коллаборативных РС способствует улучшению качества рекомендаций за счёт дополнительной информации. Для составления рекомендаций, на основе сходства с другими пользователями становится возможным создание индивидуальных рекомендаций для новых пользователей [23], не обладающих достаточным объёмом оценочных данных по таблице «пользователь-элемент». В результате частично решается проблема разреженности данных, поскольку в РС появляется минимальный объём информации, достаточный для формирования начальных рекомендаций.

Литература

1. Vipul Vekariya, G. R. Kulkarni, Hybrid Recommender Systems: Content-Boosted Collaborative Filtering for Improved Recommendations, IEEE, International Conference on Communication Systems and Network Technologies (CSNT), 2012, pp. 649-653.
2. Chaturvedi A.K. Recommender System for News Articles using Supervised Learning, Universitat Pompeu Fabra, Barcelona, 2017, p. 8.

3. *F. Ricci, L. Rokach, B. Shapira and P.B. Kantor*, Recommender Systems Handbook, Springer Science + Business Media, 2011, 845 p., pp. 39-46.
4. *Четикова Е.Д., Савкова Е.О., Привалов М.В.* Исследование алгоритмов рекомендательных систем. Информатика и Кибернетика, Донецк: ДонНТУ, 2016. С. 57-61.
5. *G. James, D. Witten, T. Hastie & R. Tibshirani*, An Introduction to Statistical Learning with Applications in R, Springer Science + Business Media, 2014, 441p., pp. 385-400.
6. Dataset MovieLens-100k. URL: <https://grouplens.org/datasets/movielens/100k/> (дата обращения: 06.11.2017).
7. *Н. Паклин*, «Кластеризация категориальных данных: масштабируемый алгоритм CLOPE», BaseGroup Labs, 2006. URL: <https://basegroup.ru/print/203> (дата обращения: 12.10.2017)
8. Working With Text Data, // URL // http://scikit-learn.org/0.15/tutorial/text_analytics/working_with_text_data.html (дата обращения: 25.11.2017).
9. word2vec. URL: <https://code.google.com/archive/p/word2vec/> (дата обращения: 25.11.2017).
10. *Рысьмятова А., Дьяконов А.*, ВКР «Использование сверточных нейронных сетей для задачи классификации текстов», МГУ, 2016.
11. *Колесников Е., Абгарян К.*, Курсовой проект «SVD-разложение и его практические применения», МАИ, 2015.
12. *А.Ю. Фонарёв, А.Г. Дьяконов.* ВКР «Машинное обучение с категориальными признаками», МГУ им. Ломоносова, 2014, С. 6-13.
13. *Ben Rogojan*. «Boosting and Bagging: How To Develop A Robust Machine Learning Algorithm», 2017. URL: <https://hackernoon.com/how-to-develop-a-robust-algorithm-c38e08f32201> (дата обращения: 05.02.2018).
14. UCI Machine Learning Repository. URL: <https://archive.ics.uci.edu/ml/datasets.html?format=&task=&att=cat&area=&numAtt=&numIns=&type=&sort=nameUp&view=table> (дата обращения: 05.02.2018).
15. *Бильгаева Л.П., Самбялов З.Г.* «Один способ кластеризации категориальных данных с неполным обучением и его оценка // Фундаментальные исследования.» 2013. № 11-8. С. 1571-1575. URL: <http://www.fundamental-research.ru/ru/article/view?id=33381>
16. *I.E. Vorontsov, I.V. Kulakovskiy, V.J. Makeev*, «Jaccard index based similarity measure to compare transcription factor binding site models», Algorithms Mol Biol, 2013, doi: 10.1186/1748-7188-8-23.
17. *Emeli Dral*, «Гибридные рекомендательные системы», МФТИ. URL: <https://ru.coursera.org/learn/data-analysis-applications/lecture/C9Sn4/gibridnyie-riekomendatelielnyie-sistemy> (дата обращения: 13.04.2018).
18. *П. Садовников*, «Интересные алгоритмы кластеризации, часть вторая: DBSCAN», 2017 // URL // <https://habr.com/post/322034/> (дата обращения: 21.09.2017).
19. *Y. Koren*, Factorization meets the neighborhood: a multifaceted collaborative filtering model, *Proceeding of the 14th ACM SIGKDD international conference on Knowledge discovery and data mining*, AT&T Labs – Research, 2008, pp. 426–434.
20. *Шелухин О.И., Авакян С.Т.*, Повышение эффективности коллаборативных рекомендательных систем методом гибридации, Сб. трудов XII Международной отраслевой научно-технической конференции «Технологии информационного общества» г. Москва, МТУСИ. 14-15 марта 2018. В 2-х томах. Том.1. М.: ИД «Медиа Паблишер», 2018. С. 391-392.
21. *Корнеева А.А., Чжан Е.А.*, «О компьютерной диагностике электрорадиоизделий», Интернет-журнал Науковедение. 2015. №6 (31). URL: <https://cyberleninka.ru/article/n/o-kompyuternoy-diagnostike-elektroradioizdeliy> (дата обращения: 21.09.2018).
22. *Р.И. Ролгин*, Метрики оценки качества работы систем коллаборативной фильтрации, СГАУ, Материалы Международной конференции и молодёжной школы «Информационные технологии и нанотехнологии». Самара, 2016. С. 1092-1095.
23. «Рекомендательная система: введение в проблему холодного старта». URL: <https://habrahabr.ru/company/surfingbird/blog/168733/> (дата обращения: 07.10.2017).

IMPROVING THE EFFICIENCY OF COLLABORATIVE RECOMMENDATION SYSTEMS WHEN USING ADDITIONAL DATA

Samvel T. Avakyan

Student of group M111701(73) MTUCI

avsam@rambler.ru

Oleg I. Sheluhin

Doctor of Technical Sciences,

Professor of Information Security, MTUCI

sheluhin@mail.ru

Keywords: machine learning, additional information, clustering, similarity and repulsion coefficients, user attributes, non-numeric characteristics.

The article describes one of the ways to achieve the effectiveness of the collaborative recommender system (RS). Such a solution is a hybrid [4] collaborative RS. A comparative analysis of data clustering and hybridization algorithms for collaborative RS has been carried out. The histograms of the coefficients of accuracy and completeness of the considered algorithms are constructed.

СИНХРОНИЗАЦИЯ В СЕТИ КАК ЭТАП ОБЕСПЕЧЕНИЯ БЕЗОПАСНОСТИ ДАННЫХ ТИПА “BIG DATA”

Чистов Анатолий Викторович

МТУСИ, магистрант группы М091801(71)

tolyanc@yandex.ru

Руднев Алексей Николаевич

МТУСИ, к.т.н., доцент кафедры ИБ,

alexrudnev@mail.ru

Ключевые слова: частотная синхронизация, временная синхронизация, XG-PON, целостность данных, пассивная оптическая сеть, NTP, PTP, SyncE.

Рассмотрены: основные рекомендации, регламентирующие синхронизацию и способы ее осуществления в сети XG-PON; основные способы обеспечения синхронизации в сети без использования дополнительных технологий и протоколов; методы пакетной временной синхронизации, а также метод частотной синхронизации SyncE. На основе анализа характеристик и возможностей данных методов сформулированы рекомендации по осуществлению временной и частотной синхронизации в сети XG-PON.

Big data – это огромные объемы данных, размер и разнообразие которых выходят за рамки возможностей традиционных средств управления данными, для их дальнейшего сбора и анализа. Источниками big data могут являться: различные датчики и устройства, журналы работы различных технических устройств, сигналы GPS мобильных телефонов, социальные сети, записи онлайн-транзакций. Именно поэтому важно, обеспечить необходимый уровень защиты этих данных не только во время обработки, когда они собраны воедино, но и на этапе непосредственно сбора данных.

Целостность и доступность данных является важным аспектом информационной безопасности. Обеспечение целостности и доступности информации начинается на этапе проектирования сети связи, важным аспектом при проектировании любой сети связи является обеспечение частотной и временной синхронизации. Рассмотрим обеспечение синхронизации в сети на примере сети XG-PON, которая является перспективной пассивной оптической сетью доступа, пока еще не имеющая большого коммерческого применения, однако возрастающие объемы передачи информации, позволяют судить о перспективности развертывания данных сетей на территории Российской Федерации.

10-Gigabit capable passive network или XG-PON является одним из самых последних стандартов оптических сетей доступа. Главной задачей при реализации стандарта XG-PON являлось обеспечение большой пропускной способности, для обеспечения постоянно возрастающих потребностей в широкополосном доступе к различным видам услуг. Технология XG-PON обеспечивает скорость передачи 10 Гбит/с в нисходящем направлении (прием) и 2,5 Гбит/с в восходящем направлении (передача). XG-PON может работать вместе с технологией GPON используя одну и ту же оптическую распределительную сеть (ODN – Optical Distribution Network), то есть участок сети между оптическим терминалом линии (OLT – Optical Line Terminal) и оптическим терминалом сети (ONT – Optical Network Terminal), включающий в себя распределительный и абонентские участки кабеля, а также все необходимые соединители и пассивные устройства, так как XG-PON и GPON работают на разных длинах волн. Однако переход на новый стандарт не был таким успешным как ожидалось, многие провайдеры отказались от перехода на XG-PON аргументируя это тем, что требовалась замена электроники OLT и ONT, многие операторы так же аргументировали низким спросом на высокую пропускную способность [1]. На рисунке 1 представлена архитектура сети XG-PON.

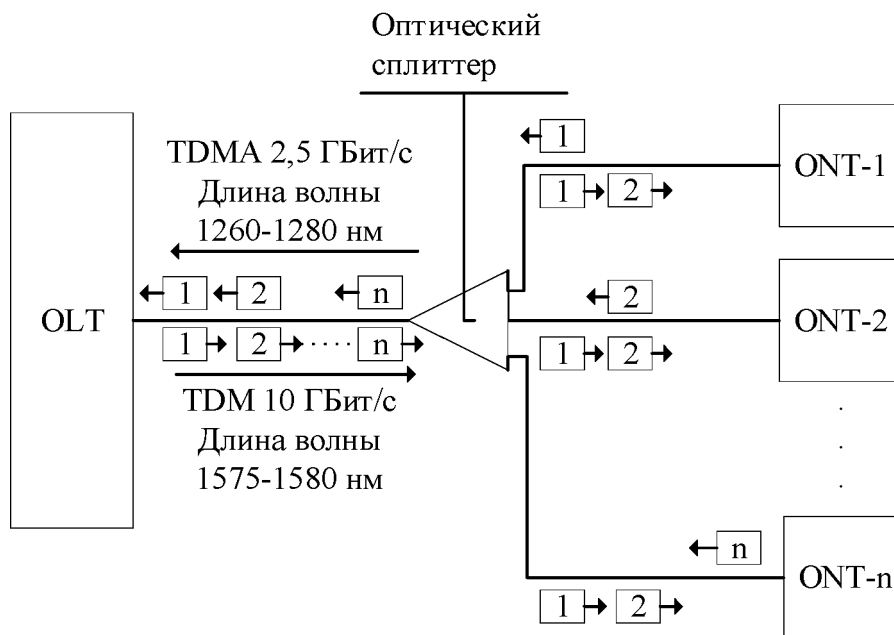


Рис. 1. Архитектура XG-PON

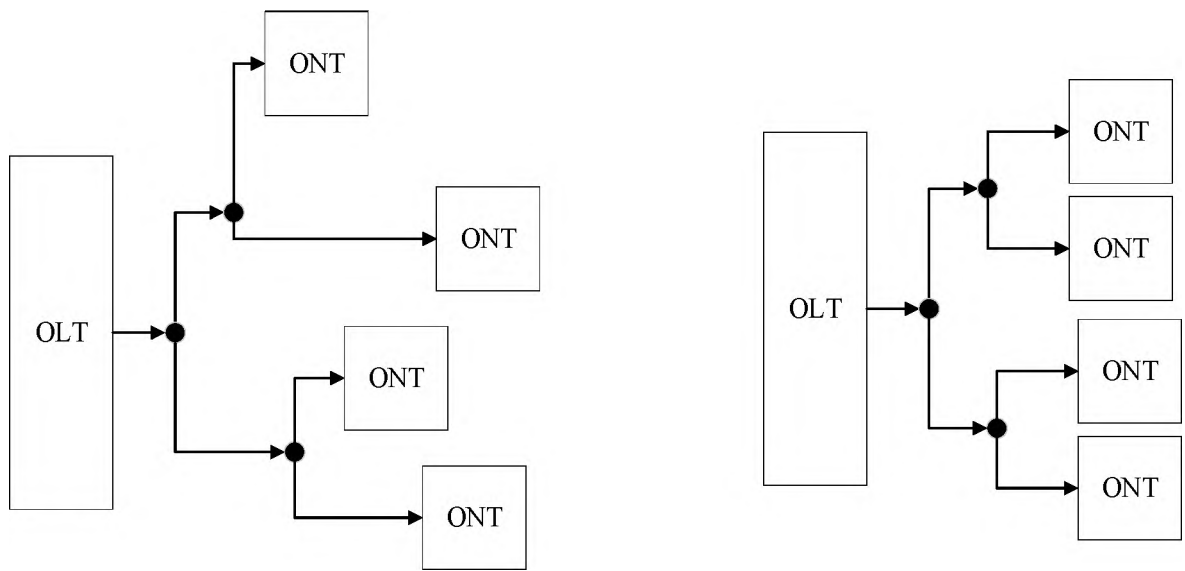
Отсутствие частотной синхронизации приводит к неправильному принятию битовых последовательностей информации, что является прямой угрозой целостности и доступности информации. Отсутствие временной синхронизации может привести к неправильной тарификации мобильной связи, в случае проведения экспериментов, где точность времени играет важную роль, а информация принимается от большого количества датчиков, физически удаленных друг от друга, может привести к неправильной интерпретации результатов эксперимента, что является угрозой целостности информации. К примеру, при экспериментах с использованием станций сотовой связи.

Существует несколько возможных подходов к синхронизации в сетях связи: плезиохронный – все устройства в сети подключены к эталонным источникам и обмена синхросигналами нет, взаимный – все устройства в сети обмениваются друг с другом синхросигналами, ведомый – ведущий, когда одно более точное устройство передает синхросигналы менее точным устройствам, и смешанный, являющийся сочетанием взаимного и ведомо-ведущего подходов [2].

Базовые способы синхронизации в сети XGPON

Ввиду топологии пассивной оптической сети P2MP (point-to-multipoint) точка-многоточка при передаче информации в восходящем потоке возможно появление коллизии (наложение кадров друг на друга), однако данная проблема была решена еще при построении сетей GPON. Передача информации в восходящем направлении осуществляется с использованием множественного доступа с разделением по времени (TDMA – Time Division Multiple Access). Синхронизация OLT и ONT требуется для поддержания необходимого выравнивания кадров что в свою очередь необходимо для постоянной скорости передачи битов (CBR – Constant Bit Rate) в восходящем направлении.

Для предотвращения коллизии в восходящем потоке так же применяется процедура ранжирования (ranging). Суть данной процедуры заключается в том, что все устройства ONT размещаются на одном виртуальном расстоянии, главная цель данной технологии в том, чтобы задержать передачу с наиболее близких узлов, таким образом что бы информация из двух любых узлов прибывала на OLT “одновременно”. Пример процедуры ранжирования представлен на рисунке 2.



До процедуры ранжирования:
От OLT до ONT физически разные
дистанции

После процедуры ранжирования:
От OLT до ONT виртуально одинаковое
расстояние

Рис. 2. Процедура ранжирования

TDMA используется в большом количестве приложении, включая сотовую связь и кабельные линии. Главным аспектом правильной работы TDMA является временные интервалы (time slots) или окна передачи, создаваемые и поддерживаемые с помощью OLT. OLT контролирует поток информации, поступающий от различных ONT/ONU с целью недопущения коллизий OLT, предоставляет так называемые гранты или проще говоря разрешения на использование определенного временного интервала для каждого отдельного ONT. Правильное выравнивание временных интервалов является главным требованием к синхронизации восходящего трафика TDMA. Графическая иллюстрация данного способа представлена на рисунке 3.

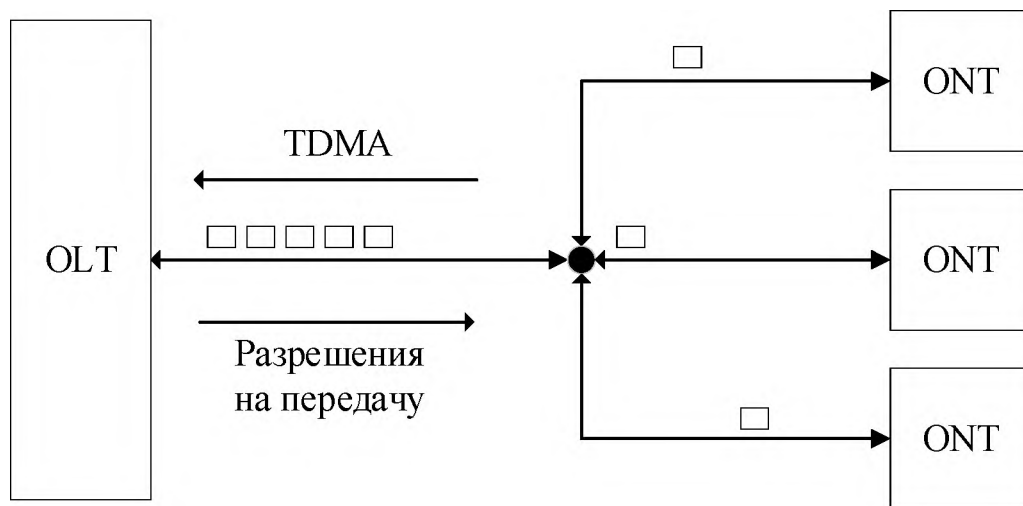


Рис. 3. Процедура предоставления временных интервалов.

Белыми квадратами на рисунке показан восходящий трафик, идущий от ONT к OLT в разные временные интервалы и объединяющийся в единый поток без наложений в оптическом сплиттере (черная точка на рисунке 3).

Режим пакетной синхронизации необходим для выравнивания скоростей передачи бит между ONT и OLT в восходящем направлении. Режим пакетной синхронизации — это

адаптивный метод восстановления времени который позволяет ONT восстанавливать время из первых нескольких служебных битов.

Длительность кадра XG-PON составляет 125 мкс, что совпадает с длительностью кадра SONET/SDH. Соответственно ONT могут восстанавливать частоту 8 КГц по аналогии с SONET/SDH на физическом уровне. С точным и стабильным эталоном синхронизации доступным OLT, данный метод должен обеспечивать большую стабильность чем метод пакетной синхронизации. Работа автомата синхронизации [3] представлена на рисунке 4.

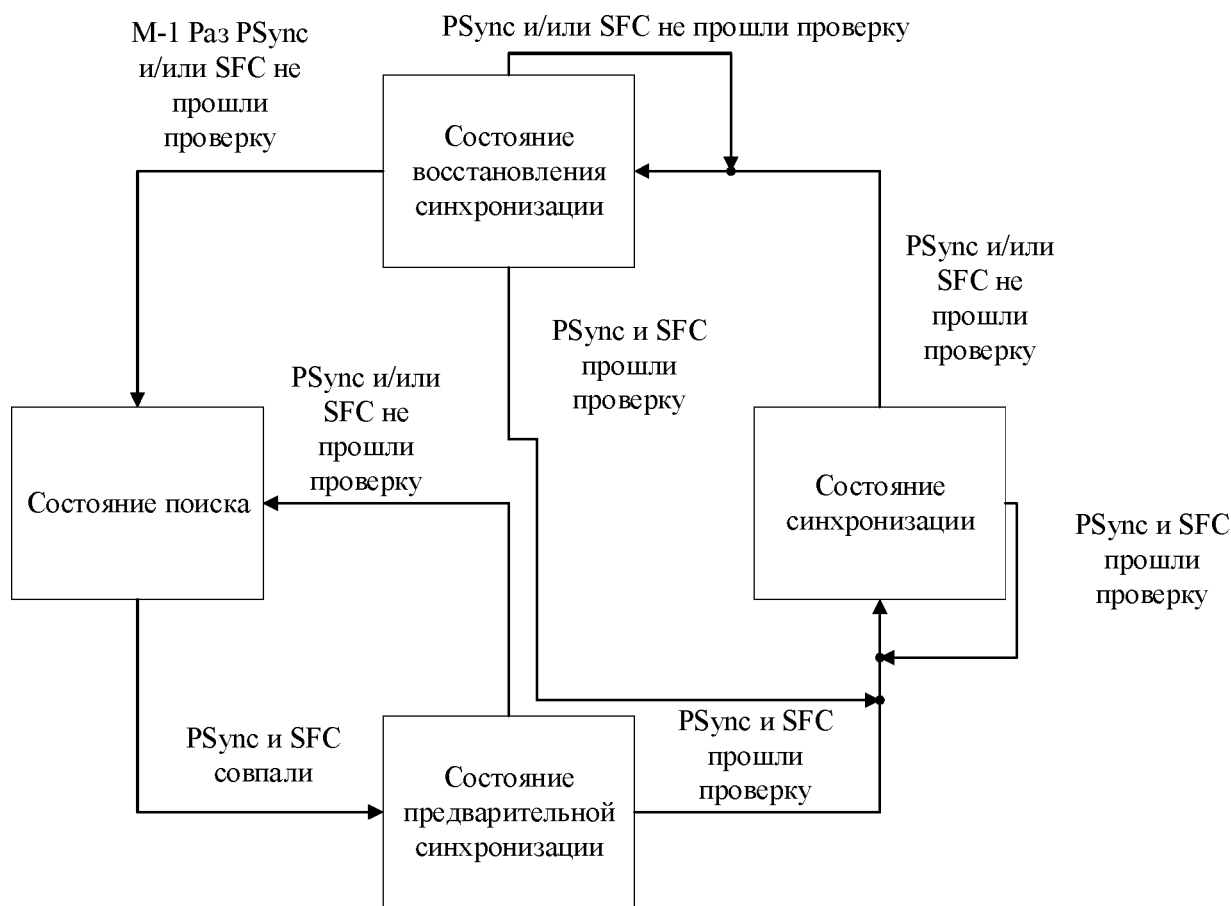


Рис. 4. Эталонный автомат конечной синхронизации ONU в нисходящем потоке.

Технология SyncE

Важным аспектом является то что для синхронизации частоты необходима синхронизация между двумя соседними устройствами, для синхронизации времени необходимо, чтобы на всех устройствах в сети было одинаковое время.

Сети Ethernet – в стандартной реализации являются несинхронными. Учитывая, что XG-PON поддерживает технологию Ethernet, вариантом для решения проблемы синхронизации в сети может служить применение технологии синхронного Ethernet [4]. Главный принцип данной технологии заключается в том, что сигналы синхронизации передаются на физическом уровне сети Ethernet. В таком случае опорный сигнал синхронизации, который возможно отследить до первичного эталонного генератора (PRC – Primary Reference Clock), вводится с помощью внешнего порта коммутатора Ethernet. Перед передачей сигнала синхронизации коммутатор Ethernet обрабатывает сигнал функцией синхронизации. Данная функция необходима для проведения фильтрации сигнала и при необходимости для удержания синхронизации в сети. Принцип работы генераторного оборудования синхронного Ethernet представлен на рисунке 5.

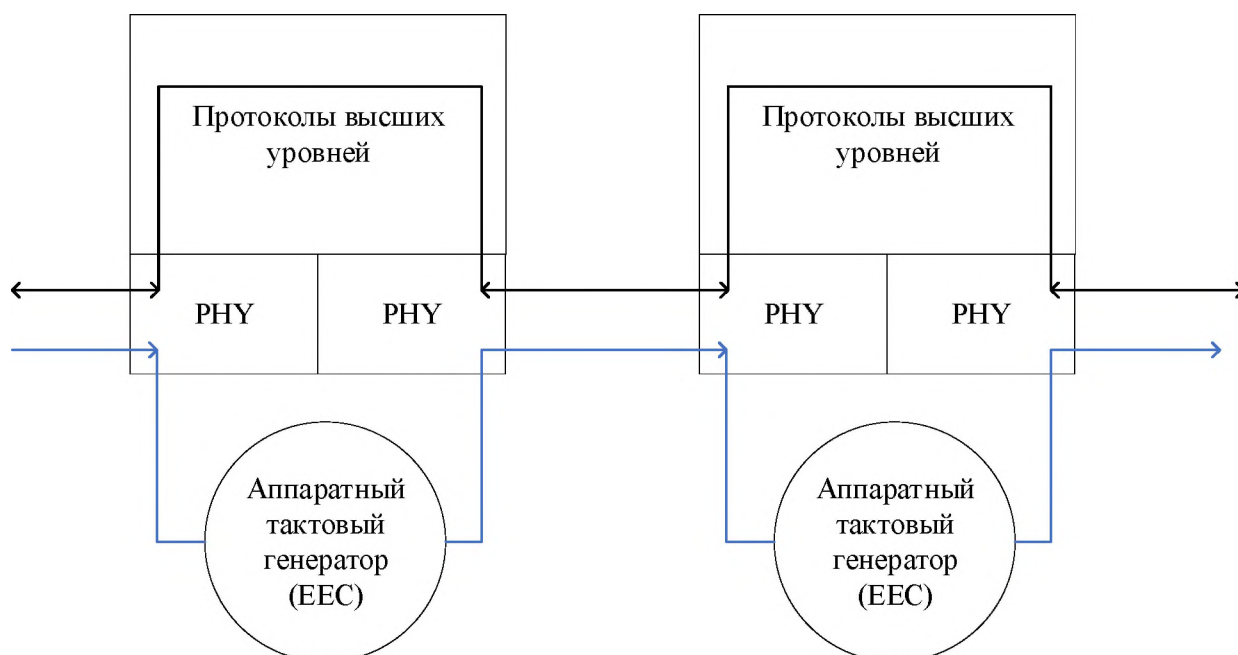


Рис. 5. Принцип работы генераторного оборудования SyncE

Тактовый генератор, поддерживающий синхронные сети Ethernet, называется аппаратным тактовым генератором синхронного Ethernet или ЕЕС (Ethernet Equipment Clock), данный генератор имеется у каждого сетевого элемента, поддерживающего режим синхронного Ethernet. Как правило, ЕЕС синхронизируется по тактам одного из входящих линейных сигналов, и в свою очередь синхронизирует все исходящие линейные сигналы.

Пакетная синхронизация NTP РТР

В случае, если синхронизация на физическом уровне невозможна, пакетные методы синхронизации являются единственным способом обеспечения синхронизации в сети [4].

Для синхронизации времени существуют специальные протоколы: NTP (Network Time Protocol) – сетевой протокол времени, РТР (Precision Time Protocol) – протокол точного времени. Данные протоколы позволяют передавать информацию о текущем времени через сеть передачи данных. Сравнение данных протоколов представлено в таблице 1.

Таблица 1. Сравнение основных технологий пакетной синхронизации

Критерий	NTP	РТР
Пиковая ошибка передачи времени	Более 1мс	Более 100 нс
Основной источник ошибок	Роутеры	Роутеры, разветвители, порты, задержка очереди, сеть и так далее
Реализация	Реализация сервера на аппаратных средствах или программном обеспечении, реализация клиента с помощью программного обеспечения	Ведущий сервер реализуется аппаратными средствами, ведомый сервер реализуется аппаратными средствами или программным обеспечением
Род работы	Клиенты получают время от сервера	Ведущий сервер раздает время ведомым
Поддержка на линии	Не существует и невозможна	Не требуется, но возможна через граничные и прозрачные часы (повышает производительность)
Относительная стоимость	Недорого	Более дорогая (более точные методы синхронизации стоят дороже)
Метрики, мониторинг и управление	Существуют, но минимально	Большие показатели диапазона для мониторинга и управления

Заключение

В ходе проведенного анализа рекомендаций МСЭ-Т и технической литературы выявлено, что частотная синхронизация в сети XG-PON может быть осуществлена с помощью циклов передачи XG-PON или с помощью технологии синхронного Ethernet (SyncE). Однако стоит заметить, что обеспечение частотной синхронизации в сети с помощью синхронного Ethernet требует установки дополнительного оборудования, что ведет к дополнительным материальным затратам и является не обязательным при наличии внутренней системы синхронизации, описанной в рекомендации G.987.3.

Временная синхронизация может быть обеспечена с помощью протоколов NTP и PTP. В данном случае важную роль играют требования к временной синхронизации в сети, протокол NTP является менее точным. Если все сетевые устройства подключены к синхронизирующему оборудованию, использующему протокол NTP, пиковая ошибка времени может достигать до 1мс. В тех случаях, когда точность времени играет важную роль рекомендуется использовать протокол PTP. Данный протокол обеспечивает точность до 100нс, однако требует установки дополнительного типа на каждое сетевое устройство в сети.

Литература

1. *Sanchez-García, Juan & Plueckebaum, Thomas.* (2016). GPON and TWDM-GPON in the context of the wholesale local access market.
2. *Стефано Брени.* Синхронизация цифровых сетей связи. М.: Мир, 2003. 417 с.
3. G.987.3: 10-Gigabit-capable passive optical networks (XG-PON): Transmission convergence (TC) layer specification.
4. G.8261: Timing and synchronization aspects in packet networks

SYNCHRONIZATION IN THE NETWORK AS A STAGE OF SECURITY “BIG DATA”

Chistov Anatoly Viktorovich

*MTUCI, Master student of group M091801(71)
tolyanc@yandex.ru*

Rudnev Alexey Nikolaevich

*MTUCI, Ph.D., assistant professor of IS department,
alexrudnev@mail.ru*

Key words: *frequency synchronization, time synchronization, XG-PON, data integrity, passive optical network, NTP, PTP, SyncE.*

The paper reviewed: *the main recommendations concerning XG-PON; The main ways to ensure synchronization in the network without using additional technologies and protocol; methods of packet time synchronization, as well as the method of frequency synchronization SyncE. Based on the analysis of the characteristics and capabilities of the protocols, recommendations were formulated regarding time and frequency synchronization in the XG-PON network.*

ИССЛЕДОВАНИЕ СУЩЕСТВУЮЩИХ СИСТЕМ ОБФУСКАЦИИ С ЦЕЛЬЮ ЗАЩИТЫ ПРОГРАММНЫХ АЛГОРИТМОВ

Зайцев Евгений Сергеевич
студент группы М091801(72) МГУСИ
z.evg.serg@gmail.com

Сераджи Сани Исмаилович
студент группы М091801(72) МГУСИ
snowdulum@gmail.com

Кальфа Александр Алексеевич
д.ф-м.н., профессор кафедры МСиУС МГУСИ
kalfa.alex@yandex.ru

Беленькая Марина Наумовна
доцент кафедры МСиУС МГУСИ
mn.belenkaya@mail.ru

Ключевые слова: обфускация, виды обфускации, методы обфускации, защита исходного кода, запутывание исходного кода.

Индустриальное программирование сопровождается необходимостью защиты разработанного исходного кода с целью предотвращения возможности кражи злоумышленником реализованных наработок и защиты интеллектуальной собственности программного обеспечения, в том числе программных алгоритмов. Одним из способов такой защиты исходного кода является его запутывание (обфускация). В статье рассматриваются существующие системы обфускации для различных языков программирования.

Обфускация – это такое изменение исходного кода программы, при котором сохраняется ее поведение, однако реверсивная инженерия или восприятие человеком измененного кода становится крайне затруднительно [1].

Основными сферами применения обфускации к исходному коду являются следующие:

1. Защита интеллектуальной собственности программного продукта (например, используемых в нем алгоритмов);
2. Улучшение безопасности исходного кода;
3. Предотвращение несанкционированного доступа;
4. Защита от фальсификации программного продукта.

Обфускации может быть подвергнут любой исходный код, написанный на поддерживаемом обфускатором языке программирования. В качестве примера можно привести программы-архиваторы - системы, предназначенные для сжатия данных с целью экономии ресурсов систем хранения. Алгоритмы сжатия, применяемые в таких продуктах, обладают такими показателями как скорость, степень и качество сжатия и представляют ценность для компании, в связи с чем недопустим факт утечки исходных алгоритмов. При применении обфускации к исходному коду программы-архиватора снижается риск кражи алгоритмов и повышается устойчивость к реверсивной инженерии данного программного продукта, что позволяет уверенно распространять его копии в коммерческих целях.

С распространением программного обеспечения на клиентские машины владельцы теряют контроль над своим программным обеспечением. В этом случае клиентские программы подвержены атакам со стороны вредоносного хоста. Это может быть пользователь со злым умыслом или другое вредоносное программное обеспечение на хосте. Примером такого рода атак может быть пользователь, который пытается взломать загруженное приложение. В этом контексте «взломать» может означать извлечение секретных ключей, кражу интеллектуальной

собственности (например, алгоритмы) или внесение злонамеренных изменений в код приложения, например, использовать его без оплаты. Некоторые методы, такие как анализ неисправностей или поиск сохраненных или встроенных ключей, легко применимы в этой среде и очень эффективны, когда программы не защищены дополнительно. Хотя защита от вредоносного кода или вредоносного хоста - это разные проблемы, стороны часто используют сходные методы для достижения тех же целей. Авторы вредоносных программ обычно хотят исключить анализ и идентификацию своего кода. Они достигают этого, используя методы запутывания. Те же самые методы могут использоваться законными программистами для защиты их программного обеспечения от анализа. Делая это, они замедляют глобальную атаку на программное обеспечение, то есть атаку, использующую все экземпляры целого «семейства» программ. Рисунок 1 иллюстрирует четыре фазы этой атаки [2].

Взломанные копии

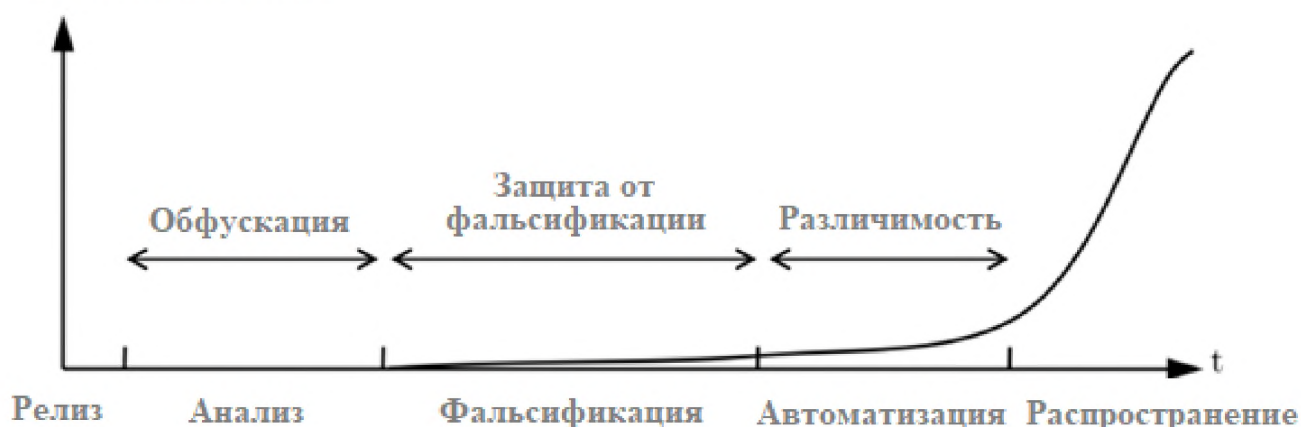


Рис. 1. Взлом программного приложения обычно подразумевает в последовательном порядке: анализ, фальсификацию, автоматизацию атаки и распространение «трещины» для запуска глобальной атаки.

Во-первых, злоумышленник должен выполнить анализ, чтобы получить знания о внутренностях программы, прежде чем вмешиваться в нее. Когда вмешательство выполняется на локальном экземпляре программы, злоумышленник может автоматизировать свою атаку, чтобы она работала и на других экземплярах. И, наконец, атакующий код или взломанное приложение будет распространяться через Интернет, чтобы иметь глобальное влияние.

Недавно в 2013 году произошел прорыв с теоретической точки зрения. Был представлен первый алгоритм-обфускация программы-кандидата (то есть градуированное кодирование) для всех цепей и было показано, что он может достичь убедительного свойства безопасности: неразличимости. Идея состоит в том, чтобы кодировать схемы с помощью полилинейных карт. Это вдохновило многих последующих исследований, которые направлены на обеспечение обфускации подходов с доказуемой безопасностью. Рисунок 2 демонстрирует взрыв таких исследований запутывания с пунктирной линией [3].



Рис. 2. Распространенность статей, посвященных обфускации

Таким образом, проблема обфускации хоть и является давней, но не теряет своей значимости в силу своей актуальности, поэтому на рынке продолжают появляться все новые программные продукты, использующие различные методы обфускации исходного кода.

Существующие обфускаторы

В таблице 1 представлена сравнительная характеристика распространенных существующих систем обфускации по основным параметрам, включая используемые методы обфускации, поддерживаемые операционные системы, стоимость решения, используемые дополнительные методы защиты программного продукта, возможную интеграцию с различными IDE.

Как видно из таблицы, некоторые решения используют инкрементальный метод обфускации. Данный метод предполагает наличие обфусцированного исходного кода и наличие обновления/изменения для него. При необходимости обновить исходный код, потребуются заново обфусцировать весь код, что является крайне медленной операцией для больших программ. Инкрементальная обфускация призвана обеспечить совместимость между обновлениями и исходной обфусцированной версией кода. Данный метод будет использовать те же имена классов, методов, переменных, те же методы шифрования/дешифрования строк, которые применялись при обфускации исходной версии программы. Таким образом, потребуются обфускация лишь малой части кода, а не всего продукта в целом. Такая обфускация применяется при выходе обновлений программного продукта и так как она относится скорее к обеспечению совместимости двух версий исходного кода, данный метод не был вынесен в отдельный столбец.

Также из таблицы можно наблюдать наиболее распространенные дополнительные методы защиты, поставляемые вместе с решениями. Такими методами являются шифрование строк и водяные знаки. Помимо этого, большинство решений под фреймворк .NET имеют возможность интеграции со средой разработки Visual Studio.

Стоит заметить, что даже незначительная обфускация представления или управления (например, удаление только комментариев или вставка простого цикла в разные участки исходного кода) позволяет разработчикам указывать в описании системы обфускации использование того или иного метода обфускации. В связи с этим необходимо детально изучать каждый отдельный продукт и обращать внимание на конкретные изменения в коде, производимые данным продуктом. В противном случае защищенность исходного кода от реверсивной инженерии может оказаться минимальной.

Таблица 1. Сравнительная таблица решений

Название решения	Обфускация представления	Обфускация управления	Другие виды обфускации	Только под Windows	Поддерживаемые языки программирования	Стоимость	Бесплатная/пробная версия	Наличие поддержки	Дополнительные методы защиты
PreEmptive Solutions Dotfuscator			Инкрементальная		C#, VB .NET, Managed C++	Неизвестна			Шифрование строк, водяной знак
NTCore Phoenix Protector			-		.NET	-			-
Eazfuscator.NET			-		.NET	99\$+			Шифрование строк
Eziris .Net Reactor			Инкрементальная		C#, J#, F#, VB C++ Delphi .NET	179\$/279\$			Шифрование строк
Skater .NET Obfuscator			-		.NET	79\$+			Шифрование строк
SecureTeam Agile.NET obfuscator			Инкрементальная		.NET	795\$+			Шифрование строк
SecureTeam Jfuscator Code Protection			Инкрементальная		Java	599\$+			Шифрование строк
Semantic Designs obfuscator			-		C/C++ C# Java Ada PHP PLSQL	Неизвестна			Шифрование строк, водяной знак
Stunnix C and C++ Obfuscator			-		C/C++	539\$+			Водяной знак
StarForce C++ Obfuscator			-		C/C++	Неизвестна			Шифрование строк
GuardSquare ProGuard			Инкрементальная		Java	-			-

Разные решения могут поддерживать различные методы обфускации. Также системы обфускации, как правило, поставляются дополнительно с другими системами защиты программного продукта, например, такими, как шифрование строк, системы лицензирования, водяной знак. При совместном использовании данные системы обеспечивают наибольшую защиту программного продукта.

Наличие поддержки системы обфускации также является важным параметром, поскольку совершенствуются как методы обфускации, так и методы реверсивной инженерии, поэтому своевременное внесение коррективов в методы обфускации повышает защищенность программного продукта.

Из таблицы видно, что ряд решений ориентирован на обфускацию исходного кода, использующего фреймворк .NET. Фреймворк .NET является патентованной технологией корпорации Microsoft и официально рассчитан на работу под операционными системами семейства Microsoft Windows, поэтому существующие решения обфускации, рассчитанные под данный фреймворк, предполагают работу под данной операционной системой. Также данный фреймворк определяет круг языков программирования, на которых он используется: C++, C#, Visual Basic .NET, F#. Помимо обфускаторов, ориентированных под фреймворк .NET, также существует ряд обфускаторов, которые запутывают код на таких языках программирования, как JavaScript, C, C++, C#, Java, PHP, PLSQL, Ada, Visual Basic и других. Некоторые из них являются кроссплатформенными.

Были рассмотрены решения как с закрытым, так и открытым исходным кодом. Зная открытый исходный код и, соответственно, примененные методы и алгоритмы обфускации, можно разработать программу, которая будет реверсировать произведенные изменения и таким образом получить исходную версию обфусцированного данным решением программного продукта. Однако, учитывая многообразие существующих обфускаторов, предположить, что обфусцированный код был изменен каким-то конкретным из них, становится крайне затруднительно. Поэтому решения с открытым исходным кодом также высокоэффективны, поскольку они применяют те же методы обфускации, что и проприетарные решения, но крупные корпорации предпочитают зарекомендовавшие себя решения с закрытым исходным кодом.

Заключение

Были рассмотрены распространенные решения как под фреймворк .NET, так и независимые от него. Как проприетарные, так и свободные решения используют схожие методы обфускации и пользуются большим спросом среди разработчиков программных продуктов. В зависимости от преследуемых целей, может быть достаточно и решений с открытым исходным кодом, ибо их эффективность сопоставима с платными решениями. Для достижения наилучшей защищенности исходного кода, необходимо использовать как можно больше видов обфускации в программном продукте, а также дополнительные методы защиты, например, шифрование строк. Только в этом случае применение реверсивной инженерии будет наиболее затруднено. Однако следует учитывать тот факт, что обфускация снижает быстродействие программного продукта, и подвергать обфускации лишь наиболее важные участки исходного кода.

Литература

1. *Gregory, Wroblewski*. General Method of Program Code Obfuscation / Wroblewski. Gregory. Wroclaw, 2002. 112 p.
2. *J. Cappaert*. Code Obfuscation Techniques for Software Protection / PhD dissertation, Leuven University, 2012. 132 p.
3. *Hui Xu, Yangfan Zhou, Yu Kang, Michael R. Lyu*. On Secure and Usable Program Obfuscation: A Survey / [Электронный ресурс]. Режим доступа: <https://arxiv.org/pdf/1710.01139.pdf> (дата обращения: 17.03.2019)

OBFUSCATION SYSTEMS IN SOFTWARE ALGORITHMS PROTECTION

Evgeny S. Zaytsev

Student of group M091801(72), MTUCI

[*z.evg.serg@gmail.com*](mailto:z.evg.serg@gmail.com)

Sunny I. Seraje

Student of group M091801(72), MTUCI

[*snowdulum@gmail.com*](mailto:snowdulum@gmail.com)

Alexander A. Kalfa

MTUCI, Doctor of Physics and Mathematics

[*kalfa.alex@yandex.ru*](mailto:kalfa.alex@yandex.ru)

Marina N. Belenkaya

MTUCI, associate professor of MNaCS department

[*mn.belenkaya@mail.ru*](mailto:mn.belenkaya@mail.ru)

Keywords: *obfuscation, obfuscation system, obfuscation solution, source code security, source code confuse.*

Industrial programming is accompanied by the need to protect the developed source code in order to prevent the attacker from stealing the implemented developments and protecting the intellectual property of the software, including program algorithms. One way to protect the source code is to obfuscate it. The article discusses the existing obfuscation systems for various programming languages.

ОБЕСПЕЧЕНИЕ ОТКАЗОУСТОЙЧИВОСТИ ВС НА ОСНОВЕ СТАТИЧЕСКОГО ОТКАЗОУСТОЙЧИВОГО ПЛАНИРОВАНИЯ ЗАДАЧ

Кибедов Руслан Арифович
студент группы М091801(70) МТУСИ
kibedov.r96@gmail.com

Турута Евгений Николаевич
МТУСИ, д.т.н., профессор кафедры МКиИТ
eturuta@yandex.ru

Ключевые слова: отказоустойчивая кластерная вычислительная система, метод отказоустойчивого планирования задач, кластер высокой доступности, структура программного обеспечения отказоустойчивого кластера, алгоритм отказоустойчивого планирования задач в кластерной вычислительной системе.

Дано определение отказоустойчивого кластера. Представлен метод отказоустойчивого планирования задач в кластерной вычислительной системе. Определены основные ограничения на распределения задач в методе отказоустойчивого планирования задач. Рассмотрены два возможных варианта обеспечения высокой доступности кластера достигаемой за счёт использования концепции отказоустойчивого планирования задач в узлах вычислительной системы.

Реализация кластерной вычислительной системы является лучшим способом (на данный момент времени) обеспечения бесперебойной работы любых вычислительных, сетевых, информационных систем.

Отказоустойчивый кластер (англ. High-Availability cluster, HA cluster — кластер высокой доступности) — кластер (группа серверов), спроектированный в соответствии с методиками обеспечения высокой доступности и гарантирующий минимальное время простоя за счёт аппаратной избыточности[3]. Без кластеризации сбой сервера приводит к тому, что поддерживаемые им приложения или сетевые сервисы оказываются недоступны до восстановления его работоспособности. Отказоустойчивая кластеризация исправляет эту ситуацию, перезапуская приложения на других узлах кластера без вмешательства администратора в случае обнаружения аппаратных или программных сбоев. Процесс перезапуска известен как аварийное переключение. В рамках этого процесса программное обеспечение кластеризации может дополнительно настроить узел перед запуском приложения на нём (например, импортировать и смонтировать соответствующие файловые системы, переконфигурировать сетевое оборудование или запустить какие-либо служебные приложения)[1].

Задача обеспечения отказоустойчивости вычислительной системы формулируется следующим образом: требуется разработать и реализовать в системе рациональную процедуру статического перераспределения задач при отказах вычислительных модулей, т.е. такую, которая обеспечивала бы выполнение поставленных требований к отказоустойчивости системы, а также к значениям заданных частных показателей качества функционирования и показателей стоимости ВС при заданных временных и ресурсных ограничениях.

Для этого необходимо сформировать рациональное отказоустойчивое размещение задач (ОУ-РмЗ), а так же рациональные планы размещения задач (РЗ) для всех работоспособных s-состояний, т.е. такие, которые обеспечивают выполнение поставленных требований. Если ставится задача оптимизации значений каких-либо показателей, то ОУ-РмЗ и соответствующие планы РЗ называются оптимальными.

Далее будем обозначать эту процедуру РСПЗ/ОУ.

Отказонеустойчивую систему A^0 будем называть прототипной для создаваемой отказоустойчивой системы A^{FT} , если эти системы выполняют одно и тоже задание и для них одинаковы типы вычислительных модулей, общих ресурсов и системы связи, режим функционирования и начальное РЗ (если оно известно)[1].

Существует две важных для практики постановки задачи РСПЗ/ОУ:

1. Найти рациональную процедуру статического ПЗ/ОУ при условии, что известно множество H^0 вычислительных модулей (и, возможно, начальное РЗ D^0) для прототипной отказонеустойчивой системы, выполняющей известное задание $\Gamma = \Omega^E$, которое предстоит выполнять и отказоустойчивой системе;

2. Найти одновременно рациональную процедуру статического ПЗ/ОУ и множество H^{FT} вычислительных модулей, требуемое для реализации этой процедуры в отказоустойчивой ВС, полагая, что прототипной отказонеустойчивой системы не существует и, следовательно, H^0 и D^0 неизвестны.

В данной работе рассматривается первая постановка задачи. В рамках этой постановки решается задача построения рационального статического ПЗ/ОУ, максимизирующего функциональную мощность системы в искаженных работоспособных состояниях.

Требуется построить процедуру РСПЗ/ОУ, которая обеспечивает максимальное значение функциональной мощности (ф.м.) системы (минимальное значение потерь) в каждом состоянии s_ω^r из заданного множества S_ω^ω искаженных работоспособных состояний при заданных временных и ресурсных ограничениях.

Известно:

- $H^0 = H = \{M_1, \dots, M_i, \dots, M_n\}$ – множество всех вычислительных модулей системы;
- n – число вычислительных модулей в системе;
- $\Omega^E = \{U_1, \dots, U_j, \dots, U_L\}$ – множество всех задач, выполняемых системой в начальном состоянии (начально множество задач);
- L – число всех задач начального множества;
- v_j – объём памяти, требуемый для хранения программного кода задачи U_j и ее выполнения;
- τ_j – время выполнения задачи U_j ;
- b_j – вес задачи U_j , характеризующий ее важность для системы;
- Начальное распределение задач по n вычислительным модулям системы, являющееся основным размещением задач, и представленное матрицей (рис. 1).

	M_1	M_2	...	M_i	...	M_n
U_1	0	1	...	0	...	0
U_2	1	0	...	0	...	0
...						
U_j	0	d_{ji}^0
...						
U_L	0	0	...	1	...	0

Рис. 1. Матрица начального распределения задач

Где: $d_{ji}^0 = 1$, если задача U_j назначена в начальном состоянии для решения в вычислительном модуле M_i и размещена в нем, $d_{ji}^0 = 0$ в противном случае, каждая задача назначена в один и только в один вычислительный модуль, т.е. в каждой строке – только одна единица;

Обозначим:

- Ω_i^0 – множество задач, размещенных в вычислительном модуле M_i в начальном состоянии и называемых собственными задачами вычислительного модуля M_i ;
- Ω_i^w – множество задач, размещенных в вычислительном модуле M_i в искаженном состоянии S_{ω} ;
- $S_{\omega}^{nr} = \{s_{\omega}^{nr}\}$ – множество искаженных работоспособных состояний системы, включающее все состояния, соответствующие отказам любого числа вычислительных модулей, не превышающего заданного значения d ;
- Дополнительные затраты, необходимые для размещения резервных копий задач и их выполнения в тех вычислительных модулях, в которых их размещение и выполнение не предусмотрено начальным (основным) РЗ, которые задаются матрицей дополнительных затрат C (рис. 2):

	M_1	M_2	...	M_i	...	M_n
U_1	c_{11}		...	c_{1i}	...	c_{1n}
U_2						
...						
U_j			...	c_{ji}	...	
....						
U_L	c_{L1}		...	c_{Li}	...	c_{Ln}

Рис. 2. Матрица дополнительных затрат

Где: $c_{j\bar{i}}$ – дополнительные затраты, необходимые для реализации размещения копии задачи U_j в вычислительный модуль M_i и ее решения в этом вычислительном модуле в

соответствующих искаженных состояниях (дополнительная память, интерфейсы, каналы связи и др.), при условии, что при основном РМЗ данная задача не была размещена в вычислительном модуле M_i ;

При заданном основном РМЗ D^0 полагаем в матрице C : $c_{ij} = 0$ для тех пар (задачи U_j , модуль M_i), которым соответствуют значения $d_{ij}^0 = 1$ в матрице D^0 ;

Максимально допустимые (граничные) значения:

- ΔC_{rs}^s - суммарные (для всех модулей системы) максимально допустимые дополнительные затраты на резервирование задач в состоянии s , т.е. на размещение и выполнение в вычислительных модулях системы дополнительных копий тех задач, которые должны выполняться в состоянии s , несмотря на отказ модулей, выполнявших их в начальном состоянии;

- T_{2i}^s - максимально допустимое суммарное время выполнения всех задач модуля M_i ;

- V_i^m - максимальная ёмкость памяти вычислительного модуля M_i .

Зона разрешенной передачи задач $H^s = \{H_1^s, H_2^s, \dots, H_j^s, \dots, H_L^s\}$,

т.е. набор подмножеств H_j^s , где H_j^s - подмножество вычислительных модулей, в которых разрешено размещать задачу U_j - при отказах ее собственных вычислительных модулей (т.е. таких, в которых эта задача размещена в соответствии с начальным РМЗ D^0) или

Зона запрещенной передачи задач $H^x = \{H_1^x, H_2^x, \dots, H_j^x, \dots, H_L^x\}$,

где $H_j^x = \frac{H_j^s}{H_j^s}$ - подмножество вычислительных модулей, в которых запрещено размещать задачу U_j при отказах её собственных вычислительных модулей.

Эти зоны задаются заказчиком системы и определяются технологическими соображениями.

Требуется решить следующую задачу, включающую две подзадачи:

А) построить планы рационального РМЗ для всех искаженных работоспособных состояний заданного множества.

Для заданного множества $S^w = \{s_\omega^t\}$ искаженных работоспособных состояний (определяемого заданным максимально допустимым числом d отказавших вычислительных модулей) найти планы $D^{w\Gamma}$ распределения задач (РМЗ), выполняемых в каждом состоянии $s_\omega^t \in S^w$, при которых достигается максимальное значение функциональной мощности (ф.м.) E_ω^t в состоянии s_ω^t при заданных граничных значениях определенных показателей, набор которых выбирается (из перечисленных ниже) в зависимости от требований к системе:

- T_{2i}^s - суммарного времени выполнения всех задач в каждом работоспособном вычислительном модуле M_i ;

- Γ_i^w - загрузки каждого работоспособного вычислительного модуля M_i ;

- ΔC_{rs}^s - суммарных дополнительных затрат на резервирование задач в состоянии s_ω^t ;

- ΔV_i^s - доступного объема памяти вычислительного модуля M_i , который может быть использован для размещения копий задач, дополнительно назначаемых этому вычислительному модулю для выполнения в состоянии s_ω^t ;

Б) Найти отказоустойчивое размещение задач (ОУ-РМЗ) по всем вычислительным модулям системы, позволяющее реализовать все найденные планы $D^{w\Gamma}$.

Возможны два варианта перераспределения задач начального множества Ω^0 при построении планов рационального РМЗ для состояний множества S^w .

Вариант 1: в каждом состоянии s_ω^t перераспределяются (т.е. передаются для выполнения в неотказавших вычислительных модулях, либо отбрасываются) только собственные задачи отказавших вычислительных модулей, при этом собственные задачи неотказавших

вычислительных модулей продолжают выполняться в "своих" вычислительных модулях (т.е. в тех вычислительных модулях, где они выполнялись при начальном РЗ).

Вариант 2: в каждом состоянии s_ω^t заново перераспределяются все задачи начального множества Ω^t .

Для упрощения изложения далее отбросим индекс t в обозначении искаженного работоспособного состояния и относящихся к нему параметров и будем обозначать это состояние $s_\omega^t = s_\omega$.

Вариант 1:

Для каждого искаженного работоспособного состояния $s_\omega = s_\omega^t$:

Найти такой план $D^\omega = \{d_{ji}^\omega\}, j = 1, \dots, L, i = 1, \dots, n$ распределения собственных задач отказавших вычислительных модулей по не отказавшим (точнее, план распределения дополнительных копий этих задач), который обеспечивает для данного состояния $s_\omega \in S_\omega^t$ максимальное значение функциональной мощности(1).

$$E_\omega^f = \sum_{U_j \in \Omega_\omega^f} \sum_{M_i \in H_\omega} d_{ji}^\omega b_j \rightarrow \max \quad (1)$$

где Ω_ω^f – множество собственных задач отказавших вычислительных модулей,

H_ω – множество не отказавших вычислительных модулей для состояния s_ω при заданном наборе временных и ресурсных ограничений.

Будем считать, что веса задач заданы величиной потерь, т.е. вес b_j каждой задачи выражается величиной потерь r_j , возникающих вследствие отбрасывания задачи U_j (т.е. прекращения ее выполнения системой).

В этом случае максимизация значения E_ω^f , т.е. суммы весов тех задач, которые сохранены в системе, означает максимизацию величины тех потерь, которые не произойдут при переходе системы в данное искаженное состояние ("непотерянных потерь") и, следовательно, - минимизацию потерь, которые будут иметь место при переходе системы в данное состояние.

Поэтому в дальнейшем рассматриваем только задачу максимизации функциональной мощности (как для варианта 1, так и для варианта 2).

Набор ограничений выбирается из следующих возможных.

Временные ограничения(2):

$$\Delta T_{\Sigma i}^\omega = \sum_{U_j \in \Omega_\omega^f} \tau_j d_{ji}^\omega \leq \Delta T_{\Sigma i}^*, \quad i = 1, \dots, g_\omega, \quad (2)$$

где g_ω – число не отказавших вычислительных модулей в состоянии s_ω ;

$\Delta T_{\Sigma i}^*$ - ресурс времени для размещения в вычислительный модуль M_i копий задач отказавших вычислительных модулей(3).

$$\Delta T_{\Sigma i}^* = T_{\Sigma i}^* - T_{\Sigma i}^0 \quad (3)$$

где $\Delta T_{\Sigma i}^*$ - заданное граничное значение суммарного времени выполнения всех задач в работоспособном вычислительном модуле M_i ;

$T_{\Sigma i}^0$ - суммарное время выполнения всех собственных задач модуля M_i (размещенных в нем в соответствии с начальным РЗ).

Ресурсные ограничения на суммарные дополнительные затраты, требуемые для размещения во всех неотказавших вычислительных модулях системы резервных копий тех задач, которые назначаются для выполнения в системе в данном работоспособном искаженном состоянии(4):

$$\Delta C_\omega = \sum_{U_j \in \Omega_\omega^f} \sum_{M_i \in H_\omega} d_{ji}^\omega c_{ji} \leq \Delta C_\omega^*; \quad (4)$$

Для случая, когда все вычислительные модули системы идентичны, принимаем, что для любого вычислительного модуля M_i справедливо $c_{ji} = c_j$.

Величина ΔC_{ω}^* для каждого состояния $s_{\omega} \in S_{\omega}^*$ может быть получена из общей заданной граничной величины затрат ΔC^* на построение результирующего отказоустойчивого размещения задач путем выделения некоторой доли этих затрат на резервирование задач в каждом состоянии s_{ω} . Можно считать, что $\Delta C_{\omega}^* = \alpha_{\omega} \Delta C^*$, где $\alpha_{\omega} \leq 1$ – коэффициент, определяющий долю общих допустимых затрат, выделяемую для размещения копий задач в данном состоянии s_{ω} и принять, что этот коэффициент α_{ω} пропорционален вероятности $P(s_{\omega})$ перехода системы в данное состояние s_{ω} .

Ресурсные ограничения на доступный объем памяти каждого работоспособного вычислительного модуля M_i , требуемый для размещения резервных копий задач, назначенных этому вычислительному модулю для выполнения в данном состоянии(5).

$$\Delta V_{\omega i} = \sum_{U_j \in \Omega_{\omega}^f} d_{ji}^{\omega} v_j \leq \Delta V_{\omega i}^*, \quad M_i \in H_{\omega}, \quad i = 1, \dots, g_{\omega}, \quad (5)$$

где v_j – объем памяти, требуемый для хранения программного модуля задачи U_{jj} и ее выполнения.

Значение $\Delta V_{\omega i}^*$ определяется как(6):

$$\Delta V_{\omega i}^* = (V_i^m - V_i^0) / N(S_i^{\omega r}), \quad (6)$$

где V_i^m – заданный максимальный объем памяти вычислительного модуля M_i

V_i^0 – объем памяти вычислительного модуля M_i , занятый программными модулями задач, размещенных в данном вычислительном модуле в соответствии с основным (начальным) размещением задач (т.е. сумма объемов памяти программных модулей "собственных" задач вычислительного модуля M_i),

$N(S_i^{\omega r})$ – число таких состояний s_{ω} , в которых данным вычислительный модуль M_i не отказал (только в этих состояниях ему могут предаваться задачи отказавших вычислительных модулей).

Ограничение на набор значений двоичных переменных(7) (учитывающее, что в каждом данном состоянии s_{ω} каждая задача U_j из множества собственных задач отказавших вычислительных модулей может быть либо назначена для решения только в один вычислительный модуль, либо отброшена, т.е. не назначена ни в один вычислительный модуль):

$$\sum_{i=1}^n d_{ji}^{\omega} \leq 1, \quad \forall U_j \in \Omega_{\omega}^f, \quad (7)$$

В каждом плане $D^{\omega} = \{d_{ji}^{\omega} | j = 1, \dots, L, i = 1, \dots, n\}$, те значения d_{ji}^{ω} , индекс i которых соответствует отказавшим вычислительным модулям, полагаются $d_{ji}^{\omega} = 0$.

Вариант 2:

При решении задачи в варианте 2 - перераспределение всех задач начального множества Ω^0 в каждом состоянии s_{ω} , необходимо учесть следующие отличия от варианта 1:

1. В формулах для целевой функции и для ограничений рассматривается начальное множество задач Ω^0 вместо множества Ω_{ω}^f ;

2. При заданном основном РМЗ D^0 полагаем в матрице затрат C : $c_{ji} = 0$ для тех пар (задача U_j , модуль M_i), которым соответствуют значения $d_{ji}^0 = 1$ в матрице D^0 ;

3. В ограничении по памяти: общий резерв памяти вычислительного модуля составляет весь объем его памяти, т.е. $\Delta V_i = V_i^m$ и $\Delta V_{\omega i} = \Delta V_i \delta_{i\omega} = V_i^m \delta_{i\omega}$

4. В ограничении по времени - ресурс времени вычислительного модуля M_i равен: ΔT_{Σ} .

Заключение

Таким образом, можно реализовать кластер высокой доступности в любых информационных, вычислительных и т.д. системах. Идея данного метода заключается в рационализации распределения задач, нежели в его хаотичности. Данный метод также можно изменить для более слабых систем с введением распределения задач по градиенту важности, чтобы изначально решались важные задачи, а задачи, который можно было отбросить в конце.

Литература

1. *Турута Е.Н.* Организация распределенных задач в вычислительных системах, обеспечивающая их отказоустойчивость // Автоматика и вычисл. Техника, 1985. №1. С. 5-14.
2. *Gary J. Nutt.* Centralized And Distributed Operating Systems. Prentice-Hall, Inc. 1992.
3. *Mark Baker, Amy Apon.* Cluster Computing and Applications. University of Portsmouth. 2000.
4. *Flynn M.* Very high-speed computing system // Proc. IEEE. 1966. No. 54, pp.1901-1909.

HIGH-AVAILABILITY CLUSTER WITH USING FAILOVER PLANNING METHOD

Ruslan A. Kibedov

Student of group M091801(70), MTUCI

Kibedov.r96@gmail.com

Turuta N. Evgeniy

MTUCI, PhD., professor of MKiIT department

eturuta@yandex.ru

Keywords: *failover cluster, method of failover planning task, high-availability cluster, program-structure failover cluster, failover algorithm of planning task.*

Here is presented the definition of a failover cluster. Present the method of failover task planning in a cluster computing system. The main restrictions on the distribution of tasks in the method of failover task planning. Two possible options for ensuring high cluster availability achieved through the use of the concept of failover task planning in the nodes of a computer system are considered.

ИССЛЕДОВАНИЕ СВОЙСТВ ДВУМЕРНЫХ ДИСКРЕТНЫХ ПРЕОБРАЗОВАНИЙ В КОМПЬЮТЕРНОМ ЗРЕНИИ

Пьяикова Татьяна Павловна
студентка группы М091801(70) МТУСИ
pyankova.tanya96@mail.ru
Яшина Марина Викторовна
МТУСИ, д.т.н., профессор
yash-marina@yandex.ru

Ключевые слова: компьютерное зрение, быстрые алгоритмы, обработка изображения, дискретное преобразование Фурье, дискретное косинусное преобразование.

Область компьютерного зрения может быть охарактеризована как молодая, разнообразная и динамично развивающаяся. И хотя существуют более ранние работы, можно сказать, что только с конца 1970-х началось интенсивное изучение этой проблемы, когда компьютеры смогли управлять обработкой больших наборов данных, таких как изображения. Методы обработки изображений с помощью преобразования – это методы, которые используются в таких областях преобразования изображений, как дискретное преобразование Фурье, дискретное косинусное преобразование, Вейвлет-преобразование и т.д. Их высокая эффективность заключается в сжатии изображений, в восстановлении изображений, в передискретизации изображений и в геометрических преобразованиях. В статье рассматриваются данные методы, акцентируя особое внимание на их сравнительных характеристиках и взаимосвязях.

Под компьютерным зрением понимают теорию и технологию создания искусственных компьютерных систем, которые осуществляют обнаружение, классификацию и отслеживание объектов. Информацию они получают из изображений, которые могут быть представлены определённым видеорядом, изображением с различных камер или трехмерными данными, отсканированными изображениями и т.п.

Цель компьютерного зрения заключается в формировании полезных выводов относительно объектов и сцен реального мира на основе анализа изображений, полученных с помощью датчиков.

Задачи компьютерного зрения:

- Распознавание
- Движение
- Восстановление сцены
- Восстановление изображения

Где используется компьютерное зрение? Одна из самых перспективных сфер применения компьютерного зрения - дополненная реальность. Так называют технологию, которая предусматривает наложение виртуальных элементов (например, текстовых подсказок) на картину реального мира. Примером могут быть, например, мобильные приложения, которые позволяют получить информацию о доме, направив на него камеру телефона или планшета. Дополненная реальность уже применяется в программах, сервисах и устройствах, но пока находится только в начале пути.

Без преувеличения можно утверждать, что цифровая обработка изображений появилась на свет с введением Кули и Тьюки в 1965 году алгоритма быстрого преобразования Фурье для вычисления дискретного преобразования Фурье. Эта публикация сразу же привела к бурному

росту активности во всех отраслях цифровой обработки сигналов и изображений и их приложений.

Вторая волна в этом процессе была обусловлена введением в инженерную коммуникацию и цифровую обработку изображений в 1970-х годах преобразования Уолша-Адамара и преобразования Хаара и разработкой большого семейства быстрых преобразований с алгоритмами типа БПФ. В то время когда преобразования Уолша-Адамара и Хаара уже были известны в математике, другие преобразования, например, довольно популярные в то время наклонные преобразования, были разработаны «с нуля». Данный период закончился введением дискретного косинусного преобразования, которое вскоре получило широкое признание в качестве лучшего выбора среди всех доступных в то время преобразований и привело к появлению стандартов сжатия JPEG и MPEG для изображений, аудио и видео.

Третья большая волна действий в области преобразований для обработки сигналов и изображений была вызвана введением в 1980-х годах ряда преобразований, получившее название «Вейвлет-преобразования». Основной мотивацией было достижение лучшего локального представления сигналов и изображений по сравнению с «глобальным» представлением, характерного для дискретного преобразования Фурье (ДПФ), дискретно-косинусное преобразование (ДКП), Уолша-Адамара и других быстрых преобразований (БПФ), доступных в то время.

В настоящее время быстрые преобразования составляют основной инструментарий цифровой обработки изображений.

Преобразования

А. Сжатие изображения

1.1 Дилемма: компрессионная дискретизация или сжатие. В идеале оцифровка изображения, то есть преобразование непрерывных сигналов от датчиков изображения в цифровые сигналы, должна выполняться таким образом, чтобы предоставить максимально сжатое цифровое представление изображения при условии, что качество воспроизведения изображения удовлетворяет заданным требованиям. Из-за технических ограничений оцифровка изображения чаще всего выполняется в два этапа: дискретизация, то есть преобразование сигнала датчика изображения в набор действительных чисел, и скалярное квантование, то есть округление, этих чисел в набор фиксированных квантованных величин.

Дискретизация изображения в общем случае может рассматриваться как получение коэффициентов расширения сигнала изображения по набору базисных функций дискретизации. Чтобы сделать набор этих коэффициентов представления как можно более компактным, следует выбирать такие базисные функции дискретизации, которые обеспечивают наименьшее количество коэффициентов расширения сигнала, достаточных для восстановления изображения с заданным требуемым качеством.

Существует как минимум два примера практической реализации принципа общей дискретизации посредством измерения коэффициентов изображения в области преобразования: визуализация с помощью кодирующей апертуры и магнитно-резонансная томография (МРТ). При визуализации изображений с помощью кодирующей апертуры изображения распознаются посредством двоичных масок, которые реализуют двоичные базисные функции. При визуализации МРТ измеряются коэффициенты разложения в ряд Фурье данных сенсора.

1.2 Методы преобразования сжатия изображений. Потребности в сжатии изображений были основными мотивами цифровой обработки изображений. Начатые ещё в 1950-х годах с

различных методов прогнозирующего кодирования, к началу 1970-х годов исследования сжатия изображений начали сосредотачиваться в основном на том, что мы сейчас называем «кодированием с преобразованием», которое в буквальном смысле реализует принцип сжатой дискретизации.

Принцип кодирования изображения с помощью преобразования иллюстрируется блок-схемами, приведенными на рисунке 1.

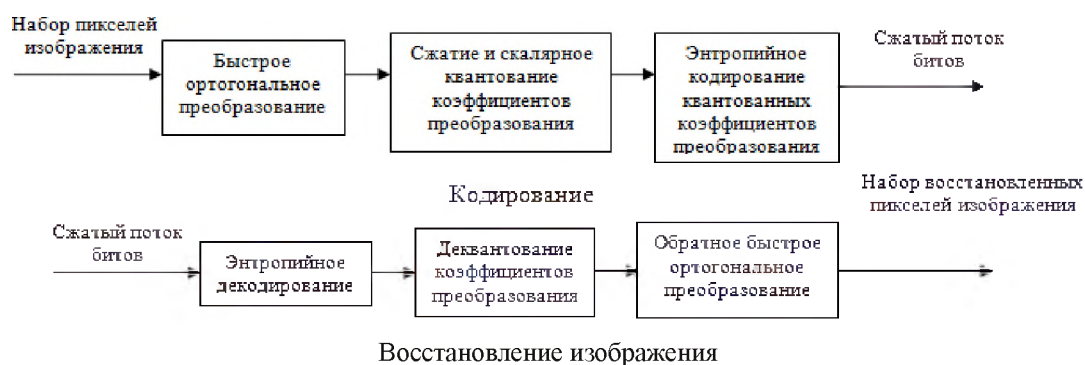


Рис. 1. Блок-схемы кодирования и восстановления изображения с помощью преобразования.

Согласно этим схемам набор пикселей изображения сначала подвергается быстрому ортогональному преобразованию. Затем коэффициенты преобразования низкой интенсивности отбрасываются, тем самым существенно уменьшая объем данных. Это главный источник сжатия изображений. Отметим, что отбрасывание определенных коэффициентов преобразования изображения означает замену изображений на их полосу ограниченную, в рамках выбранного преобразования, аппроксимациями. Остальные коэффициенты преобразования подвергаются, один за другим, оптимальному неравномерному скалярному квантованию, которое минимизирует среднее количество уровней квантования коэффициентов преобразования. Для восстановления изображения из сжатого потока битов оно должно быть подвергнуто соответствующим обратным преобразованиям.

Изображения обычно содержат много объектов, и их глобальные спектры представляют собой смесь объектных спектров, тогда как спектры отдельных фрагментов или блоков изображения гораздо более специфичны, и это позволяет более простым способом отделить наиболее интенсивные спектральные составляющие от наименее интенсивных. Это наглядно показано на рисунке 2, где для сравнения представлены глобальные и блочные спектры мощности ДКП тестового изображения. Хотя, как видно на рисунке, спектральные коэффициенты глобального спектра ДКП разрежены более (разреженность спектра данного изображения, то есть доля спектральных коэффициентов, которые содержат 95% полной энергии спектра, равна 0,095), чем, в среднем, те изображения, у которых размер фрагментов составляет 16×16 пикселей (средняя разреженность фрагментов спектров равна 0,14). Спектры отдельных фрагментов, безусловно, более специфичны, а их компоненты, отвечающие за небольшие объекты и границы объектов, имеют гораздо большую энергию, чем в глобальном спектре, и не будут утеряны при выборе 95% наиболее интенсивных компонентов.

непрерывных сигналов из их выборок при условии, что количество выборок бесконечно велико.

Идеальная интерполяционная способность дискретной *sinc*-интерполяции была продемонстрирована при всестороннем сравнении различных методов интерполяции в экспериментах с многократным вращением на 360. Результаты экспериментов, показанных на рисунке 3, ясно свидетельствуют о том, что, в отличие от других методов, в том числе интерполяции сплайнами более высокого порядка, дискретная *sinc*-интерполяция не вносит каких-либо заметных искажений в повернутое изображение.

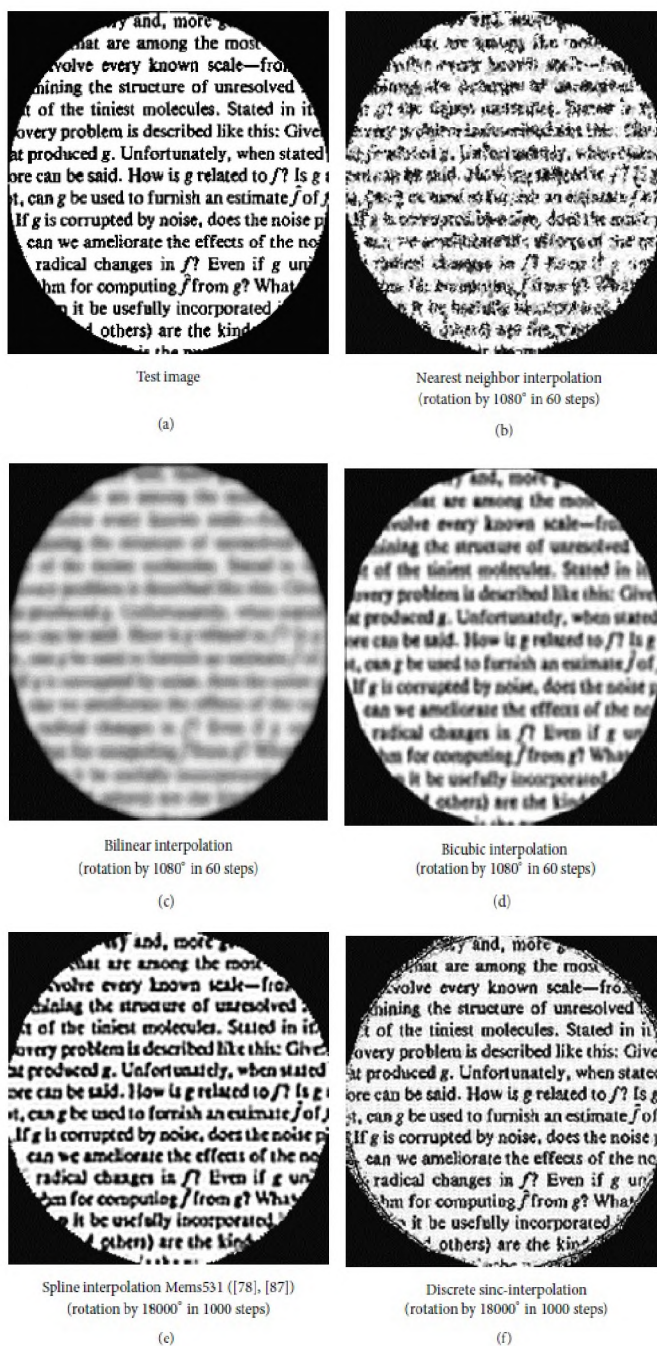


Рис. 3. Сравнение дискретной *sinc*-интерполяции с другими методами интерполяции: результаты нескольких поворотов изображения.

[original image, bilinear intrp, bicubic intrp, discr-sinc intrp]; Nit=75

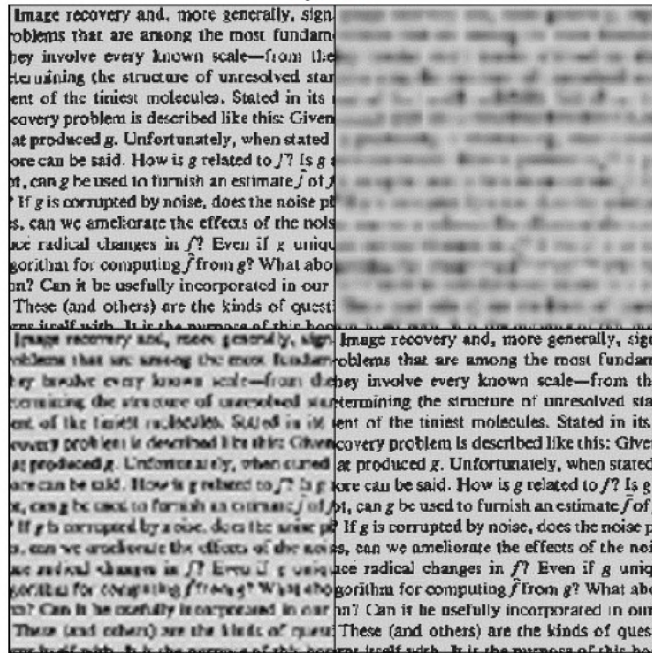


Рис. 4. Сравнение дискретной sinc-интерполяции с билинейной и бикубической интерполяциями в итеративном увеличенном/уменьшенном изображении с параметром масштаба $\sqrt{2}$.

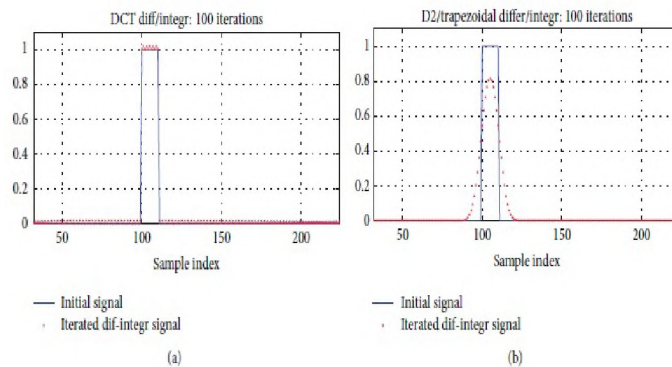


Рис. 5. Сравнение результатов итеративного чередования 100 дифференцирований и интегрирования тестового прямоугольного сигнала с использованием алгоритмов на основе ДКП (слева) и стандартных численных алгоритмов (справа, фильтра дифференцирования с функцией рассеяния точек $[-0.5, 0, 0.5]$ и алгоритма интегрирования по правилу трапеций).

Заключение

Главная отличительная особенность преобразований, благодаря которой они являются такими эффективными в цифровой обработке изображений, это их способность к уплотнению энергии. В представлениях обычных изображений в виде наборов упорядоченных пикселей некоторые пиксели, например те, которые принадлежат границам объекта, являются более важными, чем другие, и в каждом конкретном изображении всегда есть некоторые пиксели, которые не имеют такого значения и могут быть исключены из представления изображения и восстановлены из оставшихся «важных» пикселей. Но проблема заключается в том, что никто не знает заранее, какие пиксели являются «важными» в изображении, а какие нет.

Ситуация складывается совершенно иначе в представлении изображения в области преобразования. Для ортогональных преобразований, которые обладают хорошей способностью к сжатию энергии, львиная доля общей «энергии» изображения (сумма квадратов коэффициентов преобразования) сосредоточена в небольшой доле коэффициентов преобразования, индексы которых, как правило, заранее известны для данного типа изображений или могут быть легко обнаружены. Именно эта особенность преобразований называется их способностью к уплотнению энергии. Это позволяет заменять изображения на изображения «ограниченные по полосе», с точки зрения конкретного преобразования, аппроксимации, то есть аппроксимации, определяемые достаточно малой долей коэффициентов преобразования изображения.

Литература

1. *Беликова Т.П., Кронрод М.А., Чочиа П.А., Ярославский Л.П.* Цифровая обработка фотоснимков поверхности Марса, переданных АМС “Марс-4” и “Марс-5” // Косм. исслед. – 1975. Т.13, вып. 6. С. 898-906.
2. *Ярославский Л.П.* Введение в цифровую обработку изображений. М.: Сов. радио, 1979. 312 с.
3. *Буслаев А.П., Дорган В.В., Кузьмин Д.М., Приходько В.М., Травкин В.Ю., Яшина М.В.* Распознавание изображений и мониторинг состояния дорог, автотранспортных потоков и безопасности движения. Вестник МАДИ(ГТУ). № 4, 2005. 8 с. С. 102-109. ISBN 5-79620061-5.
4. *Ярославский Л.П.* Обработка изображений в медицинской интроскопии / Цифровая оптика в медицинской интроскопии. М.: ИППИ РАН, 1992. С. 4-17.
5. *Yaroslavsky L.P.* Fast Transforms in Image Processing: Compression, Restoration, and Resampling, Vol. 2014, article ID 276241, 23 p.
6. *Яшина М.В., Толмачев А.А.* Методы распознавания образов для оценки характеристик пешеходных потоков // Т-Сотт: Телекоммуникации и транспорт. 2017. Т. 11. № 8. С. 45-51.

STUDY OF PROPERTIES OF TWO-DIMENSIONAL DISCRETE TRANSFORMATIONS IN COMPUTER VISION

Tatyana P. Pyankova

*Student of group M091801(70), MTUCI
pyankova.tanya96@mail.ru*

Marina V. Yashina

*MTUCI, Doctor of Technical Sciences, Professor
yash-marina@yandex.ru*

Keywords: *computer vision, fast algorithms, image processing, discrete Fourier transform, discrete cosine transform.*

The area of computer sight can be characterized as young, various and dynamically developing. And though there are earlier works, one may say, that intensive studying of this problem when computers could operate processing of big data sets, such as images began only with the end of the 1970th. Transform image processing methods are methods that work in domains of image transforms, such as Discrete Fourier, Discrete Cosine, Wavelet, and alike. They proved to be very efficient in image compression, in image restoration, in image resampling, and in geometrical transformations and can be traced back to early 1970s. The paper reviews these methods, with emphasis on their comparison and relationships, from the very first steps of transform image compression methods to adaptive.

КЛАССИФИКАЦИЯ УГРОЗ ДЛЯ ИНТЕРНЕТА ВЕЩЕЙ

Пономарев Алексей Александрович

студент группы М091801(72)

alex1996@mail.ru

Шевелев Сергей Владимирович

МТУСИ, к.т.н., доцент кафедры МСиУС

Ключевые слова: интернет вещей, безопасность, информация, IoT, информационные технологии.

Рассмотрена проблема обеспечения безопасности информации в Интернете вещей. Предложенная декомпозиция проблемы по типу взаимодействующих устройств и характеру взаимодействия. Рассмотрены серверные и встроенные («embedded») конечные системы, их специфика и множество возможных угроз.

Введение

Интернет вещей (англ. Internet of Things , IoT) - это сеть, состоящая из взаимосвязанных физических объектов (вещей) или устройств, которые имеют встроенные датчики, а также программного обеспечения, позволяющего осуществлять передачу и обмен данными между физическим миром и компьютерными системами, с помощью использования стандартных протоколов связи.

Кроме датчиков, сеть может иметь исполнительные устройства, встроенные в физические объекты и связанные между собой через проводные и беспроводные сети. Эти взаимосвязанные объекты (вещи) имеют возможность считывания и приведения в действие, функцию программирования и идентификации, а также позволяют исключить необходимость участия человека, за счет использования интеллектуальных интерфейсов [1].

Основной концепцией IoT является возможность подключения всевозможных объектов (вещей), которые человек может использовать в повседневной жизни, например, холодильник, кондиционер, автомобиль, велосипед или даже кроссовки. Все эти объекты (вещи) должны быть оснащены встроенными датчиками или сенсорами, которые имеют возможность обрабатывать информацию, поступающую из окружающей среды, обмениваться ею и выполнять различные действия в зависимости от полученной информации. Примером внедрения такой концепции является система «умный дом» или «умная ферма» [2].

В течение последнего десятилетия Интернет вещей плавно вошел в нашу жизнь благодаря появлению систем беспроводной связи, таких как RFID, Wi-Fi, 4G, IEEE 802.15.x, которые чаще всего используются в основе приложений мониторинга и контроля. Сегодня системы Интернета вещей используют не только для частных сетей, но и на производствах, фабриках, заводах, предприятиях и даже в государственных учреждениях. Основная проблема использования сетей IoT заключается в том, что они не имеют защиты от воздействий со стороны злоумышленника.

Это может привести, в лучшем случае, к причинению вреда имуществу пользователя, а в худшем - его здоровью и жизни.

Например, устройства контроля и управления электрической сетью могут быть захвачены злоумышленником с помощью любого устройства, имеющего доступ к сети Интернет, и соответствующего программного обеспечения.

Получив полный или частичный контроль над устройством злоумышленник может осуществить отключение или порчу электрических приборов, в том числе критически необходимых приборов (систем жизнеобеспечения в больницах, систем мониторинга на производстве, охранных систем и т.д.), создать короткие замыкания в сети и даже вызвать пожар или аварию, если речь идет о производстве. Именно поэтому возникает актуальная

проблема исследования безопасности Интернета вещей и, в частности, безопасности пользователя, его имущества и личной информации, которая передается обрабатывается и хранится в сетях IoT.

Анализ последних исследований и публикаций

Кевин Эштон, который является соучредителем Auto-ID Center, первым ввел термин «Интернет вещей», именно так он назвал свой доклад для Procter & Gamble в 1999 году [3]. Это была попытка внедрить новую идею радиочастотной идентификации (RFID) в цепь поставок производственных товаров, а в результате привлекло внимание к самой идее подключения к сети новых типов устройств.

Интернет вещей как явление носит глобальный характер, а значит, при создании собственной отечественной инфраструктуры IoT, очень важно унифицировать методы его взаимодействия с остальным миром. Организацией с которой в первую очередь следует наладить взаимодействие в этом вопросе является IoT European Research Cluster. Ее целью является рассмотрение потенциала решений на основе IoT, координации исследовательской деятельности и достижения консенсуса относительно путей реализации IoT в Европе. Разработками в сфере исследований и стандартизации Интернета вещей занимаются многие страны на уровне национальных инициатив, например ANSI (США), BSI (Великобритания), ETSI (Европа), а также на международном уровне: ITU, ISO, IEC.

Кроме того, большое количество объединенных устройств создают проблему масштабируемости. В работе [4] приведена информация о рынке IoT, его перспективах, темпах развития, а также анализ возможных рисков, которые появятся с полной интеграцией IoT в жизни человека. В статье [5] приведен анализ основных проблем безопасности Интернета вещей.

Для достижения полного признания со стороны пользователей определение и достижение необходимого уровня безопасности и защищенности для Интернета вещей являются обязательными. Кроме этого, должны быть гарантированы безопасность и анонимность данных пользователя, их конфиденциальность и целостность, а также надежность механизмов аутентификации и авторизации. Это необходимо для предотвращения несанкционированного доступа неавторизованных пользователей к системе.

Цель статьи - рассмотреть возможности интернета вещей, сформировать общее описание проблем информационной безопасности в "Интернете вещей" и рассмотреть наиболее вероятностные векторы атак.

Постановка проблемы

Энергопотребление современных чипов и стоимость их производства постоянно снижается. Это касается не только центральных процессоров, но, например, и Wi-Fi-чипов, стоимость которых за последние годы значительно снизилась. В результате устройства с доступом в интернет становятся все проще и доступнее.

Согласно ITU-T Y.2060, Интернет вещей (Internet of Things, IoT) - глобальная инфраструктура для информационного общества, которая включает передовые решения по объединению физических и виртуальных вещей на основе существующих и разрабатываемых информационных и коммуникационных технологий.

По информации Tech Navio, Интернет Вещей будет составлять все большее и большее число подключений, а общее их количество возрастет до 17 млрд. в ближайшие 5 лет. Процесс будет состоять из трех волн: сначала, объединятся устройства, которые служат потребителям, дальше IoT расширится до подключенных устройств на предприятиях, и наконец его использование станет массовым благодаря внедрению в государственных органах и органах исполнительной власти.

IoT позволит частным и общественным организациям оптимизировать управление, ускорить отклик системы производства на управляющие воздействия, и разрабатывать более новые и эффективные бизнес модели. Технология IoT пересекается со многими другими технологическими областями, включая телеметрию, телематику, M2M-коммуникации (межмашинное взаимодействие), интеллектуальные сети, интеллектуальные системы транспортировки и портативные компьютеры. По сути она является чрезвычайно сложным комплексом собрав в себе все последние наукоемкие технологии современности.

Главной тенденцией присущей IoT, так же как и его главной проблемой, является довольно быстрое увеличение числа конечных устройств подключенных к сети.

При этом, интернет вещей, растущий быстрее рынка смартфонов, строится, функционирует и внедряется в существующую сетевую инфраструктуру без должного внимания вопросам безопасности. Чтобы система была устойчива необходимо закладывать функции обеспечения ее устойчивости и безопасной эксплуатации в фундамент технологии.

В свою очередь современный Интернет вещей растет взрывообразно и без оглядки даже на элементарные сервисы безопасности, такие как шифрование трафика или защищенную авторизацию.

Основная часть

Задачу обеспечения безопасности информации в IoT можно разделить на 2 подзадачи: обеспечение безопасности конечных систем и обеспечения безопасности их сетевого взаимодействия (рис.1).



Рис. 1. Структура обеспечения ИБ в IoT

Конечные системы удобно разделить на 2 категории: серверные системы и так называемые «embedded» системы. Согласно определению NC State University's Electrical and Computer Engineering Department, «embedded» system или «встроенная система» - это система специального назначения, в которой компьютер полностью встроен в устройство, которым он управляет. В отличие от компьютеров общего назначения, встроенные системы выполняют заранее определенные задачи, как правило, с очень конкретными требованиями.

Встроенные системы на базе 8- 16- или 32-х битных микроконтроллеров широко распространены в автоматизированных системах управления технологическим процессом (АСУ ТП), бытовой технике, автомобилях и других транспортных средствах и могут быть интегрированы в сеть интернет по протоколу IEEE 802.3 Ethernet или IEEE 802.11x Wi-Fi, а при наличии памяти могут работать со стеком TCP / IP.

Говорить о едином подходе к защите встроенных систем достаточно сложно потому, что в отличие от настольных компьютеров и ноутбуков, для которых используется достаточно

ограниченный набор процессорных архитектур (в основном x86), в процессорах для встроенных систем используются многочисленные, конкурирующие архитектуры. Наиболее распространенные ARM и x86 (Intel-совместимые), в меньшей степени - PowerPC и MIPS.

При наличии у злоумышленника доступа к таким системам может быть реализована модификация схемы устройства (платы / микросхемы), то есть установление физических закладок в само устройство или в разрыв кабеля (в случае Ethernet) способных образовывать каналы утечки информации.

Причиной возникновения в конечных устройствах программных закладок, способных влиять на данные, которые передаются в следующие устройства в сети, являются неправильные процессы обновления версий прошивки, или использование сторонних прошивок нелегализованных производителей.

Основным источником угроз остается сам доступ таких устройств к сети, при этом первоочередная задача состоит в недопущении несанкционированного доступа к встраиваемым системам, так как в зависимости от сферы применения они могут быть непосредственными источниками воздействия на окружающую среду.

Чаще всего функции управления во встроенных системах не выносятся в отдельное устройство и все функциональные алгоритмы выполняются непосредственно в пределах системы. Для этого на них устанавливается операционная система (ОС), которая может быть достаточно малой, специально разработанной для использования во встраиваемых системах, или урезанной версией системы, которая обычно используется на компьютерах общего назначения.

Не менее остро стоит вопрос защиты серверных систем, которые представляют собой программно-аппаратные комплексы, выполняющие определенные сетевые службы и реализующие прием запросов от клиентов. Согласно отчету W3Techs, 35.98% серверных систем работают на базе ОС Linux (Debian, Ubuntu, CentOS, RHEL, Gentoo) 31.52% на базе BSD и других Unix-подобных системах (FreeBSD, HP-UX, Solaris), и 32.5% на базе Windows Server. Проведенный анализ использования операционных систем позволяет определить угрозы их безопасности и сгруппировать их по следующим векторам:

- Использование известных (легальных) каналов получения информации (эксплуатация ошибок в конфигурации системы безопасности, несанкционированное использование легальных учетных записей и др.).
- Использование скрытых каналов получения информации (угроза использования злоумышленником недокументированных возможностей ОС, уязвимостей нулевого дня и др.).
- Создание новых каналов получения информации с помощью встроенного вредоносного кода (использование логических бомб, троянских программ и др.).

Должного внимания заслуживают вопросы сетевого взаимодействия в IoT. По этому направлению возможна реализация следующих векторов атак на информационную систему:

1. Прослушивание сетевых данных. Сюда относится сетевая разведка и анализ сетевого трафика, использование снифферов на разных уровнях модели OSI. Использование умных устройств размывает границы локальных сетей, принадлежащих отдельным субъектам, и это может позволить злоумышленнику на основе открытых данных получить доступ к закрытым сетям.

2. Влияние на сетевые ресурсы. Это очень широкая категория к которой можно отнести различные виды спуфинга на разных уровнях модели OSI, атаки типа Man in the middle, инъекции кода, подмена доверенных субъектов (session hijacking), атаки типа отказ в обслуживании (DDoS), атаки типа brutforce, криптоанализ алгоритмов передачи. При широком распространении IoT, и недостаточном обеспечении безопасности включенных в него конечных систем, количество возможных объектов для DDos-атак, а также потенциальный ущерб от их осуществления многократно возрастет. Отказ умных бытовых приборов может приводить к неполадкам в электросети и пожарам. Отказ датчиков обратной связи в системе АСУ ТП может

привести не только к потере прибыли из-за задержки производства, но и к различным чрезвычайным ситуациям.

Кроме того, устройства подключенные к сети, являются потенциальными агентами бот-сетей и реализации атак этого вектора приведут к несанкционированному использованию вычислительных ресурсов этих устройств для рассылки спама или участия в DDoS-атаках. Кроме проблем чисто технического характера следует также упомянуть проблемы организационные и нормативные:

- Масштаб создаваемой инфраструктуры. Подключенные устройства будут генерировать очень большой трафик и для его обслуживания необходимо будет обеспечить возможности расширяемости и масштабируемости сетевой инфраструктуры.

- Сложности с документированием методов и средств обработки информации во всех конечных устройствах. Значительная часть IoT будет подотчетна не корпоративным организациям, с четкими прописанными политиками безопасности, а простым пользователям, которые подобными вопросами вообще не занимаются. Что в свою очередь приведет к очень быстрому росту бот-сетей на основе умных домов, или умных бытовых приборов, а также массового нарушения конфиденциальности персональных данных.

- Отсутствие единого подхода. Несмотря на то, что существует множество современных технологий с помощью которых можно построить IoT, широко принятых унифицированных стандартов для разработчиков все еще нет. Отсутствие стандартизированного подхода к построению IoT приведет к возникновению новых типов угроз и общей уязвимости системы.

Заключение

С точки зрения технической стандартизации, IoT можно рассматривать как глобальную инфраструктуру для информационного общества, которая сочетает передовые услуги на основе существующих и разрабатываемых информационных и коммуникационных технологий.

В ближайшем будущем IoT будет непосредственно причастно как к жизни простых людей так и к бизнесу и государственной деятельности. То есть такую сложную структуру необходимо строить с учетом современных требований к информационной безопасности.

К вопросу обеспечения защищенности информации в пределах IoT необходимо подходить комплексно и особенно уделять внимание таким аспектам как безопасность конечных информационных систем и безопасность их взаимодействия.

Конечные системы в целях декомпозиции можно разделить на встроенные и серверные системы, соответственно и подходы к их защите будут отличаться. Основной задачей этого направления является построение защищенных операционных систем, которые смогут эффективно противостоять информационным атакам, как по открытым, так и закрытым каналам передачи информации.

С точки зрения сетевого взаимодействия следует уделять внимание защите от прослушивания сетевых данных и влияния на сетевые ресурсы.

Также отдельно следует отметить наличие организационных и нормативных проблем IoT, таких как отсутствие единого подхода к стандартизации взаимодействия между отдельными информационными системами, сложности с организацией документированности информационных процессов, в частности политик безопасности, и вопрос масштабируемости.

Литература

1. Росляков А.В., Ваняшин С.В., Гребешков А.Ю. Интернет вещей. Учебное пособие. Книга, 2015. 136 с.
2. Палагута К.А., Шубникова И.С., Сафонов А.Л. Справочник модуля «Умный дом». Книга, 2014. 184 с.
3. Kevin Ashton. That 'Internet of Things' Thing [Электронный ресурс] // RFID Journal, 2009 // Режим доступа: <http://www.rfidjournal.com/articles/view?4986> (26.09.2018 г.).
4. Гольдштейн Б.С., Кучерявый А.Е. Сети связи пост-NGN. СПб.: БХВ- Петербург, 2015. 160 с.
5. Соколов М.Н., Смолянинова К.А., Якушина Н.А. Проблемы безопасности интернета вещей: обзор // Вопросы кибербезопасности. 2015. № 5(13). 34 с.
6. Орлов В.Г., Мазуркевич Д.О. Алгоритмы шифрования в публичных беспроводных сетях // Т-Сomm: Телекоммуникации и транспорт. 2011. Т. 5. № 10. С. 62-64.

CLASSIFICATION OF THREATS OF THE INTERNET OF THINGS

Ponomarev Alexey Alexandrovich

student group M091801(72)

alex1996@mail.ru

Sergey Shevelev Vladimirovich

MTUCI, PhD., associate professor of MSIS

Key words: *internet of things, security, information, It, information technology.*

The problem of information security in the Internet of things is considered. The proposed decomposition of the problem by the type of interacting devices and the nature of the interaction. Considered server-side and embedded ("embedded") systems, their specificity and many potential threats.

МЕТОДЫ И СРЕДСТВА РАЗРАБОТКИ ИНТЕРАКТИВНОЙ СИСТЕМЫ ОБУЧЕНИЯ ДЛЯ ОТСЛЕЖИВАНИЯ УРОВНЯ ПОДГОТОВКИ СТУДЕНТОВ

Доткулова Анастасия Сергеевна
МАДИ, Старший преподаватель кафедры ВМ
МТУСИ, Старший преподаватель кафедры МКиИТ
asoll05@inbox.ru

Порунов Алексей Алексеевич
студент группы М091801(70) МТУСИ
alex.porunof@yandex.ru

Ключевые слова: клиент-серверная система, мобильные приложения, расписание занятий, информационные технологии в образовании, интерактивная система обучения

Анализируются современные методы подачи и обработки информации в образовании в связи с революционными процессами в цифровой экономике. Описаны причины, побудившие к разработке интерактивной обучающей системы, а также модуля генерации расписания занятий и учета и посещения. Рассмотрены основные функции и алгоритмы обучения с внедрением интерактивных элементов, автоматизированной проверки знаний и статистических данных об успеваемости.

Новые методы получения, хранения и переработки информации, процессы информационного взаимодействия в обществе, а также изменения объема информации - основа информационной революции. Информационные революции всегда являлись новым этапом развития цивилизации.

Big Data в образовании

В мире происходит быстрый рост объема информации. Из-за роста данного объема и количества пользователей, использующих компьютеры и смартфоны, актуальной задачей является создание новых методов подачи и обработки большого количества информации (Big Data) с помощью информационных инструментариев, являющихся средствами хранения, передачи и обработки информации (хард-часть): сервера, сети и т.д. [1].

Big Data оказывает большое влияние на образование. Всё больше университетов и колледжей обращаются к технологиям анализа больших данных для улучшения различных показателей. С ростом числа ресурсов для онлайн-образования и развития так называемых massive open online courses (МООС) все подобные данные приобретают особую значимость [2].

МООС активно развиваются. Во-первых, как элемент дополнительного образования, во-вторых, как источник получения альтернативного качественного образования. МООС являются одними из представителей образования E-learning.

E-learning и его компоненты

E-learning – это обучение с помощью Интернет и мультимедиа. Электронное обучение включает в себя три ключевых компонента – технологии, люди и процессы, [3].

Управление процессом обучения осуществляет преподаватель, который контролирует прохождение обучения учениками, записывает результаты усвоения материала, добавляет необходимую литературу, видеолекции или проводит вебинары, осуществляет создание и проверку тестовых работ и принимает окончательную аттестационную работу [4, 6 - 7].

Электронное обучение предполагает значительный объем самостоятельной работы. Обычно это достигается с помощью видео и веб-конференций в виде лекций и семинаров. Групповые занятия и лабораторные работы проходят на форумах или с помощью веб-

конференций. E-learning актуальное, эффективное и перспективное направление, в котором можно создать новые инструментари для улучшения качества современного образования. E-learning дало развитие ITS (Intelligent Tutoring System).

В 1970 году J. Carbonell предположил, что компьютер может работать в роли учителя. Мысль, использовать компьютер для интеллектуального обучения студентов, начала развиваться и получила название интеллектуальная обучающая система (Intelligent Computer Assisted Instruction или Intelligent Tutoring Systems (ITS)).

Интеллектуальная обучающая система (ИОС) (Intelligent tutoring system (ITS)) – это компьютерная система, часто основанная на искусственном интеллекте, предназначенная для обучения пользователя и осуществления обратной связи с ним. Интеллектуальные обучающие системы предназначены для моделирования человеческого поведения, выполняя роль преподавателя.

Общее описание интерактивной информационной системы сопровождения образовательного процесса

Основываясь на ключевых компонентах электронного образования и особенностей ITS, авторами разработана интерактивная информационная система сопровождения образовательного процесса E-Tutor (Educational Tutoring System), предназначенная для повышения эффективности обучения путем проведения автоматизированных тестирований AutoTest, непосредственно связанного со средней успеваемостью студента и расписанием занятий.

В качестве помощника в процессе обучения авторами была разработана интерактивная система обучения и контроля знаний студента E-Tutor. Главной особенностью системы является использование интерактивных компонентов при самостоятельном изучении материала. Архитектурно проект представляет собой веб-интерфейс, а также мобильные приложения для мобильных ОС Android и IOS, что подразумевает разворачивание системы на подконтрольном сервере и последующее взаимодействие с ним через браузер либо мобильное приложение на совместимых устройствах, связывающихся с данным сервером по сети.

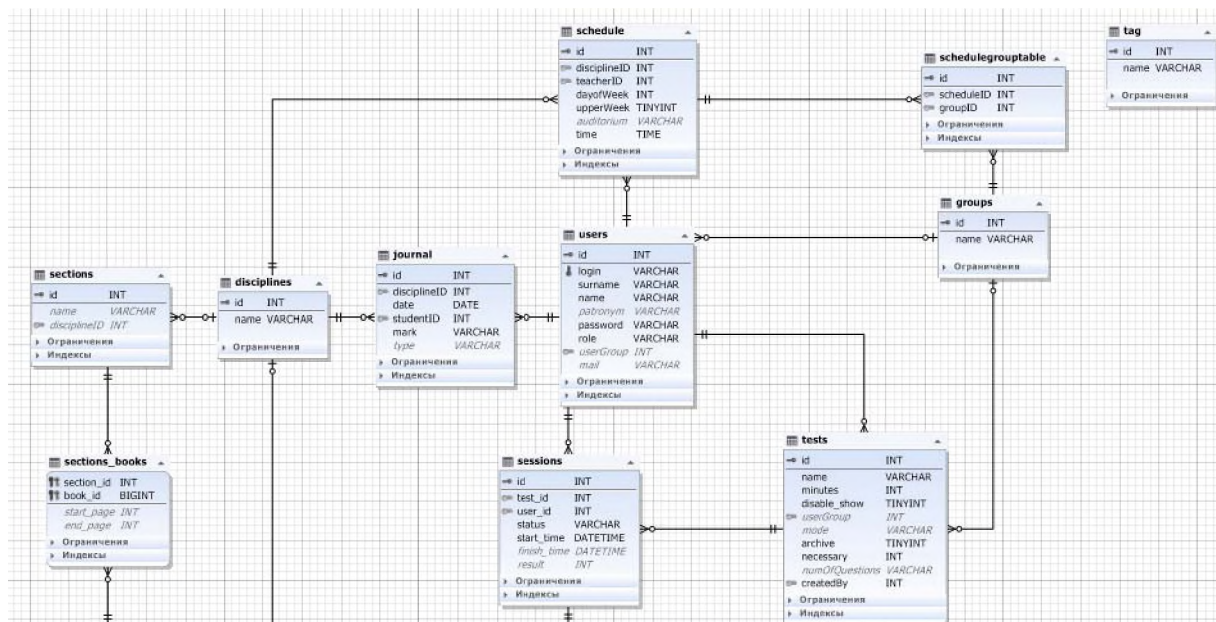


Рис. 1. Схема базы данных

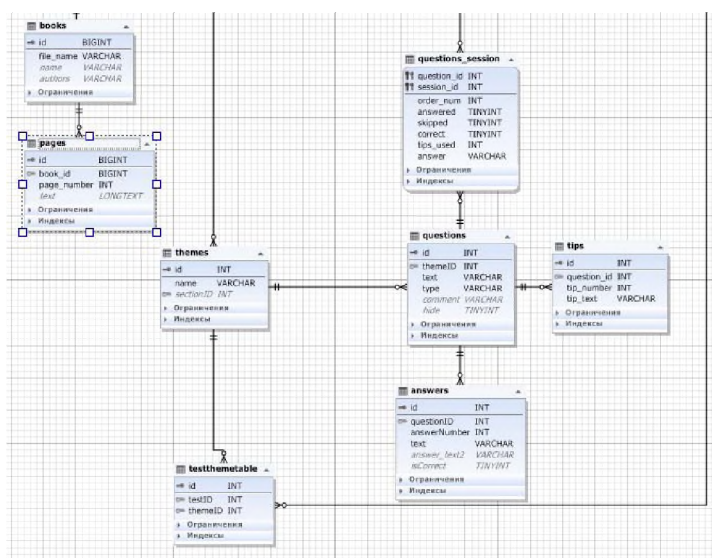


Рис. 2. Схема базы данных (Продолжение)

Система состоит из четырех взаимосвязанных модулей: интерактивной обучающей системы E-Tutor, автоматизированной проверки знаний с интерактивными элементами AutoTest- E-Tutor, электронного журнала и модуля расписания. Обучающая система помогает студенту знакомиться с новым материалом (или повторять уже изученный), проходить тестирование с запланированными подсказками и получать дальнейшие рекомендации. Проверка знаний основана на онлайн-тестировании, имеющим свои особенности, в частности, интерактивный режим «подсказка» и возможность осуществлять поиск информации по каталогу учебной и дополнительной литературы. Рассмотренные первые два модуля имеют общую часть: библиотечный каталог (представляет собой базу данных учебно-методической и вспомогательной литературы) и поисковую строку, в которую вводится запрос и осуществляется доступ к необходимому материалу из БД. Для удобства пользователя в системе предусмотрена инструкция по работе с поисковой системой, где рассматриваются функции системы, помогающие сделать результат вывода более точным.

Электронный журнал и модуль расписания направлены на облегчения доступа студента либо преподавателя к информации, которая также является важной и играющей большую роль в рамках высшего учебного заведения.

В системе E-Tutor реализовано разграничение доступа к различным функциональным возможностям между несколькими категориями пользователей. В разработанной системе присутствуют три роли:

1. Администратор — может изменять роль всех остальных пользователей и редактировать глобальные настройки веб-приложения;
2. Преподаватель — может загружать учебно-методическую и вспомогательную литературу в БД, генерировать содержание разделов дисциплин, создавать и редактировать тесты, подавлять подсказки, учебные курсы и группы, просматривать результаты прохождения всех тестов и результативность обучения в целом;
3. Учащийся — может обучаться в интерактивной системе, изучать доступную ему литературы в целом или по конкретно выбранной теме, проходить доступные его группе тесты и интерактивные курсы, просматривать свои результаты и получать индивидуальные рекомендации по дальнейшему изучению материала.

Авторизация осуществляется с помощью логина и пароля, пароль хранится в зашифрованном виде из соображений безопасности. Используется механизм сессий. Регистрация открытого типа: роль, присваиваемая при регистрации по умолчанию «учащийся», а возможность изменять ее есть только у администраторов. При регистрации, помимо логина и пароля, указывается ФИО и выбирается группа из списка доступных в системе.

Обучающая система E-Tutor глазами преподавателя

На главной странице преподавателя отображается список существующих дисциплин и их разделов, а также расписание преподавателя на текущую неделю.

Для каждой дисциплины/раздела/темы можно добавить список необходимой литературы (он будет использоваться при выводе подсказок). Учебно-вспомогательная литература выбирается из загруженных в БД материалов. Когда выбирается список литературы, имеется возможность прикрепить как источник целиком, так и определённый диапазон страниц.

Для каждой из тем можно создать интерактивный курс в формате тестирования. При создании тестовых вопросов имеется возможность выбрать один из четырех типов вопроса: один правильный ответ, несколько правильных ответов, свободный ввод и указание соответствия. Каждый из вопросов может сопровождаться набором подсказок и отсылками к литературе: на их основе пользователь может восполнять пробелы в своих знаниях.

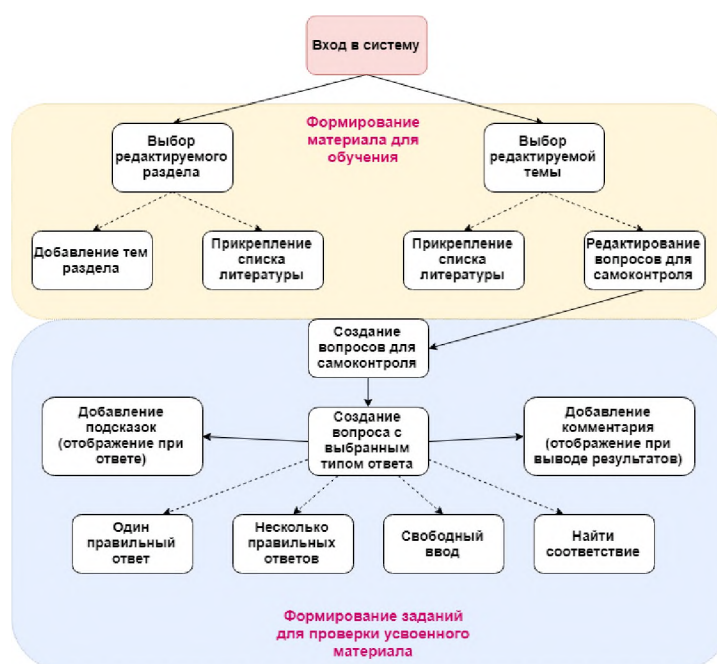


Рис. 3. Алгоритм работы преподавателя при заполнении обучающего модуля системы

При нажатии на предмет в расписании открывается окно с текущим расписанием на конкретный предмет и группу. Есть возможность проставить отсутствующих, а также оценки.

Данная возможность ориентирована на облегчение работы с успеваемостью студентов преподавателями. Преподаватель может также распечатать данный журнал, так как порой, бумажный отчет все-таки необходим.

Понедельник 15.04.2019

Занятий нет

Вторник 16.04.2019

Занятий нет

Среда 17.04.2019

Время	Предмет	Группа	Аудитория	
09:30:00	СкАОД	ББТ1701,БФИ1701,БУТ1701	501	Журнал
11:15:00	ПЯП	ББТ1801,ББТ1803,ББТ1802,ББТ1804	501	Журнал

Четверг 18.04.2019

Время	Предмет	Группа	Аудитория	
13:30:00	ГИС	ББТ1401	407	Журнал

Пятница 19.04.2019

Занятий нет

Суббота 20.04.2019

Занятий нет

Рис. 4. Интерфейс текущего расписания преподавателя

	10.04.2019	13.04.2019	16.04.2019	18.04.2019	21.04.2019	23.04.2019	28.04.2019	01.05.2019	11.05.2019
АКСЕНОВА ЯРОСЛАВА ДМИТРИЕВНА					Н				
АУШЕВ ХАМЗАТ МАГОМЕТ- БАШИРОВИЧ							Н		
БЛАШКО НИКИТА ИВАНОВИЧ						Н			
ГЕРАСИМЕНКО НИКИТА МИХАЙЛОВИЧ							Н		
ГЛЕБОВ НИКОЛАЙ СЕРГЕЕВИЧ			Н						
ГОРОХОВ ПЛАТОН ЮРЬЕВИЧ	Н								
ДЕТКОВ ПАВЛА								Н	

[Добавить тест](#)

[База вопросов](#)

[Результаты тестов](#)

[Сменить пароль](#)

Рис. 5. Интерфейс журнала

Обучающая система E-Tutor глазами студента

Учащийся при входе в систему видит тот же самый список существующих дисциплин и их разделов. Для обучения пользователь также может обратиться к библиотечному модулю системы, где ему доступны все загруженные в систему материалы, а также полнотекстовый поиск по ним.

Если же учащийся заинтересован в конкретной теме, он может выбрать её из исходного списка. После этого ему будут показаны связанные непосредственно с ней материалы, указанные ранее преподавателем при редактировании, а также пройти интерактивный курс (при

наличии). Как упоминалось ранее, курс по механике напоминает собой тестирование: пример интерфейса показан на рисунке 7.

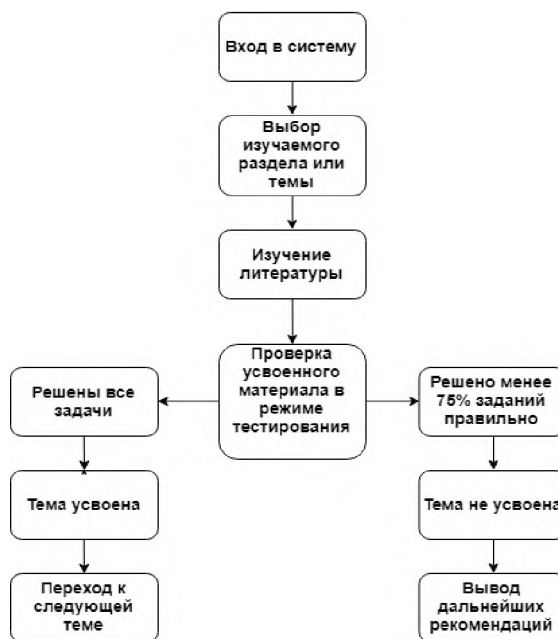


Рис. 6. Блок-схема действий студента при работе с обучающим модулем

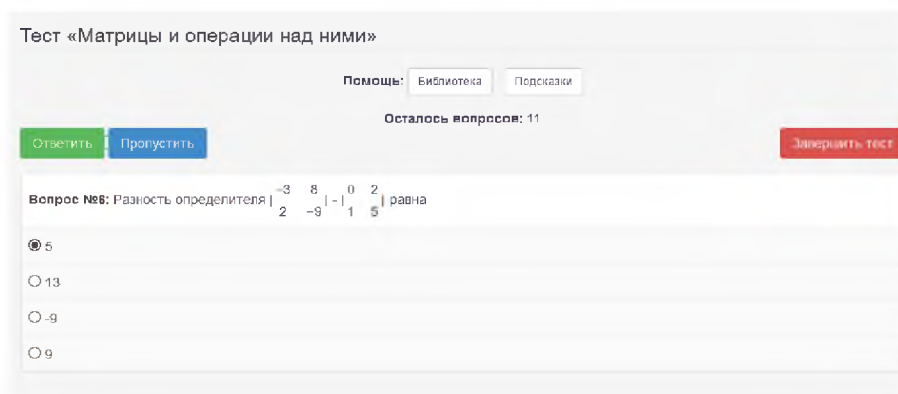


Рис. 7. Пример интерфейса интерактивного курса

По результатам прохождения курса в рамках темы (затраченного времени и правильности ответов на вопросы в них) пользователю предоставляются рекомендации о дальнейших действиях. Эти рекомендации могут включать в себя предложение перейти к следующей теме, дополнительно изучить какую-то подтему и даже вернуться к какой-то из логически более ранних тем, если системе удаётся выявить, что проблема заключается в нехватке фундаментальных относительно пройденной темы знаний.

Также, как и преподаватель, студент имеет доступ к журналу и может отслеживать свои посещения по любому предмету, которые есть у него в учебном плане.

Выводы и перспективы

Настоящую работу можно рассматривать как одну из попыток автоматизировать отдельные элементы диалога преподавателя и студента, в процессе которого последний приобретает знания, а первый – получает удовлетворение от процесса обучения, несмотря на всю тяжесть затрачиваемых усилий.

Современные исследования по искусственному интеллекту, нейронным сетям позволяют надеяться, что в данном направлении удастся продвинуться, ни в коем случае не утверждая пока, что живое общение студента и специалиста полностью уходит в прошлое.

Литература

1. Kenneth Cukier "Learning With Big Data". New York, 2014
<http://www.hmhco.com/~media/sites/home/educators/webinars/summer-session/LearningWithBigData-shortened.pdf/>
2. Аналитический доклад института информатизации ЮНЕСКО
<http://portal.unesco.org/>
3. Lewin T. Universities abroad join partnerships on the web // The New York Times. 2013. Т. 20. №. 2013. С. 21.02.
4. SeeMedia. Обзор Мирового и российского рынка электронного обучения. Июнь 2012/
5. Обучающая платформа Гоогу с использованием нейросети. Калифорния
<https://gooru.org/>
6. Доткулова А.С., Мосева М.С., Яшина М.В. Особенности систем информационного сопровождения образовательного процесса в технических университетах // Т-Comm: Телекоммуникации и транспорт. 2017. Т. 11. № 9. С. 58-64.
7. Доткулова А.С., Мосева М.С. Метод интеллектуальной верификации логических связей модулей ООП в соответствии С ФГОС ВО // Труды Северо-Кавказского филиала Московского технического университета связи и информатики. 2017. № 2. С. 287-293.

METHODS AND MEANS FOR DEVELOPING THE INTERACTIVE TUTORING SYSTEM FOR TRACKING THE LEVEL OF STUDENT PREPARATION

Anastasia S. Dotkulova

MADI, Senior Lecturer of VM department

MTUCI, Senior Lecturer of MKiIT department

asoll05@inbox.ru

Alexey A. Porunov

Student of group M091801(70), MTUCI

Alex.porunof@yandex.ru

Keywords: *client-server system, mobile applications, class schedule, information technologies in education, interactive learning system*

The article analyzes the modern methods of providing and processing information in the field of revolutionary processes in the digital economy. The reasons that led to the development of an interactive learning system are described. Consideration of the main functions and learning algorithms using interactive elements and automated knowledge tests with the conclusion of further recommendations.

РАЗРАБОТКА АЛГОРИТМА, РЕАЛИЗУЮЩЕГО МЕТОД СПЕКТРАЛЬНОГО АНАЛИЗА ПРОНИ В СРЕДЕ МАТЛАБ, И ИССЛЕДОВАНИЕ ЧУВСТВИТЕЛЬНОСТИ ЭТОГО МЕТОДА К ВОЗДЕЙСТВИЮ ШУМОВЫХ ПОМЕХ

Бушуев Эвелин Юрьевич
студент группы БЗС1602, МТУСИ
e.bushuev71521@gmail.com

Волчков Валерий Павлович
д. т. н., профессор кафедры ОТС, МТУСИ
volchkovvalery@mail.ru

Ключевые слова: спектральный анализ, спектр Прони, метод наименьших квадратов (МНК) Прони в Матлаб, дискретная экспоненциальная модель, аппроксимация, корреляционная функция, белый шум, коррелированный шум.

Рассматривается задача синтеза алгоритма спектрального анализа на основе метода наименьших квадратов (МНК) Прони и его последующая реализация в среде Матлаб. Приводятся основные этапы синтеза: методика построения адекватных дискретных моделей сигналов, оценка параметров дискретной модели Прони p -го порядка по наблюдаемой выборке размера $N > p$ и оценивание спектральных плотностей этих сигналов методом Прони. Приводятся результаты моделирования и исследования влияния шумовых помех на спектральную оценку Прони.

Введение

Одной из важнейших задач современных интеллектуальных систем обработки информации является разработка эффективных методов и алгоритмов спектрально-временного анализа процессов, наблюдаемых на выходе различных устройств регистрации [4, 5]. В качестве таких устройств, например, могут выступать датчики биомедицинских приборов, приемники эхо сигналов в радиолокаторах (или гидролокаторах) различного назначения, сейсмографические датчики, системы мониторинга земной поверхности и др. Вся получаемая при этом информация оцифровывается, чтобы потом ее можно было обрабатывать или в реальном времени, или в отложенном режиме, с помощью специализированных алгоритмов.

В данной работе рассмотрены разработка алгоритма, реализующего метод спектрального анализа Прони [1] в среде Матлаб, и исследование чувствительности этого алгоритма к воздействию шумовых помех. Данная тема научной работы весьма актуальна, поскольку m -функция, реализующая метод наименьших квадратов (МНК) Прони, который и будет использоваться для спектрального анализа сигналов в данной работе, отсутствует в пакете Матлаб Signal Processing Toolbox [3], а экспериментальное исследование влияния шумовых помех на работу алгоритма имеет научно-практическую значимость.

1. Решение поставленной задачи

Для синтеза алгоритма спектрального анализа на основе МНК Прони [1] в среде Матлаб необходимо выполнить следующее:

1. Построить адекватные дискретные модели сигналов.
2. Оценить параметры дискретной модели Прони p -го порядка по наблюдаемой выборке размера $N > p$ методом наименьших квадратов.
3. Оценить спектральные плотности сигналов методом Прони.

Ниже рассматриваются перечисленные выше этапы исследования.

1.1. Построение адекватных моделей наблюдаемых процессов

Для проведения экспериментов в качестве исходной модели непрерывного сигнала $s(t)$ выбирается детерминированный процесс, состоящий из трех квазигармонических сигналов с известной корреляционной функцией

$$\begin{aligned} R_s(\tau) &= \int_{-\infty}^{\infty} s(t)s(t+\tau)dt = \sum_{k=1}^3 R_s^{(k)}(\tau), \quad \tau \in \mathbb{R}, \\ R_s^{(k)}(\tau) &= D_s^{(k)} \exp(-\alpha_k \tau^2) \cos(2\pi f_k \tau), \end{aligned} \quad (1)$$

где $R_s^{(k)}(\tau)$ – корреляционная функция k -ой квазигармоники. Сигналу $s(t)$ соответствует спектральная плотность энергии (СПЭ)

$$\begin{aligned} G_s(f) &= \int_{-\infty}^{\infty} R_s(\tau) \exp(-j2\pi f \tau) dt = \sum_{k=1}^3 G_s^{(k)}(f), \\ G_s^{(k)}(f) &= \int_{-\infty}^{\infty} R_s^{(k)}(\tau) \exp(-j2\pi f \tau) dt. \end{aligned}$$

Учитывая (1), корреляционная функция дискретного процесса $s_i \equiv s[i] = s(\Delta i)$ в конечном числе точек на интервале длительностью T описывается выражением:

$$\begin{aligned} R_s[i] &= M[s_o, s_i] = \sum_{k=1}^3 R_s^{(k)}[i], \quad i \in J_N = \{0, \dots, N-1\}, \\ R_s^{(k)}[i] &= D_s^{(k)} \exp(-\alpha_k (\Delta i)^2) \cos(2\pi f_k \Delta i), \end{aligned} \quad (2)$$

где

$$f_1 = 15, \quad f_2 = 30, \quad f_3 = 50, \quad \alpha_1 = 78.5, \quad \alpha_2 = 201, \quad \alpha_3 = 706.5, \quad D_s^{(1)} = 0.8, \quad D_s^{(2)} = 0.7, \quad D_s^{(3)} = 0.5$$

– параметры, отвечающие за сдвиг по частоте, амплитуду или ширину спектра квазигармоники; $f_d = 1/\Delta = 2f_{\max} = 130$ [Гц] – выбранная по теореме Котельникова частота дискретизации, равная удвоенной максимальной частоте f_{\max} спектра $G_s(f)$; Δ – интервал дискретизации, $N = 2 \lfloor T f_{\max} \rfloor$. Здесь $\lfloor \cdot \rfloor$ – оператор взятия целой части числа, f_{\max} определяется по графику $G_s(f)$ на уровне $0.01 \cdot G_{\max}$, длительность T определяется по графику $R_s(\tau)$ на уровне $0.01 \cdot R_{\max}$.

С точностью до задержки, адекватную дискретную модель сигнала $s_i, i \in J_N$, отвечающую заданной корреляционной функции (2), можно сформировать по ее отсчетам, с помощью следующей процедуры:

1. По формуле (1) вычисляются N отсчетов корреляционной функции $R_s[i]$, а затем к ним добавляются N нулевых отсчетов, продолжая тем самым ее на расширенный дискретный интервал $J_{2N} = \{0, 1, \dots, 2N-1\}$.

2. К полученным отсчетам применяется дискретное преобразование Фурье (ДПФ), т. е. находится дискретный спектр корреляционной функции сигнала $P_R[k] = \text{ДПФ}(R_s[i]), k \in J_{2N}$. По теореме Винера-Хинчина этот спектр по смыслу совпадает с дискретной экстраполированной СПЭ сигнала s_i на интервале J_{2N} , т. е. $G_s[k] = P_R[k]$.

3. Из полученных отсчетов $G_s[k]$ извлекается квадратный корень и выполняется обратное дискретное преобразование Фурье (ОДПФ), т. е. $s[i] = \text{ОДПФ}(\sqrt{G_s[k]}), i \in J_{2N}$.

4. Выполняется редукция полученного сигнала $s[i]$, при которой его последние N отсчетов отбрасываются.

В результате (рис. 1а) получаем отсчеты искомого сигнала $s_i, i \in J_N = \{0, \dots, N-1\}$, у которого корреляционная функция совпадает с заданной (2).

Учитывая (2), мощность сигнала $R_s[0]$ определяется следующим образом:

$$R_s[0] = \sum_{k=1}^3 R_s^{(k)}[0] = \sum_{k=1}^3 D_s^{(k)} = 0.8 + 0.7 + 0.5 = 2.$$

В качестве случайной помехи будем использовать непрерывный аддитивный белый гауссовский шум (АБГШ) $\beta(t)$ с интенсивностью b и корреляционной функцией $R_n(\tau) = \int_{-\infty}^{\infty} n(t)n(t+\tau)dt = b \cdot D(\tau)$, где $D(\tau)$ – дельта-функция Дирака. Известно [3], что дискретизация такого шума по теореме Котельникова приводит к дискретной случайной последовательности β_i , отсчеты которой взаимно некоррелированы и распределены по гауссовскому закону с нулевым математическим ожиданием $M[\beta_i] = 0$ и дисперсией (мощностью) $M[\beta_i^2] = R_n[0] = b$, что математически записывается следующим образом:

$$R_n[i] = M[\beta_0, \beta_i] = b \cdot \delta_i, \quad \delta_i = \begin{cases} 1, & i = 0 \\ 0, & i \neq 0 \end{cases} \quad (3)$$

Таким образом, зашумленный непрерывный процесс $x(t)$ описывается выражением

$$x(t) = s(t) + \beta(t), \quad t \in \mathbb{R},$$

а его дискретный аналог x_i выражением

$$x_i = s_i + \beta_i, \quad i \in J_N. \quad (4)$$

Корреляционная функция дискретного процесса (4) имеет следующий вид:

$$R_x[i] = M[x_0, x_i] = R_s[i] + b \cdot \delta_i.$$

Отношением сигнал/шум в аддитивной смеси (4) называется величина

$$\rho = R_s[0]/b \quad (\text{в разгах}), \quad \rho_{db} = 10 \log_{10}(R_s[0]/b) \quad (\text{в децибелах}) \quad (5)$$

Дискретные отсчеты белого гауссовского шума β_i с корреляционной функцией (3) формируются в Матлаб с помощью датчика гауссовских случайных чисел командой *randn*, которая на выходе выдает реализации независимых отсчетов шума.

На рис. 1б и 1в представлены графики зашумленного процесса (4) при разных отношениях сигнал/шум.

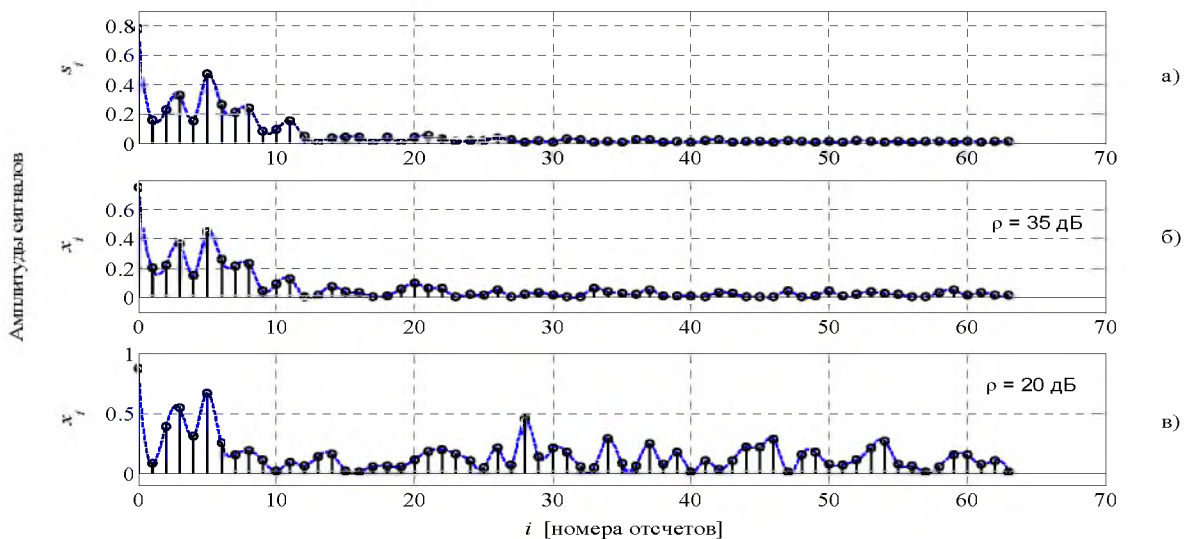


Рис. 1. а) исходный сигнал, сформированный по отсчетам своей корреляционной функции; б, в) результаты суммирования исходного сигнала и АБГШ при отношениях сигнал/шум $\rho = 35$ дБ и $\rho = 20$ дБ

Помимо АБГШ, будем использовать в эксперименте коррелированный шум (КШ) $\gamma(\tau)$. Коррелированный шум на практике может возникать либо как непреднамеренная помеха приему сигнала со стороны близлежащих по диапазону частот радиостанций (например, в сотовых системах связи некоторые удаленные соты могут работать на одной и той же центральной частоте), либо как преднамеренная помеха, когда сигнал хотят подавить (такая ситуация возникает при радиоэлектронной борьбе с противником). Для синтеза КШ за основу возьмем корреляционную функцию

$$R_\gamma(\tau) = e^{-\alpha|\tau|} \cos(2\pi f_0 \tau), \quad \tau \in \mathbb{R}, \quad (6)$$

приведенную в п. 7 таблицы корреляционных функций на стр. 494 в [2], где f_0 – центральная частота в спектре процесса $\gamma(\tau)$, α – параметр, определяющий ширину спектра шума. Дискретный аналог корреляционной функции (6) имеет вид:

$$R_\gamma[i] = e^{-\alpha \Delta i} \cos(2\pi f_0 \Delta i), \quad i \in J_N. \quad (7)$$

СПЭ КШ может пересекать СПЭ сигнала, и максимальное действие КШ будет, когда его спектр находится в центре спектра сигнала. Другими словами, для большей чувствительности ошибок метода Прони к действию шума для эксперимента выберем центральную частоту f_0 , совпадающую с центральной частотой сигнала $f_0 = f_n / 2 = f_d / 4 = 32.5$, где $f_n = f_d / 2$ – частота Найквиста.

Используя ансамбль реализаций АБГШ $\mathcal{B} = \{\beta_i, i \in J_N\}$ и корреляционную функцию (7), ансамбль коррелированных шумов $\mathcal{T} = \{\gamma_i, i \in J_N\}$ можно получить следующим образом:

1. Составляется корреляционная матрица

$$\mathbf{R}_\gamma = (R_\gamma(i, j)), \quad R_\gamma(i, j) = R_\gamma[|i - j|], \quad i, j = 0, \dots, N-1. \quad (8)$$

2. Из полученной корреляционной матрицы (8) извлекается матричный квадратный корень, а затем вычисленная матрица умножается на векторы $\boldsymbol{\beta} = (\beta_0, \dots, \beta_{N-1})^T$ из ансамбля \mathcal{B} . В результате получаем векторы $\boldsymbol{\gamma} = (\gamma_0, \dots, \gamma_{N-1})^T$ ансамбля \mathcal{T} коррелированных шумов:

$$\boldsymbol{\gamma} = \mathbf{R}_\gamma^{1/2} \cdot \boldsymbol{\beta}.$$

Аналогично (4), зашумленный таким шумом сигнал описывается выражением

$$y_i = s_i + \gamma_i, \quad i \in J_N. \quad (9)$$

Корреляционная функция дискретного процесса (9) имеет следующий вид:

$$R_y[i] = M[y_0, y_i] = R_s[i] + R_\gamma[i], \quad R_\gamma[i] = R_\gamma(0, i).$$

Отношение сигнал/шум процесса (9) определяется аналогично (5) по формулам

$$\rho = R_s[0] / R_\gamma[0], \quad \rho_{db} = 10 \log_{10}(R_s[0] / R_\gamma[0]) \text{ [дБ]},$$

где мощность шума $R_\gamma[0]$ определяется следующим образом $R_\gamma[0] = M[\gamma_i^2] = 2\alpha$.

На рис. 2 представлены графики зашумленного процесса (9) при разных отношениях сигнал/шум.

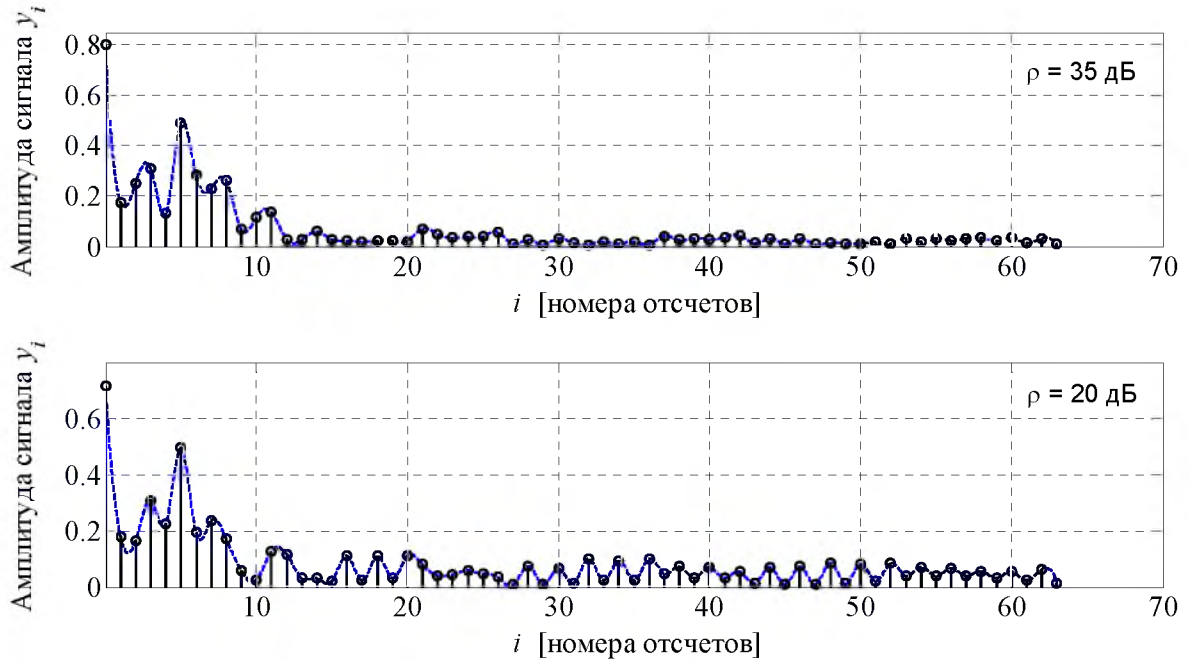


Рис. 2. Результаты суммирования исходного сигнала и КШ при отношениях сигнал/шум $\rho = 35$ дБ ($\alpha = 3.25 \cdot 10^{-4}$) и $\rho = 20$ дБ ($\alpha = 10^{-2}$)

1.2. Метод наименьших квадратов Прони

Пусть известна дискретная реализация $x_n = x[n]$, $n \in J_{N+1} = \{1, \dots, N\}$, некоторого комплексного сигнала, который может быть детерминированным или случайным. В частности, сигнал может быть и вещественным, например, любым из описанных в разделе 1.1. Тогда согласно [1], мы можем аппроксимировать последовательность $x[n]$, $n \in J_N$, дискретной экспоненциальной моделью Прони p -го порядка:

$$\hat{x}[n] = \sum_{k=1}^p h_k z_k^{n-1}, \quad 1 \leq n \leq N, \quad N > 2p, \quad (10)$$

$$h_k = A_k \exp(j\theta_k), \quad z_k = \exp[(\alpha_k + j2\pi f_k)T],$$

где h_k – комплексные амплитуды, z_k – комплексные экспоненты модели Прони, которые необходимо подобрать так, чтобы последовательность $\hat{x}[n]$ наилучшим образом аппроксимировала реализацию сигнала $x[n]$ в смысле некоторого критерия оптимальности.

Однако, одновременное нахождение порядка p и параметров $\{h_k, z_k\}$, где $1 \leq k \leq p$, которые минимизируют сумму квадратов ошибки

$$\rho_{\text{ош}} = \sum_{n=1}^N |\varepsilon[n]|^2, \quad \varepsilon[n] = x[n] - \hat{x}[n]$$

представляет собой трудную нелинейную задачу. Используя вариант метода Прони, описанный в [1, разд. 11.4], можно определить субоптимальное решение, обеспечивающее удовлетворительные результаты. При таком подходе задача экспоненциальной подгонки сводится по сути дела к задаче полиномиальной факторизации

$$\sum_{m=0}^p a[m] z^{p-m} = \prod_{k=1}^p (z - z_k) \quad (11)$$

и связанной с ней задачей оценивания параметров $a[m]$ модели авторегрессии p -го порядка

$$\sum_{m=1}^p a[m]x[n-m] = e[n], \quad p+1 \leq n \leq N, \quad (12)$$

Для ее решения используется ковариационный метод линейного предсказания [1], в котором параметры $a[m]$ определяются по критерию минимума суммы квадратов ошибок линейного предсказания $e[n]$, т.е.

$$a[m]: \sum_{n=p+1}^N |e[n]|^2 \rightarrow \min. \quad (13)$$

Если значения параметров $a[m]$ были определены по критерию (13) методом наименьших квадратов (МНК), а затем найдены корни z_1, \dots, z_p полинома (11), то экспоненциальная аппроксимация (10) становится линейной относительно оставшихся неизвестных параметров h_1, \dots, h_p . Минимизируя сумму квадратов ошибок по каждому параметр h_k , получаем следующее комплексное нормальное уравнение [1]:

$$(\mathbf{Z}^H \mathbf{Z}) \mathbf{h} = (\mathbf{Z}^H \mathbf{x}), \quad (14)$$

с матрицей $\mathbf{Z}^H \mathbf{Z}$ размером $p \times p$, где надстрочный символ « H » в обозначении матрицы \mathbf{Z}^H означает эрмитово сопряжение матрицы \mathbf{Z} , а $(N \times p)$ -матрица \mathbf{Z} , $(p \times 1)$ -вектор \mathbf{h} и $(N \times 1)$ -вектор отсчетов данных \mathbf{x} определяются выражениями:

$$\mathbf{Z} = \begin{pmatrix} 1 & 1 & \dots & 1 \\ z_1 & z_2 & & z_p \\ \vdots & \vdots & & \vdots \\ z_1^{N-1} & z_2^{N-1} & \dots & z_p^{N-1} \end{pmatrix}, \quad \mathbf{h} = \begin{pmatrix} h_1 \\ h_2 \\ \vdots \\ h_p \end{pmatrix}, \quad \mathbf{x} = \begin{pmatrix} x[1] \\ x[2] \\ \vdots \\ x[N] \end{pmatrix}.$$

Ниже приводится программа, реализующая описанный выше вычислительный алгоритм МНК Прони и позволяющая в среде Матлаб вычислить $\{h_k, z_k, x_k\}$ параметры и значения аппроксимации Прони (10).

```
% mnk_prony.m - данная процедура позволяет вычислить комплексные параметры  $\{h_k, z_k\}$  и
% значения  $x_k$  аппроксимации Прони (10).
% -----
% Входные данные:
% X - входная последовательность (сигнал), p - порядок модели,
% N - число отсчетов в сигнале X.
% -----
% Выходные данные:
% h, z - комплексные параметры, Xn - аппроксимирующая экспоненциальная
% последовательность (10)
% -----
function [h,z,Xn]=mnk_prony(X,p,N)
a=arcov(X,p); % Решение уравнения (12) по критерию (13) методом МНК
z=roots(a); % Факторизация полинома (11)
Z=ones(p,');
for k=1:N
    Z(k,:)=z.^(k-1); % Создание матрицы Z
end
ZH=Z'; % Эрмитово сопряжение матрицы Z
h=ZH*Z\ZH*X; % Решение уравнения (14) относительно параметров h
Xn=zeros(N,1);
for n=1:N
```

```

(10)   for k=1:p      % Формирование аппроксимирующей экспоненциальной последовательности
        Xn(n)=Xn(n)+h(k)*z(k)^(n-1);
    end
end
end
end

```

1.3. Оценка спектральной плотности сигналов методом Прони

Решение уравнения (14) относительно неизвестных параметров h_k дает возможность вычислить спектр Прони [1]. Допустим, что сумма экспонент дискретного интервала времени в уравнении (10) определяется на интервале $-\infty < n < \infty$ как односторонняя функция вида:

$$\hat{x}[n+1] = \begin{cases} \sum_{k=1}^p h_k z_k^{n-1}, & n \geq 0; \\ 0, & n < 0. \end{cases} \quad (15)$$

Тогда, полагая, что сигнал $x[n]$ действителен и все параметры затухания отрицательны, z-преобразование от (15) будет иметь следующий вид:

$$\hat{X}(z) = \sum_{k=1}^p \frac{h_k}{1 - z_k z^{-1}}.$$

В этом случае спектральная плотность энергии (СПЭ) Прони для рассматриваемой односторонней экспоненциальной модели будет определяться следующим образом:

$$G_{\text{Пр}}(f) = \left| \Delta \hat{X}(z) \Big|_{z=\exp(j2\pi f T)} \right|^2 = \left| \Delta \hat{X}(f) \right|^2, \quad -1/2\Delta \leq f \leq 1/2\Delta, \quad \Delta = 1/f_d. \quad (16)$$

2. Результаты эксперимента

Погрешность вычисления СПЭ методом наименьших квадратов Прони при помощи дискретной модели p -го порядка будем определять следующим образом:

$$\varepsilon = \| G_S - G_{\text{Пр}} \| = \sqrt{\sum_{i=0}^{N-1} [G_S[i] - G_{\text{Пр}}[i]]^2}, \quad (17)$$

где G_S – дискретная СПЭ незашумленного сигнала, полученная в п. 2 на стр. 3, $G_{\text{Пр}}$ – дискретная СПЭ сигнала, рассчитанная по МНК Прони.

Ниже представлены результаты эксперимента – на рис. 3-5 приведены спектры (СПЭ) Прони (темно-синие сплошные линии), СПЭ, построенные методом ДПФ, (красные штриховые линии) и дискретные СПЭ, полученные в п.2 на стр. 3, (черные пунктирные линии) для сигналов на рис. 1 и рис. 2. Спектры Прони построены с использованием МНК Прони, описанного в разделе 1.2, и выражения (16). Также в табл. 1 сведены погрешности вычисления СПЭ, определяемые выражением (17), для дискретной модели порядка $p = 10$.

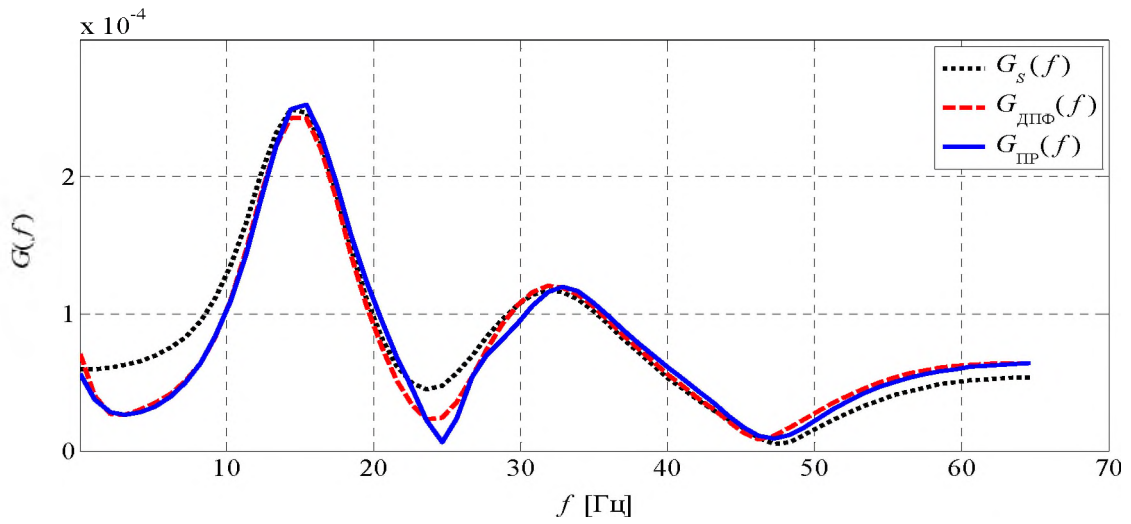


Рис. 3. Графики СПЭ незашумленного сигнала s_i

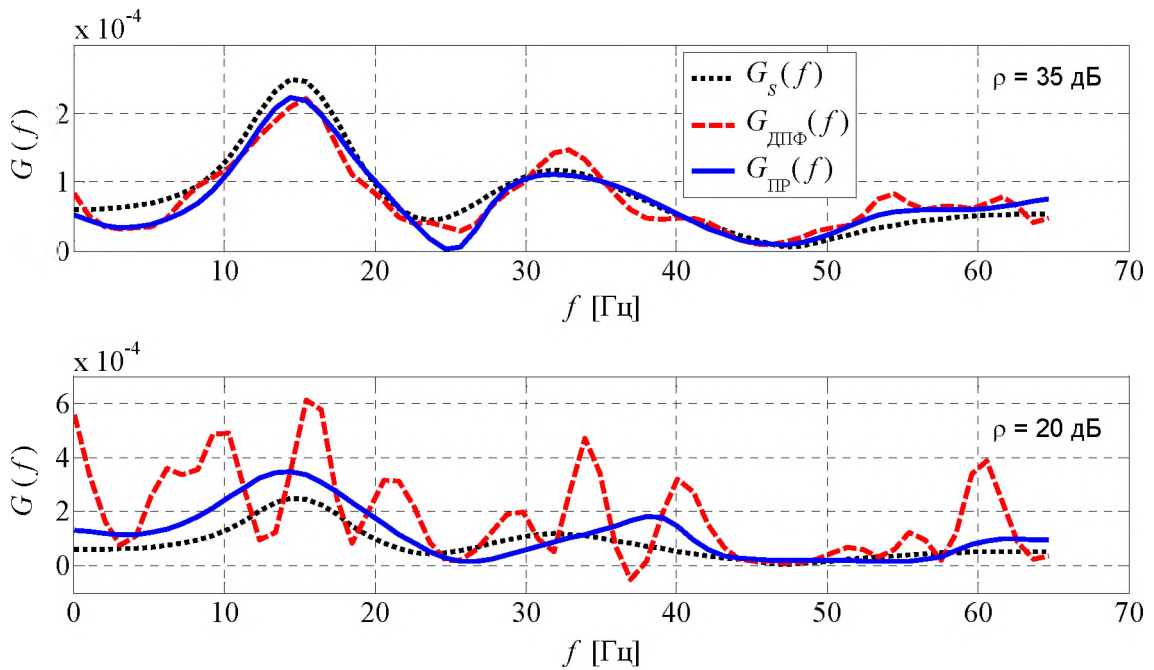


Рис. 4. Графики СПЭ процесса x_i с АБГШ при отношениях сигнал/шум $\rho = 35$ дБ и $\rho = 20$ дБ

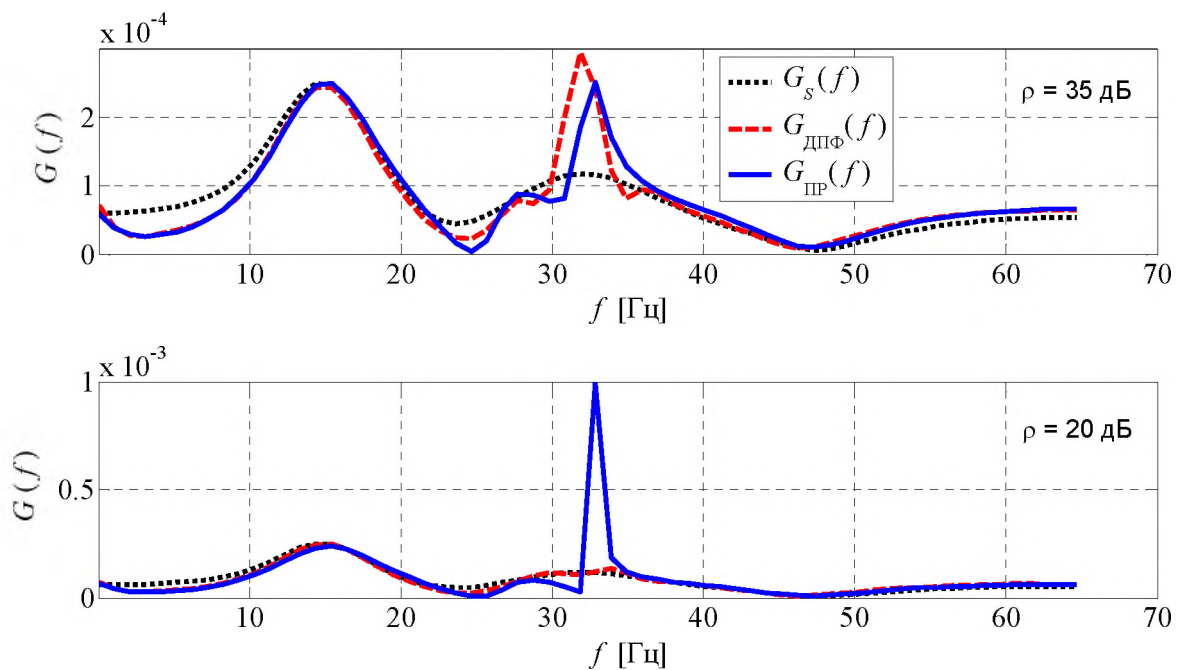


Рис. 5. Графики СПЭ процесса y_i с КШ при отношениях сигнал/шум при $\rho = 35$ дБ ($\alpha = 3.25 \cdot 10^{-4}$) и $\rho = 20$ дБ ($\alpha = 10^{-2}$)

Таблица 1. Погрешности вычисления СПЭ для модели порядка $p = 10$.

Процесс	s_i	x_i (влияние АБГШ)		y_i (влияние КШ)	
ρ , дБ	—	35	20	35	20
Погрешность ε	$1.3 \cdot 10^{-4}$	$1.5 \cdot 10^{-4}$	$4.4 \cdot 10^{-4}$	$1.8 \cdot 10^{-4}$	$2.2 \cdot 10^{-3}$

Выводы

1. Разработана методика построения адекватной дискретной модели сигнала (без шумов и с аддитивным зашумлением) по заданным корреляционным функциям сигнала и шумовых помех.

2. Разработана программа в среде Матлаб, реализующая вычислительный алгоритм МНК Прони и позволяющая провести анализ оценок СПЭ сигналов по их дискретной выборке.

3. Проведенный анализ показал:

а) Алгоритм МНК Прони позволяет получить для сигнала без шумов достаточно хорошую оценку СПЭ с погрешностью $\varepsilon = 1.3 \cdot 10^{-4}$.

б) Белый шум, действующий на сигнал, вызывает ошибки оценивания СПЭ вдоль всего диапазона частот. Если же шум коррелированный и сосредоточен по спектру в области центральной частоты сигнала, то он приводит к максимальным ошибкам оценивания СПЭ именно в этой центральной области, причем эти ошибки будут больше, чем в случае белого шума. Ближе к краям диапазона ошибки от влияния коррелированного шума уменьшаются и могут быть меньше, чем ошибки от белого шума (при одинаковых реализациях белого шума и полученных из них реализациях коррелированного шума и одинаковом отношении сигнал/шум). Если же говорить о суммарной ошибке аппроксимации вдоль всего диапазона частот, то при указанных условиях коррелированный шум оказывает большее влияние, чем белый шум.

в) Шум любого рода негативно влияет на работу метода Прони, поэтому существует необходимость модификации метода для уменьшения влияния шумов на точность аппроксимации.

Литература

1. Марпл-мл. С.Л. Цифровой спектральный анализ и его приложения / под ред. Рыжасака И.С. / пер. с англ. Хабарова О.И., Сидоровой Г.А. М.: Мир, 1990. 584 с.
2. Тихонов В.И. Статистическая радиотехника. 2-е изд., перераб. и доп. М.: Радио и связь, 1982. 624 с.
3. Сергиенко А.Б. Цифровая обработка сигналов. СПб.: Питер, 2003. 604 с.
4. Волчков В.П., Уваров С.С. Аппроксимация узкополосных случайных процессов с помощью комплексной рекуррентной m -модели скользящего окна второго порядка // Т-Comm: Телекоммуникации и транспорт. 2015. Т. 9. № 3. С. 54-61.
5. Волчков В.П., Поборчая Н.Е., Шлома А.М. Параметрический спектральный анализ случайных сигналов с использованием рекуррентных циркулянтных моделей скользящего окна // Системы синхронизации, формирования и обработки сигналов. 2014. Т. 5. № 4. С. 115-118.

DEVELOPING AN ALGORITHM, REALIZING THE PRONY'S METHOD SPECTRAL ANALYSIS IN MATLAB, AND STUDY OF THE SENSITIVITY OF THIS METHOD TO THE EXPOSURE OF NOISE INTERFERENCE

Evelin Y. Bushuev

Student of group BZS1602, MTUCI

e.bushuev71521@gmail.com

Valery P. Volchkov

Dr. Sci. Tech., professor of chair of General Communication Theory, MTUCI

volchkovvalery@mail.ru

Section «General Communication Theory»

Keywords: *spectral analysis, Proni's spectrum, least squares Proni's method (LSM) in Matlab, discrete exponential model, approximation, correlation function, white noise, correlated noise.*

The article deals with the problem of synthesizing a spectral analysis algorithm based on the least squares Proni's method (LSM) and its subsequent implementation in Matlab. The main stages of the synthesis are presented: the method of constructing adequate discrete models of signals, estimating the parameters of the discrete Prony's model of the p^{th} order using an observable sample of size $N > p$, and estimating the spectral densities of these signals by the Prony's method. The results of modeling and studies of the effect of noise interference on the Prony's spectral estimate are presented.

ПРИМЕНЕНИЕ ДИНАМИЧЕСКОГО ПРОГРАММИРОВАНИЯ В ВЫРАВНИВАНИИ ПОСЛЕДОВАТЕЛЬНОСТЕЙ ДНК

Келимов Эмиль Рустемович,
МТУСИ, Студент группы М111801(74)
semkel07@gmail.com

Гузев Алексей Валерьевич,
МТУСИ, к.т.н., доцент кафедры МСцУС
a_guzeev@mtuci.ru

Ключевые слова: биоинформатика, динамическое программирование, выравнивание последовательностей, наибольшая общая последовательность, алгоритм Нидлмана-Вунша, алгоритм Смита-Ватермана.

Подробно изложены основные алгоритмы динамического программирования, которые используются в биоинформатике, такие как алгоритм Нидлмана-Вунша и алгоритм Смита-Ватермана. Выделены основные причины возникновения использования данных алгоритмов с точки зрения биологии. Рассмотрено определение выравнивания последовательностей. Описаны различия между локальным и глобальным выравниванием. Обсуждаются результаты алгоритмов, разработанных на языке Python.

Генетические базы данных содержат чрезвычайно большое количество “сырых” данных. В одном только геноме человека содержится более трех миллиардов пар оснований ДНК[1]. Из-за этого, перспективно использование строковых алгоритмов в информатике, чтобы проводить поиск и находить значимые связи в гене.

Геном живых существ состоит из ДНК, которая состоит из двух цепочек. При этом каждая из цепочек состоит из четырех элементов [1]: аденин (А), тимин (Т), гуанин (G), цитозин (С). Сами цепи связаны по принципу комплементарности. По этому принципу аденин может образовать связь только с тиминем, то есть аденин (А) комплементарен тимину (Т), а гуанин (G), соответственно, комплементарен цитозину (С). Таким образом, зная последовательность одной цепи ДНК, состоящей из множеств А, Т, G, С, можно восстановить вторую последовательность ДНК. Следовательно, последовательность ДНК можно рассматривать как строку, состоящей из символов А, Т, G, С.

В основе биоинформатики лежит сравнение последовательностей ДНК при помощи методов выравнивания [2]. Выравнивание последовательностей (англ. *Sequence alignment*) — это процесс, с помощью которого сравниваются две последовательности путем поиска сходных участков в строках. Для решения этой задачи используется динамическое программирование. Динамическое программирование — это алгоритмический метод, который используется в анализе последовательностей. Смысл метода в том, что он позволяет решить некую большую проблему путем её декомпозиции. Если у рекурсивных задач временная сложность составляет $O(2^n)$, то в это же время у динамического программирования временная сложность составляет $O(n^2)$ или $O(n^3)$, как мы убедимся дальше.

Рассмотрим основные причины возникновения различий в сходных последовательностях [1]:

- Однонуклеотидный полиморфизм представляет собой замену одного символа в одной из последовательностей. В качестве примера на рисунке 1 (а) отображены две последовательности, которые различаются на один символ, одна последовательность имеет нуклеотид G другая С;

- Инделы (вставка/удаление) представляют собой вставленную или удаленную подпоследовательность. На рисунке 1 (б) продемонстрирована вставка подпоследовательности CGT;

- Микросателлиты, в отличие от инделов, микросателлиты являются набором коротких повторяющихся символов, также их называют тандемными повторами. Длина микросателлитов составляет не более 15 символов [1]. На рисунке 1 (в) в качестве тандемного повтора взята подпоследовательность CGT, которая повторяется 3 раза;

- Минисателлиты – это тандемные повторы, длина которых может составлять до 100 символов.



Рис. 1. Типы различий в сходных последовательностях: а) Однонуклеотидный полиморфизм; б) Индел; в) Микросателлиты и минисателлиты.

Предположим, нам необходимо найти расстояние между двумя словами ‘КОШКА’ и ‘КРОШКА’. Для этого воспользуемся расстоянием Хемминга, но с одной особенностью, сам факт различия между двумя символами между разрядами будем считать за 1, а разницу в количестве разрядов заполним пробелами.

Таким образом, поразрядно сравниваем два слова:

КРОШКА

КОШКА_, то расстояние Хэмминга будет 5, хотя визуально можно понять, что, добавив пробел во втором слове после символа ‘К’, расстояние будет 1. Пусть второе значение называется оптимальным значением выравнивания, а ‘К_ОШКА’ называется оптимальной последовательностью.

Постановка задачи

На вход даны две строки: строка $S = S[1], S[2] \dots S[n]$ и строка $T = T[1], T[2] \dots T[m]$. Задача заключается в том, что необходимо найти оптимальную последовательность символов строк S и T . Для этого, пусть $V(n, m)$ будет оптимальным значением выравнивания двух строк $S[1] \dots S[n]$ и $T[1] \dots T[m]$. В данном случае необходимо найти оптимальную последовательность путем вычисления всех значений $V(i, j)$, где $1 \leq i \leq n$ и $1 \leq j \leq m$.

Задача о наибольшей общей последовательности

Самой простой формой анализа последовательностей является решение задачи о наибольшей общей последовательности [2], где используются только операции вставки и удаления. Подпоследовательность строки S – это строка, которая может быть получена из строки S путем удаления некоторых символов без изменения порядка оставшихся символов. Соответственно, общая подпоследовательность строк S и T – это подпоследовательность, которая содержится как в строке S , так и в строке T .

Например,

$S = CGGGATCGC$

$T = CGCGTTAC$, то общими подпоследовательностями наибольшей длины будет ‘CGCGC’. Последовательности S и T будут использоваться в качестве примера и в других алгоритмах.

Если $V(i, j)$ – длина наибольшей общей последовательности между $S[1] \dots S[i]$ и $T[1] \dots T[j]$, то пусть $V(0, 0) = V(i, 0) = V(0, j) = 0$ для всех $1 \leq i \leq n$ и $1 \leq j \leq m$. В таблице 1 приведен псевдокод [2] алгоритма заполнения таблицы по следующей логике:

- Если значения $S[i]$ и $T[j]$ совпали, то при заполнении ячейки $V(i, j)$ выбирается значение ячейки $V(i - 1, j - 1) + W$, где W оценка за совпадение двух символов;
- Если символ $S[i]$ не представлен в общей последовательности для $S[1] \dots S[i]$ и $T[1] \dots T[j]$, то берется значение $V(i - 1, j)$;
- Если символ $T[j]$ не представлен в общей последовательности для $S[1] \dots S[i]$ и $T[1] \dots T[j]$, то берется значение $V(i, j - 1)$.

В таблице 2 указана реализация данного алгоритма, где $W = 1$. Как можно увидеть расстояние между двумя последовательностями составляет 5. Серым цветом выделен путь движения по таблице от (n, m) до $(0, 0)$, а желтым цветом места совпадения символов в исследуемых последовательностях, что в сумме дает общую последовательность CGCGC.

Таблица 1. Псевдокод алгоритма нахождения наибольшей общей последовательности

```

for i = 0 .. N
    V(i, 0) = 0
for i = 1 .. M
    V(0, j) = 0
for i = 1 .. M
    for j = 1 .. N
        V(i, j) = max {
            V(i - 1, j)
            V(i, j - 1)
            V(i - 1, j - 1) + W if S[i] = T[j]
        }
        if V(i, j) = V(i - 1, j)
            ' ← if V(i, j) = V(i, j - 1)
            ' ↖ if V(i, j) = V(i - 1, j - 1) + W
    
```

Таблица 2. Нахождение наибольшей общей последовательности

										0	
										0	
			1	1	1	1	1	1	1	1	
			1	2	2	2	2	2	2	2	
			1	2	2	2	2	2	3	3	
			1	2	3	3	3	3	3	4	4
			1	2	3	3	3	4	4	4	4
			1	2	3	3	3	4	4	4	4
			1	2	3	3	4	4	4	4	4
			1	2	3	3	4	4	5	5	5

Типы выравнивания

Перед тем как рассмотреть другие алгоритмы необходимо выделить типы выравниваний. Различают два вида выравнивания: глобальное и локальное выравнивание.

Глобальное выравнивание [2] (англ. *Global Alignment*) – это метод выравнивания, в котором необходимо найти оптимальное значение выравнивания между двумя последовательностями на всей их длине. При глобальном выравнивании предполагается найти оптимальный путь от (0, 0) до (n, m).

Например, возьмем две последовательности ДНК:

S = TTCGAGTCAGСТААТ
 T = TTCCGСТАТCAGTTСGAG

Тогда результатом глобального выравнивания будет:

S' = TTCGAGT - -CAGСТА- АТ
 ||| | ||| |
 T' = TTCCGСТАТCAGTTСGAG

Локальное выравнивание (англ. *Local Alignment*) – это метод выравнивания, который не предполагает, что две последовательности имеют сходство на все длине. Этот метод находит локальные регионы с самым высоким уровнем сходства между двумя последовательностями и выравнивает эти регионы без учета выравнивания остальных областей. При локальном выравнивании необходимо найти оптимальный путь между вершинами (i, j) и (i', j') .

Результатом локального выравнивания будет:

S' = ----- TTCGAGTCAGСТААТ
 ||| |||
 T' = TTCCGСТАТCAGTTСGAG -----

Алгоритм Нидлмана-Вунша

В 1970 в статье был опубликован метод [3], который демонстрирует метод наилучшего выравнивания двух строк, путем максимизации количества совпадения символов и, соответственно, минимизации количества пробелов. Алгоритм Нидлмана-Вунша это метод глобального выравнивания. В отличие от алгоритма нахождения наибольшей длины в данном алгоритме появляются штрафы за вставку пробела и несовпадения.

В функции оценки пусть W будет начислением для совпадения двух символов, μ - штраф за несовпадения, а d - штраф за вставку пробела. В таблице 3 представлен псевдокод алгоритма [2].

Таблица 3. Псевдокод алгоритма Нидлмана-Вунша

```

for i = 0 .. N
    V(i, 0) = -i * d
for i = 1 .. M
    V(0, j) = -j * d
for i = 1 .. M
    for j = 1 .. N
        V(i, j) = max {
            V(i - 1, j) - d           (1)
            V(i, j - 1) - d           (2)
            V(i - 1, j - 1) + W       (3)
            V(i - 1, j - 1) - μ       (4)
        }
        B(i, j) = {
            UP, если (1)
            LEFT, если (2)
            DIAG, если (3,4)
        }
    
```

Смысл алгоритма состоит в том, что мы пытаемся найти оптимальное значение для i -го символа S и j -го символа T, для этого необходимо рассмотреть 3 варианта:

- Если значение $V(i, j)$ получено от $V(i - 1, j - 1)$, то это будет диагональное направление со штрафом μ при несовпадении или начислением W при совпадении;
- Если значение $V(i, j)$ получено от $V(i, j - 1)$, то направление будет горизонтальным (вставка пробела в T), штраф составляет d ;

• Если значение $V(i, j)$ получено от $V(i-1, j)$, то направление будет вертикальным (вставка в S), штраф составляет d .

Пусть штраф за несовпадение $\mu = 1$, а штраф за пропуск $\alpha = 2$, а начисление за совпадение $W = 1$, то в таблице 4 продемонстрирован результат алгоритма с использованием данных коэффициентов, где серым цветом выделен оптимальный путь.

Таблица 4. Построенная таблица по алгоритму Нидлмана-Вунша

											0
			-2	-4	-6	-8	-10	-12	-14	-16	-18
		-2	1	-1	-3	-5	-7	-9	-11	-13	-15
		-4	-1	2	0	-2	-4	-6	-8	-10	-12
		-6	-3	0	1	-1	-3	-5	-5	-7	-9
		-8	-5	-2	1	2	0	-2	-4	-4	-6
		-10	-7	↑4	-1	0	1	1	-1	-3	-5
		-12	-9	-6	-3	-2	-1	2	0	-2	-4
		-14	-11	-8	-5	-4	-1	0	1	-1	-3
		-16	-13	-10	-7	-6	-3	-2	1	0	0

Используя метод время, затраченное на заполнение матрицы, составляет $O(nm)$. Чтобы убедиться в этом, рассмотрим псевдокод на таблице 3. Инициализации первой строки и первого столбца составляет $O(n + m)$. Этот алгоритм требует заполнения таблицы размером $(n + 1) \times (m + 1)$. При этом заполнение таблицы сопровождается просмотром предыдущих значений таблицы 3 раза, что является константой c , таким образом сложность заполнения таблицы составляет $O(nm)$. Итого, можно сказать, что сложность алгоритма составляет $O(nm) + O(n + m) = O(nm)$. В алгоритме используется две таблицы: одна таблица для хранения значений, вторая же для хранения направлений оптимального пути. Тогда пространственная сложность составляет $O(2nm)$.

Алгоритм Смита-Ватермана

В статье [4], которая была опубликована в 1981, был предложен новый метод выравнивания, и он очень схож с методом Нидлмана-Вунша. Если в алгоритме Нидлмана-Вунша необходимо было выравнивать последовательности от одного конца до другого, то алгоритм Смита-Ватермана находит область наибольшего сходства между двумя подпоследовательностями на промежутке от (i, j) до (i', j') . Алгоритм Смита-Ватермана — это алгоритм локального выравнивания. В алгоритме не предполагается штрафовать за вставку пробела, а добавление веса равным 0 в формуле позволяет “свободно” перемещаться по таблице. В таблице 5 продемонстрирован псевдокод алгоритма [2], разница по сравнению с предыдущим алгоритмом составляет в инициализации таблицы, а также учет максимального значения оптимального пути и его позиция. Как и предыдущие алгоритмы, алгоритм Смита-Ватермана имеет временную и пространственную сложность $O(nm)$.

Таблица 5. Псевдокод алгоритма Смита-Ватермана

```

for  $t = 0 \dots N$ 
     $V(t, 0) = 0$ 
for  $t = 1 \dots M$ 
     $V(0, j) = 0$ 
for  $t = 1 \dots M$ 
    for  $j = 1 \dots N$ 
         $V(t, j) = \max \begin{cases} V(t-1, j) - d \\ V(t-1, j) - d \\ V(t-1, j-1) + W \\ V(t-1, j-1) - \mu \end{cases}$ 
        if  $V(t, j) > \text{max\_value}$ 
             $\text{max\_value} = V(t, j)$ 
             $\text{position} = (t, j)$ 
    
```

В таблице 6 можно увидеть реализацию алгоритма, где серым цветом выделено локальное выравнивание по подпоследовательности CGC.

Таблица 6. Построенная таблица по алгоритму Смита-Ватермана

			1					1		1
				2	1	1			2	
			1		1			1		3
				2	1	2			2	1
					1		1	1		1
								2		
							1		1	
			1					1		1

На основе описанных алгоритмов были реализованы алгоритмы на языке Python. Результат алгоритмов можно увидеть на рисунке 2. На вход программы поступает две строки, а на выходе получаются таблицы и выравненные последовательности, соответствующие тому или иному алгоритму. Стоит отметить гибкость алгоритма Нидлмана-Вунша, изменяя коэффициенты можно получить абсолютно другой результат. Так, если в примерах штраф за пропуск был больше чем за несовпадение, то на рисунке 3 показан результат, когда штраф за несовпадение $\mu = 3$ больше чем за пропуск $d = 1$. При сравнении похожих последовательностей, которые обладают одной и той же функцией, выравнивание позволяет выявить биологические изменения. Такой вывод алгоритма, который отображен на рисунке 3, будет бессмысленным, так как выравнены будут только одинаковые символы, а все остальное по пробелам. Соответственно, не получится выявить функциональные различия между последовательностями, например, цвет глаз.

```

LCS table:
0 0 0 0 0 0 0 0 0 0
0 1 1 1 1 1 1 1 1 1
0 1 2 2 2 2 2 2 2 2
0 1 2 3 3 3 3 3 4 4
0 1 2 3 3 3 4 4 4 4
0 1 2 3 3 4 4 4 4 4
0 1 2 3 3 4 4 5 5 5
All LCSs: {'CGCGC', 'CGGTC', 'CGGAC'}

Needleman-Wunch table:
0 -2 -4 -6 -8 -10 -12 -14 -16
-2 1 -1 -3 -5 -7 -9 -11 -13
-4 -1 2 0 -2 -4 -6 -8 -10
-6 -3 0 1 1 -1 -3 -5 -7
-8 -5 -2 -1 2 0 -2 -4 -6
-10 -7 -4 -3 0 1 -1 -1 -3
-12 -9 -6 -5 -2 1 2 0 -2
-14 -11 -8 -5 -4 -1 0 1 1
-16 -13 -10 -7 -4 -3 -2 -1 0
-18 -15 -12 -9 -6 -5 -4 -3 0
Needleman-Wunch sequence:
CGGGATCGC
CGCG-TTAC

Smith-Waterman table:
0 0 0 0 0 0 0 0 0 0
0 1 0 0 0 0 0 1 0 1
0 0 2 1 1 0 0 0 2 0
0 1 0 1 0 0 0 1 0 3
0 0 2 1 2 0 0 0 2 1
0 0 0 1 0 1 1 0 0 1
0 0 0 0 0 0 2 0 0 0
0 0 0 0 0 1 0 1 0 0
0 1 0 0 0 0 0 1 0 1
Smith-Waterman sequence:
CGGGATCGC-----
-----CGCGTTAC

```

Рис. 2. Вывод программы обработки двух строк

```

Needleman-Wunch table:
0 -1 -2 -3 -4 -5 -6 -7 -8
-1 1 0 -1 -2 -3 -4 -5 -6
-2 0 2 1 0 -1 -2 -3 -4
-3 -1 1 0 2 1 0 -1 -2
-4 -2 0 -1 1 0 -1 -2 -3
-5 -3 -1 -2 0 -1 -2 0 -1
-6 -4 -2 -3 -1 1 0 -1 -2
-7 -5 -3 -1 -2 0 -1 -2 0
-8 -6 -4 -2 0 -1 -2 -3 -1
-9 -7 -5 -3 -1 -2 -3 -4 -2
Needleman-Wunch sequence:
CGGGATCG---C
C--G--CGTTAC

```

Рис.3. Результат алгоритма при других коэффициентах

Заключение

Все три алгоритма довольно схожи по временной и пространственной сложности $O(nm)$. Но стоит отметить проблему для алгоритма Нидлмана-Вунша, который предполагает движение от нижнего правого угла таблицы в левый верхний угол. При этом остальные углы не будут задействованы из-за больших штрафов, таким образом расходуется лишнее время и память для заполнения этих ячеек.

Выделить преимущества одного алгоритма над другим довольно проблематично, хоть они и выполняют одну и ту же функцию, все зависит от контекста исследования. Если необходимо выделить функциональные изменения сходных последовательностей, то лучше использовать алгоритм Нидлмана-Вунша. Если же необходимо найти сходные участки в негомологичных последовательностях, то лучше использовать алгоритм Смита-Ватермана. Например, последовательности ДНК, называемые гомеобоксами [1], сильно консервативны и эти последовательности одинаковы у растений, грибов, животных. И для того, чтобы найти гомеобокс, используется алгоритм Смита-Ватермана.

Литература

1. Terry Brown A. (2006). Genomes 3. -third edition. 713 с.
2. Jones N.C., Pevzner P.A. An Introduction to Bioinformatics Algorithms. 2004. С. 172-183.
3. Needleman S.B., Wunsch C.D. (1970). A general method applicable to the search for similarities in the amino acid sequence of two proteins // J. Mol. Biol. 48, pp. 443-453.
4. Smith, T. F., Waterman, M. (1981). Identification of common molecular subsequences // J. Mol. Biol. 147, pp. 195-197.

USAGE OF DYNAMIC PROGRAMMING IN THE DNA SEQUENCE ALIGNMENT

Kelimov Emil Rustemovich

*MTUCI, student of group M111801(74)
semkel07@gmail.com*

Guzeev Alexey Valerievich,

*MTUCI, PhD, Associated professor of the department of multimedia networks and
communication services
a_guzeev@mtuci.ru*

Keywords: *bioinformatics, dynamic programming, sequence alignment, Longest Common Sequence, Needleman-Wunsch algorithm, Smith-Waterman algorithm.*

This article describes details of basic algorithms in dynamic programming such as Needleman-Wunch algorithm and Smith-Waterman algorithm that are used in bioinformatics. Highlighted the main reasons for using these algorithms from the point of view of biology. Also, the definition of sequence alignment is described. The differences between local alignment and global alignment are described. The results of algorithms are discussed which was created in Python.

ПРИМЕНЕНИЕ ФРЕЙМОВ ВЕЙЛЯ-ГЕЙЗЕНБЕРГА В ЗАДАЧАХ ИНТЕРПОЛЯЦИИ И ЭКСТРАПОЛЯЦИИ СИГНАЛОВ

Мирошниченко Антон Валерьевич

студент группы М61801 МТУСИ

Mirosh.A.V@yandex.ru

Волчков Валерий Павлович

МТУСИ, д.т.н., профессор кафедры ОТС

volchkovvalery@mail.ru

Ключевые слова: фрейм Вейля-Гейзенберга, WH-фрейм, интерполяция, ряд Фурье, экстраполяция, авторегрессия.

Для решения задач цифрового спектрального анализа нестационарных сигналов предлагается использовать жесткие фреймы Вейля-Гейзенберга, которые более адекватно описывают нестационарные процессы по сравнению с известными классическими моделями. В рамках работы приводится краткое описание алгоритма разложения в жесткий WH-фрейм, а также исследование возможности применения данного алгоритма в задачах интерполяции и экстраполяции сигналов.

Введение

На современном этапе развития техники большое распространение получили различные информационные датчики, сигналы с выходов которых должны быть проанализированы для принятия тех или иных решений. Они применяются повсеместно, начиная от датчиков простых популярных систем «Умный дом» и заканчивая сложными системами радиолокации. Все эти устройства объединяет необходимость спектрально-временной обработки сигналов.

Основным алгоритмом спектрального анализа является разложение в ряд Фурье, однако данный алгоритм обладает недостатком – он позволяет эффективно исследовать стационарные случайные или периодические сигналы. Исследование же нестационарных процессов, например, сигналов с выхода различных радаров, сейсмических датчиков, систем мониторинга и др. вызывает большие трудности. Для их преодоления в работах [1- 3, 5] предлагалось использовать специальные базисы Вейля-Гейзенберга (WH-базисы), функции которого ортогональны и хорошо локализованы в частотно-временной области, чего нельзя сказать о базисных гармонических функциях Фурье. Однако и в этом случае есть ограничения, связанные со степенью желаемой локализации базисных функций, а также структурой и количеством сдвигов этих функций в частотно-временной области. Эти ограничения можно ослабить, если от WH-базисов перейти к WH-фреймам и на их основе синтезировать новые эффективные алгоритмы обработки сигналов.

Данная работа посвящена исследованию возможности применения WH-фреймов для решения важного класса практических задач, связанных с интерполяцией и экстраполяцией сигналов. Если интерполяция с использованием дискретного преобразования Фурье (ДПФ) широко используется и дает хорошие результаты, то экстраполяция имеет серьезные ограничения, поскольку продолжение ряда Фурье во времени носит периодический характер. Известные алгоритмы экстраполяции, базирующиеся на полиномах и авторегрессионных моделях, также не всегда дают удовлетворительный результат [4, 6].

Например, полиномиальная модель хороша, когда необходим анализ сигналов близких по форме к полиномам различной степени, но при применении ее к сигналам других видов она дает очень большую погрешность, поэтому ее практическое применение ограничено. Авторегрессионная модель не так чувствительна к форме исходного сигнала, но уже при незначительном удалении от крайнего известного значения функции дает существенную

погрешность. Поэтому разработка других методов экстраполяции и интерполяции, позволяющих преодолеть указанные недостатки, является актуальной.

В данной работе предлагаются и исследуются алгоритмы экстраполяции и интерполяции, базирующиеся на WH-фреймах. Сравняются ошибки интерполяции и экстраполяции с использованием разложений Фурье и WH-фреймов.

Краткие сведения о разложении в WH-фрейм

Приведем краткий алгоритм разложения дискретного сигнала $s[t]$ в жесткий WH-фрейм, более подробно данный вопрос был рассмотрен в [1-3].

Пусть на вход системы поступает дискретный сигнал $s[t]$ с частотой дискретизации f_d и конечным числом отсчетов $t \in J_N = \{0, 1, \dots, N-1\}$. Применим к сигнальному вектору $\mathbf{s} = (s[0], \dots, s[N-1])^T \in \mathbb{C}^N$ фреймовый оператор $\mathbf{U}^* \in \mathbb{C}^{P \cdot N \times N}$, в результате чего получаем вектор $\mathbf{x} = \mathbf{U}^* \mathbf{s}$, который носит название фреймового частотно-временного спектра. При этом сигнал \mathbf{s} может быть точно восстановлен из фреймового спектра \mathbf{x} по формуле $\mathbf{s} = \mathbf{U} \mathbf{x}$, если выполняются два условия: \mathbf{U} жесткий фрейм, а $\mathbf{V} = \mathbf{U} \mathbf{U}^* = \mathbf{I}_N$ - единичная матрица.

Синтез оптимального жесткого WH-фрейма [2, 3] включает следующие шаги.

1. Определим формирующую функцию:

$$g_0(t) = (2\sigma)^{1/4} \exp(-\pi\sigma t^2) \quad (1)$$

где σ - параметр, задающий ширину эталонного усеченного гауссиана.

Для удобства дальнейших вычислений модифицируем функцию (1).

$$g_c(t) = \begin{cases} g_0(t), & \text{если } t \in [0, \dots, N/2] \\ g_0(N-1-t), & \text{если } t \in [N/2+1, \dots, N-1] \end{cases} \quad (2)$$

2. Введем исходные данные для WH-фрейма:

$P \geq 2$ - коэффициент перенасыщения;

$M = kP, k \in \mathbb{N}$ - количество частотных сдвигов, кратное P ;

$K = NP/M$ - количество временных сдвигов;

$L_0 = M/P$ - временной шаг, на который сдвигаются функции фрейма.

Параметры P , M и L_0 должны быть целыми положительными числами.

3. Сформируем эталонную матрицу Габора размером $N \times NP$, согласно выражению:

$$\mathbf{G} = \{G_{k,l}[n]\} = \{g_c(n - lL_0) \exp(j2\pi kn / M)\} \in \mathbb{C}^{N \times NP} \quad (3)$$

где $n \in J_N = \{0, \dots, N-1\}$, $k \in J_M$, $l \in J_K$.

4. Найдем три вспомогательные матрицы.

Матрица перестановки \mathbf{P} .

$$\mathbf{P}_{i,j} = \begin{cases} 1, & j = L_0(i \bmod(K)) + \lfloor i / K \rfloor \\ 0, & \text{иначе} \end{cases} \quad (4)$$

где $\lfloor \cdot \rfloor$ - оператор округления вниз.

Диагональная матрица $\mathbf{\Lambda}$

$$\mathbf{\Lambda}_{i,j} = \begin{cases} 1/\sqrt{\sigma_i}, & i = j \\ 0, & \text{иначе} \end{cases} \quad (5)$$

где σ_i i -ое собственное число матрицы $\mathbf{V} = \mathbf{G} \mathbf{G}^*$.

Матрица \mathbf{F} , состоящая из расположенных на диагонали матриц Фурье \mathbf{F}_0 , размером $K \times K$, которые задаются как

$$\mathbf{F}_{0,p,q} = \frac{1}{\sqrt{K}} (\exp(j2\pi pq / K)). \quad (6)$$

Размерность матриц \mathbf{P} , $\mathbf{\Lambda}$ и \mathbf{F} равна $N \times N$.

5. Приводим эталонную матрицу Габора к оптимальному жесткому WH-фрейму с помощью матричного оператора

$$\mathbf{U}_{opt} = \mathbf{H}\mathbf{G} = \mathbf{P}^T \mathbf{F} \mathbf{\Lambda}^{-1} \mathbf{F}^* \mathbf{P} \mathbf{G} \in \mathbb{C}^{N \times NP} \quad (7)$$

Дискретные функции, стоящие по столбцам матрицы (7), позволяют вычислять фреймовый частотно-временной спектр любого сигнала $\mathbf{s} \in \mathbb{C}^N$ и по нему однозначно восстанавливать этот сигнал.

Разложение нестационарного сигнала в WH-фрейм

Рассмотрим пример разложения сложного нестационарного сигнала в WH-фрейм и восстановления исходного сигнала из частотно-временного спектра.

Пусть, с помощью конечного числа N отсчетов, задан некий сигнал $s[t]$, $t \in J_N = \{0, 1, \dots, N-1\}$, например, приведенный на рис. 1. В нашем случае $N = 64$.

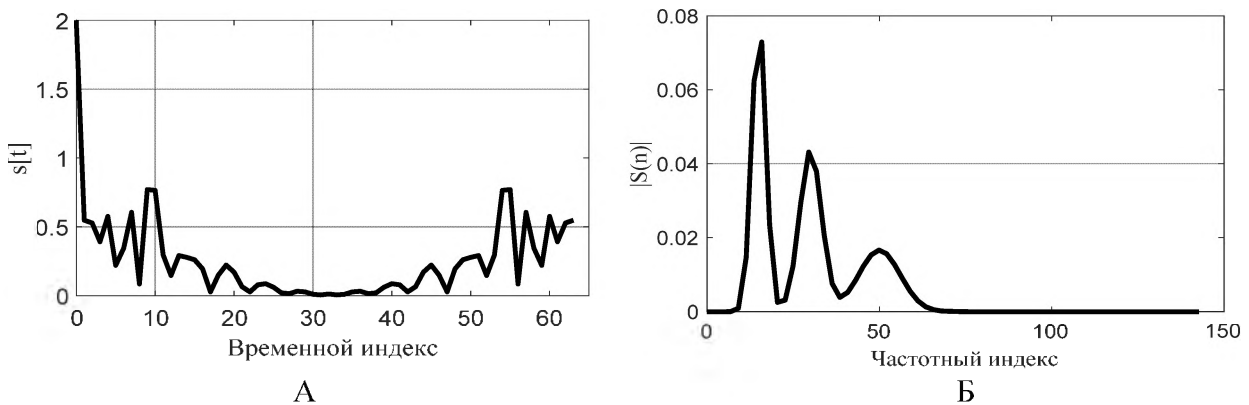


Рис. 1. Исходный сигнал $s[t]$: А – временная осциллограмма, Б – амплитудный спектр

Введем исходные данные для расчета WH-фрейма:

$\sigma = 2.5/M^2 = 0.0098$ - параметр гауссиана;

$P = 2$ - коэффициент перенасыщения;

$M = 16$ - количество частотных сдвигов;

$K = 8$ - количество временных сдвигов;

$L_0 = N/K = 8$ - временной шаг, на который сдвигаются функции фрейма.

С помощью алгоритма (1-7) синтезируем матрицу в WH-фрейма \mathbf{U}_{opt} .

Проведем разложение сигнала $\mathbf{s} = (s[0], \dots, s[N-1])^T$ в фреймовый частотно-временной спектр с помощью преобразования

$$\mathbf{x} = \mathbf{U}_{opt}^* \mathbf{s}.$$

Для восстановления исходного сигнала из его фреймового спектра воспользуемся выражением

$$\mathbf{s} = \mathbf{U}_{opt} \mathbf{x}.$$

На рис. 2 приведены графики фреймового частотно-временного спектра сигнала и восстановленного по нему исходного сигнала для $s[t]$, приведенного на рис.1, при разложении в WH-фрейм с приведенными выше параметрами.

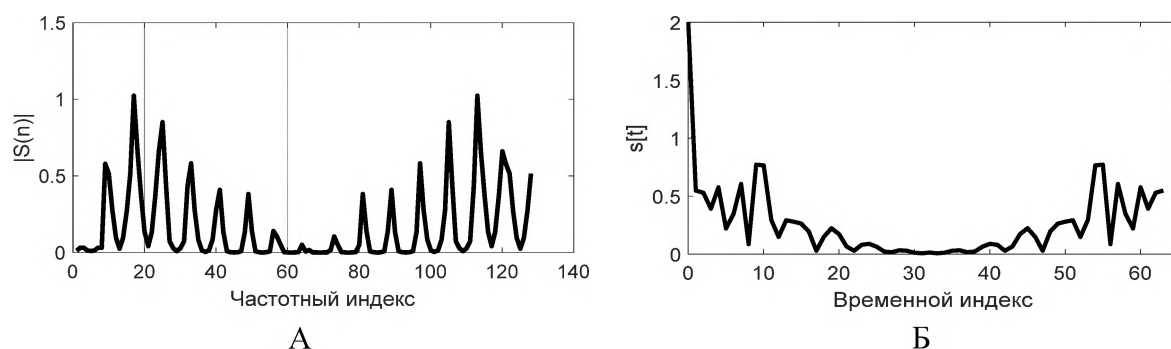


Рис. 2. Результаты разложения сигнала $s[t]$ в WH-фрейм: А – фреймовый частотно-временной спектр, Б – восстановленный сигнал

Решение задачи интерполяции с помощью WH-фрейма

Рассмотрим решение задачи интерполяции сигнала с помощью WH-фрейма.

Пусть задан дискретный вещественный сигнал $\mathbf{s} = (s[0], \dots, s[N-1])^T$, состоящий из конечного числа отсчетов, взятых через интервал времени Δt . Требуется найти промежуточные отсчеты сигнала, иными словами, увеличить частоту дискретизации в a раз.

Решим задачу интерполяции сигнала приведенного на рис. 1, при $a = 2$, с помощью WH-фрейма и ряда Фурье.

Разложив сигнал \mathbf{s} в ряд Фурье [4] и WH-фрейм, мы получим два спектра размерностью N и NP соответственно. Для решения задачи интерполяции необходимо добавить нулевые отсчеты в середину частотно-временного спектра, согласно выражению

$$x'[n] = \begin{cases} x(n), & \text{если } n = 1, \dots, 0.5N' \\ 0, & \text{если } n = 0.5N'+1, \dots, aN - 0.5N', \\ x(n + 0.5aN), & \text{если } n = aN' - 0.5N'+1, \dots, aN' \end{cases}$$

где N' - количество отсчетов в частотно-временном спектре, которое равно N - при разложении в ряд Фурье, NP - при разложении в WH-фрейм.

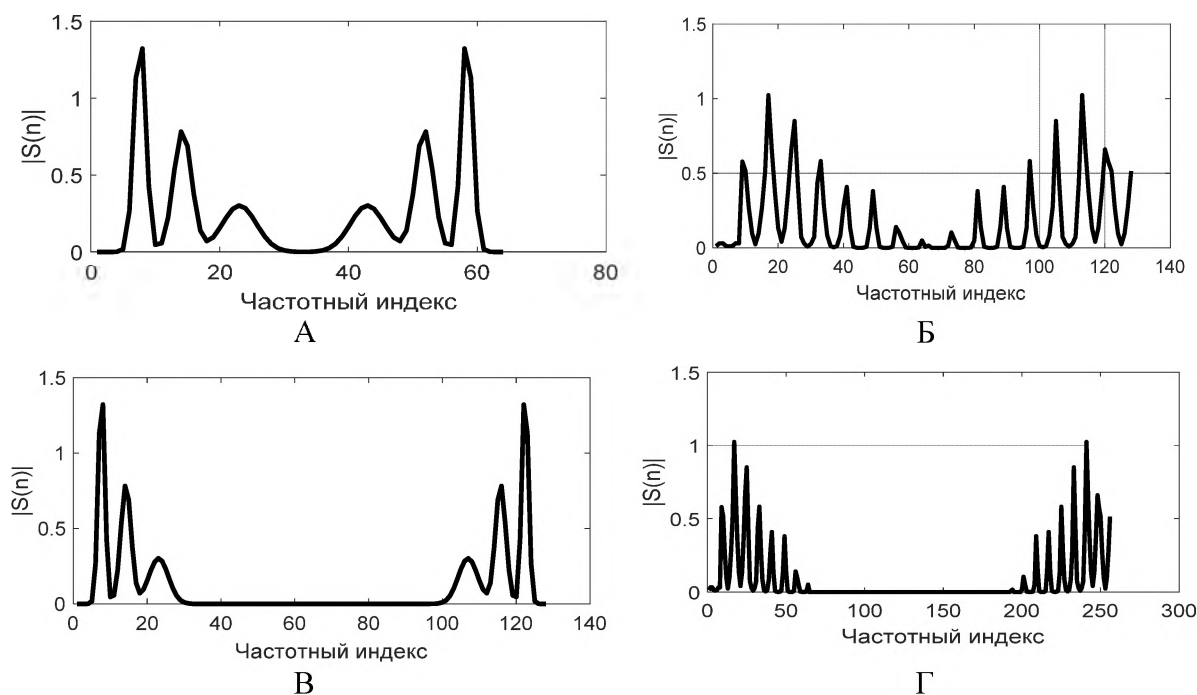


Рис. 3. Частотно-временные спектры сигнала \mathbf{s} : А – при разложении в ряд Фурье до добавления нулевых отсчетов, Б – при разложении в WH-фрейм до добавления нулевых отсчетов, В – при разложении в ряд Фурье после добавления нулевых отсчетов, Г – при разложении в WH-фрейм после добавления нулевых отсчетов

Восстановим сигнал по измененному частотно-временному спектру x' и найдем среднеквадратичное отклонение интерполированного сигнала от исходного при интерполяции рядом Фурье и WH-фреймом.

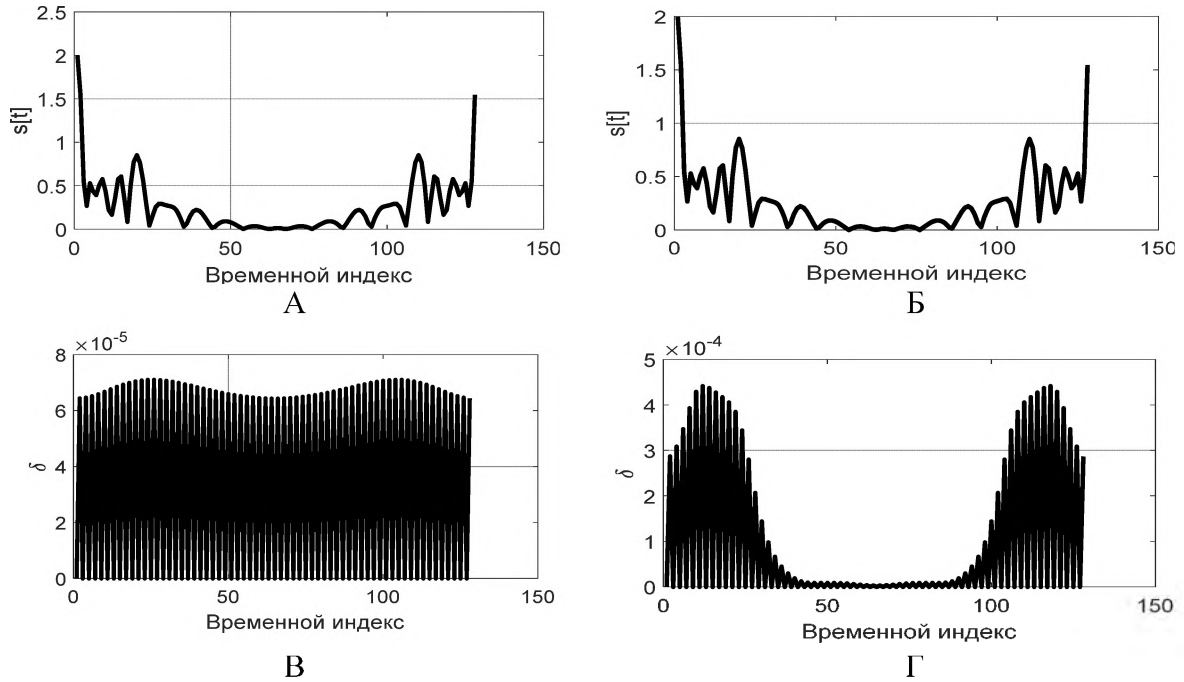


Рис. 4. Графики интерполированного сигнала s' и ошибки интерполяции $|s'[t] - s[t]|$: А – с помощью ряда Фурье, Б – с помощью разложения в WH-фрейм, В – ошибка интерполяции при интерполяции рядом Фурье, Г – ошибка при интерполяции WH-фреймом

Рассчитаем среднеквадратичную погрешность интерполяции как [5]

$$\delta = \sqrt{\frac{1}{N} \sum_{i=1}^N (x'[i] - x[i])^2}.$$

Для приведенного примера среднеквадратичная погрешность интерполяции составила: $4.74 \cdot 10^{-5}$ – для интерполяции с помощью ряда Фурье и $1.73 \cdot 10^{-4}$ – для интерполяции с помощью WH-фрейма. Отметим, что ошибка интерполяции с использованием WH-фрейма оказалась незначительной, хотя и больше, чем в случае использования ДПФ. Она зависит формы сигнала и выбранных параметров фрейма. С ростом значения $M=16$ при фиксированном N эта ошибка уменьшается, приближаясь к ошибке интерполяции метода Фурье.

Решение задачи экстраполяции с помощью WH-фрейма

Рассмотрим решение задачи экстраполяции сигнала с помощью WH-фрейма.

Сформируем эталонный сигнал $s_0[t]$, состоящий из 80 отсчетов. Преобразуем его в исходный сигнал $s[t]$, уменьшив число отсчетов до 64, после чего, зная сигнал $s[t]$, экстраполируем 16 неизвестных отсчетов и сравним с эталонным сигналом $s_0[t]$.

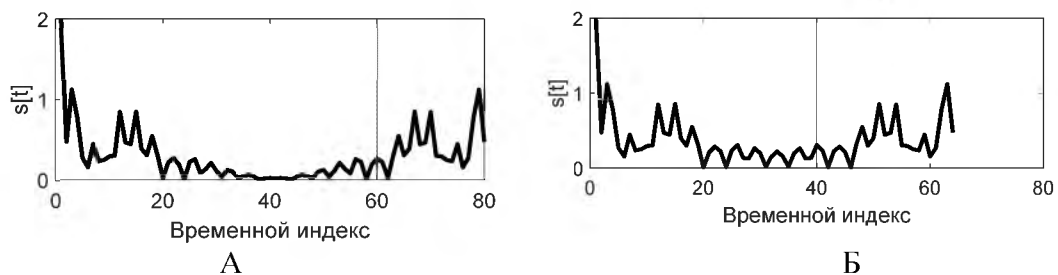


Рис. 5. Графики: А – эталонного $s_0[t]$, Б - исходного сигналов $s[t]$

Разложим сигнал $s[t]$ в WH-фрейм с параметрами, описанными в предыдущих разделах, и получим фреймовый частотно-временной спектр $x[n]$. Дополним нулевыми отсчетами $x[n]$ согласно алгоритму:

1. Определим коэффициент экстраполяции

$$q = \lceil L_e / L_0 \rceil,$$

где L_e - количество отсчетов t , которые должны быть экстраполированы, L_0 - временной шаг, на который сдвигаются функции фрейма.

2. Добавим нулевые отсчеты в частотно-временной спектр согласно выражению:

$$x'[n] = \begin{cases} 0, & \text{если } n \in [a(L_0 + 1), \dots, a(L_0 + q)] \\ x[aL_0 + b], & \text{иначе} \end{cases},$$

где $a \in [0, \dots, K - 1]$, $b = \text{mod}(n, (L_0 + q)) = n - (L_0 + q) \lfloor n / (L_0 + q) \rfloor$.

На рисунке 6 приведены графики исходного частотно-временного спектра $x[n]$ и спектра с добавлением нулевых отсчетов $x'[n]$.

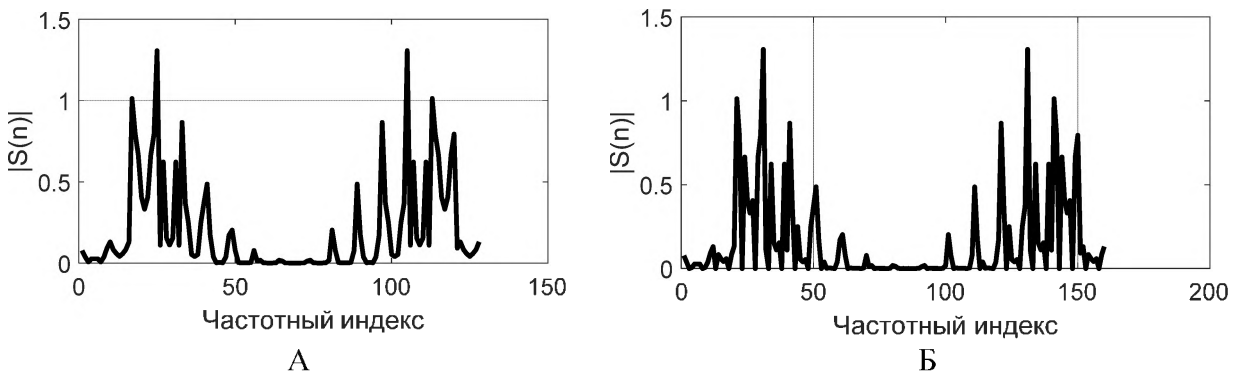


Рис. 6. Графики: А – исходного частотно-временного спектра $x[n]$, Б - спектра с добавлением нулевых отсчетов $x'[n]$

Восстановим сигнал $s[t]$ по его частотно-временному спектру $x'[n]$. Полученный сигнал $s'[t]$ будет иметь на L_e отсчетов больше, чем исходный.

Для уменьшения погрешности сформируем экстраполированный сигнал добавив к известным N отсчетам новые, согласно выражению:

$$s_{\text{экс}}[i] = \begin{cases} s[i], & \text{если } i \in [1, N] \\ s'[i], & \text{если } i \in [N + 1, N + qL_0] \end{cases}.$$

Этим действием мы провели операцию экстраполяции сигнала с помощью WH-фрейма. Сравним его с экстраполяцией с помощью ряда Фурье. Воспользовавшись свойством периодичности сигнала, представленного ДПФ спектром на интервале $J_N = \{0, \dots, N - 1\}$, операция его экстраполяции сводится к периодическому продолжению:

$$s_{\text{экс}}[i] = s[k], \quad \text{где } k = \text{mod}(i, N), \quad i \in [1, N + L_e].$$

На рис 7. приведены графики эталонного и экстраполированных сигналов, а также графики погрешностей экстраполяции.

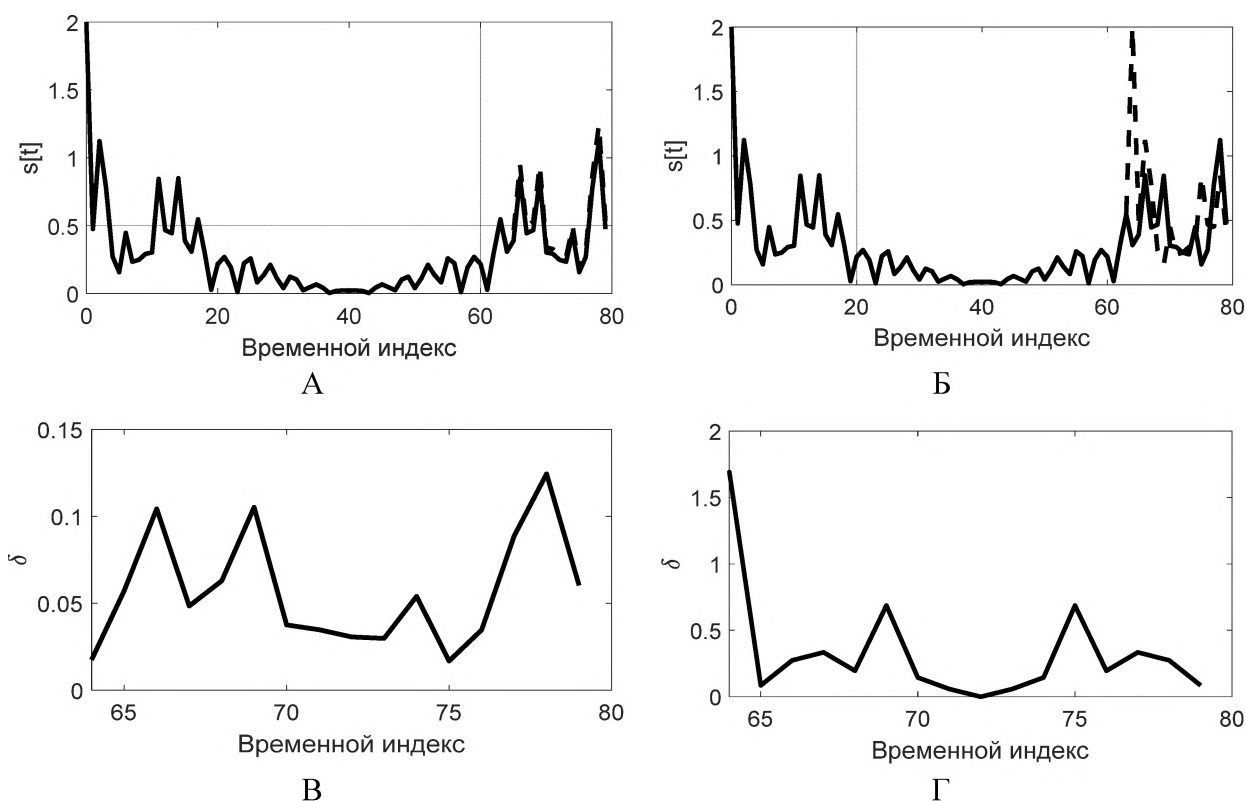


Рис. 7. Графики: А, сплошная линия – эталонный сигнал $s_j[t]$, А, пунктирная линия – сигнал экстраполированный WN-фреймом, Б, сплошная линия – эталонный сигнал $s_j[t]$, Б, пунктирная линия – сигнал экстраполированный рядом Фурье, В – график погрешности экстраполяции $|s_j[t] - s_{экс}[t]|$, при экстраполяции WN-фреймом, Г – график погрешности экстраполяции $|s_j[t] - s_{экс}[t]|$, при экстраполяции рядом Фурье

Среднеквадратичная погрешность экстраполяции на экстраполируемом участке (от $N+1$ до $N+qL_0$), составила: 0.0650 - при экстраполяцией WN-фреймом, 0.5196 - при экстраполяцией рядом Фурье. Таким образом, ошибка экстраполяции с использованием WN-фрейма оказалась в среднем почти на порядок меньше, чем в случае использования ДПФ, что говорит о хороших экстраполирующих свойствах фреймов.

Выводы

1. WN-фреймы позволяют довольно хорошо интерполировать сигналы, однако, в среднем выигрыша по сравнению с интерполяцией методом Фурье не наблюдается. Причем погрешность интерполяции с помощью ДПФ является постоянной во времени и не зависит от мгновенной амплитуды интерполированного сигнала, в то время как погрешность интерполяции WN-фреймом зависит от мгновенной амплитуды сигнала и выбранных параметров WN-фрейма. Следовательно, WN-фреймы позволяет производить более качественную интерполяцию сигналов с большим мгновенным перепадом мощности, а интерполяция методом Фурье показывает лучшие результаты при постоянной средней мощности сигнала.

2. WN-фрейм позволяет хорошо экстраполировать нестационарные сигналы, в то время как ряд Фурье очень плохо приспособлен для этой цели и дает значительные ошибки. Это особенно проявляется, когда сигнал не стационарен и имеет большие перепады мгновенных значений.

3. Хорошие экстраполяционные свойства комплексных WH-фреймов могут быть использованы для решения важных практических задач обработки сигналов, где применение традиционных методов экстраполяции не дает желаемого результата

Литература

1. Волчков В.П. Алгебраический подход к синтезу хорошо локализованных сигнальных фреймов на конечном дискретном интервале // В сборнике: «28-я Международная Крымская конференция "СВЧ-техника и телекоммуникационные технологии" (КрыМиКо'2018) Материалы конференции» 9. 2018, Том 3. С. 433-437.

2. Волчков В.П., Мирошниченко А.В. Синтез алгоритмов спектрального анализа сигналов на основе фреймов Вейля-Гейзенберга // Фундаментальные проблемы радиоэлектронного приборостроения. 2018. Т. 18. № 4. С. 1001-1006.

3. Волчков В.П. Синтез комплексных сигнальных фреймов Вейля-Гейзенберга с хорошей частотно-временной локализацией. Инфокоммуникационные и радиоэлектронные технологии. Севастополь: 2019. Т. 2. № 1. С. 86-99.

4. Марпл-мл. С.Л. Цифровой спектральный анализ и его приложения: Пер. с англ. М.: Мир, 1990. 584 с.

5. Волчков В.П., Санников В.Г. Алгебраический подход к оптимальному синтезу вещественных сигнальных базисов Вейля-Гейзенберга // Системы синхронизации, формирования и обработки сигналов. 2018. Т. 9. № 4. С. 19-25.

6. Волчков В.П., Уваров С.С. Аппроксимация узкополосных случайных процессов с помощью комплексной рекуррентной m -модели скользящего окна второго порядка // Т-Comm: Телекоммуникации и транспорт. 2015. Т. 9. № 3. С. 54-61.

APPLICATION OF WEYL-HEISENBERG FRAME IN THE PROBLEMS OF INTERPOLATION AND EXTRAPOLATION OF SIGNALS

Anton V. Miroshnichenko

MTUCI, Student

Mirosh.A.V@yandex.ru

Valery P. Volchkov

MTUCI, Doctor of Science

volchkovvalery@mail.ru

Keywords: *Weyl-Heisenberg frame, WH-frame, interpolation, Fourier series, extrapolation, autoregression.*

To solve the problems of digital spectral analysis of non-stationary signals, it is proposed to use Weil-Heisenberg rigid frames, which more adequately describe non-stationary processes in comparison with the known classical models. In the framework of the work, a brief description of the decomposition algorithm in a hard WH-frame is given, as well as a study of the possibility of using this algorithm in problems of interpolation and extrapolation of signals.

ПРИМЕНЕНИЕ OLAP-ТЕХНОЛОГИЙ ДЛЯ АНАЛИЗА МНОГОМЕРНЫХ ДАННЫХ В КОНТАКТ-ЦЕНТРЕ

Усачева Д.И., Шишкин М.О.

МТУСИ, студенты группы М091801(72)

Гадасин Д.В.

МТУСИ, к.т.н., доцент кафедры МСцУС

dengadiplom@mail.ru

Гузеев А.В.

МТУСИ, к.т.н., доцент кафедры МСцУС

Ключевые слова: системы поддержки принятия решений, OLAP-технологии, многомерный анализ данных, анализ данных контакт-центра, интеллектуальный анализ данных.

Сформулирована цель работы. Приведены примерные выходные данные. Рассмотрено понятие систем поддержки принятия решений, их архитектура. Произведен обзор общих задач OLAP-систем. Приведены конкретизированные задачи для разрабатываемого продукта. Рассмотрены задачи и понятие интеллектуального анализа данных. Дано описание и проведен сравнительный анализ некоторых методов многомерного анализа данных. Приведено описание некоторых алгоритмов. Сформулированы дальнейшие цели и требования к разрабатываемому продукту.

Введение

Развитие компьютерных технологий и информатизация общества привели к значительному увеличению объемов данных. В настоящее время у большинства людей на планете имеется по меньшей мере 1 компьютер или ноутбук в семье. Смартфоны – самые популярные мобильные устройства в мире, и это ни у кого не вызывает сомнений, достаточно лишь взглянуть на данные, которые можно найти в отчетах We Are Social о глобальном состоянии цифровых технологий на 2019 год [1]:

- Сегодня в мире 5,11 миллиарда уникальных мобильных пользователей, что на 100 миллионов (2%) больше, чем в прошлом году;
- В 2019 году аудитория интернета насчитывает 4,39 миллиарда человек, что на 366 миллионов (9%) больше, чем в январе 2018 года;
- В социальных сетях зарегистрировано 3,48 миллиарда пользователей. По сравнению с данными на начало прошлого года этот показатель вырос на 288 миллионов (9%);
- Сегодня 3,26 миллиарда человек заходят в социальные сети с мобильных устройств. Это на 10% больше, чем в прошлом году, когда с мобильных в соцсетях «сидело» на 297 миллионов человек меньше.

А мобильный трафик, по данным исследований и прогнозов компании Cisco, стремится к трём эксабайтам в месяц.

Большой объем информации позволяет выполнить более точные расчеты и делать подробный анализ, однако поиск решений в таком изобилии данных превращается в крайне сложную задачу.

Обзор предметной области

С задачами анализа огромного количества данных каждый день сталкивается множество компаний, особенно в сфере услуг – чем больше знаний о желаниях потребителя, тем выше продажи. И операторы связи не являются исключением.

В работе рассматривается компания по предоставлению услуг связи, а именно – служба поддержки клиентов, которая подразумевает собой call-центр. Звонки попадают на автоответчик, распределяются и направляются к специалистам в конкретной области. Таким же образом осуществляется поддержка в онлайн-чате.

Этап развития центров поддержки, на котором было создано универсальное ПО для автоматизации работы, контроля деятельности call-центра и накопления разнообразной статистической информации о работе call-центра, позади. Сейчас ведется разработка аналитических инструментов, позволяющих оперативно анализировать информацию и принимать решения о ходе дальнейшей работы call-центра.

Целью работы является разработка системы поддержки принятия решений для анализа трафика контакт-центра оператора связи.

Анализ данных в рассматриваемом случае может проводиться:

- для выявления кратковременных или долгосрочных тенденций с целью улучшения качества предоставляемых услуг;
- для поиска наиболее популярных запросов с целью ускорить обработку наиболее востребованных обращений;
- для выявления скрытых потребностей клиентов;
- для индивидуализации предлагаемых услуг.

Статистическая информация является одним из основных источников для стратегического планирования бизнеса. Применение средств, агрегирующих информацию за годы и позволяющих быстро просматривать данные в любом желаемом разрезе, значительно упрощает решение стратегических задач, в том числе выполняемых менеджерами и аналитиками call-центра. Например, изучая статистическую информацию по распределению звонков, работники call-центров принимают решения об изменении штата сотрудников в соответствии с географическими, сезонными колебаниями и тенденциями изменения нагрузки, распределения графика работы сотрудников по времени суток и т. д.

В результате необходимости упростить задачу поиска решений появился класс программных систем, призванных облегчить работу по анализу данных. Такие системы принято называть системами поддержки принятия решений (СППР). На данный момент существует огромное количество СППР, разработанных и внедренных в различных областях человеческой деятельности. Темпы разработок таких систем постоянно возрастают с ростом спроса на них [1].

СППР – компьютерные автоматизированные системы, целью которых является помощь людям, принимающим решение в сложных условиях для полного и объективного анализа предметной деятельности.

СППР представляют собой комплекс методов и средств, таких как:

- СУБД, реализующие концепцию хранилищ данных. Они являются источником информации для средств обработки данных. Вид базы данных зависит от вида всей системы и алгоритмов анализа данных;
- Средства оперативной аналитической обработки (OLAP - On-Line Analytical Processing);
- Средств интеллектуального анализа данных (ИАД), которые выполняют такие задачи, как классификация и кластеризация данных, построение обобщенных показателей эффективности и качества, а также прогнозирование.

OLAP-системы и анализ данных

Особое место занимают средства OLAP. Системы поддержки принятия решений обычно обладают средствами предоставления пользователю агрегатных данных для различных выборок из исходного набора в удобном для восприятия и анализа виде. Как правило, такие

агрегатные функции образуют многомерный (нереляционный) набор данных, нередко называемый гиперкубом или метакубом. Оси гиперкуба содержат параметры, а ячейки — зависящие от них данные. Вдоль каждой оси данные могут быть организованы в виде иерархии, представляющей различные уровни их детализации. Благодаря такой модели данных пользователи могут формулировать сложные запросы, генерировать отчеты, получать подмножества данных [2].

Типовыми задачами средств OLAP являются:

- обеспечение полноты и достоверности хранимых данных - «очистка» введенных данных, включая проверку на непротиворечивость;
- обеспечения доступа к сложным многомерным данным в любом заданном разрезе (формирование сложных запросов);
- обеспечение отображения сложных многомерных данных в удобном для восприятия виде.

OLAP представляет данные в виде многомерных кубов, называемых OLAP-кубами. Осями там служат основные параметры, а на пересечении осей — данные, количественно характеризующие какой-либо процесс (рис. 1).

	Интернет	СМС	Звонки
Москва	10000	5000	7000
Самара	5000	3000	6000
Тамбов	4000	3000	5500

Рис. 1. Пример трехмерного OLAP-куба

Основная таблица хранилища данных — таблица фактов. Она содержит сведения о событиях и объектах, совокупность которых будет потом анализироваться. Для многомерного анализа данных важны таблицы фактов с как можно более подробными данными [2].

Создание хранилища данных — одна из поставленных подзадач, т.к. в рассматриваемом контакт-центре нет аналогичной технологии накопления данных. В процессе создания хранилища данных используются службы DTS MS SQL Server, с помощью которых проектируется архитектура хранилища данных, организуется импорт, экспорт и преобразование данных.

Разрабатываемая система будет содержать всю информацию о трафике контакт-центра: информация о клиентах, географические данные, количественные показатели спроса, популярности услуг, отзывы и пожелания клиентов, интенсивность трафика и др.

Следует отметить, что OLAP-функциональность может быть реализована различными способами, начиная с простейших средств анализа данных в офисных приложениях и заканчивая распределенными аналитическими системами, основанными на серверных продуктах.

Сложность реализации конечного продукта также зависит от выбранных методов многомерного анализа данных. В настоящее время существует множество различных алгоритмов и их модификаций, однако при поиске решений каждой конкретной задачи нужно учитывать особенности данных и желаемого результата. В зависимости от конечной цели, будут меняться и средства ее достижения. Чаще всего используются гибриды нескольких методов, это позволяет достичь наибольшей точности и минимизировать потери.

К основным задачам интеллектуального анализа данных относятся:

- классификация (распознавание) данных и ситуаций;
- кластеризация данных - разбиение массивов данных на группы (кластеры) по близости структур значений признаков;
- построение обобщенных показателей эффективности и качества;
- прогнозирование.

Основными методами анализа многомерных данных, используемыми для ИАД, являются:

- факторный анализ;
- кластерный анализ;
- дискриминантный анализ и классификация.

Далее проводится обзор различных методов, сформулированы их достоинства и недостатки.

Факторный анализ

Факторный анализ — многомерные методы, применяемые для изучения взаимосвязей между значениями переменных. В его основе лежит предположение, что известные переменные зависят от меньшего количества неизвестных переменных и случайной ошибки.

Одним из самых популярных методов факторного анализа является метод главных компонент, который позволяет с минимальными потерями информации эффективно уменьшить размерность данных.

Один из способов вычисления главных компонент – вычисление сингулярного разложения матрицы данных или собственных векторов и собственных значений ковариационной матрицы исходных данных, формула (1).

$$M = UV^T \Sigma \quad (1)$$

$$\Sigma = \begin{pmatrix} \sigma_1^2 & \dots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \dots & \sigma_m^2 \\ 0 & 0 & 0 \end{pmatrix} \quad (2)$$

Матрицы U порядка m и V порядка n — две унитарные матрицы, состоящие из левых и правых сингулярных векторов соответственно (а V^* — это сопряженно-транспонированная матрица к V).

Σ — неотрицательная матрица размера $m \times m$, формула (2), у которой элементы, лежащие на главной диагонали, являются сингулярными числами.

Причем сингулярные числа будут располагаться по убыванию, т.е. в порядке уменьшения значимости факторов. Чаще всего последние числа практически равны нулю, что нам позволяет отбросить все несущественные зависимости и уменьшить размерность данных.

Кластерный анализ

Кластеризация – задача в мире статистической обработки данных, основная цель которой разделить множество объектов на группы таким образом, чтобы внутри каждой группы оказались наиболее похожие друг на друга объекты, а члены разных групп были бы максимально разными. Получившиеся группы в кластерном анализе принято называть кластерами.

Задач и методов в кластерном анализе достаточно много, перечислять их все не имеет смысла, т.к. множество из них применяются в других областях науки.

Перед кластерным анализом к имеющимся выборкам данных чаще всего применяют факторный анализ, это позволяет улучшить точность деления по группам, исключив некоторые незначимые переменные в векторе признаков объектов.

Формальная постановка задачи выглядит следующим образом:

Пусть I – множество объектов, каждый из которых представлен неким конечным вектором признаков $i = \{p_1, p_2, \dots, p_m\}$, G^k – множество, состоящее из k групп. Зададим функцию расстояния между двумя элементами $\rho(i_1, i_2)$, зависящую от признаков объектов. Требуется разделить элементы конечной выборки $I^n = \{i_1, i_2, \dots, i_n\} \subset I$ на подмножества. Каждый кластер должен содержать элементы, близкие по метрике ρ , и отличные от элементов другого кластера. Каждому элементу $i_j \in I^n$ будет присвоен номер соответствующего кластера $g_j \in G^k$.

Функция для нахождения расстояния должна удовлетворять условиям:

1. $\rho(i_1, i_2) \geq 0$, $\rho(i_1, i_2) = 0$ тогда и только тогда, когда $i_1 = i_2$;
2. $\rho(i_1, i_2) = \rho(i_2, i_1)$;
3. $\rho(i_1, i_3) < \rho(i_1, i_2) + \rho(i_2, i_3)$, если $i_1, i_2 \neq i_3$ не лежат на одной прямой.

Алгоритм кластерного анализа – это функция $f: I \rightarrow G$, которая любому объекту $i \in I$ ставит в соответствие номер кластера $g \in G$.

k -средних – самый распространенный и простой метод кластерного анализа [3].

Относится к классу стохастических алгоритмов. В зависимости от различных типов данных, он может немного видоизменяться, но в общем случае происходит следующее:

1. Все точки множества, разделяются на k кластеров;
2. На каждом шаге вычисляется центр μ_j каждого кластера и объекты

перераспределяются по группам, расстояние до середины которых наименьшее, формула (3);

$$g_j = \arg \min_{g \in G} \rho_g(i, \mu_j), \quad j = 1, \dots, n; \quad (3)$$

3. Условие завершения – на очередной итерации ни один элемент не меняет свой кластер.

В целом алгоритм стремится минимизировать суммарное квадратичное отклонение объектов в каждом кластере от его центра, формула (4).

$$V = \sum_{j=1}^k \sum_{i \in G_j} (i - \mu_j)^2. \quad (4)$$

Где k – количество кластеров, G_j – полученные кластеры, причем $j = 1, 2, \dots, k$, μ_j – центры масс векторов признаков элементов i в кластере G_j .

При применении этого алгоритма нужно учитывать факторы, влияющие на его корректную работу. Один из них: число кластеров необходимо определить заранее, а неправильный выбор может очень плохо повлиять на качество кластеризации. Конечно, существует не один способ улучшения данного алгоритма, но они обычно отвечают на несколько отличающиеся от традиционной кластеризации вопросы.

Алгоритм КНП – расширяется как алгоритм Кратчайшего Незамкнутого Пути и относится к графовым алгоритмам кластеризации [3]. Главное отличие: в процессе работы строится граф, вершины – объекты, а ребра – расстояния между ними $R_{xy} = \rho(i_x, i_y)$. Выделить из достоинств таких методов можно наглядность и простоту усовершенствования.

В процессе обработки входных данных, алгоритм строит граф, состоящий из $n - 1$ ребер. Они соединяют все n точек выборки и обладают наименьшей суммарной длиной из всех возможных. Отсюда и название метода. Алгоритм выглядит следующим образом:

1. Находятся пара точек с наименьшим расстоянием, формула (5), между ними и соединяются ребром $e_{x,y}$.

$$\rho_{xy} = \min_{i \in I} \rho(i_x, i_y). \quad (5)$$

2. Ищется новая изолированная точка, ближайшая к некой неизолированной точке, и снова они соединяются ребром.

3. Повторяется до тех пор, пока не останется ни одной изолированной точки.

4. Удаляются $k-1$ самых длинных ребер граф «распадается» на k маленьких графов, которые олицетворяют кластеры.

Выбранное k может существенно повлиять на качество кластеризации, но есть возможность заранее определить оптимальный вариант. Для этого, после составления дерева, нужно упорядочить все ребра в порядке убывания и отложить их на графике. В результате в начале части графика будет заметен резкий скачек вниз, показывающий количество наиболее точно выделяемых групп.

Алгоритм Кратчайшего Незамкнутого Пути достаточно прост как для понимания, так и для реализации, но так же, как и k -средних имеет ряд особенностей применимости. При наличии разреженного фона или «узких проходов» между кластерами, алгоритм может показать не очень хорошие результаты. Так же у КНП достаточно высокая трудоемкость.

В Таблице 1 приведено сравнение нескольких методов кластерного анализа.

Таблица 1. Сравнение алгоритмов кластеризации

	Вычислительная сложность	Форма кластеров	Входные данные	Результаты
k-means	$O(nki)$	Гиперсфера	Число кластеров k	Центры кластеров, матрица принадлежности
КНП	$O(n^2)$	Произвольная	Число кластеров k	Древовидная структура кластеров
Иерархический	$O(n^2)$	Произвольная	Число кластеров k ; порог расстояния	Бинарное дерево кластеров
c-means	$O(nki)$	Гиперсфера	Число кластеров k ; степень нечеткости	Центры кластеров, матрица принадлежности
Послойная	$O(\max\{n, m\})$	Произвольная	Последовательность порогов расстояния	Древовидная структура кластеров с разными уровнями иерархии

Классификация

Классификация – задача распределения новых объектов к одному из заранее известных классов по принципу схожести его формализованных признаков с признаками класса. Также объекты могут одновременно принадлежать нескольким классам, что определяет пересекающуюся задачу. Каждый элемент в задачах классификации описывается вектором в N -мерном пространстве признаков, т.е. каждое измерение является описанием конкретного признака объекта.

В отличие от кластерного анализа, алгоритм классификации не может начать работу, пока его не обучили. Для этого создают «обучающую выборку», в которой содержатся объекты, уже распределенные по классам. Разметка обучающей выборки может проводиться разными способами: вручную, привлекая специалистов исследуемой области, или же с помощью алгоритмов, решающих задачу кластеризации, которая была рассмотрена выше.

Классификаторы бывают двух различных типов: активного и ленивого. В процессе обучения ленивый классификатор не делает ничего, он просто хранит размеченные тренировочные данные. Работать же начинает только с появлением новых, немаркированных элементов. В отличие от него, активный строит некую классификационную модель, с помощью которой потом будет классифицировать новые данные.

Рассмотрим самый популярный метод – kNN (k Nearest Neighbor или k Ближайших Соседей). Это один из самых простых и самых популярных алгоритмов классификации. Его используют в задачах регрессии, и с него начинают знакомство с областью Machine Learning [4].

Метод kNN относится к ленивому типу. Работу алгоритма можно разделить на 2 этапа:

1. Поиск k ближайших маркированных точек;
2. Определение класса для новых данных, основываясь на классах соседей.

На первом этапе kNN находит такое число точек, которое выбирается на основе особенностей конкретной задачи. Выбор ближайших соседей происходит по некоей дистанционной метрике или коэффициенту подобия.

На втором этапе происходит классификация. Алгоритм решает в какой класс отнести новый элемент. Если все соседи в одной группе – проблем не возникает. Когда ближайшие соседи принадлежат разным кластерам, становится немного сложнее, для решения такой проблемы существуют множество подходов, некоторые из них:

- Правило большинства;
- Взвешенное большинство;
- По центру кластера (k-средних);

Несмотря на свою простоту в понимании и реализации, kNN слишком ресурсозатратный для большого объема данных. Так же на адекватную работу может отрицательно повлиять сильная зашумленность данных, не правильный выбор метрики и характеристики с отличающимся количеством значений. Следует учитывать и то, что алгоритм хранит тренировочные данные, следовательно, требует гораздо больше места, чем активные классификаторы.

Заключение

На данном этапе идет сбор данных для дальнейшего тестирования и доработки разрабатываемой СППР, а также выбор наиболее подходящих методов анализа данных.

Хочется отметить, что проектируемая СППР будет решать следующие задачи:

- быстрый доступ к данным, посредством графического интерфейса пользователя;
- возможность работы с информацией многим пользователям одновременно;
- многомерное концептуальное представление данных;
- расчеты и вычисления по нескольким измерениям, иерархиям и/или членам;
- представление информации в графическом виде;
- получение отчетных документов;
- экспорт результатов анализа в MS Excel;

Литература

1. Международное агентство We Are Social, статья «The state of digital in april 2019: all the numbers you need to know», [Электронный ресурс] – Режим доступа: <https://wearesocial.com/blog/2019/04/the-state-of-digital-in-april-2019-all-the-numbers-you-need-to-know> Дата обращения: 18.03.2019.
2. Кудрявцев Ю. OLAP-технологии: обзор решаемых задач и исследований, 2008.
3. Аминева З.М. Использование OLAP-технологий для эффективного анализа данных ЛПУ // Всероссийская научная конференция "Информационные технологии интеллектуальной поддержки принятия решений", Уфа-Ставрополь, Россия, 2018.
4. Montavon G. Neural Networks. Tricks of the Trade/ Montavon G., Orr G.B., Müller K.-R. Springer, 2012. 753 с.
5. Hastie, T., Tibshirani, R., Friedman, J. The Elements of Statistical Learning, 2nd edition. Springer, 2009. 533 p.

APPLICATION OF OLAP-TECHNOLOGIES FOR ANALYSIS MULTI-DIMENSIONAL DATA IN CONTACT CENTER

Usacheva D.I., Shishkin M.O.

Students of group M091801(72), MTUCI

Gadasin D.V.

MTUCI, PhD., associate professor of MSiUS department

dengadiplom@mail.ru

Guzeev A.V.

MTUCI, PhD., associate professor of MSiUS department

Keywords: *decision support systems, OLAP technology, multidimensional data analysis, contact center data analysis, data mining.*

Formulated the purpose of the work. The approximate output is given. The concept of decision support systems, their architecture is considered. An overview of common tasks OLAP-systems. Concrete tasks for the developed product are given. The tasks and the concept of data mining are considered. A description is given and a comparative analysis of some methods of multivariate data analysis is carried out. A description of some algorithms. Formulated further objectives and requirements for the developed product.

РАЗРАБОТКА ПОДСИСТЕМЫ РАСПОЗНАВАНИЯ ПРОСТЫХ ДИНАМИЧЕСКИХ ЖЕСТОВ С ПОМОЩЬЮ 3DCNNLSTM

Артемов Михаил Денисович
студент группы М151701(70) МГУСИ

artemov_mikle@mail.ru

Воронова Лилия Ивановна
МГУСИ, д.ф.-м.н., проф., зав. каф. ИСУиА
voronova.lilia@yandex.ru

Ключевые слова: 3DCNNLSTM; LSTM; FC; 3DCNN; распознавание простых динамических жестов.

В статье описана подсистема разрабатываемой в Московском Техническом Университете Связи и Информатики (МТУСИ) интеллектуальной информационно-коммуникационной системы социальной доступности для людей с ограниченными возможностями по слуху на основе методов машинного обучения – «Surdotelephone». На основе последних исследований и разработок в области систем распознавания жестового языка, рассмотрена архитектура и способ обучения построенной 3DCNNLSTM модели, которая будет применяться для распознавания простых динамических жестов. В статье описан набор данных, содержащий 25 классов простых динамических жестов, разделённый на обучающий, проверочный и тестовый наборы данных, исходные условия, а также проведён эксперимент, по результатам которого обученная модель нейронной сети получила достоверность предсказания равную 11,11%, а перекрёстная энтропия в среднем составила 3.3.

На кафедре Интеллектуальных систем в управлении и автоматизации МТУСИ разрабатывается интеллектуальная информационно-коммуникационная система социальной доступности для людей с ограниченными возможностями по слуху на основе методов машинного обучения – «Surdotelephone» [1,2]. Её подсистемой является 3DCNNLSTM (3-dimensional convolutional neural network with long short-term memory) модель нейронной сети, разработанная авторами, которая служит для распознавания простых динамических жестов.

3DCNNLSTM архитектура состоит из последовательности свёрточных слоёв, слоя долгосрочной кратковременной памяти и последовательности полносвязных слоёв на выходе нейронной сети, поэтому она получила название 3DCNNLSTM (3-dimensional convolutional neural network with long short-term memory) [3,4]. На рисунке 1 представлена последовательность слоёв, составляющих архитектуру 3DCNNLSTM.

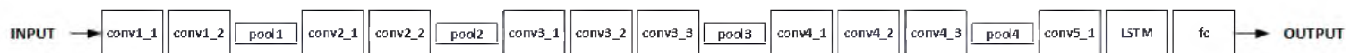


Рис. 1. Последовательность слоёв, составляющих архитектуру 3DCNNLSTM

На входе 3DCNNLSTM располагается свёрточный слой «conv1_1». Он принимает последовательность из 24 кадров видеозаписи размером 128×128, которые одновременно обрабатываются трёхмерными фильтрами слоя и проходят через ReLU-функцию активации. В результате образуются карты признаков, которые поступают на вход следующего свёрточного слоя «conv1_2», который образует новые карты признаков. Всего в архитектуре сети содержится 11 трёхмерных свёрточных слоёв.

Для уменьшения пространственных размеров карт признаков, а также снижения сложности вычислений при формировании карт признаков и вероятности переобучения сети между некоторыми свёрточными слоями располагаются слои субдискретизации. Попадая на эти слои, карты признаков обычно уменьшаются в размерах, за счёт прохода по ним скользящего окна, которое при помощи специальной функции объединяет захваченные

значения в одно. В архитектуре 3DCNNLSTM содержится всего 4 таких слоя все они используют скользящие окна одинакового размера, которые передвигаются с одинаковым шагом, а для объединения значений используют функцию вычисления максимального значения. Так, например, 128 карт признаков размером 128×128 и глубиной 24, попадая из свёрточного слоя «conv1_2» на слой субдискретизации «pool1», уменьшаются и образуют 128 карт признаков размером 64×64 и глубиной 12.

Параметры описанных слоёв 3DCNNLSTM представлены в таблице 1.

В таблице используются следующие обозначения depth – глубина скользящего окна; height – высота фильтра; width – ширина фильтра; in_channels – количество карт признаков, поступающих на входе; out_channels – количество карт признаков, получаемых на выходе; channels – количество каналов, охватываемое одним скользящего окна. str_depth – размер шага скользящего окна в глубину; str_width – размер шага скользящего окна в ширину; str_height – размер шага скользящего окна в высоту; str_channels – количество каналов, охватываемое одним скользящего окна.

Таким образом, из исходной последовательности из 24 изображений размером 128×128 формируется набор из 512 карт признаков размером 8×8.

Таблица 1. Параметры свёрточных слоёв и слоёв субдискретизации 3DCNNLSTM

№	Слой	Размеры фильтра [depth, height, width, in_channels, out_channels]	Размеры скользящего окна [depth, height, width, channels]	Шаг скользящего окна [str_depth, str_width, str_height, str_channels]	Функция	Размер карт признаков на выходе слоя [depth, height, width, channels]
1	conv1 1	[1, 2, 2, 3, 64]	-	[1, 1, 1, 1]	ReLU	[24, 128, 128, 64]
2	conv1 2	[1, 2, 2, 64, 128]	-	[1, 1, 1, 1]	ReLU	[24, 128, 128, 128]
3	pool1	-	[1, 2, 2, 2, 1]	[2, 2, 2, 1]	Max	[12, 64, 64, 128]
4	conv2 1	[1, 2, 2, 128, 256]	-	[1, 1, 1, 1]	ReLU	[12, 64, 64, 256]
5	conv2 2	[1, 2, 2, 256, 256]	-	[1, 1, 1, 1]	ReLU	[12, 64, 64, 256]
6	pool2	-	[1, 2, 2, 2, 1]	[2, 2, 2, 1]	Max	[6, 32, 32, 256]
7	conv3 1	[1, 2, 2, 256, 512]	-	[1, 1, 1, 1]	ReLU	[6, 32, 32, 512]
8	conv3 2	[1, 2, 2, 512, 512]	-	[1, 1, 1, 1]	ReLU	[6, 32, 32, 512]
9	conv3 3	[1, 2, 2, 512, 512]	-	[1, 1, 1, 1]	ReLU	[6, 32, 32, 512]
10	pool3	-	[1, 2, 2, 2, 1]	[2, 2, 2, 1]	Max	[3, 16, 16, 512]
11	conv4 1	[1, 2, 2, 512, 512]	-	[1, 1, 1, 1]	ReLU	[3, 16, 16, 512]
12	conv4 2	[1, 2, 2, 512, 512]	-	[1, 1, 1, 1]	ReLU	[3, 16, 16, 512]
13	conv4 3	[1, 2, 2, 512, 512]	-	[1, 1, 1, 1]	ReLU	[3, 16, 16, 512]
14	pool4	-	[1, 2, 2, 2, 1]	[2, 2, 2, 1]	Max	[2, 8, 8, 512]
15	conv5 1	[1, 2, 2, 512, 512]	-	[2, 1, 1, 1]	ReLU	[1, 8, 8, 512]

Следом за свёрточными слоями размещается LSTM-слой (LSTM cell), он содержит 512 блоков памяти, задача которых сохранять состояние слоя. Условно этот слой можно представить, как скрытый слой в полносвязной сети, как показано на рисунке 2.

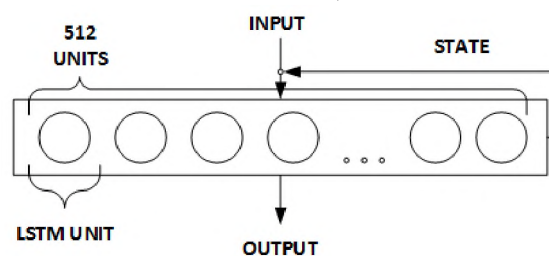


Рис. 2. LSTM ячейка

Для последовательностей, подаваемых на вход, LSTM-слой формирует зависимости. В нашем случае этот слой будет формировать зависимости для последовательности, состоящей из 256 векторов по 128 значений. Поэтому перед тем как поступить на вход этого слоя карты признаков объединяются в единый массив размером 256×128 .

За LSTM-слоем расположились четыре слоя полносвязной сети, изображённые на рисунке 3 («fc» блок на рисунке 1). Первый из них имеет 512 нейронов, он принимает значения из LSTM-слоя. Следующий слой имеет 256 нейронов и для расчёта их значений использует ReLU-функцию активации. Далее следует слой из 128 нейронов, на котором используется та же функцию активации. Выходной слой сети имеет 27 нейронов (количество классов видеозаписей) и применяет softmax-функцию активации для вычисления их значений.

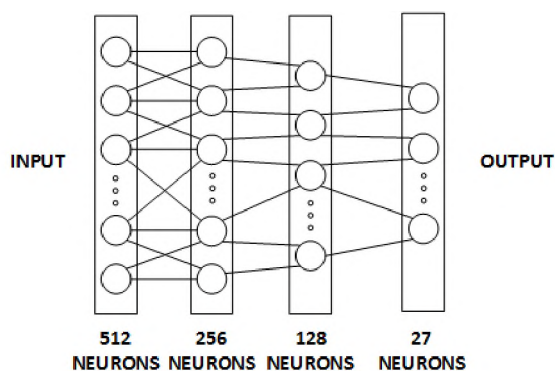


Рис. 3. Полносвязные слои 3DCNNLSTM модели

Для обучения и проверки работы нейронной сети выбран набор данных «The 20BN-jester Dataset V1» [5]. Этот набор данных включает 148 092 записей, представляющих JPG-кадры видео, на которых люди перед камерой ноутбука или веб-камерой выполняют руками заранее определённые жесты. Всего насчитывается 25 классов с жестами: «Swiping Left» («смахивание влево»), «Swiping Right» («смахивание вправо»), «Swiping Down» («смахивание вниз»), «Swiping Up» («смахивание вверх»), «Pushing Hand Away» («отталкивание руки прочь»), «Stop Sign» (сигнал «стоп») и т.д., один класс, говорящий об отсутствии жеста («No gesture») и класс «Doing other things», который говорит о том, что какое-то действие выполняется, но оно не может быть отнесено ни к одному из классов. Используемый набор данных предварительно проклассифицирован и разделён на наборы данных для обучения, валидации и тестирования.

Во время подготовки к обучению формируются записи, в которые записываются по 24 изображения исходного набора, размеры которых приведены к масштабу 128×128 , каждой записи назначается один из 27 классов.

Перед обучением значения фильтров свёрточных слоёв нейронной сети, LSTM-блоков и весов связей между слоями перцептрона инициализируются в значения по умолчанию, а в конце каждой эпохи обучения полученные значения записываются в файловую систему и следующая эпоха использует эти значения и обучение продолжается.

В ходе обучения записи из обучающего набора синхронно поступают на вход нейронной сети. Весь процесс обучения логируется и на каждой 1000ой записи происходит проверка точности обучаемой модели на данных из соответствующего набора.

Среди основных метрик нейронной сети выделяют: достоверность, точность, перекрёстную энтропию

В нашей работе для оценки точности предсказания модели служит перекрёстная энтропия (значения которой также логируются). Чем ближе значение энтропии к нулю, тем выше точность предсказания нейронной сети. Для минимизации ошибки 3DCNNLSTM мы используем алгоритм оптимизации Adam, начальная скорость обучения которого равна 0,1, но с каждой эпохой мы уменьшаем это значение на порядок, чтобы более точно настроить веса модели и уменьшить значение перекрёстной энтропии.

При оценке нейронной сети мы также проверяем её достоверность на тестовом наборе данных. Достоверность определяет имеет ли ожидаемый класс видеозаписи наибольшую вероятность в предсказании обученной модели.

Во избежание переобучения внутри LSTM-слоя, к результату, выдаваемому его блоками, во время обучения добавляется небольшая ошибка. При валидации добавляемая ошибка равна нулю и не оказывает влияние на предсказание сети.

Обучение модели 3DCNNLSTM производилось на графическом процессоре NVIDIA Quadro M4000 при этом среднее время обработки 1000 видеозаписей составляло 10 минут, откуда средняя скорость составляет 0,028 видеозаписей/секунду или 0,67 кадров/секунду.

На рисунке 4 представлен результат пятой эпохи обучения. На этой эпохе скорость обучения составляла 0,00001. Из графиков видно, что высокую достоверность обученная модель принимает в 16 примерах из 144 примеров проверочного набора данных (11,11%), а среднее значение перекрёстной энтропии составляет $\sim 3,30$.

Таким образом, разработанная модель переобучилась и не обеспечивает ожидаемой точности предсказания. Для устранения этой проблемы было принято решение приостановить обучение сети и внедрить методы тонкой настройки, которые ранее не были применены в должной мере. Среди них будут использованы: инициализация начальных значений весов при помощи метода Xavier или He, батч-нормализация (batch normalization), расширение обучающего множества (data augmentation) и ансамбли (ensembles). Кроме того, возможно, будет использована ранняя остановка (early stopping), которая относится к оптимизации гиперпараметров.

Итак, в статье описана архитектура и способ обучения построенной 3DCNNLSTM модели, для распознавания простых динамических жестов. При обучении нейронной сети использовался набор данных, содержащих 25 классов с простыми жестами, но при этом запись жестов производилась в цвете и на сложном, неоднородном, а иногда и изменяющемся фоне. По этим причинам результаты обучения и валидации дали значение перекрёстной энтропии около 3.30, что не обеспечило приемлемой достоверности предсказания для того, чтобы конкурировать с современными нейронными сетями. Тем не менее, данная сеть является перспективной основой для дальнейших разработок, связанных с распознаванием динамических последовательностей кадров в видеопотоке.

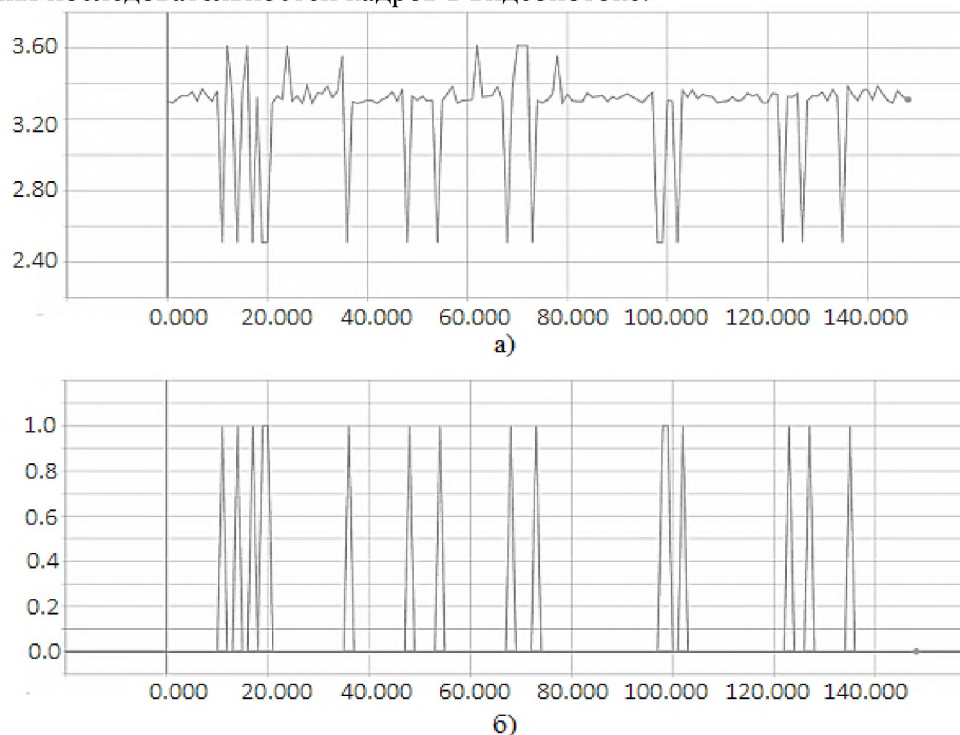


Рис. 4. а) Значения перекрёстной энтропии; б) Значения достоверности предсказания при проверке 3DCNNLSTM модели со скоростью обучения 0,00001

Литература

1. *Voronov V.I., Genchel K.V., Artemov M.D. and Bezumnov D.N.* “Surdotelephone” project with convolutional neural network// Статья конференции: 2018 Systems of Signals Generating and Processing in the Field of on Board Communications, М., 2018. С. 1-6.
2. *Mikhail D. Artemov, Lilia I. Voronova* Design of a architecture of the program complex for equipment control using gestures»// Статья конференции: Technology & Entrepreneurship in Digital Society (TEDS), М., 2018.
3. Gesture recognition using end-to-end learning from a large video database. 2017. URL: <https://medium.com/twentybn/gesture-recognition-using-end-to-end-learning-from-a-large-video-database-2ecfb4659ff> (дата обращения: 23.02.2019)
4. *Eleni Tsironi, Pablo Barros u Stefan Wermter* Gesture Recognition with a Convolutional Long Short-Term Memory Recurrent Neural Network. Proceedings of the European Symposium on Artificial Neural Networks// Статья конференции: Computational Intelligence and Machine Learning (ESANN), Hamburg – Germany, 2016. С. 213-218
5. Dataset «The 20BN-jester Dataset V1» [Электронный ресурс]. URL: <https://20bn.com/datasets/jester> (дата обращения: 23.02.2019)

DEVELOPMENT OF A SUBSYSTEM RECOGNITION SIMPLE DYNAMIC SIGN WITH 3DCNNLSTM

Mikhail D. Artyomov

*Student of group M151701(70), MTUCI
artemov_mikle@mail.ru*

Liliya I. Voronova

*MTUCI, Dr.Sc., professor, Head of ISUA department
voronova.lilia@yandex.ru*

Keywords: *3DCNNLSTM; LSTM; FC; 3DCNN; recognition of simple dynamic gestures.*

The article describes a subsystem of intellectual information-communication system of social accessibility for people with hearing impairments developed by the machine learning methods “Surdotelephone” developed at the Moscow Technical University of Communications and Informatics (MTUCI). Based on the latest research and development in the field of sign language recognition systems, the architecture and method of training the model constructed by 3DCNNLSTM, which will be used to recognize simple dynamic gestures, is considered. The article describes a data set containing 25 classes of simple dynamic gestures, divided into training, test and test data sets, initial conditions, and also an experiment was conducted, according to the results of which the trained neural network obtained a prediction accuracy of 11.11%, and the cross entropy averaged 3.3.