

ФЕДЕРАЛЬНОЕ АГЕНТСТВО СВЯЗИ
ФЕДЕРАЛЬНОЕ ГОСУДАРСТВЕННОЕ БЮДЖЕТНОЕ ОБРАЗОВАТЕЛЬНОЕ УЧРЕЖДЕНИЕ
ВЫСШЕГО ОБРАЗОВАНИЯ
«СИБИРСКИЙ ГОСУДАРСТВЕННЫЙ УНИВЕРСИТЕТ
ТЕЛЕКОММУНИКАЦИЙ И ИНФОРМАТИКИ»

Вестник СибГУТИ № 4 (48) 2019

Выпускается ежеквартально, выходит с 2007 г.

Председатель редакционного совета А. Н. Фионов, д.т.н., проф.

Редакционный совет:

В. П. Бакалов, д.т.н., проф.	В. И. Носов, д.т.н., проф.
В. Б. Барахнин, д.т.н., доц.	В. П. Петров, д.т.н., проф.
В. Н. Васюков, д.т.н., проф.	С. В. Поршнев, д.т.н., проф.
В. Ю. Васильев, д.х.н.	А. С. Родионов, д.т.н., с.н.с.
Н. И. Горлов, д.т.н., проф.	А. И. Романенко, д.ф.-м.н., проф.
Н. Л. Казначеева, д.э.н., доц.	Б. Я. Рябко, д.т.н., проф.
В. С. Канев, д.т.н., доц.	И. И. Рябцев, д.ф.-м.н.
А. И. Карпович, д.э.н., проф.	Э. Сименс, д.т.н.
М. Г. Курносов, д.т.н., доц.	О. В. Стукач, д.т.н., проф.
В. В. Лебедев, д.т.н., проф.	В. К. Трофимов, д.т.н., проф.
А. В. Лихачёв, д.т.н.	А. И. Фалько, д.т.н., проф.
С. Н. Мамоиленко, д.т.н., доц.	С. В. Федоренко, д.т.н., доц.
В. Г. Мартынец, д.ф.-м.н.	А. Г. Черевко, к.ф.-м.н., доц.
А. Б. Мархасин, д.т.н., проф.	С. В. Шилдовский, д.т.н.,
О. Г. Мелентьев, д.т.н., проф.	Ю. И. Шокин, д.ф.-м.н., акад. РАН
Р. В. Мещеряков, д.т.н., проф.	В. П. Шувалов, д.т.н., проф.
И. Г. Неизвестный, д.ф.-м.н., чл.-корр. РАН	

Редакция:

А. Н. Фионов (главный редактор), М. Ю. Галкина (заведующая редакцией),
Н. А. Двуреченская (технический и литературный редактор, компьютерная вёрстка),
Т. А. Алфёрова (лингвист-корректор), Е. Л. Грозовская (секретарь)

Адрес редакции

630102, г. Новосибирск, ул. Кирова, д. 86

e-mail: vestnik@sibsutis.ru

Информация о журнале доступна в сети Internet по адресу <http://vestnik.sibsutis.ru>.

Журнал включён в Перечень ВАК российских рецензируемых научных журналов, в которых должны быть опубликованы основные научные результаты диссертаций на соискание учёных степеней доктора и кандидата наук.

Журнал зарегистрирован Федеральной службой по надзору за соблюдением законодательства в сфере массовых коммуникаций и охране культурного наследия, свидетельство о регистрации ПИ № ФС77-25835 от 29.09.2006. ISSN 1998-6920. Подписной индекс в объединённом каталоге «Пресса России» – 82519.

Отпечатано в издательском центре СибГУТИ. Бумага офсетная, формат А4. Тираж 300 экз.

© СибГУТИ, 2019 г.

СОДЕРЖАНИЕ

О. Н. Коваленко, К. С. Фадеев, Д. Н. Коваленко

Адаптивное управление сервером видеоконференцсвязи при организации
дополнительного сервиса в корпоративных сетях 3

Д. Г. Алгазина, Ю. Г. Алгазина

Динамика и равновесие в модели Курно при неполной информации 10

Д. Е. Быченков

Сравнение характеристик распределённых беспроводных систем связи
с мягкими и жёсткими решениями 16

В. П. Бакалов

Масштабирование рекурсивных цифровых фильтров с нерекурсивными
латеральными связями 23

И. В. Нечта

Полухрупкие цифровые водяные знаки, базирующиеся на спектре Фурье 33

Ю. В. Шевцова, В. С. Канев, А. Н. Полетайкин, Н. В. Кулешова

Новая математическая модель риск-менеджмента образовательной деятельности
вуза 42

В. К. Трофимов

Оценка избыточности универсального кодирования произвольного множества
источников без памяти 56

А. А. Бычков, А. И. Карпович

Российский опыт внедрения систем менеджмента качества 62

К. М. Раздобреев, М. М. Раздобреев

Оценка качества алгоритма оптимизации состава задач автоматизированного
управления 72

Д. В. Кутузов, А. В. Осовский, О. В. Стукач

Модель генерации и обработки трафика IoT параллельными коммутационными
системами 78

**А. Г. Черевко, Ю. В. Моргачев, Е. М. Ильин, А. И. Полубехин,
В. И. Флегонтов**

Оценка критического радиуса изгиба графеновых антенн 88

И. Н. Максимов, О. Г. Мелентьев, И. Е. Шевнина

Применение закодированных сверточным кодом преамбул для кадровой
синхронизации 93

Адаптивное управление сервером видеоконференцсвязи при организации дополнительного сервиса в корпоративных сетях

О. Н. Коваленко, К. С. Фадеев, Д. Н. Коваленко

В статье предложена система организации дополнительного сервиса в служебных сетях передачи данных, принимающая решение о качестве передаваемого видеопотока с помощью математического аппарата нечеткой логики. Показана необходимость гибкого управления сервером видеоконференцсвязи (ВКС) при наличии каналов служебной сети передачи данных с ограниченной полосой пропускания. Введены параметры, требуемые для принятия решения о необходимости изменения выделяемой полосы пропускания для аудио- и видеопотоков. Для составления функций принадлежности предлагается использовать метод парных сравнений.

Ключевые слова: видеоконференцсвязь, полоса пропускания, трафик, теория нечеткой логики.

Видеоконференцсвязь (ВКС) применяется во многих производственных процессах. С помощью видеоконференций организуются обучающие семинары, совещания, собеседования, и, как следствие, снижаются затраты на командировки и повышается эффективность взаимодействия распределенных сотрудников.

Однако с развитием техники в целом и улучшения качества ВКС значительно возрастают требования к пропускной способности сетей передачи данных (СПД). Также необходимо отметить, что у многих крупных распределенных компаний, например ОАО «РЖД», СПД служат для обеспечения основных технологических процессов [1]. Сети передачи данных без предварительной модернизации не смогут гарантировать ключевых качественных показателей при внедрении полномасштабных систем ВКС.

Таким образом, цель работы – определение использования ресурсов служебной СПД в периоды невысокой загрузки для организации дополнительных сервисов, не допуская при этом наличия непроизводительного трафика в период «всплеска активности» основных систем.

Для решения данной проблемы можно использовать технологию адаптивного потокового вещания (adaptive streaming), которая позволяет подстраивать скорость доставки видеoinформации под состояние сети. Также существует технология масштабируемого видеокодирования SVC (Scalable Video Coding), позволяющая передавать в одном потоке несколько подпотоков видео различного качества. SVC позволяет серверу видеоконференций подстраивать видеопоток под изменяющиеся характеристики терминалов участников, такие как процессорные ресурсы и ширина канала связи. [2].

Однако обе эти технологии предназначены для гарантированной доставки видеосигналов, т.е. относятся к видео как к основной услуге, а в нашем случае гарантия должна быть предоставлена другому классу трафика – трафику служебных программ.

Также поставленную задачу можно решить с помощью приоритизации трафика и выделения соответствующего канального ресурса трафику каждого приоритета [3]. Однако в

большинстве работ, например в [3–5], показывается, что трафик видеоконференцсвязи относится к трафику реального времени, чувствительного к джиттеру, поэтому для обеспечения QoS рекомендуется присваивать данному трафику наивысший приоритет. Если высший приоритет установлен для видео, то пострадает основной трафик.

Предлагается адаптивное управление сервером видеоконференций в зависимости от заданных параметров с целью организации ВКС как дополнительного внутреннего сервиса в служебных сетях.

В качестве параметров должны быть использованы уровни загрузки ключевых узлов и каналов, качество активных конференций, иерархия проводимых собраний, а также нужно учесть прогнозируемый объем трафика. На основании этих данных может быть принято решение о том, можно ли начинать новую конференцию и с каким качеством; выданы ли рекомендации по переключению разрешения или необходимости отключения видеопотока; досрочного окончания всех совещаний для недопущения сбоя основных технологических процессов. Управляющие команды могут применяться автоматически, переключая режимы работы сервера ВКС, или носить рекомендательный характер для администратора сервера при ручном управлении.

Однако при реализации такой модели одновременный анализ нескольких разнородных параметров осложняет принятие оперативных решений. Поэтому в рамках данной статьи описывается применение агрегированной оценки о требуемой полосе пропускания для проведения ВКС, полученной с помощью аппарата нечеткой логики. В отличие от существующих методов расчета необходимых ресурсов сети [6, 7], теория нечетких множеств позволяет учитывать как параметры, измеренные с высокой точностью, так и субъективные параметры, выраженные, например, высказываниями «хорошая разборчивость» или «нечеткое изображение». В условиях часто изменяющихся условиях функционирования телекоммуникационной сети это позволяет уменьшить количество измерений, необходимых для принятия решения, а следовательно, увеличить быстроту реакции системы управления ВКС на изменения.

Обобщенно схема управления системой ВКС с использованием математического аппарата теории нечеткой логики представлена на рис. 1.

Блок системы управления (СУ), построенной на базе нечеткого вывода, получает данные от блоков заданных параметров качества (ЗПК), прогнозирования загрузки (ПЗ) и системы мониторинга (СМ), проводит необходимые вычисления и выдает управляющие сигналы блоку ВКС.

Блок ЗПК содержит информацию о заданных допустимых параметрах качества и ограничениях. Принимая данные о характеристиках текущих конференций, он проводит необходимые сравнения и сопоставления и передает на вход блока СУ результаты вычислений.

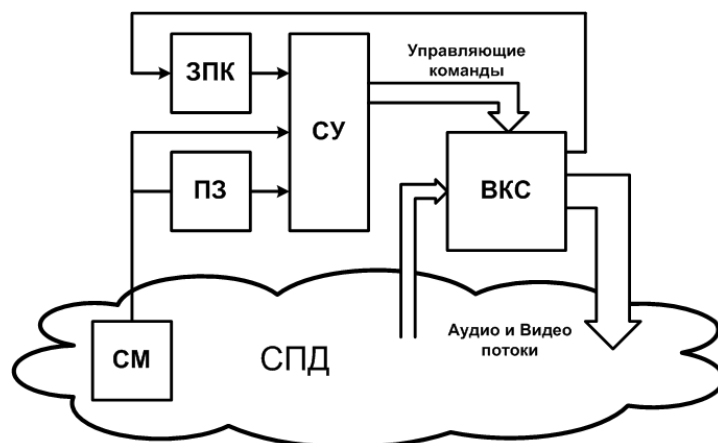


Рис. 1. Схема управления сервером ВКС

Блок ПЗ с помощью методов прогнозирования и экстраполяции оперирует данными на основе статистики загрузки СПД за прошлые периоды, полученной от блока СМ. Результаты передаются на вход блока СУ.

Блок СМ представляет собой систему мониторинга сети СПД. Такая система существует в любой корпоративной сети и отвечает за оперативный сбор информации с коммутационного оборудования ключевых узлов. Дополнительно эта система ведет контроль работы этих узлов и в заданных случаях организует оповещение технических служб. Данные по выбранным узлам передаются в блок СУ для дальнейшей обработки.

Блок ВКС представляет собой непосредственно сервер ВКС, организующий маршрутизацию аудио- и видеопотоков. Также этот блок собирает статистику по потерям, задержкам и джиттеру активных конференций и передаёт эти данные на вход блока ЗПК.

Основной задачей системы управления сервером ВКС является предоставление услуги на заданном уровне качества. Однако в зависимости от ситуации для разных конференций могут быть заданы различные приоритеты и уровни допустимого качества. Таким образом, необходима единая система оценки параметров с гибкой системой результирующих рекомендаций.

При создании такой системы определенную сложность представляет количественная оценка субъективных показателей.

Для обобщенной оценки качества предоставляемой услуги ВКС целесообразно воспользоваться рекомендациями Международного союза электросвязи (МСЭ) Y.1221 [8], Y.1540 [9], Y.1541 [10]. В рекомендации Y.1540 определяются объективные показатели качества, которые следует контролировать в IP-сетях. Рекомендация Y.1541 вводит понятие классов обслуживания (Network Quality of Service Classes), привязывая их к различным видам пользовательских услуг и приложений. В этой же рекомендации приводятся численные значения для параметров, определенных в Y.1540, которые также сопоставляются с классами обслуживания (табл. 1). Рекомендация Y.1221 дает подробное описание и схемы измерений величин этих показателей обслуживания, минимизируя тем самым возможность разночтений при измерениях.

Таблица 1. Классы обслуживания по рекомендации МСЭ Y.1541

Сетевая характеристика	Классы обслуживания					
	Класс 0	Класс 1	Класс 2	Класс 3	Класс 4	Класс 5
Задержка	100 мс	400 мс	100 мс	400 мс	1 с	Н
Джиттер	50 мс	50 мс	Н	Н	Н	Н
Вероятность потерь	1×10^{-3}	1×10^{-3}	1×10^{-3}	1×10^{-3}	1×10^{-3}	Н

Рекомендация Y.1541 устанавливает следующее соответствие между классами качества обслуживания и приложениями [6]:

- 1) класс 0 – приложения реального времени, чувствительные к джиттеру, характеризующиеся высоким уровнем интерактивности (VoIP, видеоконференции);
- 2) класс 1 – приложения реального времени, чувствительные к джиттеру, интерактивные (VoIP, видеоконференции);
- 3) класс 2 – транзакции данных, характеризующиеся высоким уровнем интерактивности (например, сигнализация);
- 4) класс 3 – транзакции данных, интерактивные;
- 5) класс 4 – приложения, допускающие низкий уровень потерь (короткие транзакции, массивы данных, потоковое видео);
- 6) класс 5 – традиционное применение сетей IP.

Таким образом, согласно Y.1541 качественное оказание услуги передачи голоса и видео доступно по классу обслуживания 0 или 1. При прочих условиях качество связи будет очень низкое, а значит, поток видео и голоса лучше вообще отключить.

Из табл. 1 видно, что для обеспечения класса обслуживания 0 необходимо обеспечение низкой задержки, низкого джиттера и низких потерь. Для обеспечения обслуживания по классу 1 допустимый диапазон задержки может быть увеличен до средних значений.

Для работы блока СУ на базе нечеткого вывода необходимо оценить значения входных параметров и определить соответствие этих значений заданным требованиям. Для этого математический аппарат нечеткой логики оперирует такими понятиями, как лингвистическая переменная, терм-множество и функция правдоподобия нечеткого множества [12].

Таким образом, можно сформировать полный набор лингвистических переменных и правил для построения системы нечеткого вывода. Для удобства все переменные сведены в табл. 2.

Таблица 2. Переменные системы нечеткого вывода

X	x1	Имя переменной	Задержка
		Терм-множество	{низкая, средняя, высокая}
		Пределы значений	[0, 450] мс
	x2	Имя переменной	Джиттер
		Терм-множество	{низкий, высокий}
		Пределы значений	[0, 100] мс
	x3	Имя переменной	Вероятность потерь
		Терм-множество	{низкая, высокая}
		Пределы значений	[10 ⁻⁴ , 10 ⁻²]
Y	Имя переменной	Задержка	
	Терм-множество	{низкая, средняя, высокая}	
	Пределы значений	[0, 450] мс	

Система правил в табличной форме будет иметь вид, приведенный в табл. 3.

Таблица 3. Табличный вид системы лингвистических правил

Номер правила	Задержка	Джиттер	Потери	Качество
1	Низкая	Низкий	Низкие	Высокое
2	Средняя	Низкий	Низкие	Среднее
3	Низкая	Высокий	Низкие	Низкое
4	Средняя	Высокий	Низкие	Низкое
5	Высокая	Высокий	Низкие	Низкое
6	Любая	Любой	Высокие	Низкое

В теории нечетких множеств лингвистические переменные описываются пятеркой величин (N, T, U, G, M) , где N – имя переменной; T – терм-множество, элементы которого задаются нечеткими множествами на универсальном множестве U ; G – синтаксические правила, порождающие название термов; M – семантические правила, определяющие функции принадлежности нечетких термов, порожденных синтаксическими правилами из G .

Для каждого имеющегося множества термов $T = \{t_1, t_2, \dots, t_m\}$ и универсального множества $U = \{u_1, u_2, \dots, u_n\}$ нечеткое множество t_j для задания лингвистического терма t_j на универсальном множестве U представляется в виде [7]:

$$t_j = \left(\frac{\mu_{1j}(u_1)}{u_1}, \frac{\mu_{1j}(u_2)}{u_2}, \dots, \frac{\mu_{1j}(u_n)}{u_n} \right), \quad j = \overline{1, m}. \quad (1)$$

Таким образом, для работы системы необходимо определить семантические правила для каждой переменной. В терминах теории нечетких множеств следует задать функции принадлежности для каждого терма, определив степени принадлежности элементов множества U к элементам из множества T , т.е. найти $\mu_j(u_i)$ для всех j и i .

Согласно общей теории нечетких множеств существуют прямые и косвенные методы задания функции принадлежности для выбранных лингвистических переменных.

Прямые методы для одного эксперта состоят в непосредственном задании функции, позволяющей вычислять значения. Такие методы осложнены необходимостью строгого однозначного ранжирования проявления заданных свойств у рассматриваемых объектов. При этом так как степень принадлежности рассматривается на реальном множестве, а не в абсолютном смысле, то интенсивность принадлежности можно определить, исходя из попарных сравнений элементов [14].

Воспользуемся методом парных сравнений. Этот метод используется в случае, если универсальное множество конечно и имеет небольшую размерность.

Для каждой пары элементов универсального множества U оценим преимущество одного элемента над другим по отношению к свойству заданного нечеткого множества. Для оценки интенсивности проявления некоторого свойства используется девятибалльная шкала Саати:

$$\{0, 1, 2, 3, 4, 5, 6, 7, 8, 9\} \{1/k\}, k = 1 \dots 9, \quad (2)$$

где 0 – два объекта в отношении некоторого свойства незначимы (нет значимости);

1 – очень слабая значимость;

3 – слабая значимость;

5 – более или менее значимы;

7 – сильная значимость;

9 – абсолютная значимость;

2, 4, 6, 8 – промежуточные значения, предназначены для случаев неуверенности экспертов, но чаще всего не используются.

С помощью 9-балльной шкалы Саати составим матрицу парных сравнений для функции принадлежности нечеткого множества «низкая задержка» (табл. 4).

Таблица 4. Матрица парных сравнений для функции принадлежности нечеткого множества «низкая задержка»

Задержка, мс	0	50	100	150	200	250	300	350	400	450
0	1	1	1	3	7	9	9	9	9	9
50	1	1	1	3	7	9	9	9	9	9
100	1	1	1	3	5	9	9	9	9	9
150	1/3	1/3	1/3	1	3	5	5	5	7	9
200	1/7	1/7	1/5	1/3	1	1	3	3	5	7
250	1/9	1/9	1/9	1/5	1	1	1	3	3	5
300	1/9	1/9	1/9	1/5	1/3	1	1	1	1	3
350	1/9	1/9	1/9	1/5	1/3	1/3	1	1	1	3
400	1/9	1/9	1/9	1/7	1/5	1/3	1	1	1	1
450	1/9	1/9	1/9	1/9	1/7	1/5	1/3	1/3	1	1

Найдем собственный вектор матрицы A как $W = \{w_1, w_2, \dots, w_n\}$ из системы уравнений [15]:

$$\begin{cases} AW = \lambda_{\max} W, \\ w_1 + w_2 + \dots + w_n = 1 \end{cases} \quad (3)$$

Вычислим степени принадлежности по координатам собственного вектора:

$$\mu(u_i) = w_i, i = 1, n. \quad (4)$$

Находя собственный вектор матрицы согласно (4), получим значения степеней принадлежности для данного нечеткого множества (табл. 5).

Таблица 5. Степени принадлежности для нечеткого множества «низкая задержка»

Задержка, мс	0	50	100	150	200	250	300	350	400	450
μ	1.000	1.000	0.956	0.482	0.233	0.166	0.107	0.097	0.079	0.062

Аналогично рассчитываются степени принадлежности для нечеткого множества «средняя задержка» и «высокая задержка», а также для джиттера и вероятности потерь. Таким образом, задаются функции принадлежности всех входных переменных, на основании которых вычисляется оценка задержки.

Данный пример расчета показывает простоту перехода от субъективных показателей к числовым значениям при использовании математического аппарата нечеткой логики. Разработка нечеткой системы оценки практически не требует точных знаний о значениях показателей качества и привлечения многочисленных экспертов. Также не требуется составление сложных функций описания работы системы в целом. Все возможные неточности при проектировании системы могут быть скорректированы благодаря возможности самообучения системы на основе накопленных статистических данных.

Аналогичным образом могут быть преобразованы и в дальнейшем сопоставлены любые другие необходимые параметры.

Все рассчитанные функции принадлежности используются блоками ЗПК и СУ, где с помощью классических методов нечеткой логики производятся необходимые сравнения и вычисления, принимаются решения и выдаются рекомендации или управляющие команды для системы управления видеоконференцией. Таким образом, можно осуществить адаптивное динамическое управление услугой предоставления ВКС как дополнительного сервиса.

Предложенная модель управления позволяет обеспечивать адаптивное выделение ресурсов сети для организации ВКС с учетом загрузки сети трафиком технологических процессов. Применение аппарата нечеткой логики позволяет упростить принятие решения о распределении ресурсов сети в условиях изменяющейся конфигурации сети передачи данных и отсутствия статистических данных.

В дальнейшем планируется предлагаемую методику распределения ресурсов сети между технологическим трафиком и трафиком ВКС на основе теории нечетких множеств апробировать на конкретной сети.

Литература

1. Коваленко О. Н., Коваленко Д. Н. ВКС как дополнительный сервис корпоративных сетей // Материалы V МНПК «Актуальные направления научных исследований: от теории к практике». Чебоксары: ЦНС «Интерактив плюс», 2015. № 3 (5). С. 266–267.
2. Флюстер Н. ВКС. Скачок в будущее // Технологии и средства связи. 2014. № 2. С. 34–36.
3. Меликов А. З., Пономаренко Л. А., Паладюк В. В. Телетрафик: модели, методы, оптимизация. Киев: ИПК Политехника, 2007. 256 с.
4. Яновский Г. Г. Качество обслуживания в сетях IP // Вестник связи. 2008. № 1. С. 8–24.
5. Пшеничников А. П., Али Раад А. М. Оценка времени установления соединений при предоставлении услуг IPTV с использованием платформы IMS // T-Comm – Телекоммуникации и транспорт. 2013. № 7. С. 99–101.

6. Вишневский В. М., Семенова О. В. Математические методы исследования систем поллинга // Автоматика и телемеханика. 2006. Вып. 2. С. 3–56.
7. Заринова Э. Р. К методам исследования процесса обмена сообщениями при предоставлении видеоконтента // Материалы всероссийской конференции «Информационно-телекоммуникационные технологии и математическое моделирование высокотехнологичных систем». М.: РУДН. 2014. С. 86–87.
8. ITU-T Recommendation Y.1221 (2002). Traffic control and congestion control in IP based networks.
9. ITU-T Recommendation Y.1540 (2002). Internet protocol data communication service – IP packet transfer and availability performance parameters.
10. ITU-T Recommendation Y.1541 (2002). Network performance objectives for IP based services.
11. Соколов Д. П. Нечеткая система оценки качества // Технологии и средства связи. 2009. № 4. С. 26–28.
12. Курилов О. С. Объективный анализ качества речи в IP-телефонии // Технологии и средства связи. 2002. № 4. С. 22–28.
13. Новак В., Перфильева И., Мочкорж И. Математические принципы нечеткой логики: пер. с англ. М.: Физматлит, 2006. 347 с.
14. Яхьяева Г. Э. Нечеткие множества и нейронные сети. М.: БИНОМ, 2006. 316 с.

*Статья поступила в редакцию 14.02.2019;
переработанный вариант – 24.05.2019.*

Коваленко Ольга Николаевна

к.т.н., доцент, доцент кафедры «Телекоммуникационные, радиотехнические системы и сети» ОмГУПС (644050, Омск, пр. К. Маркса, д. 35), e-mail: onkovalenko@mail.ru.

Фадеев Константин Сергеевич

к.т.н., доцент, и.о. заведующего кафедры «Телекоммуникационные, радиотехнические системы и сети» ОмГУПС, e-mail: fadeev_spi@mail.ru.

Коваленко Дмитрий Николаевич

ведущий технолог Новосибирского ИВЦ – структурного подразделения ГВЦ – филиала ОАО «РЖД» (644020, Омск, ул. Леконта, д. 4), e-mail: dnkovalenko@mail.ru.

**Adaptive control via video conference server while organizing additional service
in corporate networks**

O. Kovalenko, K. Fadeev, D. Kovalenko

In this article, organization of additional service system in office data transmission networks making a decision on quality of the transferred video flow by means of a mathematical apparatus of fuzzy logic is presented. The necessity of flexible server control of VCS in the presence of service data network channels with limited bandwidth is shown. The parameters demanded for making decision on the necessity for changing the selected bandwidth for audio-and video flows are entered. To draw up accessory functions a method of paired comparisons is offered.

Keywords: video conference, bandwidth, traffic, theory of fuzzy logic.

Динамика и равновесие в модели Курно при неполной информации

Д. Г. Алгазина, Ю. Г. Алгазина

Игровое моделирование рационального поведения агентов на рынке олигополии при неполной их информированности предполагает многократное повторение игры и построение траектории принятых решений. Важной является задача построения такой траектории, которая гарантированно приводит к равновесию и эффективна с точки зрения предположений относительно текущей информированности агентов в условиях ограниченности их когнитивных возможностей, возможностей агентов в определении величины шагов, скорости сходимости. Представлена модель динамического поведения на рынке Курно в классе линейных функций спроса и издержек агентов. Используя минимальную информацию о рынке и наблюдая только сложившуюся на рынке цену, агенты в динамике на основе модели коллективного поведения уточняют свои объемы выпуска. Получены достаточные условия сходимости динамики к равновесию Курно–Нэша с меняющимися от игры к игре, но равными в каждой отдельной игре шагами агентов. Обсуждаются особенности и возможные решения по улучшению динамики. Проведен также сравнительный анализ условий на величины шагов для сходимости ряда траекторий.

Ключевые слова: модель Курно, неполная информированность, повторяющиеся игры, уточнение выпусков, равновесие.

1. Введение

Теоретико-игровое моделирование рационального некооперативного поведения агентов при неполной их информированности предполагает осуществление рефлексии (предсказывания ходов оппонентов), чтобы находить лучшие ответы на ходы оппонентов (см., например, [1–4]). Теория коллективного поведения дополняет теорию игр тем, что дает такие механизмы, которые без предсказания выбора оппонентов и при достаточно слабых предположениях об информированности игроков позволяют прийти к равновесию игры, т.е. когда ходы всех игроков являются наилучшими ответами на ходы оппонентов (см., например, [5, 6]).

В представленной работе подходы теории коллективного поведения к решению рефлексивных некооперативных игр рассмотрены на классической игровой модели олигополии Курно [7–10 и др.]. Содержательное описание изучаемой проблемной ситуации заключается в следующем. Наблюдая текущее состояние рынка, агент может убедиться в том, что его объемы выпуска продукции не являются оптимальными. К такому выводу могут прийти не один, а несколько или сразу все конкурирующие друг с другом объемами выпуска агенты. Естественно, что у каждого из них возникает желание уточнить свой выпуск так, чтобы он был оптимальным ответом на выбор конкурентов. Если это удастся сделать всем агентам, то выбранные объемы выпуска будут равновесными, так как агенты не будут заинтересованы, чтобы при отсутствии кооперации в одиночку изменить их. Решение игры требует ее многократного повторения. Поэтому важным для рынка является изучение такой траектории уточнения от игры к игре выпусков каждым агентом, которая бы гарантированно приводила к равновесию.

Суть рассмотренной в статье траектории заключается в следующем. Каждый агент в каждой текущей игре, наблюдая цену товара, сложившуюся в предыдущей игре, и выбирая в качестве текущей цели соответствующий ей оптимальный объем своего выпуска, корректирует свой объем выпуска в предыдущей игре, делая от него «шаг» в направлении текущей цели. Траектория основана на модели коллективного поведения, агенты используют минимальную информацию о рынке. Такой информацией выступает цена товара.

В качестве альтернативной рассматривается траектория, в которой агенты в качестве текущей цели используют свои текущие представления о предельных издержках конкурентов и корректируют свои представления в предыдущей игре, делая от них «шаг» в направлении текущей цели [10]. Полагается, что исходные представления агентов недостоверны. Представлен сравнительный анализ траекторий.

Основной результат работы: сформулированы и математически доказаны достаточные условия на величины шагов для сходимости динамик к равновесию Курно–Нэша. Этот результат обобщает известные аналитические решения рефлексивной игры одновременного порядка для классической модели олигополии Курно с произвольным числом агентов.

2. Модель рынка олигополии и адаптивная динамика поведения агентов

На рынке n фирм-агентов конкурируют объемами выпуска однородной продукции. Каждый агент продает произведенный им выпуск q_i по единой рыночной цене $p(Q)$, которая определяется общим объемом выпуска $Q = \sum_{i=1}^n q_i$. Фирмы рациональны, их действия направлены на максимизацию прибыли:

$$\Pi_i(p(Q), q_i) = p(Q)q_i - \phi_i(q_i) \rightarrow \max_{q_i > 0}. \quad (1)$$

Цена $p(Q)$ и полные издержки фирм $\phi_i(q_i)$ заданы линейными функциями:

$$p(Q) = a - bQ, \quad \phi_i(q_i) = c_i q_i + d_i, \quad i = 1, \dots, n, \quad (2)$$

где a, b – параметры спроса; c_i, d_i – предельные и постоянные издержки фирм. Считается, что если $Q \geq a/b$, т.е. $p(Q) \leq 0$, то фирмы несут потери в объеме полных издержек. Кооперация фирм друг с другом, барьеры на вход и ограничения мощности отсутствуют. Весь выпуск будет реализован.

Согласно предположению Курно относительно объемов выпуска каждая фирма действует так, что не ожидает от своих конкурентов изменения объемов выпуска, даже если сама сделает это. Формально это предположение можно записать в виде условий (см., например, [9, 10]):

$$\frac{\partial Q}{\partial q_i} = 1, \quad i = 1, \dots, n. \quad (3)$$

Состояние, в которое приходит рынок, если каждый агент действует подобным образом, называется *равновесием Курно*, которое для классической модели олигополии является *равновесием Нэша*. Обозначим верхним индексом C показатели для равновесия Курно–Нэша. Будем считать, что $q_i^{(C)} > 0$ ($i = 1, \dots, n$), т.е. в равновесии нет неконкурентоспособных по выпуску агентов. Имеют место равенства [9, 10]

$$c_i + bq_i^{(C)} = p(Q^{(C)}), \quad i = 1, \dots, n. \quad (4)$$

Справедливо и обратное: если для вектора положительных действий $(q_1, q_2, \dots, q_n)$ агентов выполняется система равенств (4), то он является равновесным. Этот факт следует из того, что дифференцирование (4) по q_i приводит к равенствам (3). Таким образом, по системе равенств

(4) можно определить, является ли вектор положительных действий $(q_1, q_2, \dots, q_n)$ равновесным или не является.

Рассмотрим следующий динамический процесс.

1. Каждый из агентов независимо от других, используя наблюдаемую цену p^{t-1} и полагая, что в текущем t -м периоде все остальные агенты выберут те же действия, как и в предыдущем $(t-1)$ -м периоде, на основе (1) рассчитывает согласно своей функции реакции текущий оптимальный выпуск x_i^{t-1} по формуле

$$x_i^{t-1} = \frac{1}{2b} (p^{t-1} + bq_i^{t-1} - c_i), \quad i = 1, \dots, n. \quad (5)$$

2. Каждый агент изменяет свой выпуск за предыдущий $(t-1)$ -й период, делая от него шаг по направлению к текущему оптимальному выпуску x_i^{t-1} по формуле

$$q_i^t = q_i^{t-1} + \gamma_i^t (x_i^{t-1} - q_i^{t-1}), \quad i = 1, \dots, n, \quad (6)$$

где $\gamma_i^t \in [0; 1]$ – параметры, определяющие величины шагов.

3. По текущему выпуску агентов аукционером формируется рыночная цена товара $p^t = a - bQ^t$.

Затем процесс повторяется с п. 1.

Утверждение. В модели олигополии Курно с линейными функциями спроса и издержек агентов для того, чтобы итеративный процесс (5)–(6) сходил к равновесию Курно–Нэша при $\gamma_i^t \equiv \gamma^t$, достаточно, начиная с некоторого t , выполнение условий

$$\gamma^t \in (0, 1] \text{ при } n = 1, 2; \quad \gamma^t \in \left(0, \frac{4}{n+1}\right) \text{ при } n > 2. \quad (7)$$

Доказательство. Пусть (7) выполняется. Из формул (2), (5) и (6) имеем

$$q_i^t = q_i^{t-1} + \frac{\gamma^t}{2} \left(\frac{a - c_i}{b} - Q^{t-1} - q_i^{t-1} \right). \quad (8)$$

Учитывая, что $Q^t = \sum_{i=1}^n q_i^t$, имеем

$$Q^t = Q^{t-1} + \frac{\gamma^t}{2} \left(\sum_{i=1}^n \frac{a - c_i}{b} - (n+1)Q^{t-1} \right). \quad (9)$$

Из (9) для t -го и $(t-1)$ -го периодов получаем рекуррентную формулу $Q^t - Q^{t-1} = \frac{\gamma^t}{\gamma^{t-1}} \left(1 - \frac{\gamma^{t-1}(n+1)}{2} \right) (Q^{t-1} - Q^{t-2})$ Или

$$Q^t - Q^{t-1} = \frac{\gamma^t}{\gamma^1} (Q^1 - Q^0) \prod_{i=1}^{t-1} \left(1 - \frac{\gamma^i(n+1)}{2} \right). \quad (10)$$

Из (9) и (10) имеем $\sum_{i=1}^n \frac{a-c_i}{b} - (n+1)Q^{t-1} = \frac{2}{\gamma^1} (Q^1 - Q^0) \prod_{t=1}^{t-1} \left(1 - \frac{\gamma^t(n+1)}{2}\right)$. Из (7) следует, что $\left|1 - \frac{\gamma^t(n+1)}{2}\right| < 1$. Поэтому $\sum_{i=1}^n \frac{a-c_i}{b} - (n+1)Q^{t-1} \rightarrow 0$ и $Q^{(C)} = \lim_{t \rightarrow \infty} Q^t = \frac{1}{(n+1)b} \left(na - \sum_{i=1}^n c_i\right)$ есть равновесный по Курно–Нэшу суммарный выпуск агентов.

По (8) для t -го и $(t-1)$ -го периодов получаем

$$q_i^t - q_i^{t-1} = \frac{\gamma^t}{\gamma^{t-1}} \left(1 - \frac{\gamma^{t-1}}{2}\right) (q_i^{t-1} - q_i^{t-2}) - \frac{\gamma^t}{2} (Q^{t-1} - Q^{t-2}) \quad (11)$$

$$\begin{aligned} \text{Тогда } \lim_{t \rightarrow \infty} (q_i^t - q_i^{t-1}) &= \lim_{t \rightarrow \infty} \frac{\gamma^t}{\gamma^{t-1}} \left(1 - \frac{\gamma^{t-1}}{2}\right) (q_i^{t-1} - q_i^{t-2}) - \lim_{t \rightarrow \infty} \frac{\gamma^t}{2} (Q^{t-1} - Q^{t-2}) = \\ &= \lim_{t \rightarrow \infty} \frac{\gamma^t}{\gamma^{t-1}} \left(1 - \frac{\gamma^{t-1}}{2}\right) (q_i^{t-1} - q_i^{t-2}) = \lim_{t \rightarrow \infty} \frac{\gamma^t}{\gamma^1} (q_i^1 - q_i^0) \prod_{t=1}^{t-1} \left(1 - \frac{\gamma^t}{2}\right) = 0. \end{aligned}$$

Из (8) и (11) получаем сходимость процесса к равновесным выпускам агентов:

$$\begin{aligned} \frac{1}{\gamma^{t-1}} \left(1 - \frac{\gamma^{t-1}}{2}\right) (q_i^{t-1} - q_i^{t-2}) - \frac{1}{2} (Q^{t-1} - Q^{t-2}) &= \frac{a-c_i}{b} - Q^{t-1} - q_i^{t-1}, \\ \lim_{t \rightarrow \infty} \left(\frac{a-c_i}{b} - Q^{t-1} - q_i^{t-1}\right) &= \frac{a-c_i}{b} - Q^{(C)} - \lim_{t \rightarrow \infty} q_i^{t-1} = 0 \text{ и } q_i^{(C)} = \lim_{t \rightarrow \infty} q_i^{t-1} = \frac{a-c_i}{b} - Q^{(C)}. \end{aligned}$$

При выводе соотношений (11) мы полагали, что $\gamma^t \neq 0$. Если, начиная с некоторого t , $\gamma^t \equiv 0$, то по (8) $q_i^t = q_i^{t-1}$, т.е. процесс остановится и не придет в равновесие, если не находился в нем в начальном состоянии.

Утверждение доказано.

3. Обсуждение и выводы

1. Рекуррентные соотношения рефлексивной игры одновременного порядка [9, с. 23], можно получить из соотношений (6) при $\gamma^t \equiv 1$, т.е. агенты в каждом периоде выбирают свой наилучший ответ. Из доказанного утверждения следует, что такой процесс сходится только для случая монополии или дуополии, иначе расходится, что согласуется с выводами в [9].

2. Для итеративных игровых процессов Брауном предложена политика выбора шага γ^t , исходящая из принципа равного вклада в формируемую стратегию q_i^t оптимальных ответов x_i^t и опыта, накопленного в t сыгранных партиях [5]. Тогда $\frac{\gamma^t}{1-\gamma^t} = \frac{1}{t}$ или $\gamma^t = \frac{1}{t+1}$. Для этой

политики выполнено условие (7) и $\gamma^t \rightarrow 0$. Поэтому процесс сходится к равновесию.

3. Рядом авторов отмечается, что для рынка с рефлексорирующими агентами возникает потребность в регулировании ([6, 11] и др.), в том числе в форме «навязывания» участникам представлений о процессе через модель рефлексивного коллективного поведения (см., например, [6, с. 149]). Здесь это нашло отражение следующим образом. Всем агентам для уточнения собственных выпусков предложено использовать единую модель, в которой значение параметра «величина шага» движения к цели определяет, будет ли динамика сходящейся к равновесию или нет. С содержательной точки зрения рациональные агенты, следуя только концепции индивидуального выбора, предпочтут наилучший ответ на действия конкурентов, что

соответствует $\gamma_i^t \equiv 1$. Но этот выбор не дает (при $n > 2$) сходящуюся динамику, поэтому ее развитие может направляться выделением множества траекторий условиями на величины шагов. В нашем случае рефлексивной игры длины шагов могут меняться во времени.

4. В (10) для некоторого t^* может иметь место $\gamma^{t^*} = \frac{2}{n+1}$. Тогда условие (7) выполняется

и при $t > t^*$ будет $Q^t = Q^{t-1}$, по (9) – $Q^{t-1} = Q^{(C)} = \frac{1}{n+1} \sum_{i=1}^n \frac{a - c_i}{b}$ и в (6) в качестве текущего положения цели x_i^{t-1} будет всегда выступать его равновесное значение $q_i^{(C)}$. Процесс сходится, так как $q_i^t - q_i^{(C)} = (1 - \gamma^t)(q_i^{t-1} - q_i^{(C)})$.

5. Представляет интерес сравнение условий сходимости траектории по уточнению объемов выпуска агентов (динамика 1), которая представлена в этой статье, и траектории (динамика 2), в которой агенты, не зная достоверно предельные издержки конкурентов и наблюдая текущую цену товара, выбирают в качестве текущей цели соответствующие этой цене свои представления об их предельных издержках. Затем корректируют эти представления, делая от них «шаг» в направлении текущей цели [10, 12].

Доказано [10], что для сходимости такого итеративного процесса при $\gamma_i^t \equiv \gamma$ ($i = 1, \dots, n; t = 1, 2, \dots$) должно выполняться условие $\gamma \in \left(0, \frac{2}{n}\right)$ ($n > 1$). Из (1) и (4) следует, что при $n = 2$ нет различий между диапазонами значений параметров для обоих сходящихся процессов. При $n > 1$ имеем, что $\frac{4}{n+1} > \frac{2}{n}$.

Анализ указанных динамик показывает:

а) сходящаяся динамика 1 дает больше возможностей по выбору значений параметров. Если динамика 2 сходится, то динамика 1 сходится для тех же значений параметров. Обратное неверно;

б) вычисление динамики 1 проще и менее трудоемко;

в) сходимость динамики 1 доказана для единой длины шага у агентов в каждый момент времени, а динамики 2 – при более жестких условиях, предполагающих единую длину шага для всех агентов и всех моментов.

6. В условиях доказанного утверждения процесс (5)–(6) сходится к равновесию Курно–Нэша при любых первоначальных значениях выпуска и цены.

7. Представляет интерес исследование процессов, в которых каждый игрок может придерживаться своей собственной политики выбора величины шага γ_i^t , отличной от политики других агентов. Для таких процессов пока не существует универсального аппарата поиска аналитических решений, но представляется перспективным получение аналитических выражений условий сходимости для некоторых частных случаев процессов, учитывающих определенный содержательный смысл в том, почему в конкретном периоде тот или иной агент выбирает именно такой, а не другой шаг. В качестве альтернативного метода исследования сходимости таких процессов используется также метод численного или имитационного моделирования, составляющий базис результатов расчетов, иллюстрирующих, что в некоторых случаях разноинформированные агенты могут в динамике (при повторении игры) прийти к истинному информационному равновесию, как например в [8, 9].

Литература

1. Myerson R. Game Theory: Analysis of Conflict. Harvard University Press, 1997. 600 p.
2. Novikov D., Chkhartishvili A. Reflexion and Control: Mathematical Models. Leiden: CRC Press, 2014. 298 p.

3. *Kukushkin N.* Best Response Dynamics in Finite Games with Additive Aggregation // *Games and Economic Behavior*. 2004. № 48. P. 94–110.
4. *Sakovics J.* Games of Incomplete Information without Common Knowledge Priors // *Theory and Decision*. 2001. № 50. P. 347–366.
5. *Беленький В. З., Волконский В. А., Иванков С. А. и др.* Итеративные методы в теории игр и программировании. М.: Наука, 1974. 320 с.
6. *Васин А. А.* Модели динамики коллективного поведения. М.: МГУ, 1989. 156 с.
7. *Kamalinejad H., Majda V., Kebriaei H., Kian A.* Cournot Games with Linear Regression Expectations in Oligopolistic Markets // *Mathematics and Computers in Simulation*. 2010. V. 80, № 9. P. 1874–1885.
8. *Yang H., Zhang Y.* Complex Dynamics Analysis for Cournot Game with Bounded Rationality in Power Market // *Journal Electromagnetic Analysis and Applications*. 2009. № 1. P. 48–60.
9. *Дюсуше О. М.* Статическое равновесие Курно–Нэша и рефлексивные игры олигополии: случай линейных функций спроса и издержек // *Экономический журнал ВШЭ*. 2006. № 1. С. 3–32.
10. *Алгазин Г. И., Алгазина Д. Г.* Информационное равновесие в модели динамики коллективного поведения на конкурентном рынке // *Управление большими системами*. 2016. № 64. С. 112–136.
11. *Понькина Е. В., Маничева А. С., Комаров П. В.* Модель рассредоточенного рынка с барьерами на вход // *Изв. Алт. гос. ун-та*. 2012. № 1–2 (73). С. 104–109.
12. *Algazin G., Algazina D.* Collective Behavior in the Stackelberg Model under Incomplete Information // *Automation and Remote Control*. 2017. V. 78, № 9. P. 1619–1630.

Статья поступила в редакцию 08.05.2019.

Алгазина Дарья Геннадьевна

к.т.н., доцент кафедры прикладной информатики в экономике, государственном и муниципальном управлении АлтГУ, e-mail: darya.algazina@mail.ru.

Алгазина Юлия Геннадьевна

к.э.н., доцент кафедры информационных систем в экономике АлтГТУ, e-mail: algazina@inbox.ru.

Dynamics and equilibrium in the Cournot model with incomplete information

D. Algazina, J. Algazina

Game simulation of agents' rational behavior in the market of oligopoly with their incomplete awareness involves multiple repetition of the game and the construction of decisions made trajectory. The problem of constructing the trajectory that guarantees the equilibrium and efficiency from the point of view of assumptions concerning the current awareness of agents in conditions of their limited cognitive capabilities, the capabilities of the agents in determining the value of steps, the rate of convergence is important. A model of dynamic behavior in the Cournot market in the class of linear demand functions and costs of agents is presented. Using minimal information about the market and observing only the current market price, agents in the dynamics based on the model of collective behavior clarify their output volumes. Sufficient conditions for convergence of dynamics to the Cournot–Nash equilibrium with the steps of agents varying from game to game, but equal in each individual game are obtained. Features and possible solutions for improving the dynamics are discussed. A comparative analysis of the conditions on steps value for a number of trajectories convergence is also carried out.

Keywords: Cournot model, incomplete information, repeated games, clarifying output, equilibrium.

Сравнение характеристик распределённых беспроводных систем связи с мягкими и жёсткими решениями

Д. Е. Быченков

Рассматривается эффективность использования мягких оценок для передачи по распределённой беспроводной сети датчиков. По результатам моделирования было определено, что предложенный способ передачи имеет ощутимый выигрыш в показателях вероятности битовой ошибки в отличие от традиционного использования жёстких оценок для передачи через беспроводной канал с релеевскими замираниями.

Ключевые слова: распределённые беспроводные сенсорные сети, мягкие оценки, многолучевое распространение.

1. Введение

Кооперативная ретрансляция – это одна из наиболее широко исследуемых и перспективных технологий, которая используется для реализации функций распределённой беспроводной сети сенсоров и достижения преимуществ виртуального разнесения. В данной технологии существует источник (передатчик), ретрансляторы (один или более) и приёмник. В зависимости от конфигурации сети канал передачи источник – приёмник может отсутствовать.

Схемы ретрансляции в основном встречаются двух типов: ретрансляция без декодирования в ретрансляторах (Amplify-and-Forward, AF) и ретрансляция с декодированием на стороне ретрансляторов (Decode-and-Forward, DF) [1]. В протоколе AF, когда ретранслятор получает сигнал от источника, он будет усилен и отправлен ретранслятором в приёмник без декодирования [2]. В протоколе DF после приёма сигнала от источника ретранслятор будет декодировать, кодировать, а затем отправлять этот сигнал в приёмник [3]. Существуют также исследования по передаче ретранслятором мягких оценок (Estimate-and-Forward, EF) [4, 5], но конфигурация сети в этих работах подразумевает наличие канала передатчик – приёмник.

2. Постановка задачи

В данной работе сравниваются вероятности битовой ошибки в приёмнике для двух схем ретрансляции: с использованием жёстких решений (Decode-and-Forward) в ретрансляторе и с использованием мягких решений в ретрансляторе (Estimate-and-Forward). Эти схемы используются для передачи данных в сейсмической системе наблюдения. Элементы данной системы (датчики) располагают ограниченным объёмом памяти и вычислительным ресурсом. Поэтому они вынуждены кооперироваться для решения сложных вычислительных (в частности, оптимизационных) задач, связанных с совместной обработкой сейсмических сигналов. Эти датчики используют общий беспроводной канал связи для распределения вычислительной нагрузки и обмена блоками данных. Поскольку антенны датчиков находятся на небольшом расстоянии от поверхности земли, то это приводит к многочисленным переотражениям и многолучевому распространению сигналов. Достаточно часто вычисления должны производиться

синхронно и все датчики системы должны располагать одним блоком данных, который передаётся всем датчикам. Датчики посылают запрос необходимого блока данных в общий канал связи. После поступления в общий канал адреса запрашиваемого блока датчики, у которых имеется в памяти такой блок, начинают широковещательную передачу с целью сделать его доступным всем датчикам системы.

3. Описание модели

Рассмотрим беспроводную сеть с количеством узлов $R+2$, которые размещены произвольным образом. Один из узлов является передатчиком (источником) и один приёмником. Остальные R узлов являются ретрансляторами. Каждый узел имеет одну антенну, которая используется как для приёма, так и для передачи. Так как все узлы (датчики) для вычислений должны обладать одним и тем же блоком данных, то будем считать, что все узлы приняли данные, когда самый удалённый от передатчика узел (не имеющий возможности принимать от него напрямую сигнал) успешно их принял. Обозначим канал передачи от передатчика до i -ого ($i=1, 2 \dots R$) ретранслятора как f_i , а канал от i -ого ретранслятора до приемника как g_i . Эти каналы являются каналами с рэлеевскими замираниями и неселективными по частоте. Предположим, что f_i и g_i являются независимыми комплексными величинами, а их действительные и мнимые части имеют гауссовское распределение с нулевым средним и единичной дисперсией. Также предположим, что все комплексные коэффициенты g_i точно известны приёмнику, а f_i точно известны i -ому ретранслятору, и в течение времени передачи вектор-сигнала остаются неизменными. В качестве сигналов рассматриваются их комплексные огибающие.

Схема с изображением работы сети показана на рис. 1, из которого следует, что передача данных происходит в две стадии. На первой передатчик одновременно транслирует сигнал (вектор) $\mathbf{s}$ для R ретрансляторов в течение некоторого фиксированного интервала времени. Принятый сигнал в i -ом ретрансляторе будет иметь следующий вид:

$$\mathbf{r}_i = \sqrt{P} f_i \mathbf{s} + \mathbf{v}_i,$$

где P – мощность переданного сигнала, $\mathbf{v}_i$ – вектор принятых шумов, чьи значения являются независимыми комплексными гауссовскими величинами с нулевым средним и дисперсией σ_{SR}^2 . На второй стадии ретрансляторы в течение некоторого другого фиксированного интервала времени передают сигнал приёмнику. Для того чтобы иметь возможность одновременного приёма сигналов от всех ретрансляторов, на второй стадии используется ортогональный пространственно-временной код [6]. Приёмник одновременно принимает сигнал (вектор) $\mathbf{x}$ от всех ретрансляторов:

$$\mathbf{x} = \sum_{i=1}^R g_i \mathbf{t}_i + \mathbf{w},$$

где $\mathbf{t}_i$ – вектор переданного сигнала от i -ого ретранслятора, $\mathbf{w}$ – вектор принятых шумов, чьи значения являются независимыми комплексными гауссовскими величинами. Будем считать, что суммарная мощность переданных ретрансляторами сигналов равна мощности сигнала, переданного передатчиком на первой стадии. Вектор $\mathbf{t}_i$ будет иметь разный вид в зависимости от схемы ретрансляции.

Напомним, что в рассматриваемой системе отсутствует канал передачи между приёмником и передатчиком.

Предположим, что матрица пространственно-временного кода $\mathbf{C}$ имеет размеры $R \times T$, где R – количество ретрансляторов, а T – число тактовых интервалов кода. Тогда вектор принятых сигналов в приёмнике можно представить в виде:

$$\mathbf{X} = \mathbf{G}\mathbf{C} + \mathbf{W},$$

где $\mathbf{W} = (w_1, w_2, \dots, w_T)^T$ – вектор собственных шумов, которые являются гауссовскими некоррелированными во времени случайными комплексными величинами с нулевым средним и

дисперсией σ_{RD}^2 ; $G = (g_1, g_2, \dots, g_R)$ – вектор, составленный из комплексных коэффициентов каналов; $(\cdot)^T$ – операция транспонирования. Преобразование сигналов в декодере пространственно-временного кода может быть описано матрицей U с размерами $T \times K$, где K – количество информационных символов в кодовом блоке. В результате вектор выходного сигнала декодера имеет вид:

$$S = (Y)^T U.$$

Данный сигнал представляет собой мягкую оценку. Жёсткие оценки вычисляются методом максимального правдоподобия: из множества возможных значений сигнала (созвездия модуляции) R выбирается тот, который имеет минимальное евклидово расстояние между ним и принятой мягкой оценкой:

$$\hat{S} = \arg \min d(R, S).$$

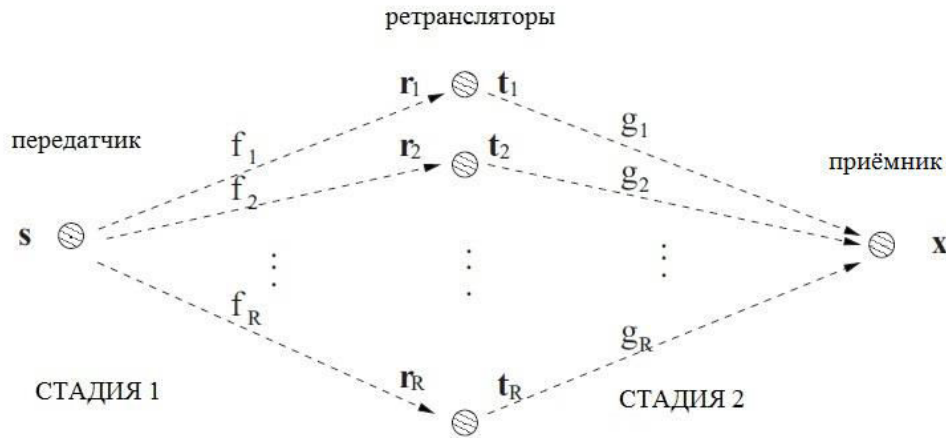


Рис. 1. Распределённая беспроводная сеть с ретрансляторами

3.1. Decode-and-Forward

В данной схеме ретранслятор обрабатывает принятый от передатчика сигнал по следующему алгоритму. Сначала вычисляется решающая статистика, которая выглядит следующим образом: принятый сигнал $\mathbf{r}_i$ умножается на комплексно сопряжённый канальный коэффициент f_i^* , где $*$ – знак комплексного сопряжения.

$$\hat{\mathbf{t}}_i = \mathbf{r}_i \cdot f_i^*.$$

Жёсткие оценки вычисляются методом максимального правдоподобия: из множества возможных значений сигнала (созвездия модуляции) S выбирается то, которое имеет минимальное евклидово расстояние между этим сигналом и принятой мягкой оценкой:

$$\mathbf{t}_i = \arg \min d(\hat{\mathbf{t}}_i, S).$$

Для простоты описания рассмотрим сначала случай с двумя ретрансляторами, в качестве созвездия модуляции используется BPSK. Формула для битовой ошибки в приёмнике будет выглядеть следующим образом:

$$P_r = \varepsilon(1 - \varepsilon)(1 - P_{ou})P(|g_1| > |g_2|) + \varepsilon(1 - \varepsilon)(1 - P_{ou})P(|g_2| > |g_1|) + \varepsilon^2(1 - P_{ou}) + (1 - \varepsilon)^2 P_{ou}, \quad (1)$$

где ε – вероятность ошибки в i -ом ретрансляторе на первой стадии ($i = 1, 2$); P_{ou} – вероятность ошибки в приёмнике на второй стадии; $P(|g_2| > |g_1|) = P(|g_1| > |g_2|) = 0.5$ – вероятность того,

что модуль канального коэффициента ретранслятора с ошибкой превысит модуль канального коэффициента ретранслятора без ошибки. Согласно [6] формулы для вероятностей битовой ошибки на второй и первой стадиях соответственно будут выглядеть следующим образом:

$$P_{out} = \frac{1}{2} \left[1 - \sqrt{\frac{q_{RD}^2}{2 + q_{RD}^2}} \left(1 + \frac{1}{2 + q_{RD}^2} \right) \right],$$

$$\varepsilon = \frac{1}{2} \left[1 - \sqrt{\frac{q_{SR}^2}{1 + q_{SR}^2}} \right],$$

где $q_{RD}^2 = \frac{P}{\sigma_{RD}^2}$ – отношение сигнал/шум по мощности в канале ретрансляторы – приёмник;

$q_{SR}^2 = \frac{P}{\sigma_{SR}^2}$ – отношение сигнал/шум по мощности в канале передатчик – ретранслятор.

3.2. Estimate-and-Forward

При наличии канала передатчик – ретранслятор с низким отношением сигнал/шум в ретрансляторе при использовании схемы с жесткими оценками (DF) будут возникать ошибки, которые невозможно будет исправить на стороне приёмника даже с учетом наличия канала ретранслятор – приёмник с высоким ОСШ. Чтобы этого не происходило, ретрансляторы должны передавать мягкие оценки. В качестве мягкой оценки используется вероятность правильного детектирования [7].

В данной схеме ретранслятор обладает информацией об отношении сигнал/шум в канале передатчик – ретранслятор и обрабатывает принятый от передатчика сигнал по следующему алгоритму. Сначала вычисляется решающая статистика, которая выглядит следующим образом: принятый сигнал $\mathbf{r}_i$ умножается на комплексно-сопряженный канальный коэффициент f_i^* .

$$\hat{\mathbf{t}}_i = \mathbf{r}_i \cdot f_i^*.$$

Затем вычисляется мягкая оценка:

$$\mathbf{t}_i = \text{sign}(\hat{\mathbf{t}}_i) \cdot \alpha \cdot \left[1 - 0.5 \cdot \exp \left(-\frac{q_{SR}^2 \cdot (\hat{\mathbf{t}}_i)^2}{2} \right) \right],$$

где α – коэффициент, выравнивающий дисперсии выходных сигналов ретрансляторов в схемах DF и EF.

4. Результаты

Экспериментальные данные были получены с помощью моделирования в программе MATLAB. Исследования проводились на основе математической модели из раздела 3 в основной полосе частот, а в качестве созвездия модуляции использовалась BPSK.

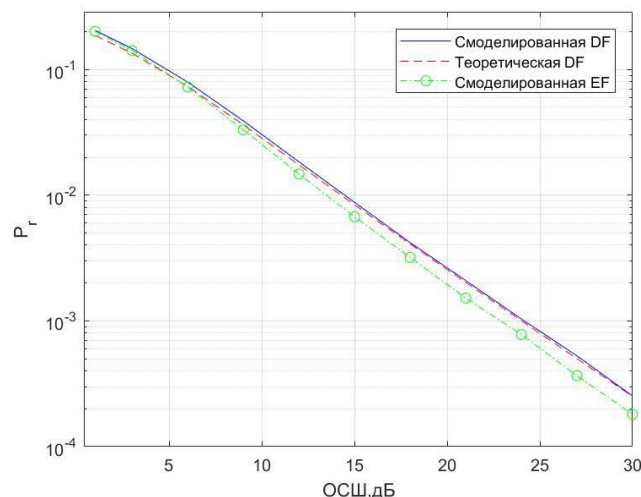


Рис. 2. Сравнение схем DF и EF для распределённой сети с двумя ретрансляторами при эквивалентных ОСШ в каналах передатчик – ретранслятор и ретранслятор – приёмник

На рис. 2 изображены теоретическая зависимость для схемы DF, вычисленная по формуле (1), и смоделированные в среде MATLAB зависимости для схем DF и EF вероятности битовой ошибки от отношения сигнал/шум в каналах передатчик – ретранслятор и ретрансляторы – приёмник. Эти каналы считаются эквивалентными. В данном эксперименте исследовалась распределённая беспроводная сеть с двумя ретрансляторами. Из данного рисунка видно, что теоретическая и смоделированная зависимости для схемы DF практически совпадают. Также можно увидеть, что схема EF даёт выигрыш приблизительно в 1.7 дБ по сравнению с DF.

На рис. 3–5 изображены результаты моделирования в среде MATLAB для беспроводной распределённой сети с четырьмя ретрансляторами. На рис. 3 приведены зависимости вероятности битовой ошибки от отношения сигнал/шум в канале передатчик – ретранслятор, в то время как отношение сигнал/шум в канале ретрансляторы – приёмник зафиксировано на уровне 10 дБ. Выигрыш от использования схемы EF составляет в этом случае примерно 2.5 дБ.

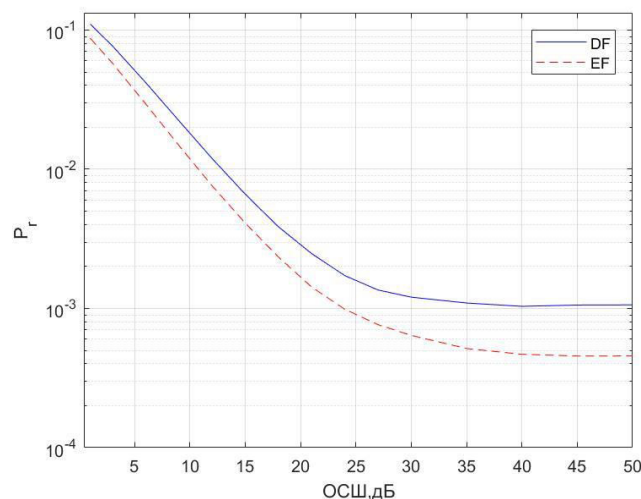


Рис. 3. Сравнение EF и DF при фиксированном ОСШ (10 дБ) в канале ретрансляторы – приёмник для распределённой сети с четырьмя ретрансляторами

На рис. 4 приведены зависимости вероятности битовой ошибки от отношения сигнал/шум в канале ретрансляторы – приёмник, в то время как отношение сигнал/шум в канале передатчик – ретранслятор зафиксировано на уровне 10 дБ. Выигрыш от использования схемы EF составляет в этом случае приблизительно 1.5 дБ.

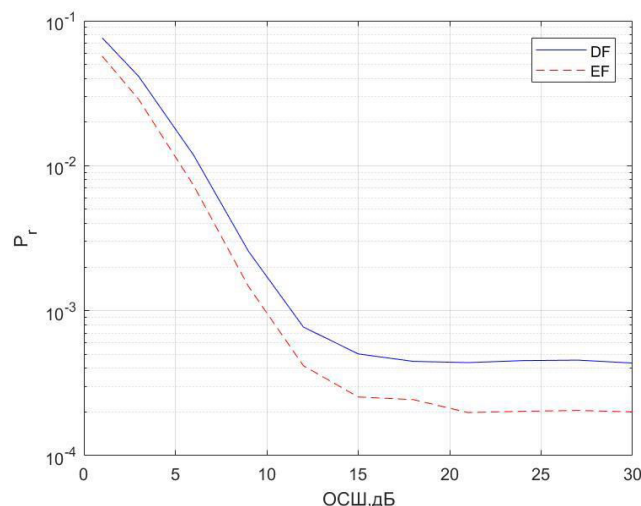


Рис. 4. Сравнение EF и DF при фиксированном ОСШ (25 дБ) в канале передатчик – ретранслятор для распределённой сети с четырьмя ретрансляторами

На рис. 5 изображены графики зависимостей вероятности битовой ошибки от отношения сигнал/шум в каналах передатчик – ретранслятор и ретрансляторы – приёмник, которые считаются эквивалентными (имеют одинаковое среднее ОСШ). Выигрыш от использования схемы EF составляет в этом случае приблизительно 3 дБ.

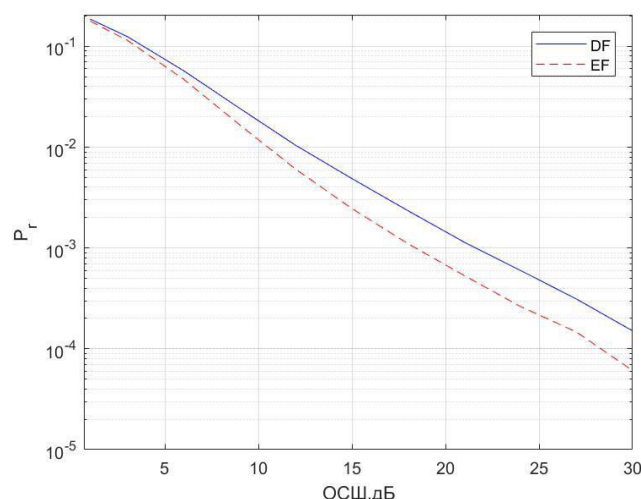


Рис. 5. Сравнение EF и DF при эквивалентных ОСШ в каналах передатчик – ретранслятор и ретрансляторы – приёмник для распределённой сети с четырьмя ретрансляторами

Литература

1. *Al-Mistarihi M. F., Mohaisen R., Sharaqa A., Shurman M. M., and Darabkh K. A.* Performance evaluation of multiuser diversity in multiuser two-hop cooperative multi-relay wireless networks using maximal ratio combining over Rayleigh fading channels // *International Journal of Communication Systems*, September 2013.
2. *Hu J., and Chen X.* Symbol Error Probability and Channel Capacity for Multiuser Diversity of Amplify-and-Forward Cooperative Networks // *Proc. of 7th International Conference on Wireless Communications, Networking and Mobile Computing (WiCOM-2011)*, September 2011, Wuhan, China. P. 1–4.

3. *Rui X., Hou J., and Zhou L.* Decode-and-forward with full-duplex relaying // *International Journal of Communication Systems*. 2012. V. 25, Is. 2. P. 270–275.
4. *Ting Z., Fang W., Jing X., Lilleberg J.* Soft Symbol Estimation and Forward Scheme for Cooperative Relaying // *IEEE 20th International Symposium on Personal, Indoor and Mobile Radio Communications*, Tokyo, Japan, September 13–16, 2009.
5. *Sneessens H. H., Vandendorpe L.* Soft Decode and Forward Improves Cooperative Communications // *2005 1st IEEE International Workshop on Computational Advances in Multi-Sensor Adaptive Processing*. V. 13–15. P. 157–160.
6. *Jafarkhani H.* Space-Time Coding: Theory and Practice. Cambridge University Press, 2005. 302 p.
7. *Ипатов В. П.* Широкополосные системы и кодовое разделение сигналов: принципы и приложения / В. П. Ипатов; пер. с англ. под ред. авт. М.: Техносфера, 2007. 487 с.

*Статья поступила в редакцию 19.04.2019;
переработанный вариант – 13.06.2019.*

Быченков Даниил Евгеньевич

аспирант кафедры теоретических основ радиотехники НГТУ (630073, Новосибирск, пр. К. Маркса, 20), e-mail: omg250@mail.ru.

Область научных интересов: телекоммуникационные системы.

**Comparison of the distributed wireless communication systems characteristics
with soft and hard decisions**

D. Bychenkov

The efficiency of using soft assessment for transmission over a distributed wireless sensor network is considered. According to the simulation results, it was determined that the proposed transmission method has a tangible gain in terms of the probability of bit error, in contrast to the traditional use of hard assessment for transmission over a wireless channel with Rayleigh fading.

Keywords: distributed wireless sensor network, soft assessment, multipath propagation.

Масштабирование рекурсивных цифровых фильтров с нерекурсивными латеральными связями

В. П. Бакалов

В статье излагаются способы масштабирования цифровых фильтров (ЦФ) с латеральными обратными связями (ЛОС). Рассматриваются модели четырех типов рекурсивных ЦФ с ЛОС при их нерекурсивной реализации. Масштабирование фильтров осуществляется с использованием импульсной характеристики h_i и АЧХ цепи. Полученные результаты позволяют, с одной стороны, предотвратить переполнение разрядной сетки ЦФ с ЛОС, а с другой – существенно повысить его помехоустойчивость и эффективность.

Ключевые слова: рекурсивные цифровые фильтры, нерекурсивные латеральные обратные связи, масштабирование, ограничение максимума сигнала, ограничение энергии сигнала, ограничение максимума АЧХ цепи.

1. Введение

Как известно, при цифровой обработке сигналов сложение чисел с фиксированной запятой может приводить в цифровом фильтре (ЦФ) к такому опасному явлению, как переполнение [1, 2]. Чтобы в ЦФ не возникало переполнения, применяют в определенных точках фильтра процедуру масштабирования множителей, которая позволяет, с одной стороны, предотвратить переполнение, а с другой – сохранить максимально возможное отношение сигнала к уровню шума округления в ЦФ.

Особенно важна проблема масштабирования в ЦФ с латеральными обратными связями (ЛОС), использование которых позволяет существенно повысить помехоустойчивость и эффективность цифровых фильтров [4–6].

В данной работе рассматриваются вопросы расчета масштабных множителей для четырех типов рекурсивных ЦФ с нерекурсивными ЛОС по условиям ограничения максимума сигнала, максимума энергии сигнала и максимума АЧХ цепи.

2. Цифровой фильтр с ЛОС 1-го типа

Рассмотрим модель двухканального одноуровневого ЦФ с ЛОС 1-го типа [3], представленную на рис. 1.

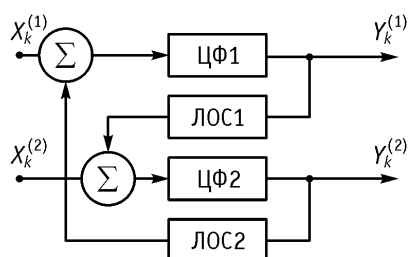


Рис. 1. Двухканальная система ЦФ с ЛОС. Тип № 1

На базе этой схемы могут быть реализованы различные типы ЦФ с ЛОС: рекурсивные, нерекурсивные и комбинированные.

Будем считать, что ЦФ реализуется в форме рекурсивных, а ЛОС реализуются в виде нерекурсивных цепей n -го порядка. Положим, что шумы квантования и округления на выходе каждого из умножителей и на выходе аналого-цифровых преобразователей (АЦП) некоррелированы между собой и некоррелированы с входными последовательностями каждого канала. Причем источники шума $e_1 \dots e_n$ имеют одинаковую дисперсию и определяются уравнениями [7]:

$$\sigma_k^2 = \sigma^2(kT) = \Delta^2/12, \quad (1)$$

а дисперсия шума на выходе ЦФ уравнением:

$$\sigma^2 = \frac{\Delta^2}{12} \sum_{k=1}^n h_k^2, \quad (2)$$

где Δ определяется числом используемых разрядов; h_k – импульсная характеристика ЦФ.

С учетом вышеизложенного на рис. 2 изображена шумовая модель ЦФ с ЛОС этого типа, где e_0 – шумовой сигнал, учитывающий шум квантования входного сигнала; e_1, e_n – шумовой сигнал, учитывающий квантование сигналов в умножителях.

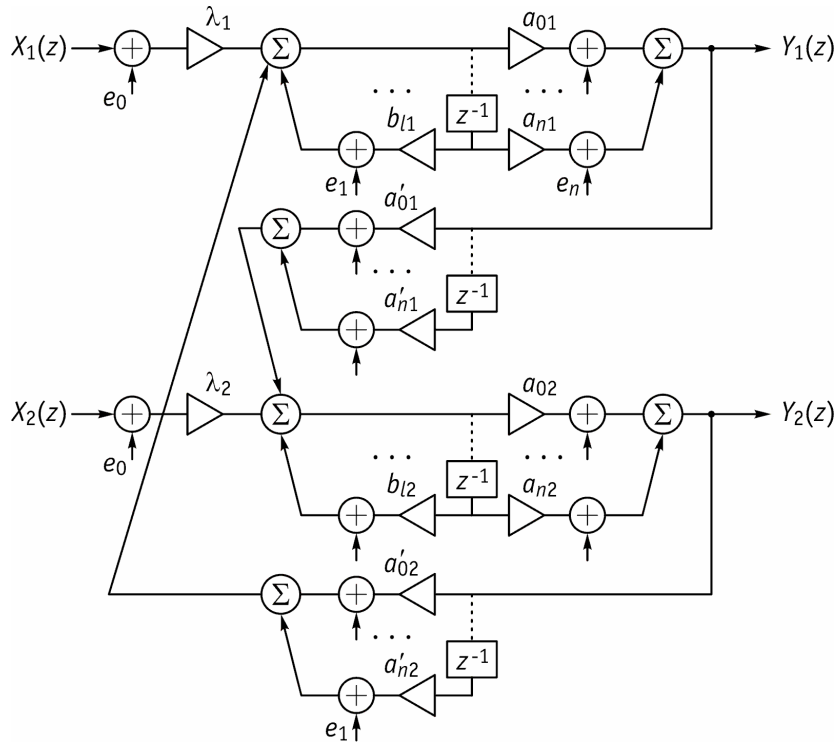


Рис. 2. Модель ЦФ с НЛОС 1-го типа

Передаточная функция ЦФ1 определяется уравнением:

$$H_1(z) = \sum_{n=0}^{N_1} a_{n1} z^{-n} / \left(1 - \sum_{l=1}^{L_1} b_{l1} z^{-l} \right), \quad (3)$$

где N_1, L_1 – количество умножителей прямой и обратной связи ЦФ1 соответственно.

Аналогично определяется передаточная функция ЦФ2:

$$H_2(z) = \sum_{n=0}^{N_2} a_{n2} z^{-n} / \left(1 - \sum_{l=1}^{L_2} b_{l2} z^{-l} \right), \quad (4)$$

где N_2, L_2 – количество умножителей прямой и обратной связи ЦФ2 соответственно.

Передаточная функция нерекурсивного фильтра НЛОС1 определяется уравнением:

$$H_{\text{ЛОС1}}(z) = \sum_{n=0}^{N'_1} a'_{n1} z^{-n}, \quad (5)$$

где N'_1 – количество умножителей нерекурсивного фильтра ЛОС1.

Передаточная функция нерекурсивного фильтра НЛОС2 определяется уравнением:

$$H_{\text{ЛОС2}}(z) = \sum_{n=0}^{N'_2} a'_{n2} z^{-n}, \quad (6)$$

где N'_2 – количество умножителей нерекурсивного фильтра ЛОС2; a_{n1}, a_{n2} – коэффициенты, характеризующие веса нерекурсивной части ЦФ1 и ЦФ2 соответственно; b_{l1}, b_{l2} – коэффициенты, характеризующие веса рекурсивной части ЦФ1 и ЦФ2 соответственно; a'_{n1}, a'_{n2} – коэффициенты, характеризующие веса нерекурсивной части ЛОС1 и ЛОС2 соответственно.

Определим масштабные коэффициенты данного фильтра с использованием импульсной характеристики h_i и максимума АЧХ цепи.

Положим, что импульсная характеристика ЦФ с ЛОС (или его части) равна $h_i(n)$, тогда выходной сигнал фильтра (или его части) $y_i(n)$ будет ограничен величиной

$$\max |y_i(n)| \leq \max |x(n)| \sum_{n=0}^{\infty} |h_i(n)|. \quad (7)$$

Если принять, что $\max |x(n)| \leq 1$, то необходимым условием отсутствия переполнения является:

$$\sum_{n=0}^{\infty} |h_i(n)| \leq 1. \quad (8)$$

Для обеспечения данного условия введем масштабные коэффициенты λ_i :

$$\lambda_i \sum_{n=0}^{\infty} |h_i(n)| \leq 1. \quad (9)$$

Откуда находим λ_i по условию ограничения максимума сигнала:

$$\lambda \leq \frac{1}{\sum_{n=0}^{\infty} |h_i(n)|}, \quad (10)$$

где $h_i(n)$ – импульсная характеристика участка ЦФ от входа до выхода i -го сумматора.

Для определения масштабного множителя по условию (10) найдем импульсные характеристики 1-го канала для данного типа ЦФ с ЛОС.

Для первого сумматора передаточная функция соответствует рекурсивной части ЦФ1:

$$H'_1(z) = \frac{1}{1 - \sum_{l=1}^{L_1} b_{l1} z^{-l}} \quad (11)$$

и передаточной функции:

$$H''_2(z) = H_2(z) H_{\text{ЛОС2}}(z), \quad (12)$$

где $H_2(z)$ и $H_{\text{ЛОС2}}(z)$ определяются формулами (4) и (6) соответственно.

Для второго сумматора импульсная характеристика соответствует передаточной функции всего фильтра $H_1(z)$, определяемой (3).

После нахождения передаточных функций $H'_1(z)$, $H''_2(z)$ и $H_1(z)$ импульсные характеристики определяются с помощью обратного z -преобразования. В частности, для первого сумматора импульсная характеристика будет равна

$$h_{1\Sigma}(n) = \frac{1}{2\pi j} \left(\oint_{|z|=1} H'_1(z) z^{n-1} dz + \oint_{|z|=1} H''_2(z) z^{n-1} dz \right), \quad (13)$$

а для второго сумматора

$$h_{2\Sigma}(n) = \frac{1}{2\pi j} \oint_{|z|=1} H_1(z) z^{n-1} dz, \quad (14)$$

где $H_1(z)$, $H'_1(z)$, $H''_2(z)$ определяются формулами (2), (11) и (12) соответственно.

После нахождения импульсных характеристик $h_i(n)$ масштабные коэффициенты для первого $\lambda_{1\Sigma}$ и второго $\lambda_{2\Sigma}$ сумматоров определяются согласно (10) и из них выбирается наименьший: $\lambda = \min \lambda_{\Sigma}$.

Аналогичным образом можно использовать импульсные характеристики h_i при масштабировании ЦФ с ЛОС при ограничении энергии сигнала:

$$\lambda \leq \frac{1}{\sqrt{\sum_{n=0}^{\infty} h_i^2(n)}}, \quad (15)$$

где $h_i(n)$ определяется согласно (13), (14).

При расчете масштабных множителей по условию ограничения усиления цепи необходимо задаться неравенством

$$\lambda = \frac{1}{\max H(\omega)}, \quad (16)$$

где $H(\omega) = |He^{j\omega T}|$.

При этом максимальное значение АЧХ цепи $H_{\max}(\omega)$ определяется на частоте, равной собственной частоте нуля или полюса передаточной функции $H(z)$ [7].

Аналогичным образом определяются масштабные множители для 2-го канала.

Учитывая, что при масштабировании обычно уменьшается отношение сигнал/шум на выходе ЦФ, определим на основании шумовой модели (рис. 2) и уравнений (1), (2) дисперсию шума на выходе ЦФ с ЛОС 1-го типа:

$$\begin{aligned} \sigma_{\text{вых}}^2 = & \frac{\lambda_1^2 (2^{-b_{\text{АЦП}}})^2}{12} \left[\sum_{n=0}^{\infty} h_1^2(n) + \sum_{n=0}^{\infty} h_3^2(n) \right] + \\ & + \frac{(2^{-b_{\text{умн}}})^2}{12} \left[(L_1 + N'_2) \sum_{n=0}^{\infty} h_1^2(n) + L_1 \sum_{n=0}^{\infty} h_3^2(n) + N'_1 \sum_{n=0}^{\infty} h_4^2(n) + N_2 \sum_{n=0}^{\infty} h_6^2(n) + N_1 \right], \end{aligned} \quad (17)$$

где $h_1(n) \div H_1(z)$; $h_3(n) \div H_1(z) \cdot H_{\text{ЛОС1}}(z) \cdot H_2(z)$; $h_4(n) \div H_2(z) \cdot H_{\text{ЛОС2}}(z) \cdot H_1(z)$; $h_6(n) \div H_{\text{ЛОС2}}(z) \cdot H_1(z)$.

Из формулы (17) следует, что дисперсия шума на выходе ЦФ с НЛОС будет определяться суммой дисперсий отдельных составляющих, причем наибольший вклад в дисперсию шума на выходе, как и в ЦФ с рекурсивным ЛОС, вносят умножители. При этом дисперсия шума на выходе от АЦП одинакова, а от умножителей – меньше, чем в ЦФ с рекурсивной ЛОС, и больше, чем в нерекурсивных ЦФ с НЛОС [8].

3. Цифровой фильтр с ЛОС 2-го типа

Рассмотрим масштабирование ЦФ с ЛОС 2-го типа (рис. 3).

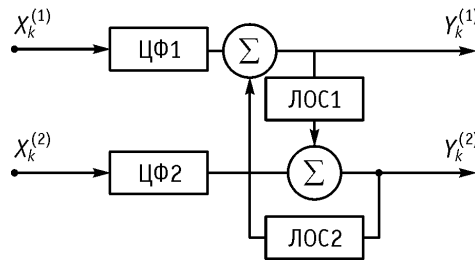


Рис. 3. Двухканальная система с ЛОС. Тип № 2

На рис. 4 представлена шумовая модель этого фильтра при рекурсивной реализации ЦФ и нерекурсивной ЛОС.

Для этого фильтра передаточная функция первого сумматора определяется формулой (11), а второго сумматора – передаточной функцией ЦФ1 и передаточной функцией $H_2''(z)$, определяемой формулой (12).

При этом импульсная характеристика первого сумматора находится согласно (11) как

$$h_{1\Sigma}(n) \div H_1'(z),$$

а второго

$$h_{2\Sigma}(n) = (h_1(n) + h_2''(n)) \div (H_1(z) + H_2''(z)), \quad (18)$$

где $H_1(z)$ определяется (3), а $H_2''(z)$ – формулой (12).

После нахождения $h_{1\Sigma}(n)$ и $h_{2\Sigma}(n)$ по формуле (10) находим масштабные коэффициенты $\lambda_{1\Sigma}$ и $\lambda_{2\Sigma}$ по условию максимума сигнала и выбираем из них наименьший.

Аналогичным образом определяются масштабные множители по условиям (15) и (16).

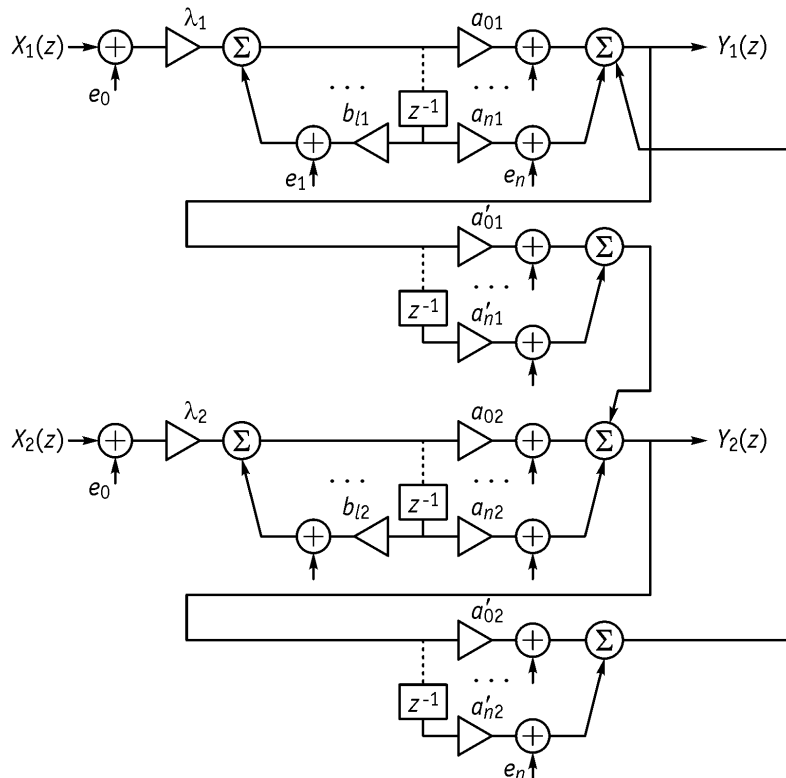


Рис. 4. Модель ЦФ с НЛОС 2-го типа

Дисперсия шума на выходе ЦФ с НЛОС 2-го типа после масштабирования определяется уравнением:

$$\sigma_{\text{вых}}^2 = \frac{\lambda_1^2 (2^{-b_{\text{АЦП2}}})^2}{12} \left[\sum_{n=0}^{\infty} h_2^2(n) + \sum_{n=0}^{\infty} h_6^2(n) \right] + \frac{\lambda_2^2 (2^{-b_{\text{АЦП2}}})^2}{12} \sum_{n=0}^{\infty} h_4^2(n) +$$

$$+ \frac{(2^{-b_{\text{УМН}}})^2}{12} \left[L_1 \sum_{n=0}^{\infty} h_1^2(n) + L_2 \sum_{n=0}^{\infty} h_4^2(n) + L_1 \sum_{n=0}^{\infty} h_5^2(n) + \right.$$

$$\left. + N_1 \sum_{n=0}^{\infty} h_7^2(n) + (N_2 + N'_1) \sum_{n=0}^{\infty} h_9^2(n) + (N_1 + N'_2) \cdot 1 \right], \quad (19)$$

где $h_1(n) \div H_1(z)$; $h_5(n) \div H_{\text{ЛОС1}}(z) \cdot H_{\text{ЛОС2}}(z) \cdot H_1(z)$; $h_4(n) \div H_2(z) \cdot H_{\text{ЛОС2}}(z)$; $h_7(n) \div H_{\text{ЛОС1}}(z) \cdot H_{\text{ЛОС2}}(z)$; $h_9(n) \div H_{\text{ЛОС2}}(z)$.

Анализ уравнения (19) показывает, что дисперсия шума на выходе ЦФ с НЛОС увеличивается по сравнению с 1-й схемой за счет увеличения дисперсии от шумов АЦП и суммарной дисперсии шума умножителей. По сравнению с нерекурсивной реализацией ЦФ с НЛОС дисперсия шума на выходе оказывается выше, а по сравнению с рекурсивной реализацией ЦФ и ЛОС – существенно ниже за счет уменьшения числа умножителей цепи НЛОС.

4. Цифровой фильтр с ЛОС 3-го типа

Рассмотрим ЦФ с ЛОС 3-го типа (рис. 5), шумовая модель которого изображена на рис. 6.

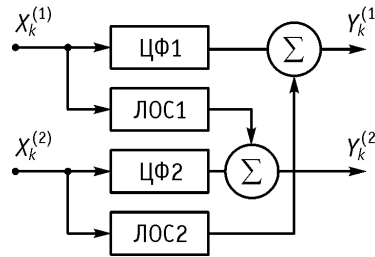


Рис. 5. Двухканальная система с ЛОС. Тип № 3

По аналогии с предыдущими схемами импульсная характеристика первого сумматора определяется рекурсивной частью $H'_1(z)$ (11), а второго – передаточной функцией $H_1(z)$ и передаточной функцией ЛОС₂ по формулам (3) и (6) соответственно.

После нахождения $h_{1\Sigma}(n) \div H'_1(z)$ и $h_{2\Sigma}(n) \div (H_1(z) + H_{\text{ЛОС2}}(z))$ множители λ определяются по формулам (10), (15), (16) в зависимости от условий масштабирования.

Дисперсия шума на выходе ЦФ с НЛОС 3-го типа после масштабирования определяется уравнением

$$\sigma_{\text{вых}}^2 = \frac{\lambda_1^2 (2^{-b_{\text{АЦП1}}})^2}{12} \sum_{n=0}^{\infty} h_1^2(n) + \frac{\lambda_2^2 (2^{-b_{\text{АЦП2}}})^2}{12} \sum_{n=0}^{\infty} h_4^2(n) +$$

$$+ \frac{(2^{-b_{\text{УМН}}})^2}{12} \left[L_1 \sum_{n=0}^{\infty} h_1^2(n) + L_2 \sum_{n=0}^{\infty} h_4^2(n) + (N_1 + N'_2) \cdot 1 \right], \quad (20)$$

где $h_1(n) \div H_1(z)$; $h_4(n) \div H_{\text{ЛОС2}}(z)$.

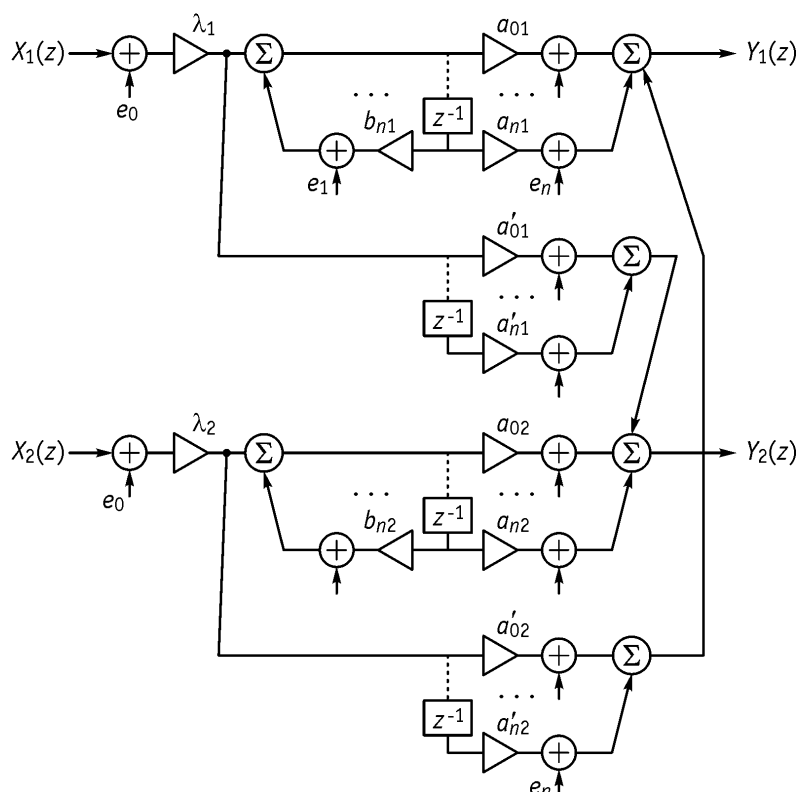


Рис. 6. Модель ЦФ с НЛОС 3-го типа

Анализ уравнения (20) показывает, что шумы АЦП в схемах 1 и 3 при одинаковых разрядных сетках примерно одинаковы и меньше, чем во 2-й схеме, а шумы умножителей в 3-й схеме существенно меньше, чем в предыдущих схемах. Это происходит за счет уменьшения путей прохождения шума, что приводит к существенному уменьшению суммарной дисперсии шума на выходе ЦФ с НЛОС 3-го типа.

5. Цифровой фильтр с ЛОС 4-го типа

В заключение рассмотрим схему ЦФ с ЛОС 4-го типа (рис. 7), шумовая модель которого изображена на рис. 8.

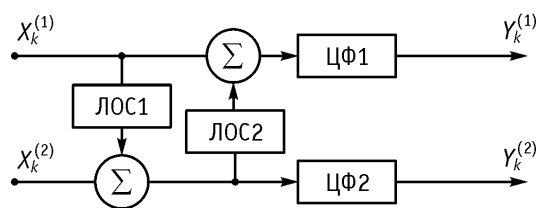


Рис. 7. Двухканальная система с ЛОС. Тип № 4

Импульсная характеристика первого сумматора определяется рекурсивной частью передаточной функции $H_1'(z)$ и передаточной функцией ЛОС $H_{\text{ЛОС2}}(z)$ по формулам (11) и (6) соответственно.

Импульсная характеристика второго сумматора определяется передаточной функцией $H_1(z)$ (3).

После нахождения импульсных характеристик для 1-го и 2-го сумматоров определяются соответствующие масштабные множители согласно (10), (15), (16).

Дисперсия шума ЦФ с НЛОС этого типа после масштабирования определяется уравнением:

$$\begin{aligned}
\sigma_{\text{ВЫХ}}^2 = & \frac{\lambda_1^2 (2^{-b_{\text{АЦП1}}})^2}{12} \left[\sum_{n=0}^{\infty} h_1^2(n) + \sum_{n=0}^{\infty} h_5^2(n) \right] + \frac{\lambda_2^2 (2^{-b_{\text{АЦП2}}})^2}{12} \sum_{n=0}^{\infty} h_4^2(n) + \\
& + \frac{(2^{-b_{\text{УМН}}})^2}{12} \left[(L_1 + N'_2) \sum_{n=0}^{\infty} h_1^2(n) + (L_2 + N'_1) \sum_{n=0}^{\infty} h_4^2(n) + \right. \\
& \left. + (L_1 + N'_2) \sum_{n=0}^{\infty} h_5^2(n) + N_1 \cdot 1 \right], \quad (21)
\end{aligned}$$

где $h_1(n) \div H_1(z)$; $h_4(n) \div H_1(z) \cdot H_{\text{ЛОС2}}(z)$; $h_5(n) \div H_1(z) \cdot H_{\text{ЛОС2}}(z) \cdot H_1(z)$.

Из уравнения (21) следует, что в данной схеме дисперсия шума АЦП при прочих равных условиях больше, чем дисперсия шума АЦП в схеме № 1 и схеме № 3, и одинакова с дисперсией шума АЦП в схеме № 2. Дисперсия шума от умножителей в схеме № 4 больше, чем дисперсия шума в схеме № 3 и схеме № 1, и меньше дисперсии шума в схеме № 2. Таким образом, наилучшей с точки зрения минимума шумов является схема № 3, а наихудшей – схема № 2.

В заключение следует сказать, что процесс масштабирования в цифровых фильтрах, как было отмечено выше, обычно приводит к существенному уменьшению отношения сигнал/шум на выходе фильтра по сравнению с отношением сигнал/шум на входе, что, в свою очередь, ухудшает помехоустойчивость ЦФ в целом [1]. Использование же принципа ЛОС, как показано в [3], позволяет существенно повысить помехоустойчивость ЦФ даже при отношении сигнал/шум на входе меньше единицы. Таким образом, при масштабировании ЦФ с ЛОС не происходит уменьшения отношения сигнал/шум на выходе фильтра, что существенно повышает его помехоустойчивость по сравнению с традиционными методами масштабирования ЦФ.

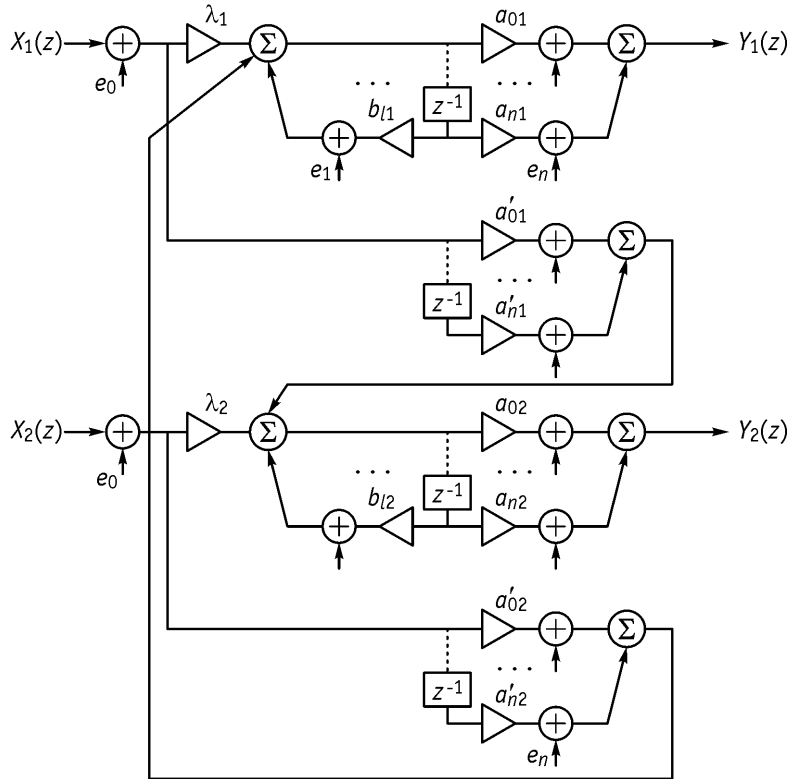


Рис. 8. Модель ЦФ с ЛОС 4-го типа

Если сравнивать по помехоустойчивости масштабируемые ЦФ с ЛОС, то лучший результат дает схема 3, затем схемы 2 и 4, а наименьший выигрыш дает схема 1.

Полученные результаты по масштабированию ЦФ с ЛОС можно обобщить и на n -канальную систему с ЛОС. Причем в зависимости от вида помех могут изменяться характер ЛОС, структура охвата каналов латеральными связями (каждый канал с каждым, по группам мешающих параметров и т.д.), подбор глубины обратной связи.

6. Заключение

Исследованы вопросы масштабирования рекурсивных одноуровневых цифровых фильтров с нерекурсивными латеральными обратными связями с использованием импульсной характеристики и АЧХ цепи.

Определены масштабные коэффициенты по условиям ограничения максимума сигнала, энергии сигнала и максимума АЧХ для 4-х типов ЦФ с ЛОС.

Показано, что при масштабировании ЦФ с НЛОС 1-го и 2-го типов определяющую роль играют передаточные функции ЦФ1, ЦФ2, $H_{\text{ЛОС}2}(z)$ и рекурсивные части передаточной функции ЦФ1. При масштабировании ЦФ с НЛОС 3-го и 4-го типов определяющую роль играют передаточные функции ЦФ1, $H_{\text{ЛОС}2}(z)$ и рекурсивная часть ЦФ1 $H_1'(z)$.

Получены уравнения для дисперсии шума ЦФ с НЛОС после масштабирования для 4-х типов фильтров. Показано, что с точки зрения помехоустойчивости наилучшей является типовая схема 3, а наихудшей – типовая схема 1.

Таким образом, полученные результаты по масштабированию рекурсивных ЦФ с НЛОС позволяют, с одной стороны, предотвратить переполнение разрядной сетки, а с другой – обеспечить максимально возможное отношение сигнала к шуму округления.

Приведенные в статье данные по масштабированию ЦФ с ЛОС с использованием импульсных и передаточных функций для типовых одноуровневых двухканальных ЦФ с ЛОС можно обобщить и на n -канальную многоуровневую систему с различной структурой охвата каналов ЛОС при их рекурсивной, нерекурсивной и комбинированной реализации.

Литература

1. Гольденберг Л. М., Матюшкин Б. Д., Поляк М. Н. Цифровая обработка сигналов. М.: Радио и связь, 1990. 256 с.
2. Рабинер Л., Голд Б. Теория и применение цифровой обработки сигналов. М.: Мир, 1978. 848 с.
3. Бакалов В. П., Субботин Е. А. Обработка сигналов в системах с латеральными связями. М.: Горячая линия – Телеком, 2017. 178 с.
4. Бакалов В. П., Черных Ю. С. Использование принципа латеральных связей для обработки дискретных сигналов // Вестник СибГУТИ. 2012. № 1. С. 43–50.
5. Бакалов В. П., Субботин Е. А. Исследование помехоустойчивости дискретных фильтров с латеральными связями // Вестник СибГУТИ. 2015. № 3. С. 89–95.
6. Бакалов В. П., Черных Ю. С. О потенциальной скорости передачи информации в инфокоммуникационных сетях с латеральными связями // Вестник СибГУТИ. 2012. № 2. С. 62–70.
7. Бакалов В. П., Крук Б. И. Теория электрических цепей. М.: Горячая линия – Телеком, 2018. 537 с.
8. Бакалов В. П., Субботин Е. А. Масштабирование цифровых фильтров с латеральными связями // Вестник СибГУТИ. 2018. № 4. С. 12–19.

Бакалов Валерий Пантелеевич

д.т.н., профессор, заведующий кафедрой теории электрических цепей СибГУТИ
(630102, Новосибирск, ул. Кирова, 86), тел. 286-80-35, e-mail: bvp@sibsutis.ru.

Scaling recursive of digital filters with nonrecursive lateral feedbacks**V. P. Bakalov**

The article states ways to scale digital filters (DF) with lateral feedbacks (LFB). Models of four types recursive of DF with LFBs with their nonrecursive implementation are considered. Filters are scaled using hi impulse response and scaling using the amplitude-frequency characteristic of the circuit. The obtained results allow, on the one hand, to prevent overfilling of the grade graticule of the DF with LFBs, and on the other hand, to significantly increase its noise immunity and efficiency.

Keywords: recursive digital filters, nonrecursive lateral feedbacks, scaling, limiting the maximum signal, maximum signal energy, maximum amplitude-frequency characteristic of the circuit.

Полухрупкие цифровые водяные знаки, базирующиеся на спектре Фурье

И. В. Нечта

Для обнаружения несанкционированных изменений исполняемых файлов вредоносными программами применяются цифровые водяные знаки (ЦВЗ). Свойство хрупкости водяного знака обеспечивает его искажение при незначительных изменениях в программе. Наличие и корректность ЦВЗ контролируется конечным пользователем программы, который заинтересован в использовании целостного (немодифицированного) программного обеспечения. Полухрупкие ЦВЗ разрушаются при изменении программы сверх установленного порога. Согласно предлагаемому методу программа рассматривается как контейнер, содержащий цифровой сигнал, который может быть восстановлен по своим отдельным командам (отсчётам). Изменения в программе меняют спектр сигнала. При анализе модифицированной программы обнаруживаются некоторые искажения исходного сигнала – изменения фаз некоторых гармоник. Предложенный подход позволяет выявлять не только факт изменения программы, но и объём добавленных команд.

Ключевые слова: хрупкие цифровые водяные знаки, стеганография, спектр Фурье.

1. Введение

Цифровые водяные знаки (далее – ЦВЗ) представляют собой двоичную последовательность, которая встраивается в объект данных – так называемый контейнер. В качестве контейнера могут выступать файлы различного формата. ЦВЗ имеют широкую область применения и в зависимости от задачи могут использоваться как для подтверждения авторства или идентификации пользователя, так и для доказательства отсутствия изменения в контейнере.

В одних случаях, например, автор фотографии встраивает ЦВЗ, в котором содержатся его паспортные данные. При обнаружении такой фотографии у постороннего лица, которое выдает указанную фотографию за свою собственную, истинный автор в ходе судебного разбирательства может извлечь и продемонстрировать водяной знак. В других случаях необходимо выявить канал утечки лицензионного продукта. Так, в каждую создаваемую копию программы встраивается ЦВЗ, содержащий данные лицензиата. При обнаружении пиратской копии продукта водяной знак извлекается и определяется его принадлежность к лицензиату.

Еще одной областью применения ЦВЗ является контроль отсутствия изменений в цифровом объекте данных. Например, пользователь скачивает браузер с официального сайта. Согласно требованиям к информационной безопасности пользователь проверяет сертификат сайта и цифровую подпись или значение хэш-функции скачанного дистрибутива. Однако после установки программы возможна её вредоносная модификация компьютерными вирусами с целью кражи паролей. В таких случаях в программу встраивается ЦВЗ, который разрушается при модификации программы. Наличие и корректность водяного знака контролируется самой программой и пользователем.

Исходя из указанных областей применения выдвигаются требования к свойствам ЦВЗ. Свойство устойчивости гарантирует, что ЦВЗ не может быть искажен или удален злоумышленником. Данным свойством должен обладать водяной знак, хранящий сведения об авторе

или лицензиате. Свойство хрупкости гарантирует, что водяной знак будет разрушен при малейших изменениях в контейнере. Например, такое свойство необходимо для контроля целостности программы.

В ряде случаев программы могут санкционировано (по желанию автора) изменяться. Например, при написании программы часто возникают ошибки в коде. Для устранения ошибок автор распространяет специальные программы, исправляющие указанные ошибки. Для обозначения существенных по объему исправлений применяются термины *software update* и *service pack*. Для незначительных исправлений (порядка нескольких десятков байт) используется термин *patch*. В этой связи возникает свойство полужхрупкости, которое гарантирует разрушение ЦВЗ, когда доля изменений превышает некоторый заданный порог.

В научной литературе известно множество методов доказательства целостности программ. Ряд методов, например [1–3], базируется на использовании аппаратных средств: Usb-токенах, смарт-картах и аппаратных модулях безопасности. Известны методы, представленные в работе [4], использующие самосканирование выполняемого кода и подсчет его контрольной суммы или хэш-функции. Наибольшей уязвимостью такой защиты обладают места хранения или сравнения подсчитанных хэш-значений с эталонным. В таких случаях применяют методы запутывания программного кода (*obfuscation*) [5–7], что крайне затрудняет анализ программы злоумышленником. В код программы добавляется множество избыточных инструкций, анализ которых может быть произведен только в ручном режиме. Существуют методы самомодификации кода [8, 9], которые направлены на затруднение выявления места, в котором хранится код, хранящий или проверяющий ЦВЗ.

В работах [10, 11] предлагается интеграция проверки ЦВЗ с полезным кодом программы при помощи нейросети. Утверждается, что анализ работы программы эквивалентен извлечению правил работы нейросети, что является NP-трудной задачей. Частичное искажение такого водяного знака сделает программу неработоспособной. Существуют методы, например [12, 13], которые держат под контролем только критически важные участки кода и используют проверку графа переходов между участками кода программы.

Все вышеприведенные методы позволяют обнаружить малое изменение в контролируемом участке кода. Применяемые функции (хэш или контрольная сумма) меняются даже при изменении одной команды программы. Целью данной работы является создание метода контроля целостности программы, который способен срабатывать при заданном пороге изменений. Такой метод может быть использован автором, когда учитывается возможность последующего незначительного изменения программы с помощью патчей.

2. Описание предлагаемого метода

2.1. Извлечение сигнала из программы

В настоящей работе предлагается рассматривать программу, которая является последовательностью инструкций, в качестве формы записи некоторого сигнала. В связи с тем, что малое изменение кода программы может повлечь за собой существенное изменение в её поведении, предполагается использовать трассировщик¹ [14] для получения списка выполненных команд программы. Последовательность команд, полученная трассировщиком, полнее характеризует поведение программы, следовательно, ЦВЗ будет более чувствительным к искажениям.

Согласно предлагаемой идее список команд программы, полученных трассировщиком, есть форма записи значений сигнала в некоторые моменты времени (отсчеты). Считается, что сигнал 2π -периодический на интервале $[-\pi; \pi]$ и является кусочно-непрерывным и кусочно-монотонным. Каждая команда программы соответствует очередному отсчету сигнала. В силу теоремы Дирихле такой сигнал может быть представлен в виде ряда Фурье.

¹ Трассировщик – это специальная программа, которая создаёт список команд, выполненных другой программой.

$$S(t) = \frac{a_0}{2} + \sum_{n=0}^{\infty} A_n \sin(nx + \varphi), \quad (1)$$

где $a_0 = \frac{1}{\pi} \int_{-\pi}^{\pi} f(x) dx$, A_n – амплитуда гармоники, φ – фазовый сдвиг гармоники.

Разложив сигнал программы в ряд Фурье, мы можем сравнить его с эталонным сигналом, не имеющим искажений. При изменении программы спектр также изменится, что будет обнаружено при сравнении. Спектр сигнала меняется пропорционально внесённым искажениям. Чем больше новых команд в программе, тем сильнее отличается спектр сигнала от эталона. Если изменения окажутся больше, чем предельно установленный порог, то водяной знак считается разрушенным и делается вывод о факте несанкционированного искажения программы.

Работа с предлагаемым водяным знаком строится следующим образом. После создания программы автор при помощи трассировщика получает список команд, выполненных в определенный момент времени работы программы. На основе команд получают сигнал и выполняют дискретное преобразование Фурье, результаты² которого затем хранятся либо отдельно, либо внутри файлов программы. На этапе эксплуатации по инициативе пользователя или в рамках периодического самоанализа программы производится аналогичное получение и разложение сигнала в ряд Фурье. По результатам сравнения мнимых частей спектра сигнала с эталонными значениями, сохраненными ранее, пользователю выводится уведомление о состоянии водяного знака (повреждён или не повреждён).

Для получения ряда Фурье необходимо определить функцию $f(x)$, задающую сигнал контейнера (программы). В нашем случае $f(x)$: $x \in N$, $f(x) \in R$, где x – шаг работы программы, в котором была выполнена одна команда³. В рамках настоящей работы рассматривается два варианта функции $f(x)$. В связи с тем, что существует множество различных команд, которые могут влиять на сигнал взаимно компенсирующим образом, предлагается выбрать одну ключевую команду Cmd_{key} , которая будет влиять на сигнал. Остальные команды не будут учитываться.

Рассмотрим первый вариант функции $f_1(x)$, который вычисляется по алгоритму F1.

Алгоритм F1.

вход (x, cmd_x, Cmd_{key}) // x – номер команды в списке, полученном трассировщиком.
// cmd_x – выполненная на шаге x команда.

выход ($Value_{sig}$) // $Value_{sig}$ – значение сигнала в отсчете x .

начало

если $x < 1$ **то** $Value_{sig} = 0$ **возврат** $Value_{sig}$ **конец если**

если $cmd_x = Cmd_{key}$

то $Value_{sig} = 1$

иначе $Value_{sig} = \max(F1(x-1, cmd_{x-1}, Cmd_{key})) - 0.1, 0$

конец если

возврат $Value_{sig}$

конец

Рассмотрим второй вариант $f_2(x)$, который вычисляется по алгоритму F2.

Алгоритм F2.

вход (x, cmd_x, Cmd_{key}) // x – номер команды в списке, полученном трассировщиком,
// cmd_x – выполненная на шаге x команда.

выход ($Value_{sig}$) // $Value_{sig}$ – значение сигнала в отсчете x .

начало

если $x < 1$ **то** $Value_{sig} = 0$ **возврат** $Value_{sig}$ **конец если**

если $cmd_x = Cmd_{key}$

то $Value_{sig} = 1$

² В настоящей работе используются только мнимые части спектра сигнала.

³ Согласно алгоритму используются команды без учёта их операндов (параметров).

```

        иначе Valuesig = 0
    конец если
    возврат Valuesig
конец

```

Для большей наглядности представим возможный вид сигнала некоторой программы на рис. 1 и 2. Как правило, ключевые команды отстоят друг от друга на значительном расстоянии (более 10 команд).

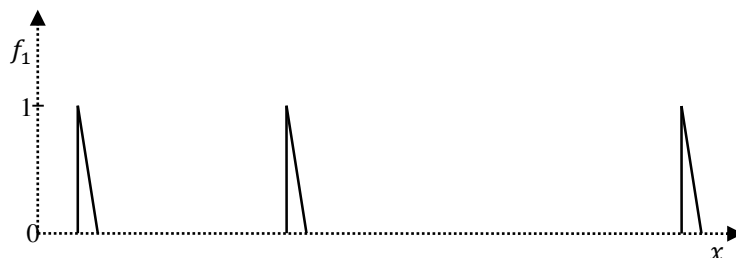


Рис. 1. Сигнал, хранимый в программе, при $f = f_1$

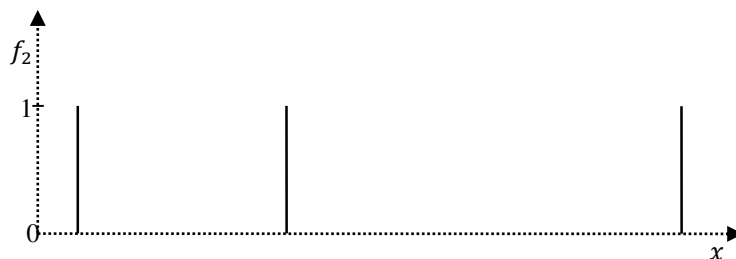


Рис. 2. Сигнал, хранимый в программе, при $f = f_2$

2.2. Сравнение спектров сигналов

Одним из важных моментов работы предлагаемого метода является сравнение спектров сигналов до и после изменения программы. Полученный спектр представляет собой множество действительных и мнимых частей сигнала. В силу теоремы Бенедикса [15] сигнал может быть конечен либо только по времени, либо только по спектру, соответственно, мы получим бесконечный спектр. Следовательно, для сравнения спектров сигналов нам необходимо избавиться от части членов ряда Фурье, получая при этом некоторое (желательно несущественное) искажение анализируемых сигналов.

Согласно предлагаемому алгоритму оставляются низкочастотные гармоники сигнала по следующим причинам:

1. Как и в случае с изображением формата jpg, высокочастотные гармоники отвечают за «мелкие детали» сигнала, низкочастотные – за «грубые». Так как малые (легальные) изменения в программе допускаются при помощи патчей, то высокочастотные гармоники сигнала будут меняться, следовательно, являются бесполезными для нашего алгоритма.

2. Так как алгоритм может быть применён к коротким фрагментам программы, то число команд и, следовательно, отсчётов сигнала может быть невелико. Тогда в силу теоремы Котельникова $F_{\text{отсч}} \geq 2 * F_{\text{max}}$ мы не сможем восстановить высокочастотные компоненты сигнала.

В результате для сравнения спектров используются мнимые части первых 20 гармоник сигнала $Im = A_i \sin(\varphi)$. Более того, из оставшихся низкочастотные гармоники считаются более значимыми, чем высокочастотные, поэтому используются весовые коэффициенты, взятые из распределения Ципфа:

$$Zipf(i) = \frac{1}{i} * \frac{1}{\sum_{i=1}^{20} \frac{1}{i}}. \quad (2)$$

Сравнение спектров сигналов осуществляется с помощью меры $Comp_{spec}(IM_1, IM_2)$, где IM – множество мнимых частей сигнала.

$$Comp_{spec}(IM_1, IM_2) = \sum_{i=1}^{19} Zipf(i) * ||im_{1i} - im_{1i+1}| - |im_{2i} - im_{2i+1}||. \quad (3)$$

Очевидно, что если мера равна нулю, то это означает отсутствие изменений в сигнале, следовательно, целостность программы не нарушена.

3. Экспериментальный анализ эффективности

Для проведения экспериментальной оценки эффективности предлагаемого метода были отобраны 30 файлов из операционной системы Windows 7 (консольные утилиты). Программы запускались в режиме трассировки и их выполненные команды записывались в trace-файлы. Из trace-файлов извлекался эталонный спектр, затем в них добавлялись случайные команды или удалялись блоки команд, имитирующие несанкционированное искажение. Имитировались следующие ситуации:

- добавление команд в начало trace-файла, имитируя работу вредоносных программ, меняющих стартовую точку входа в программу, например, вирус Virus.Win9x.CIH;
- добавление команд в случайную позицию trace-файла, имитируя искажение текущего функционала, например, проверку лицензионного ключа;
- добавление команд в конец trace-файла, имитируя вредоносные программы, дополняющие функционал программы;
- удаление блока команд из случайной позиции trace-файла, имитируя вредоносные программы, частично отключающие функционал программы.

В ходе анализа было установлено, что выбор ключевой команды из 10 самых часто встречающихся не существенно влияет на итоговый результат сравнения. С одной стороны, редко встречающаяся команда может быть неравномерно распределена по коду, что, как показано в эксперименте, создает угрозу модификации программы без искажения водяного знака. С другой стороны, если команда встречается слишком часто, то удельный вклад одного её вхождения уменьшается, что также создает возможность несанкционированной модификации кода без разрушения ЦВЗ. Среднестатистическое распределение команд в файле, рассчитанное для 30 файлов, показано в табл. 1.

Таблица 1. Распределение команд в программе

Команда	Вероятн. %	Команда	Вероятн. %	Команда	Вероятн. %	Команда	Вероятн. %
mov	18.64	pop	6.96	jb	2.20	jmp	1.11
cmp	12.76	test	4.32	lea	2.03	add	1.09
push	12.29	jz	3.05	dec	1.59	sub	0.43
inc	11.88	call	2.87	xor	1.46	jnb	0.42
jnz	10.59	movzx	2.52	retn	1.24	div	0.38

Результаты сравнения эталонных и изменённых файлов при последовательном добавлении команд для команды `pop` (извлечение данных из стека) показаны в табл. 2 и 3.

Таблица 2. Средние значения меры $Compr_{spec}$ при различном числе добавленных команд для $f = f_1$

Кол-во добавленных команд	Значение $Compr_{spec}$	
	Добавление в начало программы	
	ср. знач.	дисперсия
0	0.000	0.000
1	45.831	590.785
2	45.852	590.067
3	45.874	589.514
4	45.859	588.930
5	45.916	588.321
6	45.936	587.691
7	45.956	587.038
8	45.978	586.266
9	46.001	585.485
10	46.025	584.888

Таблица 3. Средние значения меры $Compr_{spec}$ при различном числе добавленных команд для $f = f_2$

Кол-во добавленных команд	Значение $Compr_{spec}$							
	Добавление в начало программы		Добавление в случайное место программы		Добавление в конец программы		Удаление из случайного места программы	
	ср. знач.	дисперсия	ср. знач.	дисперсия	ср. знач.	дисперсия	ср. знач.	дисперсия
0	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
1	0.069	0.029	0.045	0.002	0.000	0.000	29.699	622.835
2	0.138	0.012	0.090	0.005	0.000	0.000	29.701	623.875
3	0.207	0.026	0.141	0.016	0.000	0.000	29.641	622.465
4	0.275	0.047	0.184	0.024	0.000	0.000	29.639	623.332
5	0.344	0.073	0.229	0.037	0.000	0.000	29.657	624.086
6	0.412	0.105	0.272	0.055	0.000	0.000	29.651	624.776
7	0.479	0.143	0.319	0.078	0.000	0.000	29.663	626.462
8	0.547	0.186	0.354	0.090	0.000	0.000	29.550	623.013
9	0.615	0.235	0.397	0.124	0.000	0.000	29.585	625.410
10	0.683	0.290	0.439	0.149	0.000	0.000	29.629	625.734

Из полученных результатов видно, что меньшую дисперсию даёт функция f_2 , следовательно, она рекомендуется к использованию в данном методе для извлечения водяного знака. Вариант функции f_1 также может быть использован, но значения меры $Compr_{spec}$ сильно зависят от исходной программы. Тем не менее вариант f_2 можно считать более универсальным.

Нулевые значения $Compr_{spec}$ при добавлении в конец файла объясняются тем, что случайно добавленные команды были либо отличны от ключевой, либо воздействовали на высокочастотную часть спектра, которая не учитывается. Удаление даже незначительной доли команд существенно искажает спектр сигнала, что следует учитывать при санкционированных изменениях в программе.

На основании данных, полученных в результате эксперимента, можно утверждать следующее. Во-первых, значение $Comp_{spec}$ монотонно возрастает с увеличением добавляемых команд. Данная особенность позволяет определять объём новых команд, привнесённых в исходную программу. Следовательно, мы можем задавать порог допустимых изменений для случая санкционированной модификации программы авторскими патчами. Во-вторых, ЦВЗ защищает только тот участок программы, который находится перед последним вхождением ключевой команды. Следовательно, для эффективной защиты целесообразно использовать несколько водяных знаков, соответствующих исходному и обратному порядку команд программы.

4. Практическая реализация

В данной главе будут описаны ключевые моменты программной реализации предлагаемого метода. Будем рассматривать программу как совокупность непересекающихся блоков кода $B_{all} = \{b_1, \dots, b_n\}$ ⁴. Пусть имеется фрагмент программы $B_{prot} = \{b_i, b_{i+1}, b_{i+2}, \dots, b_{i+k}\}$, защищаемый ЦВЗ, такой что $B_{prot} \subseteq B_{all}$. Указанный фрагмент запускается, когда на вход программы подаются специальные параметры. Предлагаемый метод будет анализировать ЦВЗ, полученный только от фрагмента B_{prot} . Остальные фрагменты считаются малозначимыми и их модификация не влияет на целостность программы. Автор программы самостоятельно определяет специальные входные параметры и фрагмент B_{prot} .

Описываемая ниже реализация аналогична классической проверке контрольной суммы файла с помощью хэш-функций. Отличие настоящей реализации заключается в используемой функции и способе получения её аргументов. Предполагается, что процесс проверки ЦВЗ производится отдельной программой (модулем) и выполняется без участия пользователя.

Рассмотрим простейший вариант реализации. Для проверки целостности фрагмента B_{prot} последовательно выполняются шаги: получение списка команд фрагмента B_{prot} , получение спектра, сравнение спектра с эталонным значением. Для получения списка команд программа, содержащая ЦВЗ, запускается в режиме отладки с помощью функции `CreateProcess` с флагом `DEBUG_ONLY_THIS_PROCESS`. В память по адресу начала выполняемого кода (Original Entry Point) записывается команда прерывания для отладки (`int 3`), имеющая опкод `0xCC`, как показано далее в примере.

```
ReadProcessMemory(pi.hProcess, (void*)pStartAddress,
&OldInstruction, 1, &dwReadBytes);
NewInstruction= 0xCC;
WriteProcessMemory(pi.hProcess, (void*)pStartAddress,
&NewInstruction, 1, &dwWriteBytes);
FlushInstructionCache(pi.hProcess, (void*)pStartAddress, 1);
```

Данная команда прерывания позволит перехватить управление анализатору спектра после того как анализируемая программа загрузится операционной системой в память и статические библиотеки выполнят инициализацию функциями `DllMain`, но перед началом выполнения B_{prot} . Как только произойдет перехват, необходимо восстановить оригинальную перезаписанную команду (в примере хранящуюся в `OldInstruction`) и уменьшить регистр `EIP` на единицу (т.е. вернуться на начало выполняемого кода). Далее меняем регистр флагов для установления режима пошагового выполнения инструкций и передаем управление программе.

```
CONTEXT lcContext;
lcContext.ContextFlags = CONTEXT_FULL;
GetThreadContext(pi.hThread, &lcContext);
```

⁴ Такое утверждение правомерно, так как применяется принцип модульного программирования.

```

lcContext.EFlags |= 0x100;
lcContext.Eip --;
SetThreadContext(pi.hThread, &lcContext);
ContinueDebugEvent(debug_event.dwProcessId,      debug_event.dwThreadId,
dwContinueStatus);

```

На каждом шаге программы считываем байты выполняемой команды, адрес начала которой находится в регистре `eip` и декодируем команду согласно таблице [16]. Таким образом, мы получим список команд защищаемого фрагмента B_{prot} . Подробнее работа трассировщика описана в [17].

Последующие процессы получения спектра сигнала и сравнения его с эталоном были описаны во второй главе. Как и в случае с проверкой контрольной хэш-суммы, считается, что анализирующая программа и места хранения эталонного значения спектра не подвергаются атаке злоумышленником. Допускается, что анализ целостности ЦВЗ может выполняться по расписанию. При создании более сложных схем, например, когда проверка осуществляется в самом защищаемом модуле (программе), необходимо реализовывать самоанализ вне защищаемого фрагмента кода. Другими словами, если фрагмент программы b_j выполняет анализ ЦВЗ, то необходимо выполнение условия: $b_j \notin B_{prot}$.

Литература

1. Dyer J. G., et al. Building the IBM 4758 secure coprocessor // IEEE Computer. 2001. V. 34, № 10. P. 57–66.
2. Smith S. W., Weingart S. Building a high-performance, programmable secure coprocessor // Computer Networks. 1999. V. 31, № 8. P. 831–860.
3. Rajan H., Hosamani M. Tisa: Toward trustworthy services in a service-oriented architecture // IEEE Transactions on Services Computing. 2008. V. 1, № 4. P. 201–213.
4. Qiu J., et al. Identifying and Understanding Self-Checksumming // Defenses in Software. 2015. P. 207–218.
5. Junod P., et al. Obfuscator-LLVM software protection for the masses // 2015 IEEE/ACM 1st International Workshop on Software Protection. 2015. P. 3–9.
6. Gautam P., Saini H. A Novel Software Protection Approach for Code Obfuscation to Enhance Software Security // International Journal of Mobile Computing and Multimedia Communications (IJMCMC). 2017. V. 8, № 1. C. 34–47.
7. Schrittwieser S. et al. Protecting software through obfuscation: Can it keep pace with progress in code analysis? // ACM Computing Surveys (CSUR). 2016. V. 49, № 1. P. 4.
8. Официальный сайт программы UPX [Электронный ресурс]. URL: <https://upx.github.io/> (дата обращения: 24.03.2019).
9. Touili T., Ye X. Reachability Analysis of Self Modifying Code // 2017 22nd International Conference on Engineering of Complex Computer Systems (ICECCS). IEEE. 2017. P. 120–127.
10. Chen Z., Wang Z., Jia C. Semantic-integrated software watermarking with tamper-proofing // Multimedia Tools and Applications. 2018. V. 77, № 9. P. 11159–11178.
11. Balachandran V., Emmanuel S. Potent and Stealthy Control Flow Obfuscation by Stack Based Self-Modifying Code // IEEE Transactions on Information Forensics and Security. 2013. V. 8, № 4. P. 669–681.
12. Chen Y., et al. Oblivious hashing: A stealthy software integrity verification primitive // International Workshop on Information Hiding. Springer. 2002. P. 400–414.
13. Jacob M., Jakubowski M. H., Venkatesan R. Towards integral binary execution: Implementing oblivious hashing using overlapped instruction encodings // Proceedings of the 9th Workshop on Multimedia & security. ACM. 2007. P. 129–140.

14. Официальный сайт программы IDA Pro Free [Электронный ресурс]. URL: <https://www.hex-rays.com/products/ida/index.shtml> (дата обращения: 24.03.2019).
15. *Benedicks M.* On Fourier transforms of functions supported on a set of finite Lebesgue measure // *Journal of Mathematical Analysis and Applications*. V. 106. 1985. P. 180–183.
16. X86 Opcode and Instruction Reference [Электронный ресурс]. URL: <http://ref.x86asm.net/coder32.html> (дата обращения: 06.05.2019).
17. *Vijayvargiya A.* Writing Windows Debugger – Part 2 [Электронный ресурс]. URL: <https://www.codeproject.com/Articles/132742/Writing-Windows-Debugger-Part-3> (дата обращения: 06.05.2019).

Статья поступила в редакцию 22.05.2019.

Нечта Иван Васильевич

к.т.н., доцент, доцент кафедры прикладной математики и кибернетики СибГУТИ (630102, Новосибирск, ул. Кирова, 86), тел. (383) 2-698-359, e-mail: ivannechta@gmail.com.

Semi-fragile digital watermarks based on Fourier spectrum

I. V. Nechta

Digital watermarks are used for detection unauthorized edition of executable files by malicious programs. A property called “fragile” provides its distortion after minor changes in a program. The watermark presence and correctness controlled by endpoint user, which is obviously interested in software integrity. Semi-fragile DWM distorts when changing the program more than the set threshold. According to proposed method a program is considered as a container with digital signal that can be recovered from its samples (program commands). Any changes in the program code inevitably change the signal spectrum. During the analysis of modified program it was founded some distortion of the original signal: changing phases and amplitudes of some harmonics. The proposed approach allows to reveal not only the fact of modification, but also an amount of added code.

Keywords: fragile digital watermarks, steganography, Fourier spectrum.

Новая математическая модель риск-менеджмента образовательной деятельности вуза

Ю. В. Шевцова, В. С. Канев, А. Н. Полетайкин, Н. В. Кулешова

В статье предлагается оригинальный математический подход к решению актуальной задачи риск-менеджмента в отношении образовательной деятельности вуза в части обеспечения её качества. В качестве базового метода управления рисками применяется рисковый термометр, идентифицирующий рисковый фон образовательного процесса. Оценивание рисков осуществляется посредством нечёткой композиции полученных температурных данных с нечётким отношением соответствия рисков их ключевым индикаторам. Для построения функции принадлежности рисковой температуры с оптимальными параметрами применяется метод нечёткой кластеризации. По результатам экспертизы на основе температурных данных сформирован оптимальный комплекс мероприятий развития образовательной организации.

Ключевые слова: риск-менеджмент, ключевые индикаторы риска, рисковый термометр, температурные данные, нечёткая кластеризация, нечёткая композиция, мероприятия развития образовательной организации, минимизация рисков.

1. Введение

Образовательные организации, являясь элементом нестабильной социально-экономической среды общества, так же как и коммерческие компании испытывают необходимость перманентного управления рисками. Несмотря на то, что напрямую вопросам риск-менеджмента образовательных организаций в настоящее время уделено недостаточно внимания как со стороны научного [1, 2], так и со стороны делового сообществ, процесс управления рисками в них, безусловно, реализуется, однако, на наш взгляд, опосредованно. Так, в качестве процедур риск-менеджмента вуза можно видеть механизмы оценки качества образовательной деятельности и подготовки обучающихся, что, в свою очередь, является требованием современных ФГОС ВО (3++). «Качество образовательной деятельности и подготовки обучающихся ... определяется в рамках системы внутренней оценки, а также системы внешней оценки... Организация при проведении регулярной внутренней оценки качества образовательной деятельности и подготовки обучающихся ... привлекает юридических и (или) иных физических лиц, включая педагогических работников Организации... Внешняя оценка качества образовательной деятельности ... [осуществляется] в рамках процедуры государственной аккредитации» [3, п. 4.6].

Так или иначе, стоит признать, что концепция управления рисками образовательной организации в настоящее время находится во фрагментарном, ограниченном, слабоструктурированном состоянии. При этом на начальном этапе внедрения культуры риск-менеджмента в корпоративную среду организация должна иметь простой, универсально понятный, эффективный (с точки зрения ресурсов на внедрение и поддержание его устойчивого функционирования) инструмент. Одним из таких инструментов является рисковый термометр (РТ), который позволяет в первом приближении определить рисковый фон организации.

2. Формальное описание рискового термометра

По сути, РТ является формализацией некоторого опроса, выявляющего основные рисковые уязвимости компании. Пример рискового термометра коммерческой компании представлен в [4].

Методика РТ приводит статистически обработанные результаты опроса, извлеченные из опросного листа, к интегральному показателю, который может быть интерпретирован как рисковая температура компании. Причем важно заметить, что, как правило, вопросы опросного листа идентифицируют риски компании неявно, поэтому не требуют от респондента специфических знаний и навыков в области риск-менеджмента.

Интегральный показатель рисковой температуры компании (T) определяется по формуле:

$$T = \sum_{i=1}^n \sum_{j=1}^m k_{ij}^l x_{ij}, \quad l = \overline{1, p},$$

где x_{ij} – переменная j -го ответа респондента на i -й вопрос опросного листа: $x_{ij} = 1$, если респондент присвоил i -му вопросу j -й ответ, $x_{ij} = 0$ – наоборот; k_{ij}^l – рисковый вес j -го ответа на i -й вопрос; l – индекс рискового состояния объекта, p – число таких состояний.

Рисковые веса (k_l) устанавливаются ответам экспертно и выступают нормирующими коэффициентами, приводящими интегральные результаты к определённым показателям температуры:

$$k_l = \frac{T_l}{n}, \quad l = \overline{1, p}.$$

Так, представляя вуз некоммерческой образовательной организацией (ОО), разумно видеть следующие его температурные состояния:

- нормальное $T_1 = 36.6^\circ\text{C}$ – полное соответствие требованиям ФГОС ВО;
- лихорадочное $T_2 = 38^\circ\text{C}$ – частичное соответствие требованиям ФГОС ВО;
- катастрофическое $T_3 = 42^\circ\text{C}$ – полное несоответствие требованиям ФГОС ВО.

Таким образом, под рисковой температурой образовательной программы предлагается понимать степень её соответствия требованиям ФГОС ВО. Именно в этой связи мы не рассматриваем, как это принято для коммерческих организаций, пониженную температуру вуза (ниже 36.6°C) как неблагоприятный рисковый фон.

Кроме того, мы нашли хорошее соответствие анкеты научно-педагогических работников (НПР), используемой при проведении аккредитационной экспертизы основной образовательной программы [5], задачам измерения рисковой температуры программы. Данная анкета рассматривает ценности образовательной организации, находящиеся под риском: *процессы, персонал, системы, репутацию*, и позволяет оценить их состояние именно с точки зрения степени соответствия требованиям ФГОС ВО. Структура анкеты НПР представлена в табл. 1. Группы вопросов, адресуемых НПР, будем рассматривать как целевые факторы (ЦФ), рациональные и целенаправленные воздействия на которые, по идее данного исследования, должны привести исследуемую образовательную систему в нормальное (соответствующее требованиям ФГОС ВО) состояние.

Дифференциальная (частная) рисковая температура (T^s) по укрупненной группе показателей образовательной программы (множество групп показателей представлено в столбце «Целевой фактор» табл. 1, его мощность $q = 5$) определяется следующим образом:

$$T^s = \frac{n}{n^s} \sum_{i=1}^{n^s} \sum_{j=1}^m k_{ij}^l x_{ij}, \quad l = \overline{1, p}, \quad s = \overline{1, q}. \quad (1)$$

Таблица 1. Рисковый термометр образовательной программы бакалавриата

Целевой фактор	Вопросы НПР аккредитуемой программы	Варианты ответов	Рисковый вес ответа
I. Структура ООП.	Какие технологии при проведении занятий Вы используете?	Активные	2.38
		Интерактивные	2.29
		Другие	2.63
	Реализуется ли в Вашей ОО учебные курсы с применением информационных технологий?	Да	2.29
		Нет	2.63
		Не знаю	2.63
II. Общесистемные требования к реализации ООП.	Являетесь ли Вы штатным сотрудником?	Да	2.29
		Нет	2.63
		Внутренний совместитель	2.29
	Создана ли в Вашей ОО электронная информационно-образовательная среда?	Да	2.29
		Нет	2.63
		Затрудняюсь ответить	2.63
	Насколько Вы удовлетворены информационной наполненностью сайта программы?	Не удовлетворен	2.63
		Не в полной мере	2.38
		В большей степени	2.29
		Удовлетворен полностью	2.29
	Есть ли у Вас возможность пройти курсы повышения квалификации, обучающие семинары, стажировки?	Да	2.29
		Нет	2.63
	С какой периодичностью Вы проходите повышение квалификации?	Раз в пять лет	2.63
		Раз в три года	2.29
		Ежегодно	2.29
III. Кадровое обеспечение ООП.	Имеете ли Вы ученую степень, ученое звание?	Да. Кандидат, доктор	2.29
		Да. Доцент, профессор	2.29
		Нет	2.38
	Имеете ли Вы опыт практической работы по профилю преподаваемых дисциплин?	Да	2.29
		Нет	2.38
		Работаю в данное время	2.29
		Было давно	2.38
	Являетесь ли Вы научным руководителем магистерских программ?	Да	2.29
		Нет	2.29
	Есть ли у Вас публикации в научных рецензируемых изданиях за последние 5 лет?	Да	2.29
		Нет	2.63
		Другое	2.63
	Принимаете ли Вы участие в научных семинарах, конференциях?	Да	2.29
		Нет	2.63
		Редко	2.38
		Не знаю	2.63
IV. Материально-техническое обеспечение ООП.	Всегда ли доступна Вам вся необходимая информация, касающаяся учебного процесса, внеучебных мероприятий?	Да, всегда	2.29
		Нет, не всегда	2.38
		Затрудняюсь ответить	2.63
		Другое	2.63
	Удовлетворены ли Вы качеством аудиторий, помещений кафедр, учебных лабораторий и оборудования?	Полностью	2.29
		В большей мере	2.29
		Не в полной мере	2.38
		Не удовлетворен	2.63
	Удовлетворяет ли Вас качество фондов читального зала и библиотеки?	Не удовлетворяют	2.63
		Не в полной мере	2.38
		В большей степени	2.29
		Удовлетворяют	2.29
V. Условия организации образовательного процесса в целом.	Оцените, пожалуйста, условия организации образовательного процесса по программе в целом.	Неудовлетворительно	2.63
		Удовлетворительно	2.38
		Хорошо	2.29
		Отлично	2.29

3. Разработка модели риск-менеджмента образовательной деятельности

Стандарты ФГОС ВО модификации 3++ устанавливают, что «в целях совершенствования программы бакалавриата Организация [должна] проводить регулярную оценку качества образовательной деятельности...» [3, п. 4.6]. Нетрудно видеть, что эта процедура по своей сути соответствует мониторингу рисковому фону образовательной организации. Мониторинг рисков подразумевает оценку эффективности функционирования системы управления рисками ОО и, в соответствии с её результатами, корректировку управленческих процедур, а также поддержание непрерывности их выполнения.

Мониторинг рисковому фону может быть реализован с помощью системы обобщенных показателей деятельности компании, называемых ключевыми индикаторами риска (Key Risk Indicators, KRI) [6]. В качестве ключевых индикаторов риска должны применяться показатели, состояние которых свидетельствует об уровне рисковому фону компании. В связи с этим представляется естественным рассматривать ЦФ ОО в качестве ключевых индикаторов риска, по крайней мере, на уровне экспресс-диагностики рисковому фону.

Содержательная сущность ЦФ ОО позволила выявить характерные рисковые события, индикаторами которых они выступают (табл. 2).

Таблица 2. Рисковые события, идентифицируемые ключевыми индикаторами риска

Категория KRI	Объекты риска	Идентифицируемые риски
I. Структура ООП.	Процессы	1. Риск морального устаревания знаний. 2. Рассогласование интересов работодателей, выпускников, вузов и органов власти. 3. Сбои в работе технических систем.
II. Общесистемные требования к реализации ООП.	Процессы	1. Риск морального устаревания знаний. 2. Рассогласование интересов работодателей, выпускников, вузов и органов власти. 4. Риск зависимости от ключевого персонала. 5. Вымывание кадров. 6. Застой в научных исследованиях. 7. Нивелирование творческой и созидательной энергии НПР.
III. Кадровое обеспечение ООП.	Персонал	1. Риск морального устаревания знаний. 2. Рассогласование интересов работодателей, выпускников, вузов и органов власти. 5. Вымывание кадров. 6. Застой в научных исследованиях. 7. Нивелирование творческой и созидательной энергии НПР.
IV. Материально-техническое обеспечение ООП.	Системы	1. Риск морального устаревания знаний. 3. Сбои в работе технических систем. 6. Застой в научных исследованиях. 7. Нивелирование творческой и созидательной энергии НПР.
V. Условия организации образовательного процесса в целом.	Репутация	8. Утрата своеобразия, уникальности, колорита вуза.

При этом доля вопросов из анкеты НПР (столбец 2 табл. 1), которые идентифицируют соответствующее рисковое событие, позволили определить частотные характеристики сопоставления идентифицируемых рисков с их ключевыми индикаторами (табл. 3). Коль скоро эти

частоты характеризуют степень связи рисков с их индикаторами, значит, их в некотором приближении можно понимать как нечёткие характеристики уверенности в этом соответствии. Иными словами, матрицу размером 8×5 , представленную в табл. 3, можно рассматривать как бинарное нечёткое отношение (БНО) $RC = \{\langle r_i, c_j \rangle, \mu_{RC} \langle r_i, c_j \rangle\}$, где $r_i \in R$ и $c_j \in C$ – элементы чётких множеств рисков и KRI соответственно, $\mu_{RC} \langle r_i, c_j \rangle$ – функция принадлежности (ФП) БНО, которая определяется как отображение $\mu_{RC} : R, C \rightarrow [0, 1]$.

Таблица 3. Частотные характеристики индикации рисков

Идентифицируемые риски	Категория KRI (см. табл. 2)				
	I	II	III	IV	V
1. Риск морального устаревания знаний.	1.0	0.4	1.0	0.67	0.2
2. Рассогласование интересов работодателей, выпускников, вузов и органов власти.	0.8	0.4	1.0	0.67	0.2
3. Сбой в работе технических систем.	0.9	0.2	0	0.67	0.2
4. Риск зависимости от ключевого персонала.	0	0.2	0.4	0	0.2
5. Вымывание кадров.	0	0.2	0.4	0	1.0
6. Застой в научных исследованиях.	0	0.4	1.0	0.67	0.8
7. Нивелирование творческой и созидательной энергии НПП.	0	0.4	1.0	1.0	0.8
8. Утрата своеобразия, уникальности, колорита вуза.	0.8	0.8	0.8	0.8	1.0

Технология применения нечётких вычислений на основе БНО применительно к оцениванию образовательного процесса рассмотрена в [7]. Погрешность оценивания 89 личностных качеств обучающихся составила 8 %, что в условиях высокого уровня неопределённости, характерного для существенного присутствия в исследуемой системе человеческого фактора, можно считать удовлетворительным результатом.

Технология предполагает выполнение операции композиции БНО, которая применительно к решаемой задаче нечёткого оценивания рисков и в наиболее распространенном способе определения ФП результатного БНО методом минимакса реализуется следующим образом:

$$R_T = RC \otimes C_T, \quad \mu_{R_T} \{\langle r_i, l_k \rangle\} = \min_{c_j \in C} \left\{ \max \left\{ \mu_{RC} \langle r_i, c_j \rangle, \mu_{C_T} \langle c_j, l_k \rangle \right\} \right\}, \quad (2)$$

где левое БНО RC – условно-постоянное (содержит справочные нечёткие данные о соответствии рисков их ключевым индикаторам, представленные в табл. 3), правое БНО C_T – переменное и содержит фазифицированные оценки рисковей температуры, чёткие оценки которой рассчитаны по формуле (1) для основных категорий KRI $c_j \in C$, четверками оценок уровня температуры $l_k \in \overline{TL}$ ($k = 1, 4$), выраженных в виде лингвистических термов (см. табл. 4) в соответствии с установленной функцией принадлежности (см. рис. 1).

Таблица 4. Базисное множество лингвистических термов ($TL4$) для нечёткого оценивания рисковей температуры

Обозначение	Англоязычная нотация	Русскоязычная нотация
N	Normal	Нормальная температура
AN	Above the Normal	Повышенная температура (выше нормы)
H	High	Высокая температура
Cr	Critical	Критическая температура

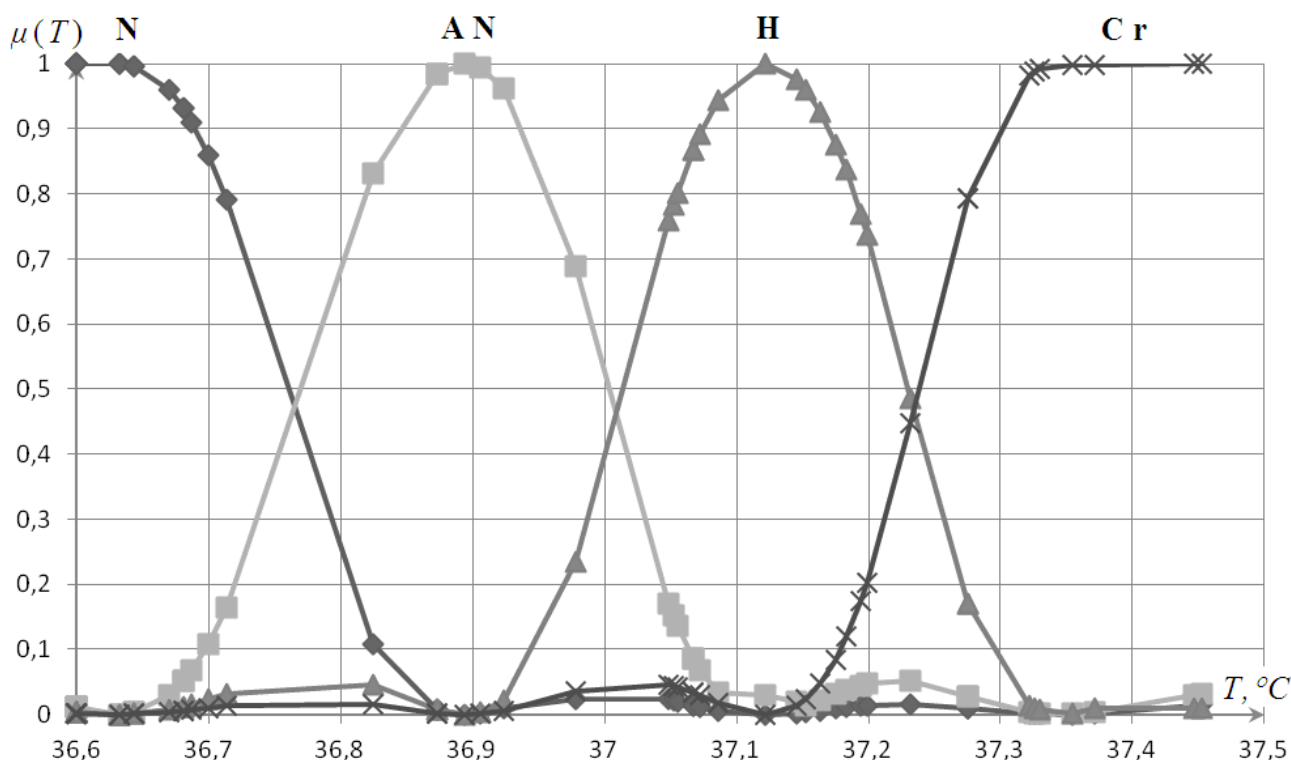


Рис. 1. Функции принадлежности лингвистической переменной «Нечёткая рисковая температура»

Операция композиции (2) не дистрибутивна и имеет своим результатом БНО R_T размером $|R| \times 4$, каждая строка которого выражает нечёткую характеристику четырех уровней выраженности риска в тех же терминах, что и уровень рискованной температуры, согласно ФП, показанной на рис. 2. В табл. 5 показано соответствие шкалы оценки вероятности наступления риска [2] и лингвистических термов из табл. 4.

Чёткая количественная оценка i -го риска r_i может быть получена посредством дефаззификации методом центра тяжести для одноточечных множеств:

$$P_{r_i} = \frac{\sum_{k=1}^4 (r_{ik} \cdot b_{ik})}{\sum_{k=1}^4 r_{ik}}, \quad (3)$$

где r_{ik} – элемент БНО R_T , соответствующий k -му терму, $k = \overline{1, 4}$; b_{ik} – чёткое значение соответствующего элемента базисного множества лингвистических термов $TL4$ (см. табл. 4), определяемое на основе ФП «Нечёткая оценка риска», график которой показан на рис. 2.

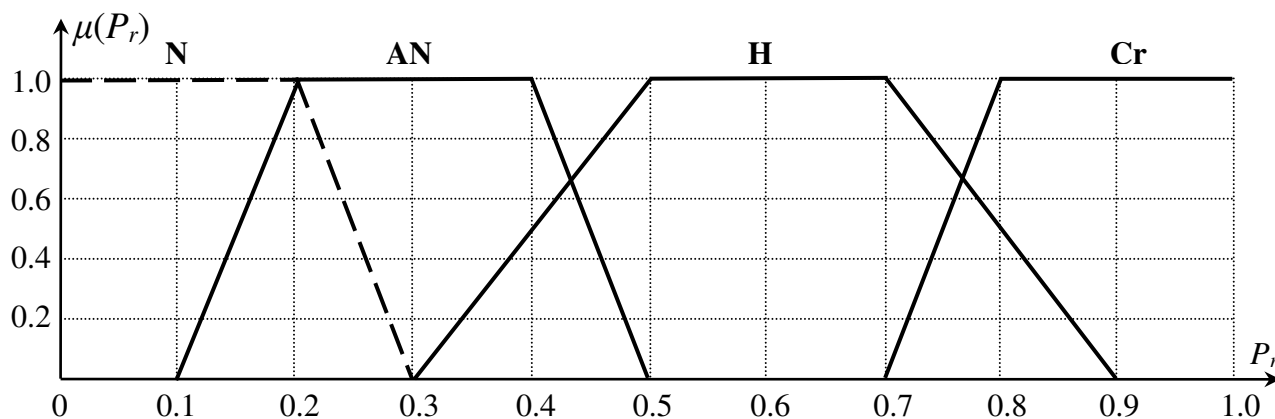


Рис. 2. Функции принадлежности лингвистической переменной «Нечёткая оценка риска»

Таблица 5. Шкала для нечёткого оценивания вероятности реализации риска [2]

Обозначение	Ранг	Интерпретация вероятности реализации риска	
N	I	Реализация риска маловероятна	< 0.20
AN	II	Средняя вероятность реализации риска	$[0.20...0.50)$
H	III	Высокая вероятность реализации риска	$[0.50...0.75)$
Cr	IV	Очень высокая вероятность реализации риска	≥ 0.75

4. Оценивание рисков на основе анкетных данных НПР

Как замечено ранее, апробация предлагаемого подхода к измерению рискованного фона вуза в целом и его отдельных образовательных программ проводилась на базе ФГБОУ ВО «Сибирский государственный университет телекоммуникаций и информатики» (СибГУТИ). Её результаты, полученные на основе анкетных данных НПР при проведении аккредитационной экспертизы основной образовательной программы в приложении к 9 направлениям подготовки, аккредитованных в СибГУТИ, представлены на рис. 3. Рисунок демонстрирует оценки рискованной температуры по группам показателей, рассчитанные по формуле (1).

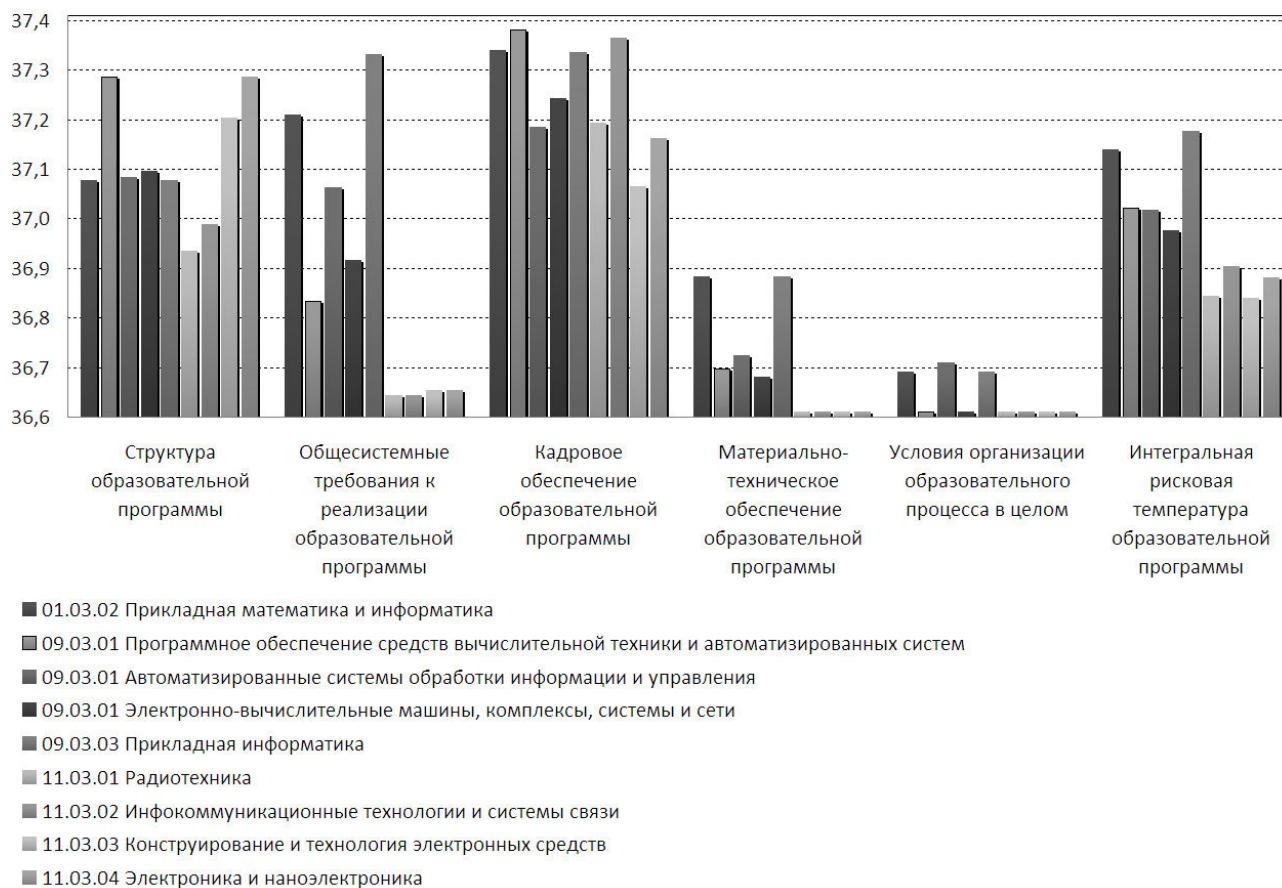


Рис. 3. Рискованная температура образовательных программ бакалавриата СибГУТИ (в легенде порядок направлений подготовки соответствует порядку рядов на диаграмме)

Функция принадлежности термов лингвистической переменной «Нечёткая рискованная температура» (см. рис. 1) была определена эмпирически с применением метода нечёткой кластеризации с-средних (с-means) [8] и пакета математического моделирования MATLAB R2018b. При этом в качестве обучающей выборки применялись полученные в СибГУТИ температурные данные образовательных программ бакалавриата (см. рис. 4).

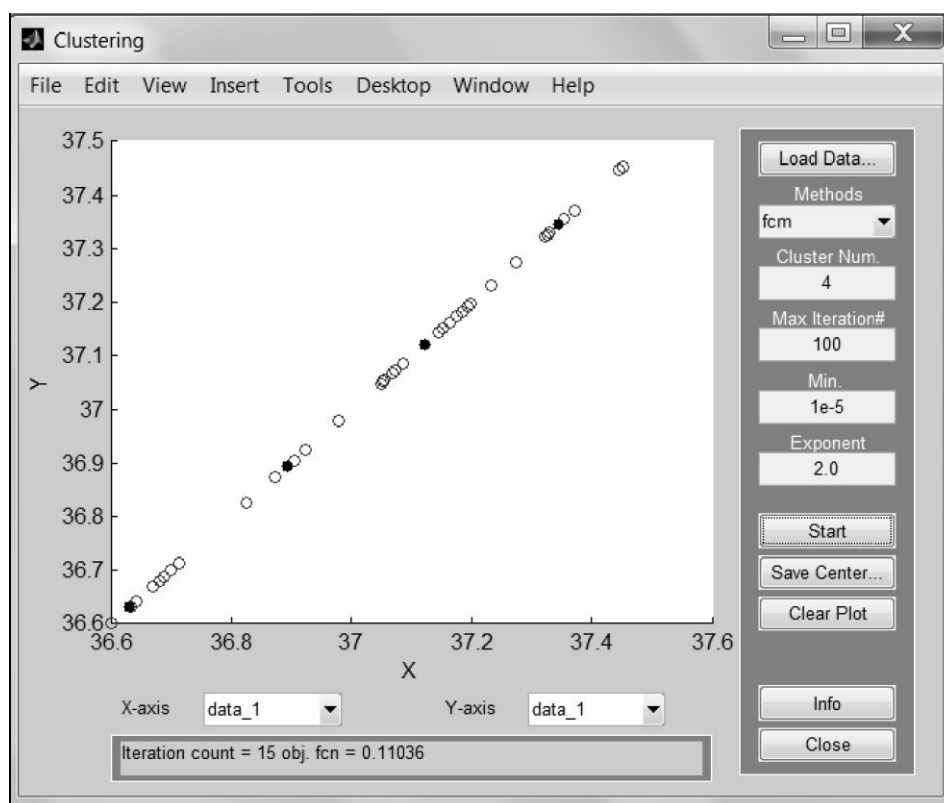


Рис. 4. Окно моделирования MATLAB при решении задачи нечеткой кластеризации температурных данных методом с-средних

В ходе нечёткой кластеризации фактически решается задача минимизации целевой функции разброса, задающей сумму расстояний от элементов исходного множества температурных данных до центра своего кластера. Характерная отличительная особенность нечёткой кластеризации состоит в том, что в результате каждый элемент множества оказывается одновременно принадлежащим нескольким кластерам (нечётким множествам) с определённой степенью принадлежности.

Рассмотрим расчёт оценок характерных рисков на примере данных по направлению подготовки 01.03.02 Прикладная математика и информатика (ряд № 1 на рис. 3). Температурные данные представлены в табл. 6, где нечёткая рисковая температура рассчитана посредством процедуры фаззификации согласно ФП на рис. 1.

Результат операции композиции (2) представлен в табл. 7. Чёткие оценки рисков в виде вероятностей их наступления (согласно шкале в табл. 4) рассчитаны посредством процедуры дефаззификации по формуле (3) согласно ФП на рис. 2.

Таблица 6. Температурные данные по направлению подготовки
01.03.02 Прикладная математика и информатика

Категория KRI	Рисковая температура	Нечёткая температура			
		N	AN	H	Cr
I. Структура ООП.	37.07	0.014	0.086	0.867	0.033
II. Общесистемные требования к реализации ООП.	37.20	0.014	0.047	0.737	0.202
III. Кадровое обеспечение ООП.	37.33	0.001	0.001	0.006	0.992
IV. Материально-техническое обеспечение ООП.	36.87	0.007	0.984	0.007	0.002
V. Условия организации образовательного процесса в целом.	36.68	0.932	0.051	0.012	0.005

Таблица 7. Оценки характерных рисков образовательного процесса

Идентифицируемые риски	Нечёткая характеристика (R_T)				Вероятность риска
	N	AN	H	Cr	
1. Риск морального устаревания знаний.	0.400	0.200	0.200	0.200	0.434
2. Рассогласование интересов работодателей, выпускников, вузов и органов власти.	0.400	0.200	0.200	0.200	0.434
3. Сбои в работе технических систем.	0.001	0.001	0.006	0.200	0.706
4. Риск зависимости от ключевого персонала.	0.007	0.086	0.007	0.002	0.325
5. Вымывание кадров.	0.007	0.086	0.007	0.002	0.325
6. Застой в научных исследованиях.	0.014	0.086	0.670	0.033	0.567
7. Нивелирование творческой и созидательной энергии НПП.	0.014	0.086	0.737	0.033	0.569
8. Утрата своеобразия, уникальности, колорита вуза.	0.800	0.800	0.800	0.800	0.467

Из табл. 7 видно, что наиболее высокие значения имеют риски 3, 6 и 7, вероятности которых согласно экспертным оценкам из табл. 5 отнесены к III рангу и характеризуются как высокие. Остальные риски относятся ко II рангу и характеризуются средней вероятностью их реализации. Такой рисковый фон может быть классифицирован как неблагоприятный и требующий принятия комплекса превентивных мер, понижающих указанные риски и соотнесенные с ними ключевые индикаторы риска.

5. Определение оптимального комплекса мероприятий развития вуза

В результате анализа первичных результатов оценивания рисков экспертами были предложены наиболее действенные мероприятия (табл. 8), которые были сгруппированы по четырем группам и оценены лингвистическими оценками согласно данным табл. 9.

Таблица 8. Матрица влияния мероприятий развития образовательной организации на целевые факторы удовлетворённости из табл. 2

Наименование мероприятия развития образовательной организации	Стоимость (C), усл. ед.	Влияние на ЦФ (KRI) (см. табл. 2)				
		I	II	III	IV	V
1	2	3	4	5	6	7
Развитие кадровой политики						
1. Формирование непрерывного графика повышения квалификации для НПП по профилю читаемых дисциплин, по вопросам внедрения новых ФГОС ВО, по информационным и педагогическим технологиям.	50	PB	PB	PM	NB	PB
2. Разработка программ профессиональной переподготовки и повышения квалификации НПП с учетом требований ФГОС.	100	PB	PB	PM	NB	PB
3. Организация программ стажировок на предприятиях в сфере телекоммуникаций и информатики.	30	NB	PB	PM	NB	PB
4. Создание системы наставничества для работы и профессионального роста молодых преподавателей.	60	NB	PM	PB	NB	PB

Окончание таблицы 8

1	2	3	4	5	6	7
Развитие материально-технической базы						
5. Модернизация учебно-лабораторного и спортивного оборудования в соответствии с современными техническими требованиями и требованиями ФГОС.	100	PB	NS	NB	PB	PB
6. Обновление основных фондов учебных аудиторий, помещений кафедр, учебных лабораторий: полная или частичная замена мебели, компьютеров, оргтехники и др. Обеспечение условий доступа в Интернет в каждой аудитории.	100	PB	NS	NB	PB	PB
7. Полное оснащение компьютерным и мультимедийным оборудованием учебных аудиторий и читальных залов библиотеки.	100	PB	NS	NB	PB	PB
8. Приобретение лицензионного программного обеспечения, необходимого для проведения лабораторных и практических занятий.	90	PB	NS	NB	PB	PB
Совершенствование научно-исследовательской деятельности						
9. Регулярное проведение научно-технических и научно-методических конференций и семинаров с целью повышения уровня профессиональной компетентности.	50	PB	PM	PB	NB	PB
10. Совершенствование системы материального стимулирования в зависимости от качества выполненных научно-исследовательских работ.	80	PS	Z	PB	NS	PM
11. Разработка механизма мотивации аспирантов и повышения эффективности организации деятельности научных руководителей.	80	NB	PM	PB	NB	PS
12. Создание единого научно-методического комплекса интеграции образования, науки и производства с участием ННПР, обучающихся всех уровней и форм обучения и представителей работодателей.	100	NB	PS	PB	PS	PB
Совершенствование образовательной деятельности						
13. Создание единого информационного внутривузовского пространства: сайт вуза, локальные сети, информационные стенды.	50	PB	PB	PS	PB	PB
14. Внедрение единой методики разработки основных профессиональных образовательных программ в соответствии с изменениями законодательства и новыми ФГОС.	70	PB	NB	NB	NB	PB
15. Создание автоматизированной системы документооборота с целью согласования учебно-методической документации между сотрудниками разных подразделений вуза.	50	NB	NB	NB	PB	PB

Таблица 9. Базисное множество лингвистических термов для нечёткого оценивания действенности мероприятий развития образовательной организации

Обозначение	Степень влияния (k)	Англоязычная нотация	Русскоязычная нотация
NB	-75 %	Negative Big	Отрицательное большое
NM	-50 %	Negative Middle	Отрицательное среднее
NS	-25 %	Negative Small	Отрицательное малое
Z	0 %	Zero	Ноль, близкое к нулю
PS	25 %	Positive Small	Положительное малое
PM	50 %	Positive Middle	Положительное среднее
PB	75 %	Positive Big	Положительное большое

Под степенью влияния определённого мероприятия развития ОО понимается степень влияния данного мероприятия на фактор риска, сопровождающаяся изменением рискованной температуры по каждой категории KRI и в идеале приводящая её к нормальному состоянию ($T_{\text{норм.}} = 36.6^\circ\text{C}$), соответствующему, напомним, полному соответствию требованиям ФГОС ВО и, соответственно, полной удовлетворённости НПП.

Кроме того, из всего вышеизложенного легко видеть следующее функциональное отображение:

$$T \rightarrow \mu(C_T) \rightarrow \mu(R_T) \rightarrow P_r.$$

Тогда задача определения оптимального комплекса мероприятий развития вуза ($A^* = (x_1^*, x_2^*, \dots, x_m^*)$) представляет собой задачу минимизации вероятности реализации рискованного события на определённом наборе данных мероприятий:

$$P_{r,A} = f(T_A^1, T_A^2, \dots, T_A^q) \rightarrow \min,$$

$$T_A^s = \begin{cases} T_{\text{исх.}} - \sum_{j=1}^m \frac{k_{a_j}^s}{100} \times x_{a_j} \times (T_{\text{исх.}} - T_{\text{норм.}}), & \text{если } \sum_{j=1}^m \frac{k_{a_j}^s}{100} \times x_{a_j} < 1, \quad s = \overline{1, q}, \\ T_{\text{норм.}} = 36,6, & \text{если } \sum_{j=1}^m \frac{k_{a_j}^s}{100} \times x_{a_j} \geq 1, \quad s = \overline{1, q}, \end{cases}$$

$$x_{a_j} = \{0, 1\}, \quad j = \overline{1, m},$$

где $k_{a_j}^s$ – степень влияния j -го мероприятия на s -й целевой фактор удовлетворённости (частный пример значений данного коэффициента представлен в табл. 9). Совместное влияние набора мероприятий на s -й целевой фактор удовлетворённости в самом простом случае может быть получено суммированием их влияния $\left(\sum_{j=1}^m \frac{k_{a_j}^s}{100} \times x_{a_j} \right)$.

Рассматриваемая оптимизационная задача не является гладкой, поэтому её решение было получено с помощью эволюционного программирования (табл. 10). Отметим, что в табл. 10 представлен один из множества альтернативных оптимальных наборов мероприятий по управлению каждым идентифицированным ранее риском (см. табл. 2.).

Таблица 10. Оптимальный набор мероприятий по управлению рискованными событиями развития образовательной организации

№ риска из табл. 7	№ мероприятия из табл. 8															Модифицированная вероятность риска ($P_{r,A}$)
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	
1	0	1	1	0	0	0	1	1	0	0	0	1	1	0	1	0.352
2	0	1	0	0	1	0	1	0	0	0	1	1	1	0	0	0.352
3	0	0	0	1	1	1	0	1	1	1	1	1	1	0	1	0.200
4	0	0	0	1	1	0	0	1	0	1	0	1	1	0	0	0.200
5	0	0	0	1	0	0	1	1	0	1	0	1	1	0	0	0.200
6	0	1	1	0	1	0	1	1	0	1	0	1	1	0	0	0.200
7	0	1	1	0	1	0	1	1	0	1	0	1	1	0	0	0.200
8	0	0	0	0	1	0	0	0	1	0	0	1	1	0	0	0.453

Итак, на данном этапе мы получили оптимальный набор мероприятий по управлению рисками развития ОО по критерию минимизации вероятности их возникновения. При этом очевидно, что для определения комплексной эффективности определённого набора меропри-

ятий необходимо помимо степени его воздействия на вероятность реализации рискованного события учитывать также и стоимость внедрения данного набора, т.е., по сути, решить двухкритериальную оптимизационную задачу следующего вида:

$$P_{r,A} = f(T_A^1, T_A^2, \dots, T_A^q) \rightarrow \min,$$

$$\sum_{j=1}^m C_{a_j} \times x_{a_j} \rightarrow \min,$$

$$x_{a_j} = \{0,1\}, \quad j = \overline{1,m},$$

где C_{a_j} – стоимость реализации j -го мероприятия (в табл. 8 представлены экспертные оценки стоимости реализации рассматриваемых мероприятий, мы намеренно даём эти оценки в условных единицах и, главным образом, имеем целью отразить отношение их порядка).

Рассмотрим решение двухкритериальной задачи поиска оптимального набора мероприятий по управлению рисковым событием «Сбои в работе технических систем» как событием, имеющем наибольшую вероятность реализации в исходном состоянии ОО. Для этого, следуя классической схеме решения задач многокритериальной оптимизации, преобразуем её во множество однокритериальных задач условной оптимизации с ограничением на бюджет (B):

$$P_{r,A} = f(T_A^1, T_A^2, \dots, T_A^q) \rightarrow \min,$$

$$\sum_{j=1}^m C_{a_j} \times x_{a_j} \leq B, \quad B \in [0; \infty],$$

$$x_{a_j} = \{0,1\}, \quad j = \overline{1,m}.$$

Результаты расчетов представлены в табл. 11.

Таблица 11. Парето-оптимальные наборы мероприятий по управлению рисковым событием «Сбои в работе технических систем» с учётом их стоимости

Бюджетное ограничение	№ мероприятия из табл. 8															Модифицированная вероятность риска ($P_{r,A}$)
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	
[0; 30)	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
[30; 50)	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0.444
[50; 80)	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0.294
[80; 150)	0	0	1	0	0	0	0	0	1	0	0	0	0	0	0	0.233
[150; 290)	0	0	0	0	0	0	0	0	0	0	0	1	1	0	0	0.210
[290; ∞)	0	0	0	0	0	0	0	1	1	0	0	1	1	0	0	0.200

Выбор конкретного набора мероприятий из Парето-оптимального множества определяется отношением ЛПР к риску и величиной бюджета на их внедрение.

6. Заключение

В статье рассмотрены результаты разработки и исследования новой математической модели регулирования рискованного фона образовательной организации. Исходными данными для исследования выступили результаты опроса преподавательского состава СибГУТИ, а также результаты экспертизы комплекса возможных развивающих мероприятий. Полученные с применением метода рискованного термометра температурные данные о функционировании образовательной организации показали существенно повышенный температурный и рискованный фон. Используемая в модели нечеткая технология позволила получить оптимальный набор

мероприятий, нормализующий рисковый фон образовательной организации, которые могут быть использованы для принятия управленческих решений с целью её развития.

Вместе с тем мы осознаем ограниченность используемого эмпирического материала и видим перспективной калибровку модели на более широком и разнообразном наборе данных, что также должно повысить точность и достоверность результатных оценок рисков и мероприятий, а также улучшить прогностические свойства модели. Однако ещё раз подчеркнём, что глобальной целью данного исследования видим развитие именно концептуального подхода к реализации процедур управления рисками образовательных организаций.

Литература

1. *Игнатова О. А., Ситников С. Г., Канев В. С.* Управление стратегическими рисками развития вуза в нестабильной среде функционирования // Интеграция образования и науки в современном университете: перспективы, проблемы, решения: сб. тез. докл. LI науч.-метод. конф. СибГУТИ, Новосибирск, 2010.
2. *Ильина Т. С., Баранова А. И., Канев В. С.* Управление рисками оценивания образовательных компетенций в высших учебных заведениях // Вестник СибГУТИ. 2017. № 1. С. 3–11.
3. Федеральный государственный образовательный стандарт высшего образования по направлению подготовки 09.03.03 Прикладная информатика (уровень бакалавриата). М: Министерство образования и науки РФ, 2017. 20 с.
4. *Вяткин В. Н., Гамза В. А.* Управление рисками фирмы: программа интегративного риск-менеджмента. М.: Финансы и статистика, 2006. 400 с.
5. Методические рекомендации по проведению аккредитационной экспертизы в отношении основных образовательных программ. М: ФГБУ «Росаккредагентство», 2015. 164 с.
6. *Катилова Н. В., Энгел С.* Практика ключевых индикаторов для операционных рисков // Управление финансовыми рисками. 2006. № 2. С. 190–204.
7. *Полетайкин А. Н., Ильина Т. С.* Нечёткий подход к оцениванию личностных качеств обучающихся вуза // ИТ-портал. 2016. № 4 (12). URL: <http://itportal.ru/science/tech/nechetkiy-podkhod-k-otsenivaniyu-li/> (дата обращения: 25.05.2019)
8. *Штовба С. Д.* Введение в теорию нечётких множеств и нечёткую логику [Электронный ресурс]. URL: <http://matlab.exponenta.ru/fuzzylogic/book1/index.php> (дата обращения: 25.05.2019).

Статья поступила в редакцию 11.07.2019.

Шевцова Юлия Владимировна

к.т.н., доцент, доцент кафедры математического моделирования бизнес-процессов СибГУТИ тел. (383) 269-82-78, e-mail: shevcova_yuliya@mail.ru.

Канев Валерий Семенович

д.т.н., доцент, заведующий кафедрой математического моделирования бизнес-процессов СибГУТИ (630102, Новосибирск, ул. Кирова, д. 86), тел. (383) 269-82-77, e-mail: kanev@ngs.ru.

Полетайкин Алексей Николаевич

к.т.н., доцент, доцент кафедры информационных технологий Кубанского государственного университета (350040, Краснодар, ул. Ставропольская, 149), тел. (861) 219-95-77, e-mail: alex.poletaykin@gmail.com.

Кулешова Наталия Владимировна

к.э.н., доцент, начальник учебно-методического отдела СибГУТИ (630102, Новосибирск, ул. Кирова, д. 86), тел. (383) 269-82-16, e-mail: natkuleshova@yandex.ru.

New mathematical model of risk management of education organization

Y. Shevtsova, V. Kanev, A. Poletaikin, N. Kuleshova

We propose new original mathematical approach to the problem of a risk management concerning ensuring quality of educational activity of higher education institution. As a basic method of risk management the risk thermometer identifying a risk background of educational process is used. Estimation of risks is carried out by means of fuzzy composition of the obtained temperature data with the fuzzy relation of compliance of risks to their key indicators. The method of fuzzy clustering is applied to construction of risk temperature's function with optimum parameters. As examination results the optimum of complex of actions for developing educational organization formed.

Keywords: risk management, key risk indicators, risk thermometer, temperature data, fuzzy clustering, fuzzy composition, developing actions of educational organization, risks minimization.

Оценка избыточности универсального кодирования произвольного множества источников без памяти

В. К. Трофимов

Получена нижняя оценка избыточности универсального кодирования произвольного множества источников без памяти, совпадающая по порядку убывания с верхней оценкой, полученной автором ранее.

Ключевые слова: кодирование, избыточность, энтропия, хранение и обработка информации, источник сообщений.

1. Введение

В настоящей работе рассматривается вопрос о кодировании блоков, порожденных источником без памяти, словами различной длины в том случае, когда источник не известен, но известно множество, которому он принадлежит.

Если источник полностью известен, то данный вопрос решен в [2, 3], если же источник полностью не известен, то оптимальное кодирование предложено в [4, 5]. При этом скорость убывания избыточности в первом случае равна c_1/n , во втором – $(c_2 \log n)/n$, где n – длина кодированного блока, c_1, c_2 – постоянные, не зависящие от n .

В [1] автором предложен метод универсального кодирования, который зависит от массивности множества источников, определяемой ε -энтропией этого множества. Как было показано в [1], в зависимости от массивности заданного множества источников скорость убывания избыточности изменяется в пределах от c_1/n до $(c_2 \log n)/n$.

Целью настоящей работы является получение нижней оценки скорости убывания избыточности, совпадающей с верхней оценкой из [1], и установление факта, является ли предложенное в [1] кодирование асимптотически оптимальным.

2. Основные определения

В данной работе мы будем использовать определения и обозначения, введенные в работе [1]. Напомним кратко основные из них. Как и прежде, буквы конечного входного алфавита $A = \{a_1, a_2, \dots, a_k\}$, $2 \leq k < \infty$ порождаются бернуллиевским источником, т.е. источником без памяти, с вероятностями $P_\theta(a_i) = \theta_i$, $i = \overline{1, k}$, $\theta_1 + \theta_2 + \dots + \theta_k = 1$.

Таким образом, источник θ однозначно определяется числами θ_i , $i = \overline{1, k}$, сумма которых равна единице. Верно и обратное утверждение. Как обычно, A^n – множество всех слов (блоков) длины n в алфавите A . В этом случае произвольная полубесконечная последовательность букв алфавита A , порождаемая источником θ , однозначно разбиваются на слова (блоки) длины n .

Каждое из слов u , $u \in A^n$ с помощью отображения φ отображается в слово $\varphi(u)$ конечного выходного алфавита B . Не уменьшая общности, можно считать, что $B = \{0, 1\}$. Мы берем только такие изображения φ , чтобы $\varphi(A^n) = \{\varphi(u), u \in A^n\}$ являлось префиксным, т.е. рассматриваются только дешифруемые кодирования [6]. Стоимость кодирования φ обозначим как $c(n, \theta, \varphi)$ [6]. Здесь и в дальнейшем:

$$c(n, \theta, \varphi) = \frac{1}{n} \sum_{u \in A^n} P_\theta(u) \cdot |\varphi(u)|, \quad (1)$$

где $P_\theta(u)$ – вероятность порождения слова u длины n источником θ , $|\varphi(u)|$ – число букв выходного алфавита в слове $\varphi(u)$.

Энтропию источника θ обозначим $H(\theta)$. По определению

$$H(\theta) = - \sum_{i=1}^k \theta_i \log \theta_i,$$

здесь и в дальнейшем $\log x = \log_2 x$, $0 \log 0 = 0$. Разность между $c(n, \theta, \varphi)$ и $H(\theta)$ обозначим через $r(n, \theta, \varphi)$ и назовем избыточностью кодирования φ . Таким образом:

$$r(n, \theta, \varphi) = c(n, \theta, \varphi) - H(\theta) \quad (2)$$

Величина $R(n, \Omega)$, определяемая равенством

$$R(n, \Omega) = \inf_{\varphi} \sup_{\theta \in \Omega} r(n, \theta, \varphi), \quad (3)$$

называется избыточностью универсального кодирования для множества источников Ω , $\Omega \subset \Omega_0$ при заданной сложности n . В равенстве (3) нижняя грань берется по всем дешифруемым кодированиям φ , определенным на множестве входных слов A^n . Если Ω содержит единственный источник θ_0 , то поведение $R(n, \theta_0)$ полностью изучено в [2, 3], где установлено, что $R(n, \theta_0)$, как правило, убывает со скоростью c/n . Если же множество источников Ω совпадает с Ω_0 , то в [4] доказано асимптотическое равенство

$$R_n(\Omega_0) \sim \frac{k-1}{2} \cdot \frac{\log n}{n},$$

здесь и в дальнейшем $f(n) \sim g(n)$, если $\lim_{n \rightarrow \infty} \frac{f(n)}{g(n)} = 1$; $f(n) \leq g(n)$, если $\lim_{n \rightarrow \infty} \frac{f(n)}{g(n)} \leq 1$.

В работе автора [1] для произвольного множества $\Omega \subset \Omega_0$ предложено кодирование, которое позволяет получить асимптотическую оценку

$$R_n(n, \Omega) \leq \frac{h_{1/\sqrt{n}}(\Omega)}{n} + \frac{c}{n}, \quad (4)$$

где $h_\varepsilon(\Omega)$ – ε -энтропия Ω [6].

В настоящей работе установлена нижняя оценка для $R_n(n, \Omega)$, асимптотически совпадающая с верхней оценкой (4), и тем самым доказано асимптотическое равенство, когда $h_{1/\sqrt{n}}(\Omega) \rightarrow \infty$, при $n \rightarrow \infty$:

$$R(n, \Omega) \sim \frac{h_{1/\sqrt{n}}(\Omega)}{n} + \frac{c}{n}. \quad (5)$$

В том случае, когда $|\Omega| = m$, $m < \infty$, то

$$\frac{\log m}{n} \leq R(n, \Omega) \leq \frac{\log m}{n} + \frac{c}{n}.$$

Наряду с минимаксным подходом, который характеризуется избыточностью универсального кодирования для множества источников Ω , определяемого равенством (3), можно использовать среднюю избыточность $\bar{R}(n, \Omega)$ универсального кодирования для множества источников Ω , определяемую равенством:

$$\bar{R}(n, \Omega) = \inf_{\varphi} \int r(n, \theta, \varphi) dw(\theta), \quad (6)$$

где $w(\theta)$ – распределение вероятности на множестве источников Ω .

Хорошо известно, что для любого множества источников Ω и произвольной меры $w(\theta)$ всегда выполняется неравенство:

$$R(n, \Omega) \geq \bar{R}(n, \theta). \quad (7)$$

Именно этим неравенством мы будем пользоваться для получения нижней оценки для $R(n, \Omega)$.

3. Оценка избыточности универсального кодирования произвольного множества источников без памяти

Для доказательства основного результата настоящего параграфа докажем ряд вспомогательных утверждений.

Для $\varepsilon = 2/\sqrt{n}$ построим минимальную ε -сеть $\Omega(2/\sqrt{n})$ для источников Ω . Тогда согласно [7] справедливы неравенства

$$\|\Omega(1/\sqrt{n})\| \geq \|\Omega(2/\sqrt{n})\| \geq \frac{1}{2} \|\Omega(1/\sqrt{n})\|. \quad (8)$$

Каждый из источников $\Omega(2/\sqrt{n})$ накроем кубом со стороной $1/\sqrt{n}$. Полученное множество источников обозначим $\bar{\Omega}(2/\sqrt{n})$, а через $\mu(\bar{\Omega}(2/\sqrt{n}))$ – меру множества $\bar{\Omega}(2/\sqrt{n})$ с плотностью $f(\theta) = c / \sqrt{\prod_{i=1}^K \theta_i}$.

Для множества источников Ω и $\bar{\Omega}(2/\sqrt{n})$ выполняется соотношение $\Omega \in \bar{\Omega}(2/\sqrt{n})$, поэтому

$$R(n, \bar{\Omega}(2/\sqrt{n})) \geq R(n, \Omega). \quad (9)$$

Используя кодирование, предложенное в [1], для $R(n, \bar{\Omega}(2/\sqrt{n}))$ и для $R(n, \Omega)$ в силу определения $\bar{\Omega}(2/\sqrt{n})$ имеем

$$R(n, \bar{\Omega}(2/\sqrt{n})) \leq \frac{h_{1/\sqrt{n}}(\Omega)}{n} + \frac{c_1}{n}; \quad (10)$$

$$R(n, \Omega) \leq \frac{h_{1/\sqrt{n}}(\Omega)}{n} + \frac{c_2}{n}. \quad (11)$$

Из [1] следует, что всякое универсальное кодирование для $R(n, \bar{\Omega}(2/\sqrt{n}))$, которое позволяет получить оценку (10), является универсальным кодированием и для множества источников Ω .

Таким образом, если мы докажем, что для $R(n, \bar{\Omega}(2/\sqrt{n}))$ выполняется асимптотическое равенство

$$R(n, \bar{\Omega}(2/\sqrt{n})) \sim \frac{h_{1/\sqrt{n}}(\Omega)}{n}, \quad (12)$$

то для $R(n, \Omega)$ будет выполняться то же самое тождество. Таким образом,

$$R(n, \bar{\Omega}(2/\sqrt{n})) \sim R(n, \Omega). \quad (13)$$

Для доказательства асимптотического равенства (12) нам необходимо получить нижнюю оценку для $R(n, \bar{\Omega}(2/\sqrt{n}))$, асимптотически совпадающую с верхней оценкой из (10).

Для получения нижней оценки $R(n, \bar{\Omega}(2/\sqrt{n}))$ воспользуемся неравенством (7). Докажем утверждение, позволяющее получить нижнюю оценку для $\bar{R}(n, \bar{\Omega}(2/\sqrt{n}))$.

Лемма. Для средней избыточности $\bar{R}(n, \bar{\Omega}(2/\sqrt{n}))$ множества источников $\bar{\Omega}(2/\sqrt{n})$ выполняется неравенство:

$$\bar{R}(n, \bar{\Omega}(2/\sqrt{n})) \geq \frac{h_{1/\sqrt{n}}(\Omega)}{n} + \frac{c}{n}.$$

Доказательство. Обозначим через $\mu(\bar{\Omega}(2/\sqrt{n}))$ меру множества источников $\bar{\Omega}(2/\sqrt{n})$ с плотностью $f(\theta) = c / \sqrt{\prod_{i=1}^K \theta_i}$. Согласно определению средней избыточности получаем:

$$\bar{R}(n, \bar{\Omega}(2/\sqrt{n})) = \inf_{\Phi} \int_{\bar{\Omega}(2/\sqrt{n})} \frac{f(\theta)}{\mu(\bar{\Omega}(2/\sqrt{n}))} r(n, \theta, \Phi) d(\theta). \quad (14)$$

Согласно теореме Шеннона [2, 6] правая часть (13) достигает своего минимума при кодировании Φ^0 , которое каждому слову u , $u \in A^n$ ставит в соответствие слово $\Phi^0(u)$ длины

$$|\Phi^0(u)| = -\log \int_{\bar{\Omega}(2/\sqrt{n})} \frac{f(\theta)}{\mu(\bar{\Omega}(2/\sqrt{n}))} P_0(u) d(\theta).$$

Так как для любой положительной функции $g(\theta)$ выполняется неравенство

$$\int_{\bar{\Omega}(2/\sqrt{n})} g(\theta) d\theta \leq \int_{\Omega_0} g(\theta) d\theta,$$

то

$$|\Phi^0(u)| \geq -\log \int_{\Omega_0} f(\theta) P_{\theta(u)} d\theta + \log \mu(\bar{\Omega}(2/\sqrt{n})).$$

Принимая во внимание равенство

$$\mu(\bar{\Omega}(2/\sqrt{n})) = c \|\Omega(2/\sqrt{n})\| / N^{k-1/2}$$

и определение $h_{1/\sqrt{n}}(\Omega)$, получаем:

$$\left| \varphi^0(u) \right| \geq -\log f(\theta) P_0(u) d\theta + h_{1/\sqrt{n}}(\Omega) - \frac{k-1}{2} \log n + c.$$

Умножив обе части на $P_0(u)$ и просуммировав по u , получаем

$$\sum_{u \in A^n} P_\theta(u) \cdot \left| \varphi^0(u) \right| \geq - \sum_{u \in A^n} P_\theta(u) \log f(\theta) P_\theta(u) d\theta + h_{1/\sqrt{n}}(\Omega) - \frac{k-1}{2} \log n + c.$$

Разделим обе части последнего равенства на n и вычтем из обеих частей неравенства энтропию $H(\theta)$. Учитывая определения, данные выше, получаем:

$$r(n, \theta, \varphi^0) \geq r(n, \theta, \bar{\varphi}^0) + \frac{h_{1/\sqrt{n}}(\theta)}{n} - \frac{k-1}{2} \frac{\log n}{n} + \frac{c}{n}, \quad (15)$$

где кодирование $\bar{\varphi}^0$ каждому слову u ставит в соответствие слово $\bar{\varphi}^0(u)$, для которого

$$\left| \bar{\varphi}^0(u) \right| = -\log f(\theta) P_\theta(u) d\theta.$$

В [5] доказано, что для любого источника $\theta \in \Omega_0$ выполняется неравенство:

$$r(n, \theta, \bar{\varphi}^0) \geq \frac{k-1}{2} \frac{\log n}{n} + \frac{c}{n}. \quad (16)$$

Из (15) с учетом (16) имеем:

$$r(n, \theta, \varphi^0) \geq \frac{h_{1/\sqrt{n}}(\theta)}{n} + \frac{c}{n}. \quad (17)$$

Из (14), используя (17), получаем

$$\bar{R}(n, \bar{\Omega}(2/\sqrt{n})) \geq \frac{h_{1/\sqrt{n}}(\Omega)}{n} + \frac{c}{n}.$$

Лемма доказана. □

4. Избыточность универсального кодирования произвольного множества источников без памяти

Основное утверждение работы можно сформулировать в виде утверждения.

Теорема. Для произвольного подмножества источников Ω , $\Omega \subset \Omega_0$ для избыточности $R(n, \Omega)$ универсального равномерного по входу кодирования выполняется асимптотическое равенство

$$R(n, \Omega) \sim \frac{h_{1/\sqrt{n}}}{n} + \frac{c}{n}.$$

Доказательство. Верхняя оценка $R(n, \Omega)$ следует из асимптотического неравенства (4), доказанного в [1]. Нижняя оценка вытекает из леммы очевидного неравенства

$$R(n, \bar{\Omega}(2/\sqrt{n})) \geq \bar{R}(n, \bar{\Omega}(2/\sqrt{n}))$$

и соотношения (13). Теорема доказана. □

5. Заключение

Получена нижняя оценка избыточности универсального кодирования через $\mathcal{E}$ -энтропию множества источников, совпадающая с верхней оценкой, доказанной в [1]. Тем самым получено асимптотически точное значение избыточности для произвольного множества источников без памяти.

Литература

1. Трофимов В. К. Универсальное кодирование произвольного множества источников без памяти // Вестник СибГУТИ. 2018. № 4. С. 30–34.
2. Шеннон К. Математическая теория связи. Работы по теории информации и кибернетике. 1963. С. 243–332.
3. Кричевский Р. Е. Длина блока, необходимая для получения заданной избыточности // ДАН СССР. 1965. Т. 171, № 1. С. 37–40.
4. Кричевский Р. Е. Связь между избыточностью кодирования и достоверностью сведений об источнике // Проблемы передачи информации. 1968. Т. 4, № 3. С. 48–57.
5. Krichevsky R. E., Trofimov V. K. The performance of universal encoding // IEEE Transactions on Information Theory. 1981. V. 27, № 2. P. 199–207.
6. Галлагер Р. Г. Теория информации и надежная связь. М.: Советское радио, 1974.
7. Витушкин А. Г. Оценка сложности задачи табулирования. М.: Гос. изд. физико-математической лит., 1959. 228 с.

Статья поступила в редакцию 09.09.2019.

Трофимов Виктор Куприянович

д.т.н., профессор, зав. кафедрой высшей математики СибГУТИ (630102, Новосибирск, ул. Кирова, 86), e-mail: trofimov@sibsutis.ru.

Evaluation of the universal coding redundancy of an arbitrary set of sources without memory

V. K. Trofimov

This article considers a lower estimation of the universal coding redundancy of an arbitrary set of sources without memory coinciding in decreasing order with the upper estimation obtained by the author in the previous work.

Keywords: coding, redundancy, entropy, storage and processing of information, message source.

Российский опыт внедрения систем менеджмента качества

А. А. Бычков, А. И. Карпович

Статья посвящена анализу российского опыта внедрения систем менеджмента качества. В статье приводится историческая ретроспектива, выделяются общие черты, достоинства и недостатки реализуемых подходов, указывается связь с наиболее распространенными мировыми практиками. В заключении авторами даются рекомендации по предотвращению наиболее типичных ошибок при внедрении систем менеджмента качества. Статья имеет методический и практический интерес для организаций, внедряющих системы менеджмента качества.

Ключевые слова: система менеджмента качества, СМК, бережливое производство, производственная система, Toyota production system, TPS, Lean.

1. Введение

В предыдущей статье [1] мы рассмотрели наиболее распространенные зарубежные модели оценки производственных систем, применяемые компаниями в рамках внедрения систем менеджмента качества (СМК). К таковым относятся:

- Global Benchmarking, GBM (оценка корпорации Toyota) и её аналоги;
- практическая программа революционных преобразований на предприятии «Система 20 ключей»;
- оценка по критериям различных премий по качеству (Деминга, Болдриджа, Европейского фонда управления качеством).

Данные модели являются широко растиражированными инструментами менеджмента качества и в силу глобализации современной экономики все чаще внедряются в различных странах в виде так называемых “best practices” (лучших мировых практик). Разумеется, не является исключением и Россия, где данный процесс стартовал с конца 90-х годов XX века – сначала на базе крупных, зачастую государственных, предприятий, но постепенно всё более и более активно распространяясь на средний и малый бизнес.

В наше время во многом прежде всего отраслевая, а не национальная принадлежность определяет основные бизнес-процессы компаний, которые благодаря доступности информации легко копируются и затем адаптируются к существующим территориальным условиям. В связи с этим можно объяснить своего рода бум применения различных инструментов и методов качества, наблюдаемый на российских предприятиях. Но так ли просто, а главное, всегда ли разумно перестраивать деятельность компании под японские или западные «рельсы»? Как известно, к хорошему быстро привыкаешь. Почему же данное утверждение не совсем справедливо касательно построения современных производственных систем на отечественных предприятиях? Проблема, на наш взгляд, кроется в ряде ошибок, допускаемых при заимствовании зарубежного опыта, речь о которых пойдет ниже.

2. Игнорирование отечественного опыта

Говоря о внедрении систем менеджмента качества, мы традиционно подразумеваем технологии управления, основанные на опыте японского автоконцерна Toyota либо их американских аналогов, систематизированном в середине XX века в виде концепции бережливого производства. Также несколько особняком стоит концепция 6 сигм, получившая известность благодаря применению (практически тех же методик) в корпорациях Motorola и General Electric. В связи с этим внедрение методов качества и бережливого производства осуществляется на базе зарубежного опыта, по зарубежной литературе и, если позволяют финансы, с привлечением зарубежных консультантов. Как следствие, воспринимается зачастую данный процесс как заимствование извне, как нечто чуждое, противоестественное нормальному ходу вещей, и как своего рода «очередная стройка светлого будущего, на руинах старого мира».

Но если задуматься, внедрение инструментов качества в начале XX века стало ответом производителей на переход к массовому производству в условиях ограниченности ресурсов при все более возрастающей рыночной конкуренции. Неудивительно, что и научная мысль, как это часто бывает в таких случаях, развивалась в наиболее экономически развитых странах. Отечественная экономика того времени также была знакома с подобными сложностями и пыталась их решать.

Приведем лишь несколько примеров, к сожалению, практически неизвестных сегодня так называемым специалистам по СМК и тем более широкой публике:

- в 1914 в России вышла книга А. В. Панкина «Научная организация труда»;
- в 1921 году опубликована Памятка «Как надо работать» директора Центрального института труда (ЦИТ) Алексея Капитоновича Гастаева, практически полностью описывающая популярную в наше время японскую систему 5S, метод «мышинного хвостика» и т.д. [2];
- Осип Аркадьевич Ерманский ввёл в научный оборот сам термин «научная организация труда», или НОТ, в 1922 году вышла его книга «Научная организация труда и производства и система Тейлора», отмеченная В. И. Лениным как «очень полезная и очень хорошая», по его словам: «Научиться работать – это теперь главная, действительно общенародная задача Советской республики» [3].

Как и бережливое производство, НОТ предполагает грамотную организацию рабочих мест и их рациональное обслуживание, совершенствование приемов и методов работы, создание наиболее благоприятных условий, соблюдение работниками дисциплины и т.д. [4].

Нетрудно заметить, насколько основные принципы НОТ близки к современным зарубежным теориям. Ниже приведем некоторые из них.

Принцип эффективности – при достижении результатов необходимо стремиться к минимальным затратам ресурсов.

Принцип комплексности – при организации труда необходимо учитывать факторы внешней и внутренней среды.

Принцип системности – любые мероприятия по изменению организации труда должны соответствовать главным целям организации.

Принцип динамичности – при изменении условий рынка и производства организация труда также должна перестраиваться.

Принцип легитимности – должно строго соблюдаться действующее законодательство и принятые нормы.

Принцип положительного отношения работников к предприятию – нужно создать условия труда, обеспечивающие преданность сотрудника своей компании.

Говоря современным языком, можно отметить, что при общем фундаменте НОТ имеет более человекоориентированный характер, нежели концепции Lean или 6 сигм, ориентированные на максимизацию прибыли.

В 1928 году в журнале «Русский колокол» была опубликована статья И. А. Ильина «Спасение в качестве», в которой озвучивается идея «русского качества» [5], по своему масштабу вполне сопоставимая с философией Кайдзен.

С середины XX века в СССР разрабатывались собственные системы управления качеством (БИП – бездефектное изготовление продукции, СБТ – система бездефектного труда и др.), что в последующем привело к появлению комплексных систем управления качеством продукции (КСУКП).

В этот же период Генрихом Альтшуллером была разработана так называемая теория решения изобретательских задач (ТРИЗ), во многом перекликающаяся с философией Кайдзен.

ТРИЗ отвечает на вопрос: каким образом можно преодолеть затруднения при построении производства? Альтшуллер систематизировал способы изобретательства и доказал существование единых принципов и сходных приемов постановки и решения задач вне зависимости от отраслевой принадлежности. В результате он сформировал единые методы и приемы постановки и решения задач.

При решении задач в ТРИЗ делается упор на увеличение степени идеальности системы, что похоже на философию Кайдзен; основные цели Lean и ТРИЗ (максимальное снижение потерь и совершенствование производства) также совпадают [4].

Другими словами, вплоть до распада СССР в стране развивалась собственная теория управления качеством, охватывающая практически все актуальные вопросы, за исключением пресловутой удовлетворенности конечного потребителя, что, как известно, было обусловлено особенностями плановой экономики.

Японии после Второй мировой войны было запрещено содержать собственную армию. Возможно, именно это условие послужило триггером для так называемого «японского чуда», позволив направить высвободившиеся ресурсы и творческий потенциал в потребительский сектор экономики, где традиционная жесткая конкуренция, вызванная высокой плотностью населения, заставила производителей обратить внимание на потребительские свойства товара.

Как бы то ни было, нужно понимать, что теория менеджмента качества стала ответом науки на вызовы индустриальной и постиндустриальной экономики, она обусловлена естественным ходом вещей и, по сути, является критерием перехода организации на новый уровень развития. При этом абсолютно не важно, работает ли эта организация в Европе, Азии, Латинской или Северной Америке.

Прояснив данный аспект, мы понимаем, что при внедрении систем менеджмента качества, бережливого производства, Кайдзен речь идет не о слепом иностранном заимствовании в духе лихих революционных преобразований, а о планомерном развитии, опирающемся в том числе на собственный опыт. Не личная воля главного руководителя, а сама рациональность требует от нас этих перемен.

К сожалению, после распада Советского Союза в стране наступил вакуум. Опыт крупных, зачастую полужакрытых предприятий остался либо неизвестен представителям зарождающегося частного бизнеса, либо отрицался ими как неэффективный пережиток старого строя. Термин «рационализатор» стал чем-то шуточным, а призыв к порядку звучал чуть ли не как возврат в тоталитарное прошлое.

Неудивительно, что на это «пустое» место пришли западные идеи, «обернутые» в красивые учебники и непонятные иностранные слова, – идеи, подтвержденные многочисленными историями успеха процветающих компаний и брендов, по образу и подобию которых российским предпринимателям предлагалось выстраивать собственный бизнес.

По данному пути пошли в крупнейших российских компаниях, например, РАО «ЕЭС», ПАО «Сбербанк», ГК «Росатом», ГК «ГАЗ» и др., на сайтах и в нормативных документах которых нет ни слова об отечественном опыте построения систем качества. Вместо этого прямым текстом говорится о необходимости заимствования наработок «страны качества» Японии [6], приводится полный перечень инструментов TPS или бережливого производства [7].

Преподнесение изменений подобным родом, по мнению авторов, закладывает камень под фундамент производственных систем, тиражируемых на российских предприятиях, что, усиливая другие, сопутствующие данному процессу проблемы, в конечном счете приводит к непониманию и отторжению реализуемых преобразований.

3. Не образ деятельности, а образ жизни

Следующий по важности фактор, обуславливающий успешность внедрения инструментов качества, заключается в том, что принципы организации труда тесно связаны и напрямую вытекают из принципов организации того или иного общества. Человек, становясь работником, остается прежде всего частью того общества, в котором он сложился как личность. Производственные процессы или правила нельзя рассматривать отдельно от среды, в которой они применяются, зачастую они являются переложением общепринятых законов и моделей поведения на производственную деятельность.

В частности, согласно современной трактовке, в основе понятия «кайден» лежат не методы совершенствования системы менеджмента качества, а философия жизни человека, заключающаяся в его непрерывном стремлении к совершенствованию. Без преувеличения можно сказать, что всё японское общество формируется на базе данного принципа, находящего различные интерпретации в системах воспитания, образования, градостроительства, экономики и т.д. В связи с этим требования инструментов TPS воспринимаются работниками как естественные и, главное, не противоречащие их внутренним убеждениям. Подобная закономерность прослеживается и в других странах, в которых сложились собственные модели систем менеджмента качества.

В России же в силу исторических причин концепция «русского качества» не получила своего развития. Более того, принципы, заложенные в зарубежных моделях, отчасти идут вразрез с доминирующим в обществе поведением, что приводит к низкой поддержке инициатив по совершенствованию в среде коллектива, вплоть до их саботажа различного уровня. В качестве примера можно привести всё ту же Японию, где с детства прививается культура следования инструкциям. С первого класса дети формулируют себе цели на год и затем анализируют их достижение, на практике используя цикл Деминга (PDCA). В прошлом году широкую огласку получил случай, когда во время чемпионата мира по футболу игроки сборной Японии убрали за собой раздевалку. Для нас это было удивительно, в то время как в Японии в принципе нормально убирать за собой. За исключением ресторанов, в Токио практически не найти заведений общественного питания, где бы за посетителями убрали посуду. Это попросту не принято.

Теперь перенесем всё вышесказанное в плоскость производства. При каких условиях человеку проще следовать требованиям бизнес-процесса, ежедневно наводить порядок на рабочем месте, обращать внимание на собственные ошибки, останавливать конвейер в случае их выявления? Отношение к работе, производственному процессу, к собственным клиентам обусловлено, разумеется, личными качествами, но только благодаря определённой идеологии подобные паттерны тиражируются в массы, переставая быть чем-то исключительным.

Для наглядности приведем сравнительный анализ отношений к тем или иным принципам, заложенным в различных системах качества.

Таблица 1. Анализ ведущих принципов систем качества

Система	Ведущий принцип	Внутренние предпосылки	Отношение в России
Кайдзен	Инициатива сотрудников	Инициатива дает возможность продвижения, решения проблем	Инициатива наказуема
Бережливое производство	Неукоснительное соблюдение инструкций	Исполнительность свойственна восточным людям	Не принято читать инструкции
	Экономия ресурсов	Дефицит и дороговизна ресурсов в условиях маленькой территории	Нет недостатка ресурсов
	Чистота и порядок	Прививается с детства	Порядок как проявление «тоталитаризма»
	Клиентоориентированность	Острая конкуренция заставляет уделять внимание клиенту	Большая часть экономики сосредоточена в квазигосударственных монопольных структурах
	Прозрачность процессов	Желание иметь четкую картину мира и разбираться в сути происходящего	Практически нет навыка мыслить бизнес-процессами
	Отсутствие замалчивания проблем	Умалчивание проблемы может пагубно повлиять на меня и общество	Наличие проблем может вызвать репрессии
6 сигм	Лидерство	Присуще американскому национальному духу	Лидер рассматривается как «выскачка»
	Решения на основе достоверных данных	Западу свойственна рациональность	«Россия – женская душа»

Разумеется, подобная картина свойственна не всему постсоветскому обществу, но наряду с этим данное сопоставление легко дает возможность оценить причину тех или иных трудностей, с которыми сталкиваются руководители российских предприятий при внедрении систем менеджмента качества. Велика вероятность того, что при реализации изменений не на что и не на кого будет опереться.

В то же время на первоначальном этапе внедрения СМК данному фактору не уделяется должного внимания. Принято считать, что «если это сработало на Toyota, сработает и у нас», достаточно быть последовательным и требовать от персонала, чтобы нововведения неукоснительно соблюдались. Проблема кроется в отсутствии оценки внутренней готовности людей к подобным преобразованиям. В худшем случае сотрудникам потребуется менять не процессы, а себя и собственные привычки, что, как известно, гораздо труднее.

4. Трудности перевода

Следствием недостатка внимания к указанным факторам является прямое заимствование чужой терминологии, что дополнительно существенно усиливает эффект чужеродности реализуемых преобразований.

В наши дни ни для кого не секрет, как тесно связан язык и образы. Менеджмент качества в данном случае, разумеется, не является исключением. Множественное заимствование иностранной терминологии требует дополнительных усилий персонала на выстраивание новых связей между образами и описывающими их понятиями. Это придает своего рода вес новым практикам и методам, но в то же время затягивает обучение и адаптацию изменений.

По мнению авторов, ключевой причиной столь активного заимствования японских и западных терминов при внедрении систем менеджмента качества и построения производственных систем является то, что, как уже было указано выше, внедрение бережливого производства в России началось с крупных компаний, и тут сработал человеческий фактор. В первую очередь именно руководители данных компаний, общаясь с представителями миро-

вого бизнеса, получили представления о современных инструментах качества и повышения производительности труда. Чтобы находиться в едином информационном поле и, как говорится, общаться на одном языке, быть на равных, им пришлось впитать в себя сложившуюся, общепринятую терминологию. Неудивительно, что став в последующем проводниками новых идей, руководители и собственники бизнеса оперируют заимствованными терминами, а не существующими аналогами. Далее согласно иерархии управления «язык» руководителей был скопирован топ-менеджментом по вполне понятным причинам: желание угодить, соответствовать, показать свой уровень компетенций, быть современным.

Вторая причина – недостаточная компетентность и банальная лень сотрудников, которым делегирована функция по внедрению менеджмента качества и бережливого производства. В подавляющем большинстве случаев при внедрении изменений берутся шаблоны документов без первичной переработки и соотнесения с существующими понятиями и накопленным опытом конкретного предприятия. Рядовому персоналу приходится лишь подстраиваться под прихоти или хайп руководства.

В связи с этим столь популярны «лины», «гембы», «кайдзены», «муды», «канбаны» и т.д. Справедливости ради, отдельного внимания заслуживает и проблема перевода.

В качестве примера можно привести ошибку, допускаемую при внедрении системы 5S. Традиционно русский перевод по аналогии с японским оригиналом принято трактовать в виде 5С:

1. Сортировка.
2. Соблюдение порядка.
3. Содержание в чистоте.
4. Стандартизация.
5. Совершенствование.

При этом суть первого этапа заключается не просто в сортировке вещей на нужные и ненужные, а в последующем удалении всего лишнего. Разница на понятийном уровне между сортировкой и устранением лишнего – колоссальная.

Как следствие, вполне ожидаемая картина, когда в качестве подтверждения успешных результатов внедрения 5S – старый металлический стеллаж, на котором аккуратно выложены различные детали и запчасти. В лучшем случае – подписанные. О том, что требуется не просто отсортировать вещи, а понять, какие из них необходимы для работы, а какие нет, и в обязательном порядке избавиться от последних, зачастую речи не идет. А без понимания данного условия все последующие усилия по внедрению СМК и совершенствованию бизнес-процессов обречены на неудачу.

5. Искажение реальной картины

Продолжая затронутую выше тему о демонстрации «положительных» результатов внедрения инструментов качества и повышения производительности, нельзя не отметить разного рода манипуляции с реальным положением дел на предприятиях. Достаточно зайти на форумы сотрудников крупных федеральных компаний, чтобы понять, насколько действительность отличается от заявлений топ-менеджмента и радужной отчетности о результатах внедрения той или иной производственной системы. Причин подобной ситуации несколько и главную роль играет разрыв в понимании идей качества между руководителями высшего и среднего звена.

Согласно всем направлениям менеджмента качества функция лидера в вопросах совершенствования компании принадлежит ее непосредственному руководителю или собственнику. Кто-то из них инициирует и в последующем благодаря четкому внутреннему убеждению продвигает программу развития, выбирая ту или иную концепцию, формулируя ясные цели, КПЭ, лично контролируя реализацию отдельных мероприятий. Вся иерархия управления компании вынуждена подстраиваться под заданный вектор – и тут возникает опасность создания видимости позитивных изменений без соответствующей проработки и поддержки на

местах. В одном из своих интервью, описывая данную ситуацию, Герман Греф говорил о спирали лидерства, подчеркивая, что для положительного эффекта от реализации программ преобразований требуется, чтобы как можно больше сотрудников компании и прежде всего менеджеров высшего звена находилось на одном уровне развития с лидером [8]. Только в этом случае идеи найдут реальную поддержку и будут приняты, как действительно необходимые для дальнейшего развития бизнеса.

При этом, если проанализировать жизненный цикл компании (рис. 1), то нетрудно догадаться, что помимо ситуации, когда организация выстраивается с нуля по существующей технологии, разного рода программы развития более характерны для этапов «Зрелость» и «Упадок» и призваны предотвратить негативные последствия от деградации системы управления.

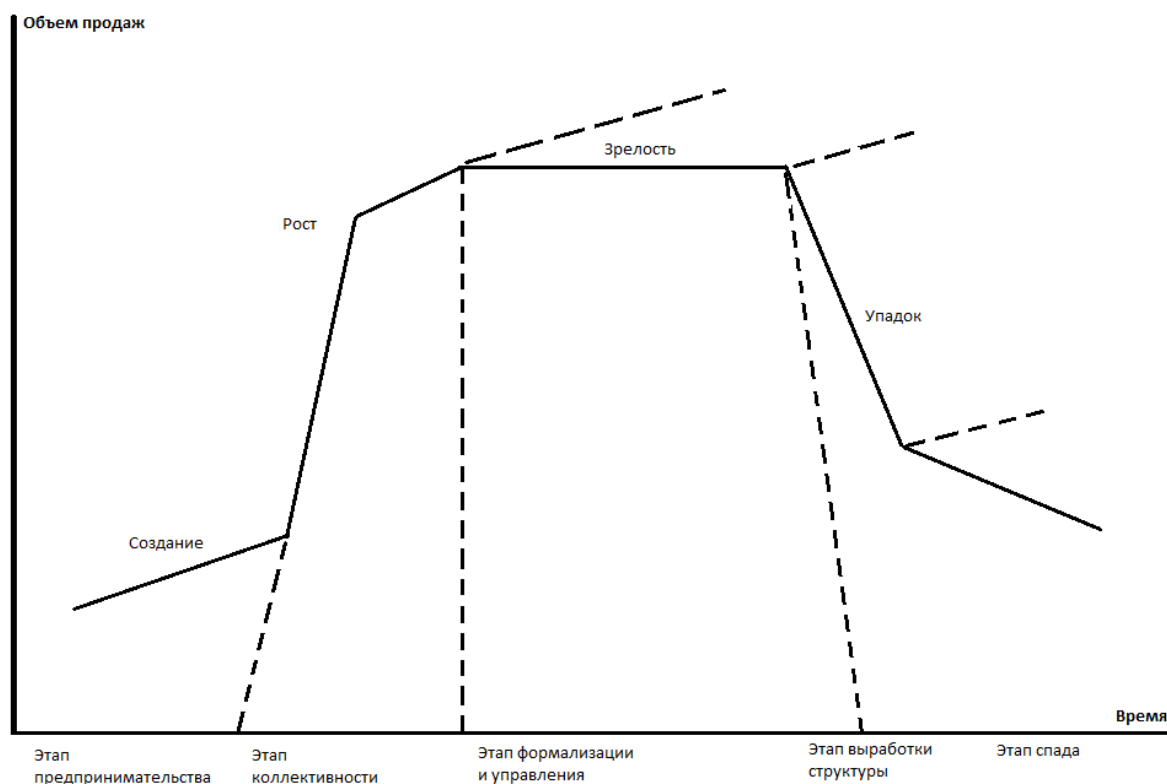


Рис. 1. Жизненный цикл организации

Как известно, основную долю менеджмента в компаниях на данных этапах составляют сотрудники с большим производственным стажем, зачастую выросшие вместе с компанией и, как это ни парадоксально, в силу прежде всего субъективных обстоятельств, более всего не желающие перемен. В данном случае выход из зоны комфорта воспринимается ими не как возможность, а как угроза потери положения, веса и собственного достатка.

Отсюда берут истоки формализм, приукрашивание и откровенное искажение действительности. Усугубляется данная ситуация тем, что в реальности у руководителя крупной компании или собственника бизнеса настолько плотное расписание, что ему не удастся регулярно посещать производство и в действительности на практике применять один из наиболее популярных инструментов менеджмента качества — «Гемба».

Как известно, рядовые сотрудники остро чувствуют подобные настроения и ведут себя соответственно. В результате растет уверенность в том, что цели по совершенствованию и повышению производительности ставят люди, не разбирающиеся в специфике конкретной работы; персонал с ностальгией апеллирует к прошлым успехам и стабильным результатам; ссылается на особенности российского менталитета либо просто, без причин и аргументов, отвергает любые улучшения [4]. Вот почему так важно доказать заинтересованность руководства в изменениях и поддержать интерес сотрудников к участию в данном процессе.

6. Работа системы в стабильном состоянии

Последним аспектом, напрямую влияющим на результаты внедрения систем качества, на котором мы бы хотели акцентировать внимание, является время. Мы живем во время невероятных скоростей, когда на смену одной технологии приходит другая, а люди эпохи потребления всё больше ориентируются на сегодняшний день, нежели задумываются о перспективе. То же самое касается и прибыли – ее стараются получить здесь и сейчас, молниеносно переводя ресурсы из одного перспективного стартапа в другой. То, на что раньше уходили годы и десятилетия, сейчас может занимать буквально часы, а знания устаревают со скоростью 2.5 года. Нам приходится приспосабливаться к жизни на высоких скоростях, и здесь возникает прямое противоречие с концепцией менеджмента качества, где во главу угла ставятся ежедневные мелкие улучшения. Мало того, что улучшать что-либо мелкими шажками сейчас нет времени, так еще требуется время на закрепление и поддержание достигнутых результатов. Другими словами, сегодня за 3–5 лет (средний срок внедрения на предприятии программ совершенствования на базе бережливого производства) персонал российских предприятий вынужден впитывать более чем восьмидесятилетний опыт автоконцерна Toyota, формируя корпоративную культуру и отношение к труду, которые выстраивались десятилетиями.

Неудивительно, что большинство изменений воспринимается сотрудниками, как принудительные, т.е. невыгодные или непонятые им, что в дальнейшем приводит к негативным последствиям.

Вне зависимости от эпохи человек остается человеком и в ситуации любых изменений по данным психологов испытывает четыре последовательно возникающих состояния:

- 1) привычка;
- 2) игнорирование;
- 3) хаос;
- 4) прорыв.

На прохождение каждой из этих стадий требуется время. Стадии игнорирования и хаоса почти всегда сопровождаются отрицательными эмоциями. Сотрудники, застрявшие на этих этапах, не поддержат перемены. Чтобы правильно закрепить достигнутый результат, требуются новые стандарты, пришедшие на смену устаревшим и неэффективным, непрерывная работа с персоналом, постоянный контроль и распространение системы на всё новые и новые бизнес-процессы – другими словами, требуются годы.

То есть без долгосрочной стратегии развития компании невозможно говорить о полноценном построении производственной системы и внедрении менеджмента качества. При этом на сегодняшний день в России доминируют жизненные стратегии с малым горизонтом планирования — пережить бы эту неделю, месяц, год, а дальше будет видно [9].

Согласно регулярным общественным опросам около 46 % россиян не знают, что с ними будет в ближайшие месяцы, еще 36 % планируют свою жизнь максимум на один-два года вперед. Бизнес в данном случае не является исключением и отражает реалии сегодняшнего дня. У большинства компаний либо отсутствует долгосрочная стратегия развития, либо она носит сугубо декларативный характер.

Подытоживая вышесказанное, напрашивается вывод, что построение систем менеджмента качества на российских предприятиях требует новых подходов, адекватных современным условиям и описанным вызовам, а именно, программа развития должна:

- разрабатываться при наличии и в соответствии с долгосрочной стратегией развития компании;
- учитывать развитие технологий Индустрии 4.0;
- обеспечивать предоставление максимально достоверных данных о результатах реализованных изменений;

- предполагать большее внимание к обоснованию среди персонала необходимости изменений, в том числе с учетом существующего отечественного опыта;
- требовать от менеджмента компаний реальной приверженности переменам, в т.ч. посредством каскадирования соответствующих КПЭ, при необходимости – быть способной на обновление команды управленцев;
- уделять особое внимание процедурам закрепления изменений;
- по возможности максимально отказаться от прямых иностранных заимствований.

Литература

1. *Бычков А. А., Карпович А. И.* Анализ зарубежного опыта оценки производственных систем в рамках внедрения систем менеджмента качества // Вестник СибГУТИ. 2018. № 3. С. 89–100.
2. *Гастаев А. К.* Как надо работать: практическое введение в науку организации труда. М.: ЛИБРОКОМ, 2011. 477 с.
3. *Ерманский О. А.* Научная организация труда и производства и система Тэйлора. М.: Государственное издательство, 1924. 395 с.
4. *Эндрю Штайн.* Философия Lean. Бережливое производство на работе и дома. М. Litres, 2017. 50 с.
5. *Иван Ильин.* Спасение в качестве // Русский Колокол. Журнал волевой идеи (сборник) [Электронный ресурс]. URL: https://books.google.ru/books?id=eFqmCgAAQBAJ&pg=PT357&dq=%D0%98%D0%BB%D1%8C%D0%B8%D0%BD+%C2%AB%D0%A1%D0%BF%D0%B0%D1%81%D0%B5%D0%BD%D0%B8%D0%B5+%D0%B2+%D0%BA%D0%B0%D1%87%D0%B5%D1%81%D1%82%D0%B2%D0%B5&hl=ru&sa=X&ved=0ahUKEwj8aDA45fjAhXx-ioKHS_LCLYQ6AEIOTAD#v=onepage&q=%D0%98%D0%BB%D1%8C%D0%B8%D0%BD%20%C2%AB%D0%A1%D0%BF%D0%B0%D1%81%D0%B5%D0%BD%D0%B8%D0%B5%20%D0%B2%20%D0%BA%D0%B0%D1%87%D0%B5%D1%81%D1%82%D0%B2%D0%B5&f=false (дата обращения: 01.07.2019).
6. *Кондратьева В. В., Кузнецов М. Н.* Показываем бизнес-процессы. М.: Эксмо, 2008. 430 с.
7. *Котельников С. А.* Производственная система Сбербанка. Знакомство с программой, 2012 [Электронный ресурс]. URL: <https://ppt-online.org/297538> (дата обращения: 01.07.2019).
8. *Греф Г.* Трансформация бизнеса в новую технологическую эпоху. [Электронный ресурс]. URL: <https://www.youtube.com/watch?v=zxOuXbjI7cg> (дата обращения: 01.07.2019).
9. *Кудрин А., Шклярук М., Комин М.* Дорога к будущему: зачем России стратегия долгосрочного развития. РБК [Электронный ресурс]. URL: <https://www.rbc.ru/opinions/economics/16/11/2016/582c12c39a79476b1a46ad4a> (дата обращения: 01.07.2019).

Статья поступила в редакцию 02.07.2019.

Бычков Александр Александрович

начальник отдела по качеству и внутреннему аудиту АО «Новосибирскэнергосбыт»,
аспирант кафедры производственного менеджмента и экономики энергетики НГТУ,
e-mail: qms@nskes.ru.

Карпович Алексей Иванович

д.э.н., профессор кафедры производственного менеджмента и экономики энергетики
НГТУ (630073, Новосибирск, пр-т К. Маркса, 20), e-mail: karpovich@corp.nstu.ru.

Russian experience in implementing quality management systems

A. A. Bychkov, A. I. Karpovich

The article considers the analysis of Russian experience in implementing quality management systems. The authors provide a historical retrospective, identify the common features, advantages and disadvantages of the implemented approaches, indicate the connection with the most common world practices. In conclusion, the authors give recommendations to prevent the most common mistakes in the implementation of quality management systems. The article has a methodological and practical interest for companies implementing quality management systems.

Keywords: quality management system, QMS, economical manufacturing, production system, Toyota production system, TPS, Lean.

Оценка качества алгоритма оптимизации состава задач автоматизированного управления

К. М. Раздобреев, М. М. Раздобреев

В статье представлена математическая модель оптимизации состава задач с учётом их информационной взаимосвязи и ресурсных ограничений. Разработан и проведен вычислительный эксперимент, результаты которого не опровергают гипотезу о том, что получаемые на основе оригинального алгоритма решения являются не только допустимыми, но и субоптимальными.

Ключевые слова: автоматизированное управление заданием, оценка качества алгоритма.

1. Общие положения

Одним из важных этапов системного анализа, определяющих функциональную структуру информационной системы автоматизированного управления, является выбор состава задач. От того, насколько обоснованно составлен комплекс задач, зависят экономическая эффективность системы и перспективы ее развития.

Анализ возможных вариантов алгоритма определения состава задач автоматизированного управления показал, что если не учитывать взаимосвязи между ними, то обсуждаемая проблема может быть решена в рамках теории линейного целочисленного программирования. Необходимость учёта информационных связей между задачами, которые можно представить в виде графа взаимосвязей задач, переводит данную постановку в класс задач нелинейного программирования, для решения которых, как известно [1, 2], общих методов не существует. Поскольку исследуемая задача является комбинаторной, решение можно получить методом прямого перебора, а количество возможных вариантов при этом составляет 2^N , где N – общее число задач. Таким образом, вычислительная сложность метода прямого перебора пропорциональна 2^N , а рассматриваемого алгоритма – пропорциональна N . В учебном пособии профессора Ю. Н. Мельникова по разработке автоматизированных информационных систем [2] даётся краткий анализ вариантов решения поставленной задачи, обеспечивающих нахождение как точного, так и приближенного решения. Однако методы, позволяющие получить точное оптимальное решение, применимы при определенных допущениях, значительно сужающих области использования предлагаемых алгоритмов, а методы поиска приближенного решения обладают существенными недостатками и не обеспечивают в ряде случаев нахождения даже допустимого решения.

2. Постановка задачи

Пусть на основе современной технологии управления определено множество задач Z_o , $N_o = |Z_o|$. Математическая модель оптимизации состава задач автоматизированного управления (с формированием рекомендуемого списка $Z_{рек}$ с количеством задач $N_{рек} = |Z_{рек}|$) имеет вид:

$$\sum_{k=1}^{N_o} \mathcal{E}_k \cdot x_k \rightarrow \max, \quad (1)$$

$$\sum_{k=1}^{N_{рек}} \Delta r_k^j \cdot x_k \leq r_{зад}^j, \quad j = \overline{1, J}, \quad (2)$$

$$x_i \leq x_k, \quad i, k \in Z_o, \quad i \neq k, \quad (3)$$

$$x_k = \begin{cases} 1, & \text{если } k \in Z_{рек}, \\ 0 & \text{в противном случае} \end{cases}, \quad (4)$$

где $\mathcal{E}_k$ – экономическая эффективность от внедрения k -ой задачи; $r_{зад}^j$ – заданное общее количество j -го вида ресурсов в системе; Δr_k^j – оценка требуемого количества ресурса j -го вида для разработки и внедрения k -ой задачи; J – количество видов ресурсов; Z_o – множество задач системы управления в количестве $N_o = |Z_o|$; $Z_{рек}$ – рекомендуемый список, $N_{рек} = |Z_{рек}|$.

Следует отметить, что условие (3) представляет математическую запись сильной взаимосвязи между задачами, для которых построен граф информационной взаимосвязи задач. Сильной взаимосвязью [2] называют такую связь типа следования от k -ой к i -ой задаче, при которой i -ая задача может быть включена в формируемый список лишь при условии включения k -ой задачи.

В статье проводится оценка алгоритма [1], основанного на сведении нелинейной задачи к многократному решению задачи линейного целочисленного программирования с обязательным перерасчетом на каждом этапе её характеристик и пошаговой модификацией исходного списка Z_o .

Алгоритм решения задачи состоит из следующих этапов (шагов).

Шаг 1. Формирование модели оптимального выбора списка задач $Z_{рек}$.

Шаг 2. Ввод первичной информации. При этом организуется диалог с пользователем, который обеспечивает ввод значений ресурсов, выделенных на разработку системы, а также параметров, характеризующих каждую задачу с учетом взаимосвязей между ними. К задачам первого ранга [2] относятся задачи, которые не используют информацию, получаемую в результате решения других задач, и не поставляют информацию другим задачам.

Шаг 3. Перерасчет характеристик задач. На этом этапе для каждой задачи заново рассчитываются (пересчитываются) параметры, характеризующие каждую задачу: экономическая эффективность $\mathcal{E}_k$ и расходы ресурсов Δr_k^j j -ого вида на разработку и внедрение k -ой задачи, $j = \overline{1, J}$. Первоначально параметры заданы в виде исходных данных для каждой задачи и характеризуют получаемую эффективность и расходы ресурсов при разработке и внедрении только данной конкретной задачи без учета их связей. По результатам триангуляции матрицы взаимосвязей между элементами графа (подграфа) определяются задачи управления, которые не поставляют информацию другим задачам, поэтому им назначается высший первый ранг. Обновлённые же параметры рассчитываются в предположении, что могут быть включены из подмножества задач первого ранга в список $Z_{рек}$ как очередная задача с учётом целевой функции (1), так и все задачи, на которые она «опирается».

Шаг 4. Решение линейной целочисленной задачи в соответствии с математической моделью (1) – (4).

Особенностью рассматриваемого алгоритма является то, что за один этап может быть включена в формируемый список одна или несколько групп взаимосвязанных задач. Группу задач составляют в соответствии с условием (3) две или более задачи, имеющие между собой

сильные информационные связи. Группа задач в алгоритме отождествляется с ведущей задачей первого ранга, которой приписываются суммарные параметры. Следует отметить, что строго оптимальное решение может быть получено в результате полного перебора вариантов элементов исходного множества.

Шаг 5. Принятие решения о включении задач в список рекомендуемых.

Если существует оптимальное решение задачи (1) – (4), то решается вопрос дополнения формируемого списка, а дальнейшее рассмотрение осуществляется на соответствующем подграфе с переходом на шаг 3, в противном случае работа алгоритма завершается, состав задач считается определенным, номера задач зафиксированы в подмножестве $Z_{рек}$.

Результат работы обсуждаемого алгоритма сравнивается с оптимальным вариантом решения, полученным методом прямого перебора.

3. Анализ статистических характеристик программного средства реализации алгоритма

При условии отсутствия в графе информационной взаимосвязи задач петель, циклов и транзитивных дуг алгоритм гарантирует [1] получение допустимого решения, близкого к оптимальному. Основными показателями, определяющими качество функционирования алгоритма, являются:

- вероятность получения оптимального решения;
- среднее отклонение допустимого решения от оптимального.

О вероятности получения оптимального решения можно судить по доле оптимальных решений в общем множестве решений, полученных в процессе многократных испытаний. Среднее отклонение допустимого решения от оптимального также может быть рассчитано по результатам эксперимента. Оптимальное решение при небольшой размерности исходной задачи, как известно, может быть получено методом прямого перебора.

Установлено [1], что характеристики работы алгоритма зависят от следующих параметров задачи: количества видов ресурсов и числа связей между задачами в процентах от максимально возможного.

Эти параметры влияют как на вероятность получения оптимального решения, так и на среднее отклонение допустимого решения от оптимального.

Поскольку предполагается, что граф взаимосвязей задач не имеет циклов, то значение параметра «доля связей между задачами», равное 100 процентам означает не то, что каждая из задач связана с любой другой задачей, а то, что количество единиц в матрице взаимосвязей между элементами графа равно $0.5 \cdot (N_0^2 - N_0)$. В реальных задачах доля ненулевых элементов в матрице инцидентности составляет порядка 20–25 %, что соответствует доле связей 40–50 %.

Поскольку алгоритм позволяет получить допустимое решение, то необходимо определить степень его близости к точному оптимальному решению. Для проведения машинного эксперимента разработана программа, позволяющая генерировать исходные данные, имеющие нормальный закон распределения.

Состав исходных данных и диапазон изменения параметров:

- матрица взаимосвязей между элементами графа, включая изменение ее размерности в диапазоне от 5 до 25 элементов;
- доля связей между задачами – в диапазоне от 0 до 100 %;
- количество видов ресурсов – в диапазоне от 1 до 25;
- оценки удельных расходов ресурсов и их запасов;
- оценки эффективности элементов графа.

Для определения значений показателей были проведены многократные испытания (более 1000).

Установлено, что доля оптимальных решений при увеличении количества видов ресурсов возрастает, обнаруживая признаки насыщения от 86 % при одном виде ресурса до 93 % при 25 видах ресурсов (рис. 1).

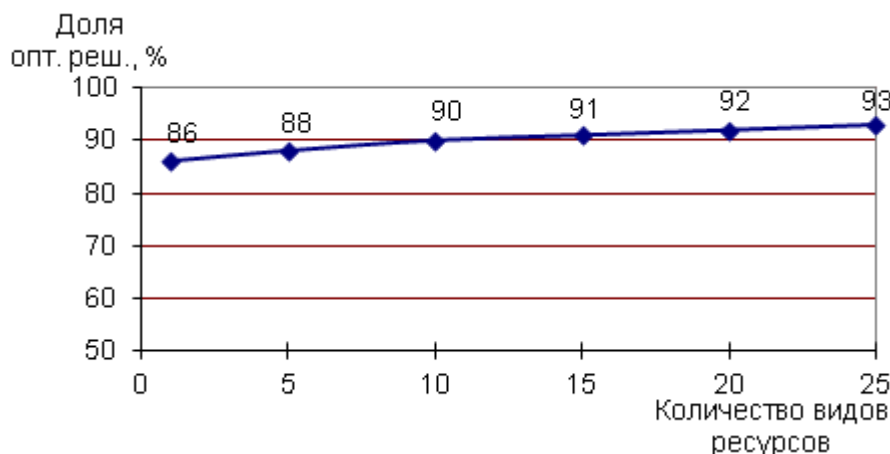


Рис. 1. Зависимость доли оптимальных решений от количества видов ресурсов (доля связей – 40 %)

Как следует из анализа результатов статистических испытаний, доля оптимальных решений изменяется также и в зависимости от числа связей между задачами от 61 % при отсутствии таковых, до практически 100 % при максимально возможном по условиям задачи варианте взаимосвязи (рис. 2).

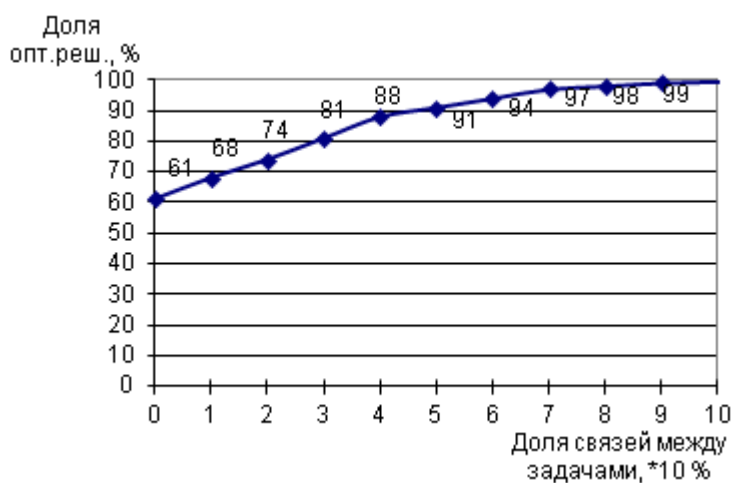


Рис. 2. Зависимость доли оптимальных решений от количества связей между задачами (количество ресурсов – 5)

Характер рассматриваемых зависимостей можно объяснить следующим образом. Множество допустимых решений включает в себя множество оптимальных решений. Следует отметить, что как при увеличении количества видов ресурсов, так и при увеличении числа связей множество допустимых решений сужается, при этом, естественно, оно стремится к множеству оптимальных решений.

Рассматриваемый алгоритм гарантирует как минимум получение допустимого решения. Отсюда ясно, что при сужении области допустимых решений возрастает вероятность нахождения оптимального решения. Так, например, при возрастании числа связей между задачами до максимально возможного значения множеств допустимых и оптимальных решений практически совпадают друг с другом и, следовательно, допустимое решение будет также и оптимальным.

Среднее отклонение допустимого решения от оптимального возрастает с насыщением от 4.1 до 13.3 % при увеличении количества видов ресурсов от 1 до 25 (рис. 3) и от 8 до 13.8 % с изменением доли связей от 0 до 90 % (рис. 4).

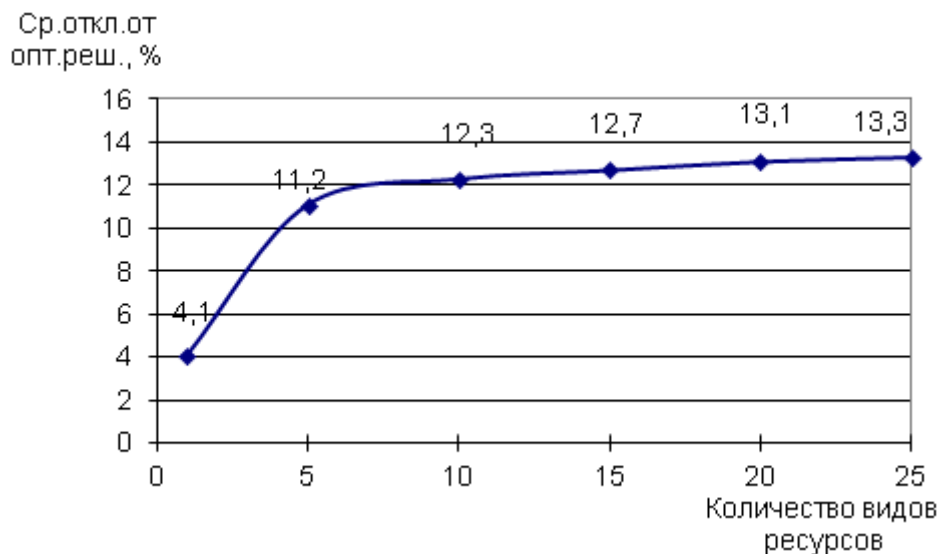


Рис. 3. Величина среднего отклонения решения от оптимального в зависимости от количества видов ресурсов (доля связей – 40 %)

При росте числа связей между задачами ошибочное включение какой-либо задачи в список рекомендуемых может привести к существенному отклонению от оптимума, поскольку с данной задачей может быть связано значительное количество других задач.

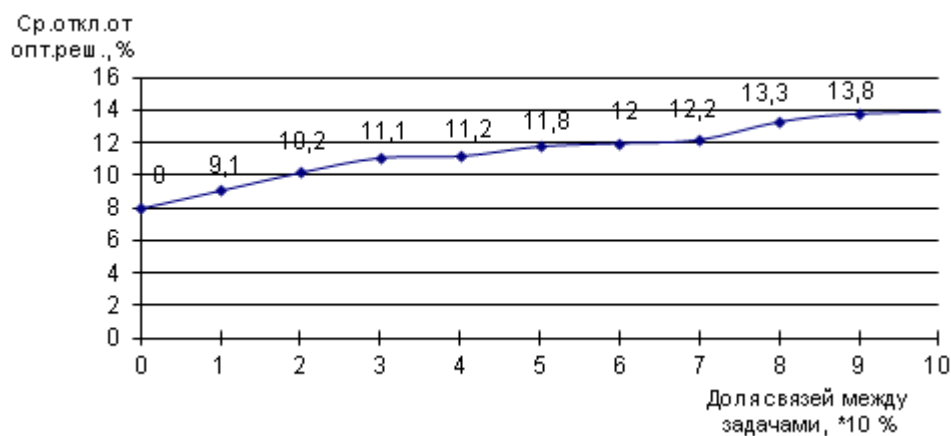


Рис. 4. Величина среднего отклонения решения от оптимального при росте количества связей между задачами (количество ресурсов – 5)

4. Заключение

В реальных информационных системах рассматриваемое количество видов ресурсов не превышает 5–10, а разреженность матрицы инцидентности графа взаимосвязей задач находится на уровне 40–50 %. Поэтому, основываясь на данных эксперимента (рис. 1, 2), можно утверждать, что вероятность получения оптимального решения составляет порядка 90 %, а возможное отклонение от оптимума в среднем не превысит 12 % (рис. 3, 4). Известно, что такие решения являются субоптимальными.

Литература

1. *Раздобреев К. М.* Алгоритм определения состава задач автоматизированного управления // Сборник научных трудов НГТУ. 1999. № 2. С. 149–153.
2. *Мельников Ю. Н.* Стадия технического задания: учебное пособие. М.: Изд. МЭИ, 1979. 60 с.

*Статья поступила в редакцию 28.05.2019;
переработанный вариант – 14.09.2019.*

Раздобреев Михаил Михайлович

к.т.н., доцент кафедры ЭОиА СГУВТ, тел. (383) 2-236-380, e-mail: razdobreev@cn.ru.

Раздобреев Константин Михайлович

к.т.н., ООО «Акватория», тел. (383) 2-236-380, e-mail: razdobreev@cn.ru.

Quality assessment of the optimization algorithm of automated control tasks composition

K. Razdobreev, M. Razdobreev

An algorithm for optimizing the composition of tasks is considered, taking into account their informational relationships and resource limitations. A computational experiment was developed and conducted. The obtained results do not refute the hypothesis that the solutions got on the basis of the original algorithm are not only admissible, but also suboptimal.

Keywords: automated control task, quality assessment of the algorithm.

Модель генерации и обработки трафика IoT параллельными коммутационными системами

Д. В. Кутузов¹, А. В. Осовский, О. В. Стукач

В статье описываются особенности генерации и обработки трафика IoT параллельными пространственными коммутационными системами. Концепция технологии интернета вещей (IoT) предполагает совмещение функций маршрутизации с функциями обработки трафика. Такое совмещение хорошо прослеживается на технологии «сеть на кристалле» (NoC). Ядром большинства систем NoC являются параллельные пространственные коммутационные системы с матрицами коммутации 5×5 . Упрощенная модель коммутатора NoC была нами реализована в виде системы массового обслуживания. Для реализации модели использовался язык программирования Python. В исследовании нами были изучены особенности обработки трафика IoT параллельной пространственной коммутационной системой, имеющей бесконечно большие входные и выходные буферы. В результате работы были получены значения максимальных и средних очередей в выходных буферных устройствах системы.

Ключевые слова: интернет вещей, параллельная коммутация, сеть на кристалле.

1. Введение

Технология интернета вещей (Internet of Things, IoT) постепенно внедряется во многие сферы человеческой деятельности. По различным прогнозам в ближайшее время ожидается включение в инфокоммуникационную структуру порядка 30 млрд умных объектов (вещей), являющихся составными единицами технологии IoT. Широкое распространение уже получили такие приложения технологии IoT, как умные дома и мониторинг здоровья [1]. Бурно развивается и программное обеспечение, позволяющее полнее использовать возможности технологии IoT, стимулирующее рынок и расширяющее сферы применения технологии.

Вместе с тем включение в информационные сети умных объектов выявило ряд особенностей обработки информационного трафика от них. Ниже мы рассматриваем наиболее распространенные модели обработки трафика IoT, особенности построения средств коммутации и предлагаем имитационную модель, позволяющую исследовать характеристики обработки трафика IoT пакетными коммутаторами.

2. Обсуждение

Исследованию особенностей генерации и регулирования трафика IoT посвящено большое количество статей. Так в работе [2] авторы утверждают, что появление большого количества умных устройств приводит к возникновению мгновенно нарастающего числа коротких сеансов. Это приводит к необходимости создания формальных моделей, гибко отражающих особенности возникающего в реальных сетях IoT-трафика.

¹ Исследование выполнено при финансовой поддержке РФФИ в рамках научного проекта № 19-07-00612.

Пример построения одной из моделей трафика IoT приведен в [3]. В статье рассмотрен поток заявок на обслуживание, возникающий на ограниченном временном интервале. По мнению авторов, моделями, наилучшим образом представляющими поток заявок в пакетных сетях передачи данных, являются так называемые модели, «не имеющие хвостов». В выводах своего исследования, авторы приводят средние значения времени задержки заявок и коэффициент вариации этой же случайной величины, которые превышают аналогичные величины, свойственные моделям с «тяжелыми хвостами».

В статье [4] авторы предлагают модель гибридного планирования приоритетов заявок для служб с различными функциями и ограничениями QoS. Согласно этой стратегии сервисам, чувствительным к задержкам, назначается высокий приоритет, и они выполняются немедленно, а сервисы, нечувствительные к задержкам, предоставляются без приоритета. Анализ такой модели обслуживания заявок показал, что её использование приводит к увеличению длины очереди ожидания обслуживания.

Одна из моделей трафика устройств IoT изложена в статье [5], где описывается применение беспроводных сенсорных сетей в медицине для контроля ЭКГ и температуры тела пациентов. Для описания модели трафика подобных устройств авторы использовали два процесса, один из них с максимальной скоростью – ЭКГ, второй – с минимальной скоростью – изменение температуры тела. На основе анализа трафика авторы пришли к выводу, что модель источника трафика имеет гауссовское распределение.

Для обработки возрастающего трафика IoT требуются технологии построения современных средств коммутации и маршрутизации, позволяющие его обрабатывать с высоким качеством. Например, в работах [6–8] описываются технологические подходы к созданию материалов и элементов, которые могут использоваться в перспективных коммутаторах трафика IoT. Ряд исследований направлен на эффективное использование буферных устройств коммутаторов для повышения пропускной способности коммутаторов [9]. Также некоторые авторы исследуют возможности и формулируют принципы построения перспективных оптических коммутаторов [10–12].

Архитектурным решением, позволяющим не только выполнять функции коммутации, но и производить обработку данных непосредственно на кристалле, является технология NoC [13–15]. В статье [16] мы показали, что подобные системы могут использоваться как основа сетей IoT, где распределение трафика между узлами совмещается с его обработкой в этих узлах, что соответствует концепции технологий IoT и 5G. Мы также предложили архитектуру параллельных коммутационных систем [17], которые могут использоваться в составе маршрутизаторов IoT, алгоритмы функционирования таких систем [18] и провели анализ их функционирования на основе классической [19] и таймированной (timed) сети Петри CPN [20].

3. Имитационная модель

Для представления системы массового обслуживания (СМО) в виде программной модели мы использовали язык программирования Python. К достоинствам этого языка можно отнести кроссплатформенность, а также наличие большого количества свободно распространяемых библиотек.

Для имитационного моделирования СМО одной из таких библиотек является queueing-tool. Эта библиотека представляет собой простой агентный симулятор систем массового обслуживания. Моделирование основано на событиях, где события состоят в числе прибытий и убытий агентов (заявок, информационных пакетов, транзактов), которые переходят из очереди в очередь в сети СМО. Однако только этой библиотеки недостаточно для создания модели СМО. Поэтому мы использовали еще некоторые другие библиотеки, такие как NetworkX, NumPy и Matplotlib. Библиотека NetworkX необходима для представления сети СМО в виде графа и выполнения действий, связанных с атрибутами этого графа, например, такими как

использование матрицы весов, визуализации графа и т.д. Библиотека NumPy также необходима для работы queueing-tool, выполняя обработку и преобразование данных. В случае если нам требуется визуализация полученных результатов в виде графиков, нам необходима библиотека Matplotlib.

Для исследования особенностей обработки трафика IoT мы использовали коммутатор с матрицей 5×5 , полагая, что обработка трафика ведется коммуникационными средствами технологии NoC или подобной ей (рис. 1). Каждый маршрутизатор NoC имеет по пять входных и пять выходных портов, условно соответствующих направлениям на север, восток, юг и запад, а также локальный элемент обработки (processing element, PE). Работа маршрутизатора состоит в том, чтобы направлять приходящие пакеты с каждого входного порта на необходимый выходной порт, и так шаг за шагом, к конечным пунктам назначения. Для реализации этой функции маршрутизатор оснащен входным буфером для каждого входного порта, коммутатором 5×5 для перенаправления трафика на желаемый выходной порт и логикой управления для обеспечения функций маршрутизации.

На рис. 2 представлена модель СМО, представляющая такой коммутатор.

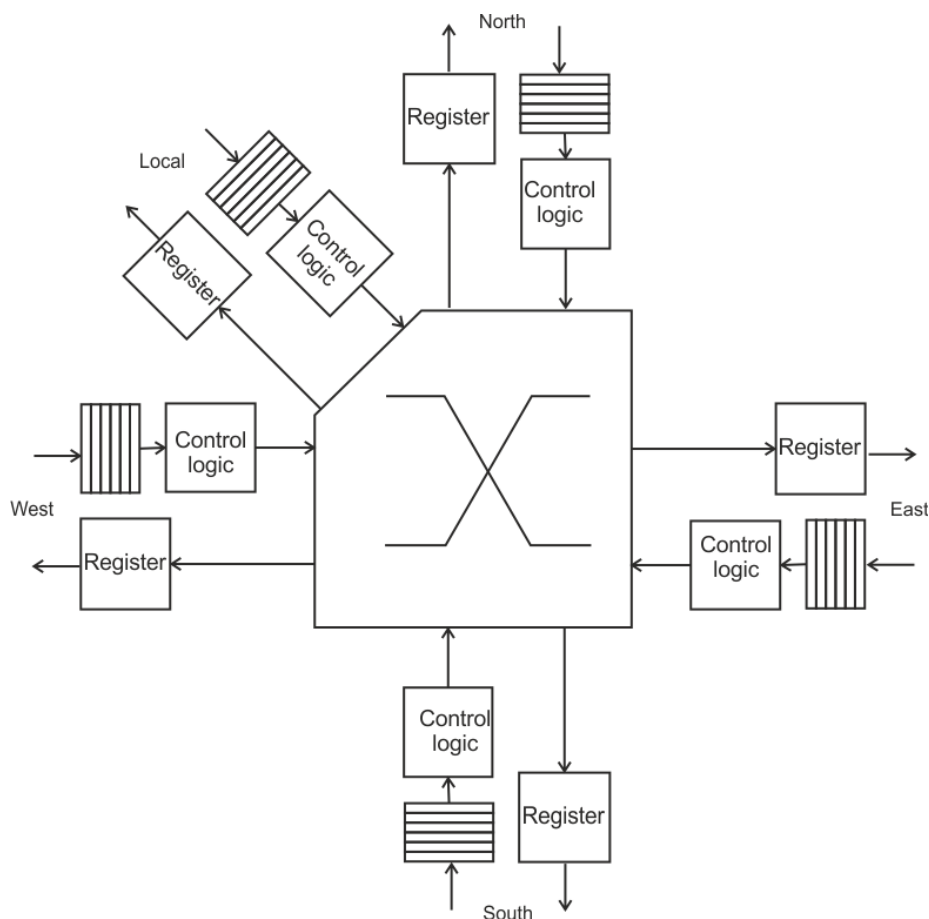


Рис. 1. Коммутатор сети обработки данных технологии NoC

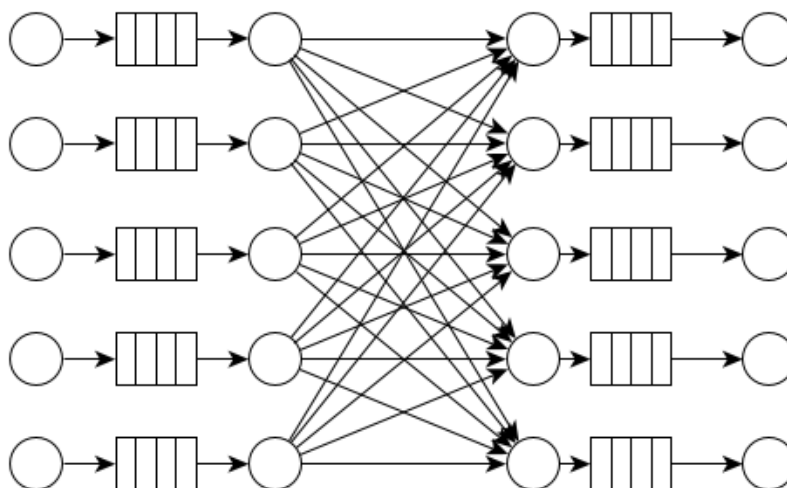


Рис. 2. Модель сети массового обслуживания, представляющая коммутатор технологии NoC

Особенностью библиотеки queueing-tool является представление очередей ребрами графа, поэтому традиционное графическое представление систем массового обслуживания (рис. 2), было нами преобразовано в граф, представленный на рис. 3.

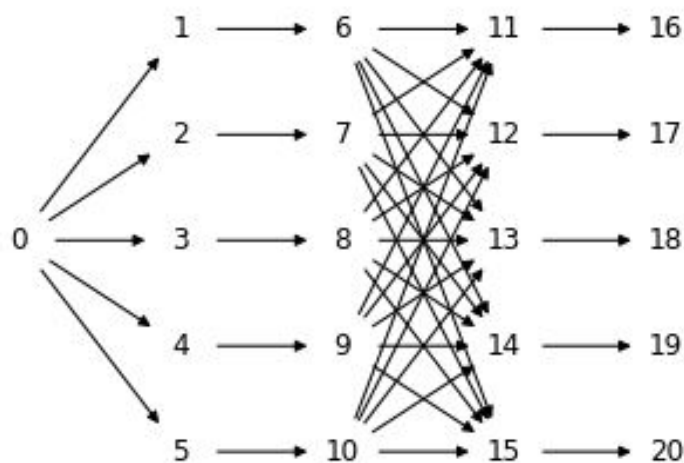


Рис. 3. Сеть СМО, использующаяся в имитационной модели

Для задания сети, представленной на рисунке, использовался следующий программный код:

```
ad = {0: [1, 2, 3, 4, 5],
      1: [6],
      2: [7],
      3: [8],
      4: [9],
      5: [10],
      6: [11, 12, 13, 14, 15],
      7: [11, 12, 13, 14, 15],
      8: [11, 12, 13, 14, 15],
      9: [11, 12, 13, 14, 15],
      10: [11, 12, 13, 14, 15],
      11: [16],
```

```

12: [17],
13: [18],
14: [19],
15: [20]
}

```

Сеть обслуживания определяется структурой `ad`, которая представляет собой словарь, где ключи – номера вершин-источников, а значения – списки номеров вершин-приемников.

Как уже сообщалось выше, помимо структуры сети нам необходимо было также задать типы ребер графа, что выполняется следующим образом:

```

ade = {0: {1: 1, 2: 1, 3: 1, 4: 1},
      1: {5: 2},
      2: {6: 2},
      3: {7: 2},
      4: {8: 2},
      5: {9: 3, 10: 3, 11: 3, 12: 3},
      6: {9: 3, 10: 3, 11: 3, 12: 3},
      7: {9: 3, 10: 3, 11: 3, 12: 3},
      8: {9: 3, 10: 3, 11: 3, 12: 3},
      9: {13: 4},
      10: {14: 4},
      11: {15: 4},
      12: {16: 4}
}

```

Переменная `ade` представляет собой словарь словарей, где ключи основного словаря, как и в предыдущей переменной `ad`, – это номера вершин-источников, ключи вложенных словарей – номера вершин-приемников, а значения вложенных словарей – тип ребра, соединяющего источник и приемник. Здесь используются ребра с типами 1, 2, 3 и 4. Кроме того, неявно используются ребра типа 0. Согласно описанию библиотеки используются следующие обозначения типов ребер: тип 1 – может принимать агенты извне сети, фактически среда, откуда поступают заявки; тип 2, 3, 4 и т.д. – очереди, свойства которых необходимо будет определить, тип 0 – терминальные ребра, его можно не задавать, они будут созданы автоматически как петли в вершинах 16 – 20.

Далее заданные нами структуры необходимо преобразовать в структуры, с которыми сможет работать библиотека:

```

g = qt.adjacency2graph(adjacency=ad, edge_type=ade)
ans = qt.graph2dict(g)

```

Функция `qt.adjacency2graph` преобразует заданные нами структуры в граф, в котором у каждого ребра/очереди есть тип, причем ребра (0, 1), (0, 2), (0, 3), (0, 4) и (0, 5) являются источниками случайных потоков заявок, поступающих извне сети, в соответствии с необходимым нам законом распределения (мы зададим его позже). Остальные ребра представляют собой очереди с теми или иными характеристиками, такими как время задержки заявки, количество устройств обслуживания и др., которые мы также зададим позже.

Функция `qt.graph2dict` – преобразует граф в словарь, формат которого использует библиотека.

Для графического представления сети обслуживания (рис. 3) можно воспользоваться функцией `draw_networkx` библиотеки `NetworkX`, но вначале необходимо задать позиции вершин графа в сетке координат, иначе в сети трудно будет ориентироваться. Определим позиции вершин для рисования графа:

```

pos = {0: (-0.25, 0),
      1: (1, 0.75), 2: (1, 0.25), 3: (1, -0.25), 4: (1, -0.75), 5: (1, -1.25),

```

```

6: (2, 0.75), 7: (2, 0.25), 8: (2, -0.25), 9: (2, -0.75), 10: (2, -1.25),
11: (3, 0.75), 12: (3, 0.25), 13: (3, -0.25), 14: (3, -0.75), 15: (3, -1.25),
16: (4, 0.75), 17: (4, 0.25), 18: (4, -0.25), 19: (4, -0.75), 20: (4, -1.25)
}

```

Прорисовываем граф и записываем его изображение в файл stuct.png:

```

nx.draw_networkx(g, pos, with_labels=True, node_size=900, node_color='w')
plt.savefig("stuct.png")

```

После того как определена структура сети обслуживания, необходимо определить параметры сети – то, как генерируются заявки (здесь $arr_f(t)$), и время обработки заявки в системе (здесь $ser_f(t)$). Время обработки заявки также может быть задано как случайный процесс с определенным законом распределения. Обе функции определяются как функции времени:

```

def arr_f(t):
    return t + np.random.normal(loc=my_lambda, scale= my_sigma)

```

```

def ser_f(t):
    return t + my_time

```

Теперь необходимо определить типы ребер и характеристики обслуживания:

```

q_classes = {1: qt.QueueServer, 2: qt.QueueServer, 3: qt.QueueServer, 4: qt.QueueServer}

```

```

q_args = {1:
    {'arrival_f': arr_f,
     'service_f': lambda t: t,
     'AgentFactory': qt.GreedyAgent
    },
    2:
    {'num_servers': 1,
     'service_f': lambda t: t
    },
    3:
    {'num_servers': 1,
     'service_f': lambda t: t
    },
    4:
    {'num_servers': 1,
     'service_f': ser_f
    },
}

```

Все ребра представляют собой структуры типа QueueServer, а характеристики определяются типом ребер (1, 2, 3 или 4).

Создаем СМО:

```

qn = qt.QueueNetwork(g=g, q_classes=q_classes, q_args=q_args, seed=13)
qn.g.new_vertex_property('pos')
qn.g.set_pos(pos)

```

Для созданной системы массового обслуживания вызываем следующие функции:

```

qn.start_collecting_data()

```

– иницилируем сбор данных от агентов сети;

```

qn.initialize(edge_type=1)

```

– инициализируем сеть, сообщая, что агенты будут поступать на ребра типа 1;

```

qn.simulate(t=time_simulate)

```

– задаем время симуляции.

Производим сбор данных от агентов:

```
qn.num_events
```

```
results = qn.get_queue_data(edge=[(vertex_source, vertex_receiver)])
```

Данные от агентов, собранные на ребре `edge` с началом в вершине `vertex_source` и концом в вершине `vertex_receiver`, будут помещены в переменную `results`.

Функция `get_queue_data` возвращает данные в виде словаря, где ключи являются идентификатором агента `agent_id`, а значения – массивом данных этого агента. Столбцы этого массива выглядят следующим образом: первый – время прибытия агента; второй – время начала работы (обслуживания) агента; третий – время выбытия агента; четвертый – длина очереди при поступлении агентов; пятый – общее количество агентов в `QueueServer`; шестой – идентификатор `QueueServer`.

После прогона модели собранные данные и СМО необходимо очищать:

```
qn.clear_data()
```

```
qn.clear()
```

4. Результаты моделирования

Для исследования особенностей обработки трафика IoT параллельными коммутационными системами мы использовали агентную модель СМО, в которой генерация агентов (фактически – заявок на обслуживание, пакетов) производится в соответствии с гауссовским распределением. Мы использовали такие параметры генерации потока заявок: среднее значение между поступлением агентов в единицах модельного времени (MTU) μ , среднеквадратичное отклонение σ , длительность обслуживания агента τ .

Интенсивность потока заявок в предлагаемой модели значительно выше, чем, например, в модели [5]. Это объясняется необходимостью изучить особенности обработки трафика в критических режимах работы маршрутизатора IoT и его ядра – параллельного матричного коммутатора.

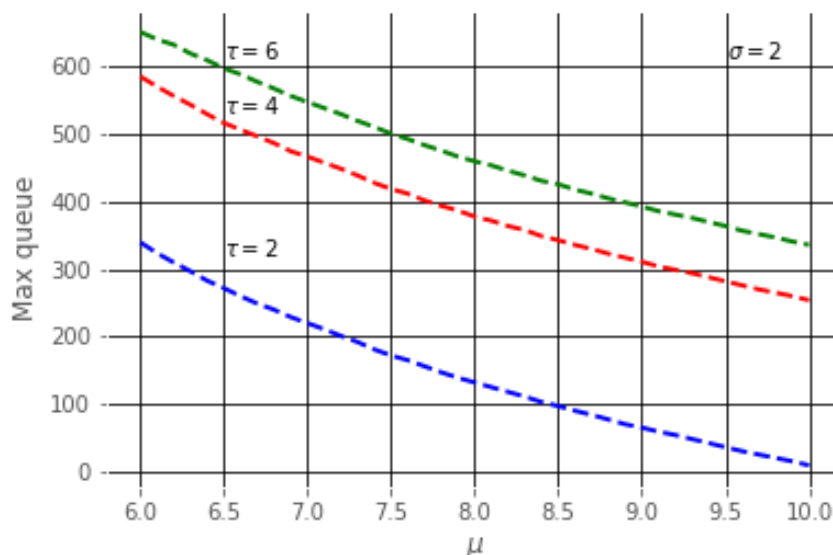


Рис. 4. Зависимость максимальной очереди от интенсивности потока заявок

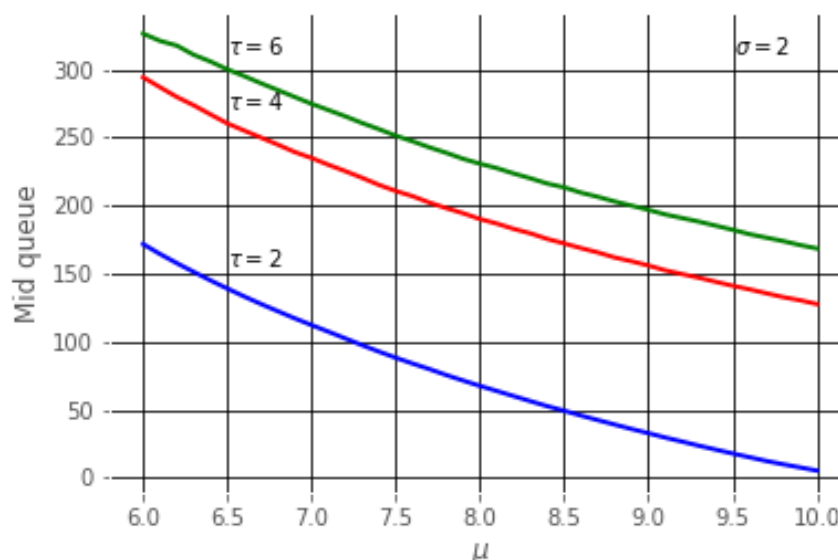


Рис. 5. Зависимость средней очереди от интенсивности потока заявок

Мы провели эксперименты и изучили максимальные и средние очереди в выходном буфере коммутатора (на рис. 3 он представлен ребром (11, 16)), изменяя интенсивность поступающего трафика. Время обслуживания каждого агента мы выбрали постоянным, что характерно для трафика, поступающего от датчиков IoT. На рисунках представлены зависимости длин очередей от среднего времени между поступлением заявок на обслуживание μ для различных значений времени обслуживания τ (при $\sigma = 2$).

5. Заключение

В этом исследовании мы показали особенности обработки потока пакетов от устройств IoT параллельными коммутационными системами. Рассмотренная здесь модель типичного коммутатора технологии NoC нами была значительно упрощена, так как реальные устройства имеют весьма сложную структуру, учесть особенности которой крайне сложно в рамках имитационного моделирования. Для генерации заявок (пакетов) мы использовали гауссовскую модель распределения вероятностей, которая является наиболее распространенной для изучения трафика IoT. Предложенная нами модель при незначительных изменениях, таких как применение других законов распределения помимо гауссовского, применении более сложных алгоритмов обслуживания, задержек в сети, может использоваться и для изучения других режимов работы подобных устройств.

Литература

1. *Perry Lea*. Internet of Things for architects: architecting IoT solutions by implementing sensors, communication infrastructure, edge computing, analytics, and security. Publicerad: 2018, Birmingham: Packt Publishing Ltd. Copyright: 2018. 505 p. ISBN 9781788470599.
2. *Wei-Hung Hsu, Qiuhui Li, Xue-Hai Han, and Chih-Wei Huang*. A Hybrid IoT Traffic Generator for Mobile Network Performance Assessment // 2017 13th International Wireless Communications and Mobile Computing Conference (IWCMC).
3. *Levakov A. K., Sokolov A. N., Sokolov N. A.* Models of incoming traffic in packet networks // T-Comm. 2015. V. 9, № 5. P. 91–94.

4. *Wen-Xiang Li; Jun Xu; Hao Jiang.* Queuing States Analysis on a Hybrid Scheduling Strategies for Heterogeneous Traffics in IOT // 2012 International Conference on Computer Science and Service System.
5. *Messier G G., Finvers I. G.* Traffic Models for Medical Wireless Sensor Networks // IEEE Communications Letters. 2007. V. 11, № 1. P. 13–15. DOI: 10.1109/LCOMM.2007.061291, URL: <https://ieeexplore.ieee.org/document/4114210>
6. *Starov D., Kutuzov D., Stukach O., Osovskiy A.* Measuring complex for studying galvanomagnetic phenomena in multigrafene layers // 2017 International Siberian Conference on Control and Communications (SIBCON), DOI: 10.1109/SIBCON.2017.7998529, URL: <http://ieeexplore.ieee.org/document/7998529/>
7. *Vytovtov K., Barabanova E., Zouhdi S.* Penetration effect in uniaxial anisotropic metamaterials // Appl. Phys. A. 2018. P. 124–137. <https://doi.org/10.1007/s00339-018-1563-z>
8. *Vytovtov K., Barabanova E.* Unusual penetration effect in ferromagnetics. Negative refraction under tangential wave incidence // Journal of Physics: Conf. Series. 2018. 1092 (1), 012164. DOI:10.1088/1742-6596/1092/1/012164
9. *Georgakopoulos G. F.* Buffered Crossbar Switches, Revisited: Design Steps, Proofs and Simulations Towards Optimal Rate and Minimum Buffer Memory // IEEE ACM Transactions on networking. 2008. V. 16, № 6. DOI: 10.1109/TNET.2007.911441. URL: <http://ieeexplore.ieee.org/document/4460578/>
10. *Vytovtov K. A., Barabanova E. A., and Podlazov V. S.* Model of Next-Generation Optical Switching System Distributed Computer and Communication Networks // 21st International Conference DCCN 2018, Moscow, Russia, September 17–21, 2018. P. 377–386.
11. *Barabanov I. O., Maltseva N. S., Barabanova E. A.* Switching cell for information transmission optical systems // 2016 International Conference on Actual Problems of Electron Devices Engineering (APEDE 2016). P. 343–347. DOI: 10.1109/APEDE.2016.7879025, URL: <https://ieeexplore.ieee.org/document/7879025>
12. *Vytovtov K. A., Barabanova E. A., Barabanov I. O.* Next-generation switching system based on 8x8 self-turning optical cell // International Conference on Actual Problems of Electron Devices Engineering (APEDE 2018). P. 306–310.
13. *Anuja Naik, Tirumale K. Ramesh.* Efficient Network on Chip (NoC) using heterogeneous circuit switched routers // 2016 International Conference on VLSI Systems, Architectures, Technology and Applications (VLSI-SATA), Bangalore, India. 10–12 Jan. 2016. DOI: 10.1109/VLSI-SATA.2016.7593043, URL: <https://ieeexplore.ieee.org/document/7593043>
14. *Wenjie Li, Yiping Gong, Bin Liu.* Performance Evaluation of Crossbar Switch Fabrics in Core Routers // 17th International Conference on Advanced Information Networking and Applications (AINA), Xi'an, China, 29–29 March 2003. DOI: 10.1109/AINA.2003.1192996, URL: <https://ieeexplore.ieee.org/document/1192996>
15. *Wen-Chung Tsai, Ying-Cherng Lan, Yu-Hen Hu, and Sao-Jie Chen.* Networks on Chips: Structure and Design Methodologies // Journal of Electrical and Computer Engineering. 2012. Article ID 509465. URL: <https://doi.org/10.1155/2012/509465>
16. *Kutuzov D. V., Osovsky A. V., Starov D. V., Motorina E. A.* Development of parallel switching facilities for 5G communication systems // Radiotekhnika. 2019. V. 83, № 3. 2019. P. 70–79.
17. *Kutuzov D., Utesheva A.* Switching Element for Parallel Spatial Systems // International Siberian Conference on Control and Communications (SIBCON-2011), Krasnoyarsk, September 15–16, 2011. P. 60–62. URL: <http://ieeexplore.ieee.org/document/6072595/>
18. *Kutuzov D., Stukach O.* Algorithms of Parallel Switching for Multistage Schemes // 2013 International Siberian Conference on Control and Communications (SIBCON), Krasnoyarsk, September 12–13, 2013. URL: <http://ieeexplore.ieee.org/document/6693642/>
19. *Kutuzov D., Osovskiy A., Stukach O.* Modeling of interconnection process in the parallel spatial switching systems // 2016 International Siberian Conference on Control and Communications (SIBCON), Moscow; 12–14 May 2016. URL: <http://ieeexplore.ieee.org/document/7491852/>

20. Kutuzov D., Osovsky A., Stukach O., Starov D. CPN-based model of parallel matrix switchboard // 2018 Moscow Workshop on Electronic and Networking Technologies (MWENT). Moscow, March 14–16, 2018. URL: <https://ieeexplore.ieee.org/document/8337180>

Статья поступила в редакцию 01.10.2019.

Кутузов Денис Валерьевич

к.т.н., доцент, доцент кафедры «Связь» Астраханского государственного технического университета, e-mail: d_kutuzov@mail.ru.

Осовский Алексей Викторович

к.т.н., доцент, главный научный сотрудник ООО «Фьюче Инжиниринг Лаб», e-mail: a_osovskiy@mail.ru.

Стукач Олег Владимирович

д.т.н., профессор департамента электронной инженерии Московского института электроники и математики им. А. Н. Тихонова НИУ ВШЭ, e-mail: ostukach@hse.ru.

IoT traffic generation and processing model with parallel switching systems

D. Kutuzov, A. Osovskiy, O. Stukach

The article describes the features of the generation and processing of IoT traffic with parallel spatial switching systems. The concept of the Internet of Things (IoT) technology involves routing functions with traffic processing functions combination. This combination is well traced on "Network on Chip" (NoC) technology. The core of most NoC systems are parallel spatial switching systems with 5×5 switching matrices. A simplified model of the NoC switch was implemented by us as a queuing system. To implement the model, the Python programming language was used. In this research, the features of IoT traffic processing with parallel spatial switching system having infinitely large input and output buffers were studied. As a result of the work, the values of the maximum and average queues in output buffer devices of the system were obtained.

Keywords: IoT, parallel routing, parallel devices, QoS, NoC.

Оценка критического радиуса изгиба графеновых антенн

А. Г. Черевко, Ю. В. Моргачев, Е. М. Ильин, А. И. Полубехин, В. И. Флегонтов

Определен параметр, характеризующий применимость гибких графеновых антенн, определена его связь с сопротивлением материала графеновых антенн постоянному току. Анализ проведен на основании эксперимента, компьютерного моделирования и мониторинга литературных данных.

Ключевые слова: графеновая антенна, печатные технологии, гибкие антенны, экологичная электроника.

1. Введение

Анализ развития рынка интернета вещей (ИВ) показывает его дальнейший экспоненциальный рост [1] (рис. 1).



Рис. 1. Тенденция потребления устройств интернета вещей

В результате в сфере ИВ будет использоваться около 3/5 всех электронных устройств уже к 2021 году.

Таким образом, с еще большей остротой возникает необходимость утилизации и переработки электронных отходов, которая стоит с начала 2000-х годов. Малоаметная в начале 2000-х годов, из-за экспоненциального роста количества доступных электронных устройств она превращается в одну из острых проблем экологии. Если оценивать количество выбрасываемых электронных устройств, которые могут быть подключены к интернету, то большая часть принадлежит к парадигме ИВ. Антенны являются одним из основных элементов ИВ. Для решения задач экологии целесообразно перейти от загрязняющих среду металлов к другим — экологичным — материалам. Таким материалом является графен, обладающий достаточно высокой проводимостью. Он дает возможность использовать экологичные подложки, например, бумагу или ткань.

Кроме того, графен — материал, на основе которого создается гибкая электроника. Для создания такой электроники может быть применена прогрессивная технология двумерной и

трехмерной печати, обеспечивающая масштабирование для удовлетворения нужд средне- и крупносерийного производства. Графеновые чернила по сравнению с металлическими аналогами кроме экологичности обладают такими преимуществами, как меньший вес, меньшая температуры высушивания, отсутствие коррозии и низкая цена [2].

Активно разрабатываются методы печати графеновых антенн на биоразлагаемых подложках, например, использование напечатанных антенн на одежде позволит увеличить расстояние передачи данных для связи до двух раз и время жизни батареи мобильных телефонов [3] и также увеличить расстояние считывания RFID меток до 5 м [4].

Несмотря на большое число публикаций по разработке гибких графеновых антенн не решен вопрос о влиянии деформаций на характеристики антенн и не определен предельный (критический) радиус изгиба антенны. Настоящая работа частично устраняет этот пробел.

2. Критические радиусы изгиба графеновых антенн

Графеновые полосковые линии, по структуре аналогичные полоску графеновой дипольной антенны (рис. 2), напечатанные на бумажной подложке, подвергались изгибу контролируемого радиуса, r . При этом измерялось их сопротивление R на постоянном токе (рис. 3). Зависимость сопротивления от радиуса изгиба $R(r)$ имеет две ясно выраженные линейные ветви (асимптотики), соответствующие большим и малым радиусам изгиба, и переходную зону (кривые 1, 2, 3 на рис. 3). Путем аппроксимации линейных ветвей прямыми линиями определялась абсцисса их пересечения и соответствующий радиус, который можно считать критическим, поскольку в этой точке производная dR/dr асимптотик зависимости $R(r)$ меняет значение (пример построения – кривая 1 на рис. 3).

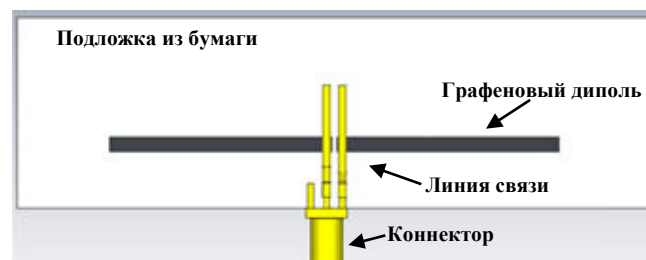


Рис. 2. Образец дипольной графеновой антенны на подложке из бумаги

Аналогично обрабатывались экспериментальные результаты для двухщелевой дипольной антенны [5] и результаты моделирования нашей дипольной антенны (рис. 2):

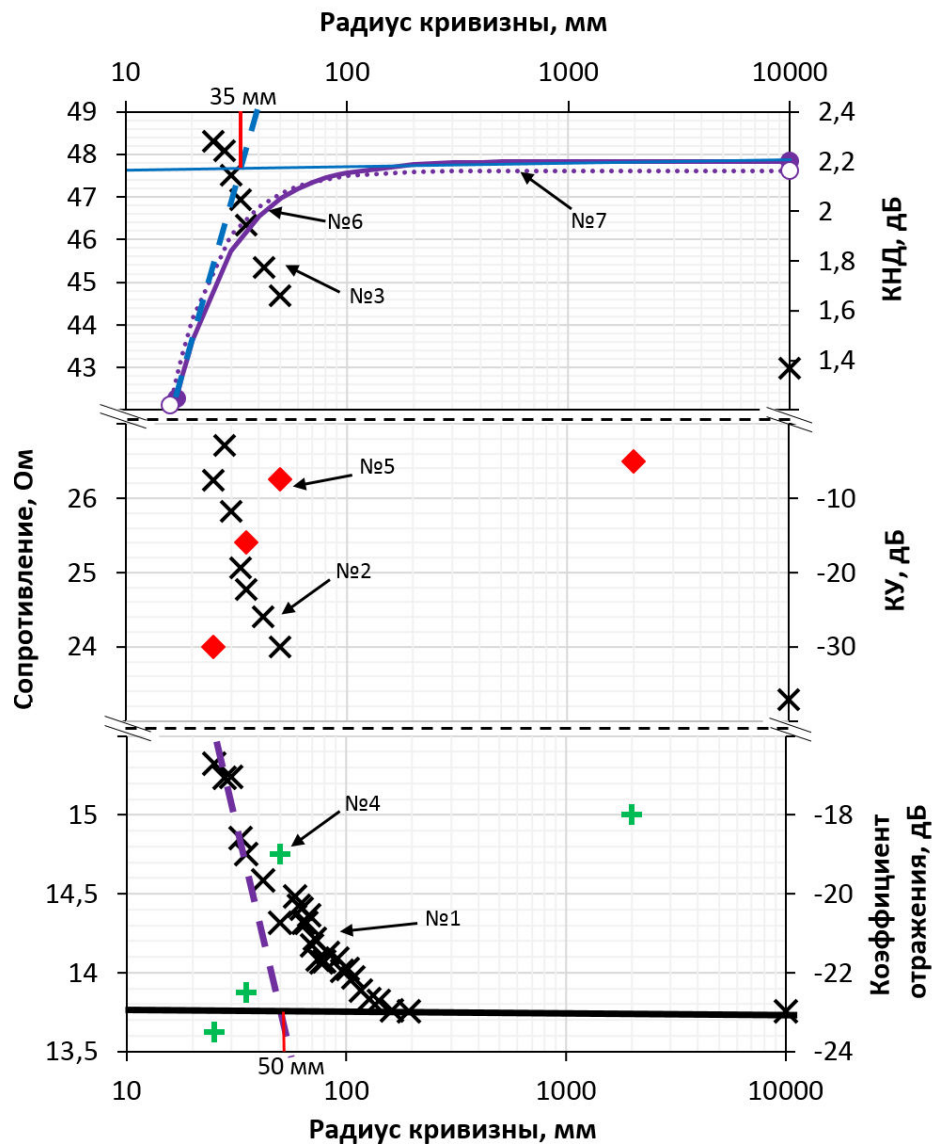
- зависимость коэффициента усиления (КУ) и коэффициента отражения (КО) напечатанной на бумаге графеновой антенны от радиуса изгиба, полученная нами путем обработки экспериментальных результатов [5] (кривые 4 и 5 на рис. 3);

- зависимость коэффициента направленного действия (КНД) дипольной графеновой антенны на подложке из бумаги и полиэстера от радиуса изгиба, полученная нами путем моделирования (кривые 6 и 7 на рис. 3).

Результаты представлены в табл. 1 и на рис. 3.

Таблица 1. Предельные(критические) радиусы изгиба графеновых антенн

Источник	Устройство	Подложка	Технология	Измеряемый параметр	r_c , мм
Наш эксперимент: Образец 1	Полосковая линия	Бумага	Трафаретная печать	Сопротивление, R	40.5
Образец 2	— ” —	Бумага	— ” —	R	50
Образец 3		Бумага		R	40
Наше моделирование: № 1	Дипольная антенна	Полиэстер		КНД	35
№ 2	— ” —	Полиэстер	— ” —	КНД	35
[5]	Двухщелевая дипольная антенна	Бумага		КУ	41
[5]	— ” —	Бумага		КО	50

Рис. 3. Зависимость параметров антенн от их радиуса изгиба, r

Х – Сопротивление полоска, + – КО, ♦ – КУ, ○ – КНД, бумажная подложка, ● – КНД, подложка полиэстер

3. Обсуждение результатов

При нашем моделировании КНД дипольной графеновой антенны, как и в случаях моделирования электромагнитных параметров графеновых антенн, описанных в литературе, ограничения, связанные с влиянием деформации на механические свойства антенны, не учитывались. Наши результаты показывают, что зависимость КНД от радиуса изгиба дает значение $r_c = 35$ мм. В то же время эксперимент по измерению сопротивления полосков графена на постоянном токе дает диапазон критических радиусов изгиба $40 \text{ мм} \leq r_c \leq 50 \text{ мм}$. Поскольку причиной изменения сопротивления в этом случае является деформация графенового полоска, т.е. изменение его механических свойств, то можно полагать, что критические радиусы гибких графеновых антенн, определяемые из их электромагнитных свойств, должны лежать в этом диапазоне. Это предположение подтверждается экспериментальными критическими радиусами, определенными по КУ и КО двухщелевой дипольной антенны [5], которые лежат в этом диапазоне.

4. Выводы

Как показал анализ, предельные возможности гибких графеновых антенн ограничиваются не их конструкцией, а механическими характеристиками графена, его устойчивостью к деформации, поскольку критический радиус изгиба антенны, связанный с механическими свойствами графена, больше критического радиуса, определяемого конструктивными особенностями антенны.

Литература

1. IoT: number of connected devices worldwide 2012-2025 | Statista [Электронный ресурс]. URL: <https://www.statista.com/statistics/471264/iot-number-of-connected-devices-worldwide/> (дата обращения: 23.10.2019).
2. Akbari M., He H., Juuti J., Tentzeris M., Virkki J., Ukkonen L. 3D Printed and Photonicallу Cured Graphene UHF RFID Tags on Textile, Wood, and Cardboard Substrates // International Journal of Antennas and Propagation. 2017. P. 1–8.
3. AUSA – Bluewater Defense & Vorbeck Introduce Wearable Antenna – Soldier Systems Daily [Электронный ресурс]. URL: <http://soldiersystems.net/2015/10/14/ausa-blue-water-defense-vorbeck-introduce-wearable-antenna/> (дата обращения: 23.10.2019).
4. He H., Akbari M., Sydanheimo L., Ukkonen L., Virkki J. 3D-printed graphene and stretchable antennas for wearable RFID applications // 2017 International Symposium on Antennas and Propagation (ISAP), 2017. [Электронный ресурс]. URL: <https://doi.org/10.1109/isanp.2017.8228746> (дата обращения: 23.10.2019).
5. Huang X, Leng T, Zhu M, Zhang X, Chen J, Chang K, et al. Highly Flexible and Conductive Printed Graphene for Wireless Wearable Communications Applications // Scientific Reports. Dec. 2015. V. 5, № 1. [Электронный ресурс]. URL: <https://doi.org/10.1038/srep18298> (дата обращения: 23.10.2019).

Статья поступила в редакцию 11.11.2019.

Черевко Александр Григорьевич

к.ф.-м.н., доцент, заведующий кафедрой физики, заведующий лабораторией физических основ телекоммуникаций СибГУТИ (630102, Новосибирск, ул. Кирова, 86), тел. 8-913-980-60-71, e-mail: perspl4@mail.ru.

Моргачев Юрий Вячеславович

инженер лаборатории физических основ телекоммуникаций СибГУТИ, e-mail: morgachev.yury@gmail.com.

Ильин Евгений Михайлович

д.ф.-м.н., профессор, ведущий аналитик инновационного технологического центра комплекса научной политики МГТУ им. Н. Э. Баумана (105005, Москва, 2-я Бауманская ул., д. 5, стр. 1), тел. 8-910-433-27-89, e-mail: evgil45@mail.ru.

Полубехин Александр Иванович

к.т.н., руководитель инновационного технологического центра комплекса научной политики МГТУ им. Н. Э. Баумана, тел. (499) 263-68-46, e-mail: polub1980@mail.ru.

Флегонтов Виталий Иванович

к.э.н, доцент, доцент кафедры экономики и финансов Одинцовского филиала МГИМО МИД РФ (143007, Московская область, г. Одинцово, ул. Ново-Спортивная, д. 3), тел. (499) 263-68-46, e-mail: Fvi55@yandex.ru.

The critical bending radius of graphene antennas estimation

A. Cherevko, Yu. Morgachev, E. Il'in, A. Polubekhin, V. Flegontov

The parameter characterizing the applicability of flexible graphene antennas is defined. Relationship of graphene antenna material resistance to direct current is determined. The analysis was carried out on the basis of the experiment, computer modeling and monitoring of literature data.

Keywords: graphene antenna, printing technology, flexible antennas, eco-friendly electronics.

Применение закодированных сверточным кодом преамбул для кадровой синхронизации

И. Н. Максимов, О. Г. Мелентьев, И. Е. Шевнина

Рассмотрены вопросы ошибок синхронизации при использовании преамбул для кадровой синхронизации. Показано, что использование преамбулы, закодированной с помощью сверточного кодера, в некоторых случаях позволяет снизить вероятность ложной синхронизации по сравнению с некодированной преамбулой. Приведена методика детектирования обратной работы и поиска начала пары элементов входного потока кодированных данных, не требующая пробного декодирования.

Ключевые слова: кадровая синхронизация, коды Баркера, обнаружение преамбулы, сверточный кодер, кодирование преамбулы.

1. Введение

В большинстве систем цифровой связи необходимым условием успешного приема информации является кадровая синхронизация (процесс, в результате которого приемник определяет момент времени начала нового кадра). Кадр – это структурированная последовательность единичных элементов, содержащая информационные и служебные поля определенных длин в строго определенной последовательности. Общеизвестным способом обеспечения кадровой синхронизации является использование преамбулы. Преамбула – это поле кадра, содержащее известную приемнику последовательность бит. Преамбула обычно размещается в начале каждого кадра. Поиск преамбулы осуществляется с помощью коррелятора, в связи с этим такие последовательности должны иметь хорошие автокорреляционные свойства. Вопросы выбора последовательностей, оптимальных с точки зрения автокорреляционных свойств, рассмотрены Баркером в работе [1]. Для подтверждения правильности принятой информации каждый кадр дополняется полем с проверочными разрядами, которые обычно вычисляются с помощью циклических кодов.

При кадровой синхронизации возможны две ошибки: пропуск преамбулы и ложная синхронизация. Пропуск преамбулы приводит к неизбежной потере кадра. Ложная синхронизация приводит к тому, что приемник отбросит принятый ложный кадр после его обработки и расчета контрольной суммы. Теоретически во втором случае можно избежать возможной потери информации, если предусмотреть достаточно большой буфер для хранения принимаемой информации и обеспечить избыток производительности процессора для её повторного анализа. На практике это не всегда возможно, поэтому мы рассмотрим случай, когда ложная синхронизация настолько же вредна, как и пропуск преамбулы.

Пусть Δ – длина преамбулы. При приеме каждого нового бита коррелятор сравнивает последовательность из Δ последних принятых бит и образец преамбулы, хранящийся в приёмнике. Количество различающихся бит в этих двух последовательностях необходимо сравнить с неким порогом d_e . Если это число меньше порога d_e , то считаем, что преамбула найдена. Очевидно, что чем ниже порог, тем выше вероятность пропуска преамбулы, а чем

выше порог, тем выше вероятность ложной синхронизации. В работе [2] авторы исследовали вопрос выбора оптимального порога при разной длине преамбулы. Из их работы, в частности, следует, что для уменьшения как вероятности пропуска преамбулы, так и вероятности ложного приема можно увеличивать длину преамбулы. Недостатком такого решения является уменьшение скорости передачи из-за увеличения объема передаваемой служебной информации.

Задача увеличения надежности обнаружения преамбулы является составной частью общей задачи повышения надежности передачи данных по каналам связи. Последняя задача успешно решается использованием корректирующего кодирования. В данной работе рассматривается использование сверточного кодирования преамбулы для улучшения качества кадровой синхронизации.

2. Математический анализ задачи

Обозначим через P_{pr} вероятность пропуска преамбулы из-за воздействия помех и P_{lo} – вероятность ложного обнаружения при случайном совпадении последовательности элементов до истинного положения преамбулы.

Пропуск преамбулы происходит тогда, когда количество ошибок в её комбинации превышает допустимый порог d_e . Для двоичного симметричного канала данная вероятность может быть найдена по следующей формуле

$$P_{pr}(\Delta, P_e, d_e) = 1 - \sum_{m=0}^{d_e} C_{\Delta}^m P_e^m (1 - P_e)^{\Delta-m}, \quad (1)$$

где Δ – длина преамбулы, P_e – вероятность ошибки по элементам.

Ложное обнаружение происходит при случайном совпадении $(\Delta - d_e)$ и более двоичных разрядов преамбулы и анализируемого потока. Пусть q – вероятность несовпадения сравниваемых двоичных разрядов. Тогда вероятность того, что количество несовпадающих разрядов на длине преамбулы не превысит допустимой величины, будет определяться следующим выражением:

$$P_{lo}(\Delta, d_e) = \sum_{m=0}^{d_e} C_{\Delta}^m q^m (1 - q)^{\Delta-m}.$$

Для двоичного симметричного канала события совпадения и несовпадения равновероятны, значит, $q = \frac{1}{2}$, и вероятность ложного обнаружения определяется выражением:

$$P_{lo}(\Delta, d_e) = 2^{-\Delta} \sum_{m=0}^{d_e} C_{\Delta}^m. \quad (2)$$

На рис. 1 приведены зависимости в координатах «вероятность пропуска – вероятность ложного обнаружения» для преамбулы длиной 16 элементов при разной вероятности ошибки и числе допустимых ошибок d_e от нуля до шести.

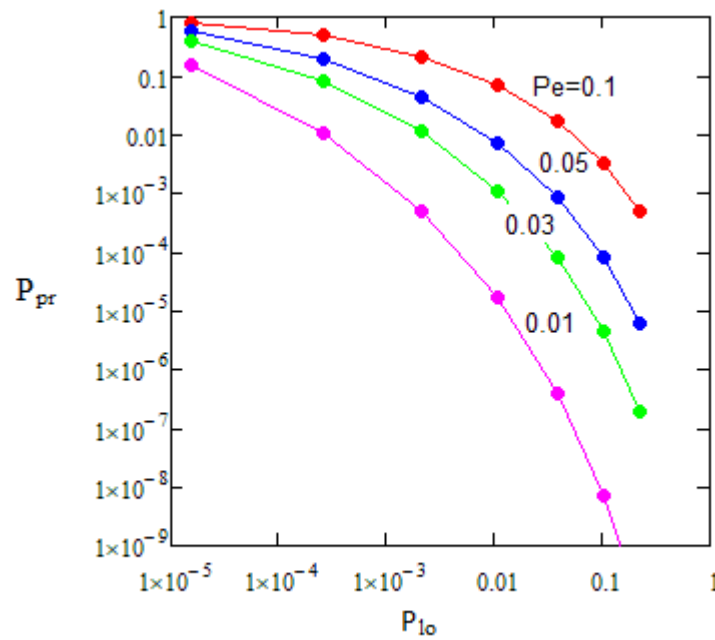


Рис. 1. Соотношение между P_{pr} и P_{lo} для $\Delta = 16$, d_e от нуля до шести и разных P_e . Каждая линия соответствует одному значению P_e . Вдоль каждой линии d_e возрастает слева направо

Рассмотрим более детально определение вероятности ложного обнаружения преамбулы на участке длиной $(L + \Delta - 1)$ элементов до истинного положения преамбулы. На рассматриваемом участке коррелятор выполнит L последовательных сравнений комбинации преамбулы с элементами входного потока. Если рассматривать только участок «чистого» шума, не содержащий в конце комбинацию преамбулы, то события отсутствия ложной синхронизации (они имеют вероятность $1 - P_{lo}$) можно считать независимыми. Тогда вероятность ложного обнаружения на всем участке будет определяться выражением:

$$P_{LO} = 1 - (1 - P_{lo})^L. \quad (3)$$

На рис. 2а приведены результаты имитационного моделирования (показаны точками) ложного обнаружения преамбулы на участке «чистого» шума длиной $(L + \Delta - 1)$ элементов (без добавления преамбулы) и оценка данной вероятности в предположении независимости событий, вычисленная по формуле (3). Моделирование приведено для последовательности Баркера длиной 13 элементов, $d_e = 1$ и $P_e = 10^{-2}$. Длины участка шума перед преамбулой выражены в длинах преамбулы.

Обязательное наличие преамбулы в конце анализируемого интервала приводит к нарушению независимости на переходном участке, когда в окно поиска начинают заходить элементы истинной преамбулы. В этом случае оценка вероятности ложного обнаружения по формуле (3) будет давать погрешность в сторону увеличения значений.

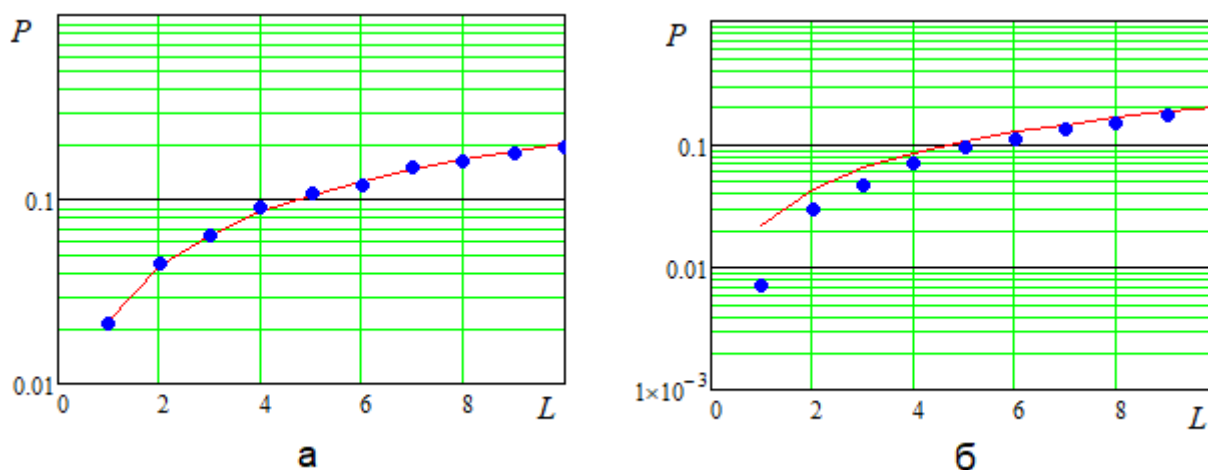


Рис. 2. Зависимости вероятности ложного обнаружения преамбулы от длины поля данных

На рис. 2б приведены аналогичные результаты моделирования и расчетов для случая, когда после L элементов шума добавлена преамбула. При проведении имитационного моделирования генерировались псевдослучайные двоичные массивы длиной L с вероятностью единиц, равной 0.5, к которым добавлялся массив преамбулы и вносились ошибки с заданной вероятностью. Далее проводился поиск преамбулы на участке длиной $L + \Delta - 1$ с шагом в один элемент. Количество испытаний для каждого значения L составляло 10 000.

Как видно из рис. 2б, вероятность ложного обнаружения при малых длинах поля данных заметно ниже оценки по выражению (3). С ростом длины поля данных выражение, полученное в предположении независимости, описывает результат с погрешностью, стремящейся к нулю. На основании имитационного моделирования для преамбул разной длины можно сказать, что при длинах поля данных, превышающих 10 – 20 длин преамбулы, результаты моделирования с высокой степенью точности аппроксимируются выражением для независимых событий.

В реальных системах длины преамбулы чаще всего составляют 13 – 49 элементов, а длина поля данных обычно превышает 1000 элементов. В указанных условиях выражение для независимых событий обеспечивает приемлемую точность.

Рассмотрим систему с одним каналом поиска и обработки преамбулы. Кадр такой системы состоит из преамбулы и поля данных, защищенного сверточным кодом и циклическим кодом для обнаружения оставшихся ошибок. После определения преамбулы система исправляет ошибки в поле данных и анализирует синдром циклического кода. Если обнаружены ошибки, поиск преамбулы возобновляется со следующего канального символа.

Для рассмотренной системы событие ложной синхронизации складывается из двух альтернатив: ложного определения преамбулы на участке $(\Delta + L)$ до истинной преамбулы или пропуск преамбулы при отсутствии ложного определения. Учитывая приведенный выше анализ, при длинах поля данных, превышающих 10 – 20 длин преамбулы, можно использовать выражения для независимых событий:

$$P_{LS} = 1 - (1 - P_{lo})^{\Delta+L} + (1 - P_{lo})^{\Delta+L} P_{pr}. \quad (4)$$

Как следует из выражения (3), для минимизации вероятности ложной синхронизации необходимо уменьшать как вероятность ложного определения, так и вероятность пропуска. Рассмотрим более подробно поведение данных вероятностей в зависимости от длины преамбулы, допустимого числа ошибок и длины информационного кадра.

Зависимости для преамбул разной длины: 16, 24, 32 и вероятности ошибки в канале $P_e = 5 \cdot 10^{-2}$ приведены на рис. 3.

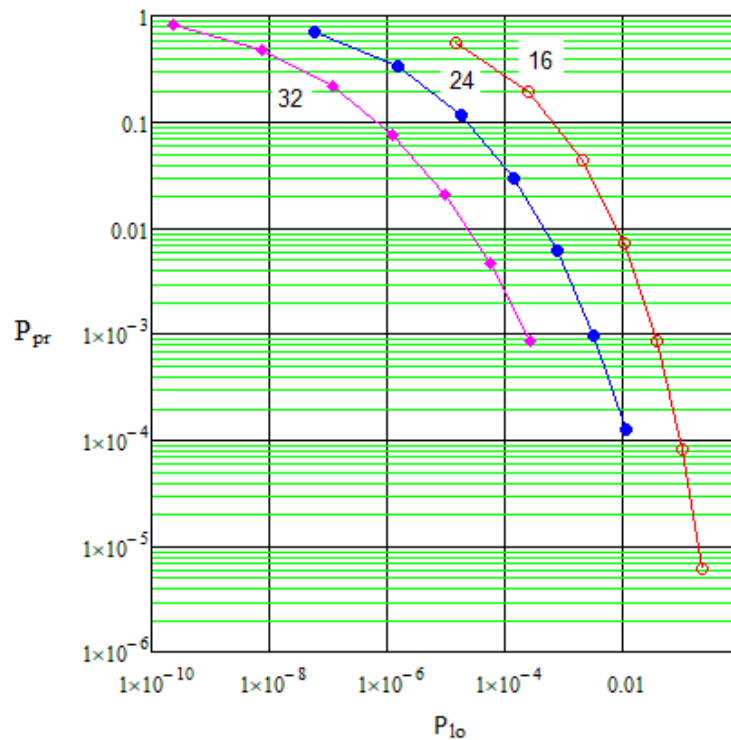


Рис. 3. Соотношение между P_{pr} и P_{lo} для $\Delta = \{16, 24, 32\}$, d_e от нуля до шести и $P_e = 5 \cdot 10^{-2}$.

Каждая линия соответствует одному значению Δ . Вдоль каждой линии d_e возрастает слева направо

Увеличение длины преамбулы заметно уменьшает вероятность ложного обнаружения, но одновременно растет вероятность её пропуска. Вероятность пропуска уменьшается при увеличении допустимого числа ошибок, при этом вероятность ложного обнаружения возрастает. Зависимости рассматриваемых вероятностей от длины информационного кадра и допустимого числа ошибок при $\Delta = 16$ и $P_e = 0.05$ показаны на трехмерных графиках рис. 4. Как видно, зависимость вероятности ложной синхронизации имеет минимум при некотором значении d_e .

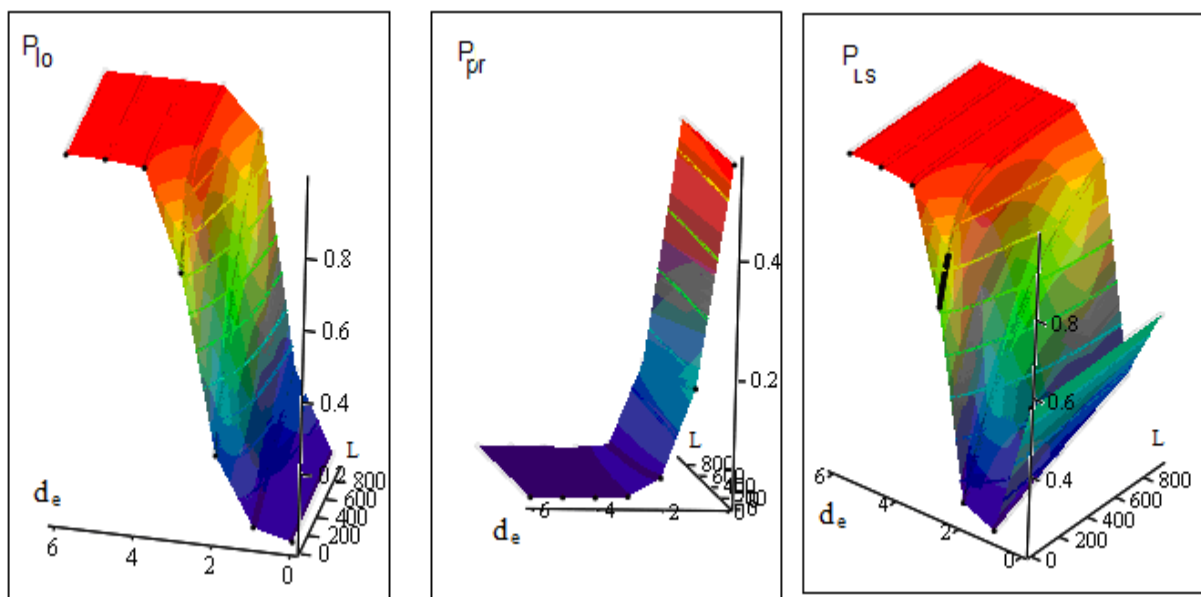


Рис. 4. Вероятности P_{pr} , P_{lo} и P_{LS} для $\Delta = 16$ и $P_e = 5 \cdot 10^{-2}$ в зависимости от d_e и L

Вероятность ложного определения преамбулы на участке чистого шума не зависит от вероятности ошибки. Как показывают результаты имитационного моделирования, приведенные выше, существующая зависимость от ошибок на переходном участке, когда в окно поиска начинают поступать элементы истинной преамбулы, уменьшается с ростом L до пренебрежимо малых значений. Для уменьшения вероятности пропуска мы предлагаем кодировать преамбулу сверточным корректирующим кодом и проводить её поиск после исправления ошибок. Решение вопросов неопределенности, возникающей при декодировании без кадровой синхронизации, рассмотрено ниже.

На рис. 5 показаны зависимости вероятности пропуска от количества допустимых ошибок для разных вероятностей ошибки в канале до и после применения сверточного кода с полиномами $g_1 = 23$ и $g_2 = 33$ при декодировании по алгоритму Витерби по жесткой схеме.

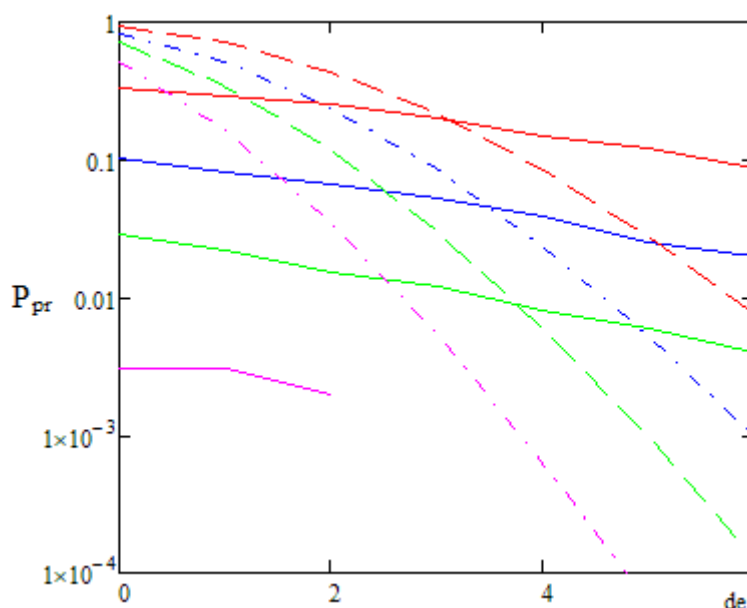


Рис. 5. Зависимости вероятности пропуска преамбулы P_{pr} для $\Delta = 24$ и $P_e = \{10, 7, 5, 3\} \cdot 10^{-2}$ в зависимости от d_e . P_e максимально для верхней линии в серии. Штриховыми линиями показаны зависимости для некодированного случая, сплошными — при применении сверточного кодирования преамбулы

Зависимости после декодирования имеют более пологий характер и указывают на выигрыш при малых значениях d_e .

Зависимости в координатах «вероятность пропуска — вероятность ложного срабатывания» для преамбул длиной 16, 24 и 32 разряда и вероятности ошибки в канале $P_e = 5 \cdot 10^{-2}$ представлены на рис. 6.

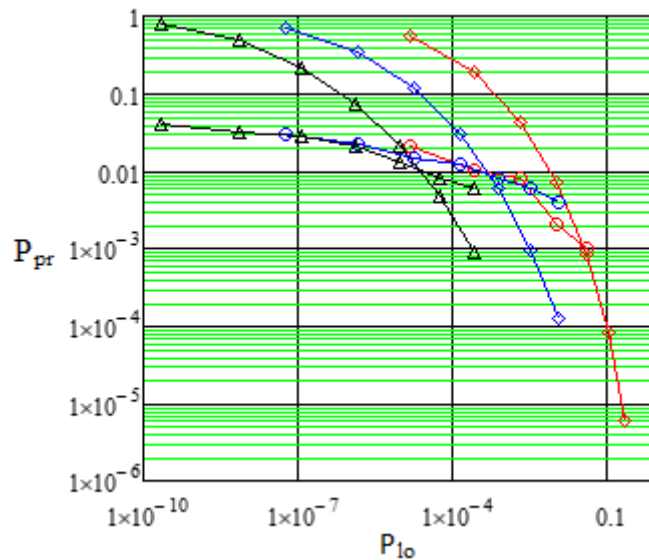


Рис. 6. Соотношение между P_{pr} и P_{lo} для $\Delta = \{16, 24, 36\}$, d_e от нуля до шести и $P_e = 5 \cdot 10^{-2}$.

Каждая линия соответствует одному значению Δ . Вдоль каждой линии d_e возрастает слева направо.

Представлены две серии линий – для кодированных преамбул и для некодированных

Поверхности вероятностей ложной синхронизации от длины информационного кадра и количества допустимых ошибок до и после применения сверточного кода показаны на рис. 7.

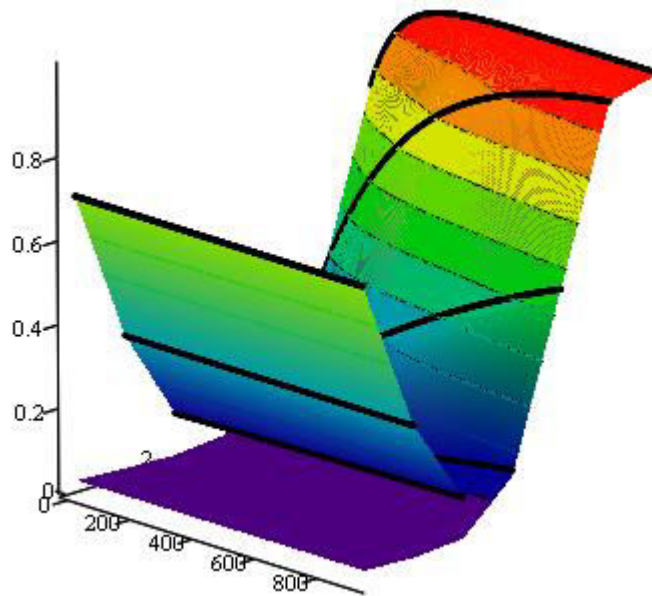


Рис. 7. Вероятности P_{LS} для $\Delta = 24$ и $P_e = 5 \cdot 10^{-2}$ в зависимости от d_e и L . Сверху располагается поверхность для некодированной преамбулы, под ней – поверхность для кодированной преамбулы

Без сверточного кода при длине информационного кадра, равной 1024 элемента, $\Delta = 24$ и вероятности ошибки в канале 0.05 минимум вероятности ложной синхронизации равен 0.132 при $d_e = 2$. После применения сверточного кода вероятность уменьшается до 0.023 при $d_e = 1$, то есть в 5.7 раз.

Рассмотрим зависимости вероятности ошибок разной кратности в преамбуле длиной 24 элемента на входе и выходе сверточного декодера при $P_e = 0.1$ и $P_e = 7 \cdot 10^{-2}$, приведённые на рис. 8. Кривая 1 отражает зависимость вероятности кратных ошибок в преамбуле без ко-

дирования, точками показаны оценки вероятностей кратных ошибок после сверточного декодера, полученные по результатам имитационного моделирования. Кривая 2 – численная оценка вероятностей кратных ошибок в комбинации преамбулы после сверточного декодера, полученная теоретически при использовании оценок параметров модели Гилберта, приведенных в таблице 1.

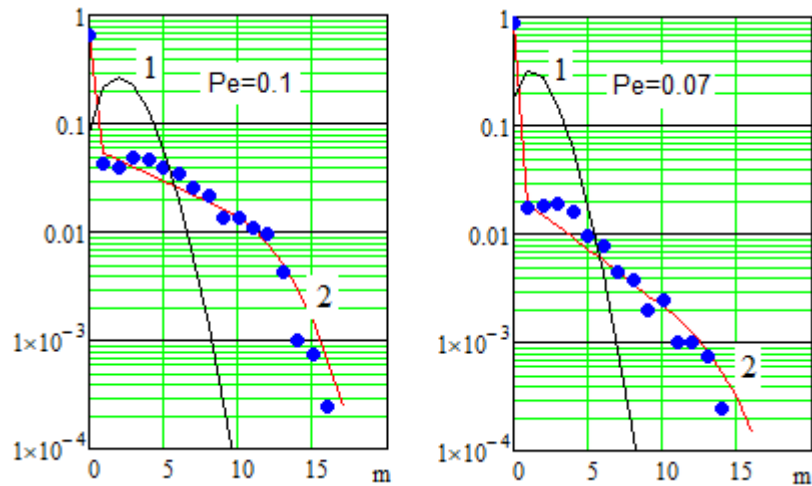


Рис. 8. Зависимости вероятности ошибок разной кратности в преамбуле длиной 24 элемента на входе и выходе сверточного декодера при $P_e = 0.1$ и $P_e = 7 \cdot 10^{-2}$

Как видим из графиков, после декодирования существенно увеличилась вероятность отсутствия ошибок, но заметно возросла вероятность большого числа ошибок в преамбуле. Скорость уменьшения вероятности заданного числа ошибок в преамбуле с ростом этого числа ошибок на выходе декодера значительно ниже, чем на входе. Такая эволюция зависимости указывает на группирующий характер ошибок после декодирования. Группирование появляется из-за того, что все одиночные ошибки исправляются, но на участках, где во входном потоке число ошибок превышает исправляющую способность кода, декодер не только не исправляет все ошибки, а может ещё и размножить их.

Рассмотренный эффект группирования определяет и характер зависимостей вероятности пропуска преамбулы от допустимого числа ошибок при использовании сверточного декодера, представленных на рис. 5 и 6.

Зависимости распределения кратных ошибок на выходе сверточного декодера достаточно точно аппроксимировались моделью Гилберта по методике, изложенной в [3] (см. рис. 8). Параметры модели Гилберта для разных значений вероятности ошибки в канале приведены в табл. 1. Напомним, что модель Гилберта имеет два состояния: «хорошее», в котором ошибок нет, и «плохое» – в котором ошибки возникают с вероятностью p . Модель характеризуется переходными параметрами: P_{gg} – вероятность сохранить «хорошее» состояние на следующем шаге, P_{bb} – вероятность остаться в «плохом» состоянии. Для потока ошибок, порождаемого этой моделью, можно вычислить такие величины, как $D_g = 1/(1 - P_{gg})$ – средняя длительность «хорошего» состояния, D_b – средняя длительность «плохого» состояния, P_d – средняя вероятность ошибок в выходном потоке. Исследование группирования ошибок на выходе декодера проводилось для сверточного кода с полиномами $g_1 = 23$ и $g_2 = 33$ (в двоичном виде $(1\ 0\ 0\ 1\ 1)$ и $(1\ 1\ 0\ 1\ 1)$). На вход кодера подавались известные псевдослучайные последовательности с преамбулой. В выходную последовательность вносились независимые

ошибки с вероятностью P_e , после чего последовательность с ошибками подавалась на декодер.

Таблица 1. Параметры модели Гилберта, имитирующей поток ошибок на выходе сверточного декодера при заданной вероятности ошибки в канале

P_e	P_{gg}	P_{bb}	p	D_g	D_b	P_d
0.1	0.988	0.933	0.5	83	15	0.075
0.07	0.997	0.9	0.49	333	10	0.015
0.05	0.999	0.88	0.495	1000	8.3	0.004
0.03	0.9998	0.79	0.63	5000	4.8	0.0005

Декодированная последовательность поэлементно суммировалась по модулю два с входной последовательностью, в результате чего формировался поток ошибок, по которому оценивались параметры модели Гилберта. Методика, изложенная в [3], основана на анализе статистики распределения серий безошибочных элементов после ошибки, оценке средних длин состояний D_g , D_b и последующем вычислении переходных вероятностей P_{gg} , P_{bb} и вероятности ошибки в плохом состоянии – p .

Из таблицы видно, что с падением вероятности ошибок в канале растет среднее значение интервала между пакетами ошибок (D_g), уменьшается длина пакета ошибок (D_b) и, соответственно, уменьшается вероятность ошибки на выходе сверточного декодера. При вероятности ошибки в канале, равной 0.01 и менее, вероятность ошибок декодирования становится пренебрежимо малой и вероятность ложной синхронизации будет определяться только вероятностью ложного обнаружения преамбулы на участке длиной $(\Delta + L)$ элементов, определяемой выражением (3).

Определив параметры модели Гилберта на выходе сверточного декодера, зависимость вероятности пропуска преамбулы можно определить аналитически, заменив в выражении (1) вероятность независимых ошибок кратности m на вероятность кратных ошибок в канале Гилберта:

$$P_{pr}(\Delta, P_{gg}, P_{bb}, p, d_e) = 1 - \sum_{m=0}^{d_e} P_{mn}(P_{gg}, P_{bb}, p, m, \Delta), \quad (5)$$

где $P_{mn}(P_{gg}, P_{bb}, p, m, \Delta)$ – вероятность того, что в преамбуле длиной Δ будет точно m ошибок, если параметры модели Гилберта равны P_{gg}, P_{bb}, p .

Методика определения вероятности кратных ошибок в канале, описываемом моделью Гилберта, представлена в [4, 5].

Имея оценки параметров модели Гилберта, характеризующей поток ошибок на выходе сверточного декодера, можно численно оценить вероятности ложной синхронизации при произвольных длинах преамбулы, поля данных и допустимого количества ошибок d_e .

3. Определение границ пар элементов кода и обратной работы

В рассматриваемой системе связи преамбула закрыта сверточным кодом, поэтому декодирование необходимо осуществлять до кадровой синхронизации. При работе сверточного декодера без кадровой синхронизации возникает неопределенность разбиения входного потока на кортежи (пары, тройки...) элементов. Для сверточного кода со скоростью $1/2$ кортеж представляет собой два канальных элемента $\{c_i^1, c_i^2\}$, сформированных на выходах сумматоров.

При когерентном приёме возможно возникновение обратной работы. Так, например, BPSK-модуляция приводит к инверсии принимаемых данных. Эту неоднозначность можно разрешить аналогичным образом, но теперь необходимо сформировать четыре последовательности S_{metr} : две – по принципу, описанному выше, а для двух других необходимо инвертировать принимаемую последовательность. Последовательность S_{metr} с минимальным числом единиц будет соответствовать правильному варианту.

Для иллюстрации сказанного на рис. 10 приведены оценки вероятностей единиц на выходе четырех рассмотренных анализаторов, полученные по результатам имитационного моделирования кода с полиномами (23, 33) при вероятности ошибки в дискретном канале $3 \cdot 10^{-2}$. Горизонтальная ось времени размечена в единичных элементах. Для обеспечения вывода анализаторов в рабочий режим перед преамбулой в кадре размещено 512 элементов, закодированных сверточным кодом. Окно каждого анализатора длиной 512 элементов инициализировано чередованием единиц и нулей.

Как видно из рисунка, на начальном участке шума, где нет закодированных данных, вероятности единиц на всех четырех анализаторах балансируют вокруг значения 0.5. С началом поступления в окна анализа закодированных данных значения на выходах анализаторов нечетных элементов начинают отклоняться в разные стороны. Анализаторы четных элементов не изменяют своего поведения. После заполнения окон элементами из канала переходной процесс заканчивается и можно принимать решение по минимальному значению вероятности единиц в окне.

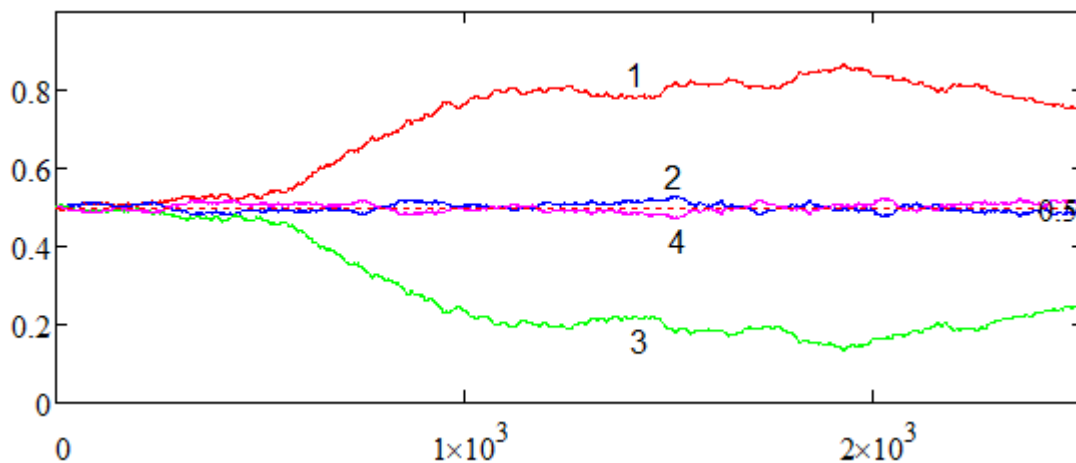


Рис. 10. Зависимости вероятности единиц на выходе анализаторов от времени:

- 1 – нечетные элементы без инверсии, 2 – четные элементы без инверсии,
- 3 – нечетные элементы с инверсией, 4 – четные элементы с инверсией

Как видно из рисунка, началом пары в данном случае будут являться нечетные элементы с инверсией.

Рассмотренная методика может быть применена для любого сверточного кода со скоростью $1/n$, при этом потребуется $2 \cdot n$ анализаторов.

4. Заключение

В работе рассмотрены компромиссы, которые возникают при использовании преамбул для кадровой синхронизации. Показано, что использование преамбулы, закодированной с помощью сверточного кодера, в некоторых случаях позволяет уменьшить вероятность ложной синхронизации в несколько раз по сравнению с некодированной преамбулой. Показано, что эффект группирования ошибок на выходе сверточного декодера при превышении его ис-

правляющей способности может быть описан моделью Гилберта, что позволяет проводить приближенную оценку вероятности ложной синхронизации численно. Кроме того, приведена методика детектирования обратной работы и поиска начала кортежа элементов входного потока кодированных данных. Методика не требует реализации пробного декодирования, что снижает вычислительные затраты при её практической реализации.

Литература

1. *Barker R. H.* Group Synchronizing of Binary Digital Systems // *Communication Theory*, Butterworth, 1953. P. 279–287.
2. *Nowbakht A., Bergmans J. W. M.* Design of optimum sync and detection patterns for frame synchronisation // *Electronics Letters*. April 2004.
3. *Мелентьев О. Г.* Теоретические аспекты передачи данных по каналам с группирующимися ошибками / под ред. В. П. Шувалова. М.: Горячая линия – Телеком, 2007. 232 с.
4. *Мелентьев О. Г.* Методика вычисления точных значений вероятностей состояний для дискретного канала, описываемого моделью Гилберта // *Труды учебных заведений связи*. 2005. № 172. С. 73–78.
5. *Мелентьев О. Г., Беляк А. Н.* Оптимизация алгоритма вычисления вероятностей поражения блока в дискретных каналах с двумя состояниями // *Материалы РНТК «Информатика и проблемы телекоммуникаций»*, Новосибирск. 2007. С. 49–53.

*Статья поступила в редакцию 19.02.2019;
переработанный вариант – 25.11.2019.*

Максимов Иван Николаевич

ассистент кафедры вычислительных систем СибГУТИ, тел. (383) 2-698-275, e-mail: ivan.maximov@mail.ru.

Мелентьев Олег Геннадьевич

д.т.н., профессор кафедры передачи дискретных сообщений и метрологии СибГУТИ, (630102, Новосибирск, ул. Кирова, 86), тел. (383) 2-698-240, e-mail: melog.aes@gmail.com.

Шевнина Ирина Евгеньевна

к.т.н., доцент кафедры передачи дискретных сообщений и метрологии СибГУТИ, тел. (383) 2-698-241, e-mail: pdsms@yandex.ru.

Using convolutional encoded preambles for frame synchronizations

O. Melent'ev, I. Maximov, I. Shevnina

The using of preambles for frame synchronization are considered. It is shown that the preamble encoded by a convolutional encoder, in some cases, can reduce the probability of false synchronization compared to a preamble without error correction coding. A technique for detecting of channel inversion and incorrect symbol identification of a pair of elements in the input stream of encoded data that does not require trial decoding is provided.

Keywords: frame synchronization, Barker codes, preamble detection, convolutional code, coded preamble.