

МИНИСТЕРСТВО ОБРАЗОВАНИЯ И НАУКИ  
РОССИЙСКОЙ ФЕДЕРАЦИИ

РЯЗАНСКИЙ ГОСУДАРСТВЕННЫЙ  
РАДИОТЕХНИЧЕСКИЙ УНИВЕРСИТЕТ

**ИНФОРМАЦИОННЫЕ  
ТЕХНОЛОГИИ**

Межвузовский сборник научных трудов

Рязань 2018

УДК 004

Информационные технологии, межвузовский сборник научных трудов. – Рязань: Рязанский государственный радиотехнический университет, 2018. – 136 с.

ISBN 978-5-7722-0301-9

Публикуются статьи о проблемах использования информационных технологий в науке и образовании.

Сборник рассчитан на научно-педагогических работников вузов и может быть использован студентами и аспирантами инфокоммуникационных и информационных специальностей.

Редакционная коллегия:

Д-р техн. наук, проф. В.П. Корячко (ответственный редактор), д-р техн. наук, проф. Я.Е. Львович (Воронежский государственный технический университет), д-р техн. наук, проф. А.Д. Иванников (Институт проблем проектирования в микроэлектронике Российской академии наук), д-р техн. наук, проф. С.В. Скворцов (РГРТУ), д-р техн. наук, проф. А.П. Шибанов (РГРТУ), канд. техн. наук А.Н. Сапрыкин (ответственный секретарь).

Рецензенты:

Кафедра ВПМ Рязанского государственного радиотехнического университета (зав. кафедрой, д-р техн. наук, проф. А.Н. Пылькин), д-р техн. наук, проф. В.А. Минаев (Московский государственный технический университета имени Н.Э. Баумана).

**ISBN 978-5-7722-0301-9**

©Рязанский государственный  
радиотехнический университет, 2018

## СОДЕРЖАНИЕ

**Аронов Л.В.**

Моделирование отношения сигнал-шум подводного открытого оптического канала передачи информации..... 7

**Баландин С.О., Сапрыкин А.Н.**

Использование генетического алгоритма при компоновке блоков электронных средств..... 11

**Буробина А.С., Сапрыкин А.Н.**

Особенности формирования новой популяции при поиске соответствий на двух множествах..... 14

**Дашкина Л.С.**

Обзор способов оценки качества кластеризации..... 18

**Дмитриева И.Н.**

Выбор моделей представления знания при исследовании интеллектуальных транспортных систем..... 22

**Евстафьев Д.А.**

Схема информационной безопасности на основе вычислительной временной визуализации изображения..... 25

**Зенина М.А.**

Исследование алгоритмов обработки видеосигнала..... 28

**Костиков М.Г., Миронов В.В.**

Автоматизация проектирования нормативов потребления коммунальных услуг..... 31

**Кузнецов А.А.**

Алгоритмы визуализации деревьев большой размерности..... 37

**Ликучев В.Ю.**

Специальная задача для группы беспилотных летательных аппаратов..... 46

**Махнач Е.С.**

Исследование методов и средств реализации системы «умный дом» . 50

**Михеева М.С.**

Исследование методов и средств технологической тренировки изделий РЭС ..... 53

**Нгуен В.Л.**

Получение экспериментальных данных с помощью системы измерения и контроля параметров дыхания ..... 56

**Никитин Р.И.**

Ускоренные испытания ..... 60

**Осин И.Г.**

Организация учета транзакций в распределенной базе данных ..... 66

**Перевезенцев Е.Е., Ромашикова В.В.**

Аспекты реализации библиотеки градиентного бустинга CatBoost ..... 69

**Пехнов А.С.**

Анализ применения библиотеки phpMorphy ..... 72

**Писака П.С.**

Сравнительный анализ оптимальных когерентных и некогерентных алгоритмов обработки информации бортовой системы телеметрических измерений при территориально-распределенных приемных станциях ..... 75

**Попов Г.А., Скворцов С.В.**

Тегирование текстов на основе технологии Word2vec ..... 80

**Попов С.В.**

Применение видеосенсорных систем в БПЛА ..... 83

**Салимгареев Р.Р., Скворцов С.В.**

Автоматизированное изменение параметров точки доступа через беспроводное подключение ..... 86

**Сапрыкин А.Н., Сапрыкина Е.Н.**

Структура хромосомы генетического алгоритма балансировки потоков данных в сетях дата-центров ..... 89

**Соболева А.И., Громов А.Ю.**

Использование библиотеки Keras для прогнозирования ухода рабочих кадров ..... 92

**Строгова Ю.А., Громов А.Ю.**

Преобразование неявных знаний эксперта в сфере банковских услуг для систем искусственного интеллекта ..... 97

**Тришин А.А., Буряков А.П.**

Повышение безопасности системы обмена информацией специального назначения..... 100

**Турбин К.Н., Петухов А.А.**

Методы оптимизации работы нейросетей ..... 103

**Фам А.М., Шибанов А.П.**

Разработка программы комбинированной системы моделирования на основе сетей Петри и GERT-сетей..... 108

**Фам В.Т., Сускин В.В.**

Использование метода градиентного спуска для определения параметров фотоэлектрического модуля с охлаждением..... 112

**Филатов К.Ю.**

Анализ свойств юзабилити ..... 117

**Цуканов И.В.**

Проблема конструирования высокостабильного неуправляемого разрядника ..... 120

**Чесалин П.Ю., Шелехин П.В.**

Использование алгоритма Random forest для решения задачи классификации ..... 123

**Чикалкин В.С.**

Адаптивный алгоритм работы системы синхронного множественного доступа ..... 126

**Ширенин И.А.**

Шаблон проектирования «Фабрика» при разработке приложений... 129

**Яковлев И.И., Баранчиков А.И.**

Разработка объектной модели представления функциональных зависимостей ..... 132

Л.В. АРОНОВ

Рязанский государственный радиотехнический университет

### МОДЕЛИРОВАНИЕ ОТНОШЕНИЯ СИГНАЛ-ШУМ ПОДВОДНОГО ОТКРЫТОГО ОПТИЧЕСКОГО КАНАЛА ПЕРЕДАЧИ ИНФОРМАЦИИ

*В статье проведено моделирование отношения сигнал-шум подводного открытого оптического канала передачи информации на примере линии построенной на основе лазерного диода NDB4116 и фотодиода S5973-02. Получены зависимости принятой мощности и отношения сигнал-шум при разных скоростях передачи данных от расстояния. Анализ проведен для различных типов океанской воды по классификации Жерлова.*

Обслуживание подводной инфраструктуры, такой как нефте- и газопроводы, опоры мостов, бурильные установки и т.д., требует применения роботизированных платформ. Это позволяет снизить риски для персонала и повысить эффективность работы. К телекоммуникационному оборудованию таких устройств предъявляются строгие требования: низкое энергопотребление, возможность передачи в реальном масштабе времени нескольких потоков видеоданных, минимально возможные масса и габариты. Выполнить эти требования позволяет использование подводных открытых оптических каналов передачи информации на основе когерентных источников излучения.

Мощность принимаемая фотоприемником определяется по формуле [1]:

$$P_{\text{ПрМ}} = P_{\text{ПрД}} \frac{D_{\text{ПрМ}}^2}{\theta_{\text{div}}^2 L^2} 10^{\left(\frac{\alpha_{\text{об}}(\lambda)L}{10}\right)} \tau_{\text{ПрД}} \tau_{\text{ПрМ}},$$

где  $P_{\text{ПрД}}$  – мощность передатчика, Вт;  $D_{\text{ПрМ}}$  – диаметр приемной апертуры, м;  $L$  – расстояние между приёмником и передатчиком, м;  $\alpha_{\text{об}}(\lambda)$  – коэффициент ослабления, дБ/м;  $\tau_{\text{ПрД}}$  – коэффициент передачи оптической системы передатчика;  $\tau_{\text{ПрМ}}$  – коэффициент передачи оптической системы приёмника;  $\theta_{\text{div}} = \frac{720\lambda}{\pi^2 D_{\text{ПрД}}}$  – угол расхождения луча [2], °;  $\lambda$  – длина волны несущей оптического передатчика, м;  $D_{\text{ПрД}}$  – диаметр передающей апертуры, м.

Наименьшее ослабление в морской воде происходит в синей части спектра, выбрав диаметр приемной и передающей апертур равным 5 см, а несущую длину волны 450 нм, рассчитаем зависимость мощности от дальности траектории (в качестве передатчика возьмем лазерный

светодиод NDB4116,  $P_{прд} = 100 мВт$ ). Учитывая, что коэффициент передачи качественной оптической системы равняется примерно 0,9, получим зависимости принятой мощности от дальности для разных типов морских вод по классификации Жерлова (рис. 1).



Рис. 1. Зависимости принятой мощности от дальности для разных типов морской воды

При этом коэффициент ослабления  $\alpha_{об}(\lambda)$  рассчитывается по методике изложенной в работе [3]. Анализ полученных зависимостей показывает, что в водах с высокой замутненностью (мутные бухты) установление связи практически невозможно, по причине сильного ослабления. В водах с низкой концентрацией хлорофилла, в т.ч. в глубинных водах ослабление на дистанции 100 метров составляет порядка 68-74 дБ, это дает возможность обеспечить обнаружение сигнала, при условии достаточной чувствительности оптического приёмника.

Мощность на выходе фотоприемника пропорциональна квадрату фототока, который равен [1]:

$$\langle i^2 \rangle = (R_0 P_{прм})^2,$$

где  $R_0$  – чувствительность фотоприемника, определяема по формуле [4]:

$$R_0 = \frac{q_e \eta}{h\nu} \text{ – чувствительность фотодиода, А/Вт,}$$

где  $\eta$  – квантовая эффективность фотодиода;

$q_e = 1,6 \cdot 10^{-19}$  Кл – элементарный заряд;  $h = 6,626 \cdot 10^{-34}$  Дж·с – постоянная Планка;  $\nu$  – несущая оптическая частота, Гц. В процессе



детектирования оптический сигнал подвергается воздействию шумов [1,4], в том числе дробового шума:

$$\sigma_{DP}^2 = 2q_e R_0 P_{\text{прм}} B,$$

где  $B$  – эквивалентная шумовая полоса, Гц. Шум темнового тока:

$$\sigma_D^2 = 2q_e I_D B,$$

где  $I_D$  – темновой ток фотодиода, А. Шум обусловленный поверхностным током (током утечки):

$$\sigma_L^2 = 2q_e I_L B,$$

$I_L$  – ток утечки, А. Тепловой шум (шум Джонсона) [4]:

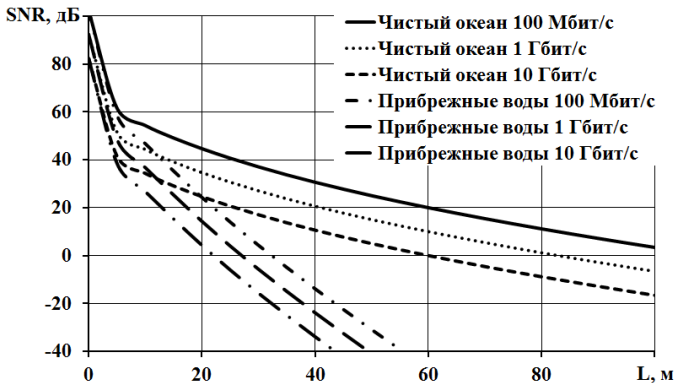
$$\sigma_T^2 = 4kTBF_T / R_H,$$

где  $k = 1,38 \cdot 10^{-23}$ ; – постоянная Больцмана;  $T$  – температура приемника, К;  $F_T$  – коэффициент шума, дБ;  $R_H$  – эквивалентное сопротивление цепи, Ом.

Результирующее выражение для расчета отношения сигнал-шум:

$$SNR = \frac{(R_0 P_{\text{прм}})^2}{2q_e (R_0 P_{\text{прм}} + I_D) B + 2q_e I_L B + 4kTBF_T / R_H}.$$

Темновой ток используемого нами фотодиода S5973-02 равен  $I_D = 0,1 \cdot 10^{-9}$  А, а чувствительность равна на несущей длине волны  $R_0 = 0,35$  А/Вт. Прочие параметры выбраны согласно [2]: коэффициент шума  $F_T = 3\text{дБ}$ , эквивалентное сопротивление цепи  $R_H = 50\text{кОм}$ . Температура приемника выбрана равной  $20^\circ\text{C}$ , т.е. абсолютная температура составляет  $T = 298\text{K}$ . Анализ полученных зависимостей показывает, что в чистом океане уровень сигнального тока на выходе будет превышать уровень шумовых токов на дистанциях до 100 метров, при скорости передачи данных до 100 Мбит/с, однако увеличение скорости передачи данных влечет за собой снижение ОСШ вследствие расширения электрической полосы приемника. В прибрежных водах, характеризующихся высокой замутненностью, дистанция на которой сигнальный ток выделяется над шумовым составляет не более 32 метров (рис. 2).



**Рис. 2.** Зависимости принятой мощности от дальности для разных типов морской воды

Увеличить ОСШ, не снижая скорость передачи данных возможно за счёт увеличения мощности передающего устройства, использования фотодиодов с повышенной чувствительностью или оптических детекторов позволяющих усиливать принятое излучение, например лавинные фотодиоды или фотоэлектронные умножители.

### СПИСОК ИСПОЛЬЗУЕМЫХ ИСТОЧНИКОВ

1. Mazin A.A.A. Characteristics of optical channel for underwater optical wireless communication system [Текст] / Ali A. Ali Mazin // IOSRJournal of electrical and electronics engineering – 2015 – Volume 10, – p. 9.
2. Abd El-Naser A. Mohamed Underwater wireless optical communications for short range typical ocean water types [Текст] / Abd El-Naser A. Mohamed, Hamdy A. Sharshar, Ahmed Nabih Zaki Rashed, Enab Salah El-dien // Canadian journal on electrical and electronics engineering. – 2012. – №7, vol. 3 – p. 344-361.
3. Аронов Л.В. Процедура расчета энергетического бюджета подводного открытого оптического канала передачи информации [Текст] / Л. В. Аронов // Новые информационные технологии в научных исследованиях материалы XXII Всероссийской научно-технической конференции студентов, молодых ученых и специалистов. Рязанский государственный радиотехнический университет. 2017. – с. 84-86.
4. Jaruwatadilok S. Underwater wireless optical communication channel modeling and performance evaluation using vector radiative transfer theory [Текст] / S. Jaruwatadilok // IEEE Journal on selected areas in communications . – 2008 . – No.9, vol. 26 – pp. 1620-1627.

**С.О. БАЛАНДИН, А.Н. САПРЫКИН**

Рязанский государственный радиотехнический университет

## **ИСПОЛЬЗОВАНИЕ ГЕНЕТИЧЕСКОГО АЛГОРИТМА ПРИ КОМПОНОВКЕ БЛОКОВ ЭЛЕКТРОННЫХ СРЕДСТВ**

*Рассматриваются основные моменты применения генетического алгоритма для решения задачи компоновки блоков электронно-вычислительных средств.*

Компоновка конструктивных модулей электронно-вычислительных средств является одной из наиболее важных задач конструкторского проектирования, возникающих на каждом из основных структурных иерархических уровней.

Однако, в настоящее время большинство промышленных САПР обходятся без компоновки блоков вовсе или накладывают на нее множество ограничений. Это вызвано высокой ресурсоемкостью и сложностью обработки необходимых массивов исходных данных.

Решение задачи компоновки блоков электронно-вычислительных средств основывается на их коммутационной схеме. Для алгоритмизации и формального решения задачи производится переход от коммутационной схемы к графу или гиперграфу [1]. В связи с данной постановкой задачи ее решение сводится к разбиению графа на заданное количество частей, что в свою очередь является NP-полной и NP-сложной задачей. В условиях современного уровня развития электронно-вычислительных средств использование стандартных алгоритмов компоновки является нецелесообразным в силу чрезмерно большого количества критериев оптимизации и количества элементов. Получение оптимального решения возможно только полным перебором различных вариантов разбиения.

Задачи компоновки в большинстве случаев представляют собой процесс принятия решений и в определенных, и в неопределенных условиях. В результате этого процесса происходит объединение конструктивных модулей низшего порядка в более крупные конструктивные модули высшего порядка. Компоновка требует наличия определенных критериев и ограничений.

Эволюционные алгоритмы получили широкое распространение, в частности, генетические алгоритмы [2, 3], которые отличаются тем, что в ходе их реализации рассматривается не одно, а сразу несколько потенциальных решений.

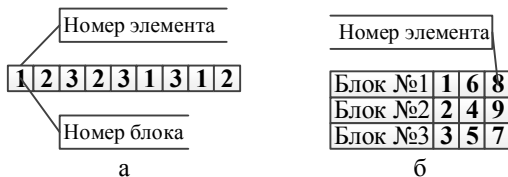
Генетические алгоритмы представляют собой один из наиболее эффективных способов решения крупномасштабных многокритериальных задач оптимизации, ведь они позволяют снизить временные и

вычислительные затраты. Алгоритм позволяет эффективно учитывать множество входных ограничений, обладая при этом малой временной сложностью. Именно поэтому он является эффективным средством для решения задач компоновки блоков ЭВС большой размерности.

Генетический алгоритм возможно применить к задаче компоновки, но для этого его необходимо модифицировать, сохранив при этом саму суть алгоритма.

В классическом генетическом алгоритме создается новая популяция из потомков текущего поколения. Однако, нет необходимости в создании новой популяции, можно использовать одну единственную, добавляя в нее новых особей и исключая «ненужных».

После определения вида функции полезности, необходимо закодировать особь для удобной и быстрой работы с ней. Были разработаны две возможные структуры особи (рис. 1).



**Рис. 1.** Виды структур особи для решения задачи компоновки:  
а) «блоки по элементам»; б) «элементы по блокам».

Первая структура состоит из одной хромосомы. Позиции генов в хромосоме являются номерами элементов, а значения генов отвечают за номера блоков, в которых элементы находятся. Во второй структуре количество хромосом равно количеству существующих блоков. Чтобы динамически не преобразовывать для каждого этапа одну структуру в другую, рациональным решением будет использовать обе структуры одновременно. Благодаря этому повысится быстродействие, наглядность и удобство работы с данными. Недостатком такого решения является повышенное потребление памяти каждой особью.

После выбора структуры необходимо определиться с основными процессами. При инициализации особи происходит заполнение структур, при этом не должно происходить переполнения блоков или повторяемости номеров элементов. Для того чтобы создавать новых особей понадобится процесс скрещивания. Однако, для того, чтобы не было случаев переполнения блоков, необходимо ввести систему контроля и управления такими случаями. Поэтому в работе используется поразрядное скрещивание (многоточечное) с появлением лишь одного потомка. На этапе скрещивания происходит непосредственная работа

со структурой «блоки по элементам», но информация от другой структуры также используется для системы контроля.

Для предотвращения состояния «застоя» в популяции понадобится мутация. Однако, как и в случае со скрещиванием, возможны случаи переполнения блоков, поэтому в работе используется одноточечная мутация (если особь была выбрана для мутирования, то меняется значение случайного гена). На данном этапе принимают участие также обе структуры. Для того, чтобы не «потерять» лучшую особь можно воспользоваться принципом «элитизма», в результате чего лучшая особь (одна) всегда будет оставаться в популяции и не будет подвергаться изменениям. Необходимой операцией также является процедура отбора полезных особей и исключения наименее полезных из популяции. Этой операцией будет являться селекция. Однако, ввиду того, что сущность селекции была изменена, необходимо модифицировать и сами сущности применяемых известных типов селекции.

Предложенный алгоритм позволяет эффективно решать задачу компоновки блоков электронных средств большой размерности и получать оптимальные результаты за полиномиальное время.

### **СПИСОК ИСПОЛЬЗУЕМЫХ ИСТОЧНИКОВ**

1. Курейчик В.М. Математическое обеспечение конструкторского и технологического проектирования с применением САПР: учебник для вузов / В. М. Курейчик. – М.: Радио и связь, 1990. – 352 с.
2. Гладков Л.А. Генетические алгоритмы [Текст] / Л.А. Гладков, В.В. Курейчик, В.М. Курейчик. –М.: Физматлит, 2010. – 317 с.
3. Курейчик В.М. Генетические алгоритмы и их применение. Таганрог: Изд-во ТРТУ, издание второе, дополненное, 2002. – 242 с.

А.С. БУРОБИНА, А.Н. САПРЫКИН

Рязанский государственный радиотехнический университет

## ОСОБЕННОСТИ ФОРМИРОВАНИЯ НОВОЙ ПОПУЛЯЦИИ ПРИ ПОИСКЕ СООТВЕТСТВИЙ НА ДВУХ МНОЖЕСТВАХ

*Рассматриваются сложности и механизмы формирования новой популяции для решения задачи нахождения соответствий на двух конечных множествах.*

Создание новой популяции – важный этап классического генетического алгоритма, являющийся основополагающим для быстрого получения наиболее точного результата. Для формирования новой популяции принято использовать генетические операторы, такие как мутация и скрещивание. Зачастую используют операцию копирования хромосом.

Для решения поставленной задачи использовались все вышеописанные генетические операторы. Выбор конкретного оператора осуществлялся с помощью задания вероятностей мутации и скрещивания и генерации случайного числа в интервале от 0 до 1. Далее действовало правило: если полученное число не превышает величины вероятности мутации, то новый генотип формируется с помощью оператора мутации, если сгенерированное значение не выходит за рамки суммарного значения вероятностей мутации и скрещивания, то формируются два новых генотипа посредством операции скрещивания, в противном случае, в новую популяцию копируется произвольно выбранный генотип родительского пула. Изменяя значения вероятностей можно варьировать скорость и точность полученного решения.

Цель оператора мутации – разнообразить многообразие рассматриваемых индивидов. В традиционном представлении классического генетического алгоритма под мутацией подразумевается незначительное изменение заданного вектора и выражается в изменении гена индивида с заданной вероятностью. Однако для рассматриваемой задачи такой подход к мутации не пригоден, так как задача решается на конечных множествах. Поэтому, в рамках поставленных условий в качестве мутации будет выступать формирование полностью случайного генотипа. Опишем правила формирования данного генотипа. Пусть имеются множества  $A$  – множество, для которого ищут соответствия,  $B$  – множество объектов, распределяемых на множество  $A$ .

Дополнительные условия формирования генотипа:

мощность множества  $A$  меньше или равна мощности множества

$B$ ;

необходимо как можно более равномерно распределить элементы множества В по элементам множества А;

каждому элементу множества А может соответствовать ограниченное число элементов множества В, указанное в условии;

каждому элементу А должно соответствовать хотя бы одно значение В.

Программная реализация оператора мутации формирует новый генотип в два этапа. На первом из множества А случайным образом выбираются элементы и им в соответствие ставится случайный элемент множества В. Этап заканчивается, когда множество А опустеет. Таким образом, каждому элементу множества А соответствует один элемент множества В. На 2 этапе происходит дораспределение элементов множества В на элементы множества А. Выбирается случайный элемент множества В, затем выбирается элемент множества А, которому он ставится в соответствие. Если число сопоставленных элементов множества А достигает определенного в условии максимума, данный элемент исключается из распределения. Формирование генотипа завершается, если опустеет множество В или всем элементам множества А будут исключены из распределения.

Цель оператора скрещивания – сформировать новые индивиды из двух случайно выбранных родительских особей. В рамках решаемой задачи было выбрано равномерное скрещивание. Суть его заключается в следующем: пусть имеются два родителя Р1 и Р2. Потомки состоят из генов родителей, стоящих на соответствующих позициях, причем порядок копирования генов в определенного потомка определяется случайным триггером, принимающим значения 0 или 1. Если значение триггера равно 0, то ген в текущей позиции родителя Р1 копируется во второго потомка, а ген родителя Р2 – в первого. Если триггер находится в положении единицы, то порядок копирования противоположный. Пример работы оператора равномерного скрещивания приведен на рисунке 1.

Данная разновидность оператора скрещивания была выбрана для рассматриваемой задачи, так как в ходе его работы потомки формируются только из значений своих родителей, что идеально подходит при использовании конечных множеств. Однако при реализации оператора для поиска соответствий на двух множествах возникли следующие проблемы. Во-первых, так как выбор родительских генотипов для скрещивания и порядок копирования в определенного потомка определяется случайным образом, то может возникнуть ситуация, когда один и тот же элемент множества В дважды попадает в один и тот же генотип. Данный пример проиллюстрирован на рисунке 2. Во-вторых,

может возникнуть ситуация, когда один потомок короче другого, что делает распределение элементов по другому множеству неравномерным. Данная ситуация является следствием решения первой проблемы.

P1	2	3	5	7	3	6
P2	4	8	9	3	2	2
Случайные числа	0	0	1	0	0	1
П1	4	8	5	3	2	6
П2	2	3	9	7	3	2

**Рис. 1.** Пример работы оператора равномерного скрещивания

П2	2	3	9	7	3	2
----	---	---	---	---	---	---

**Рис. 2.** Пример дублирования элементов в потомке

При возникновении ситуации дублирования сначала предпринимается попытка принудительного изменения порядка копирования, чтобы поместить элемент в другого потомка. Данный способ решения проблемы уместен, если изменения порядка копирования не приводит к возникновению аналогичной проблемы ни в одном из потомков. Если изменение порядка копирования не помогает, то элемент заносится в «очередь на добавление». Данная «очередь» представляет собой массив элементов, повторившихся при формировании потомков. Элементы «очереди на добавление» могут быть размещены в формируемом геноме после формирования всех хромосом.



Как отмечалось ранее, дублирование элементов может привести к укорачиванию хромосом и нарушению свойства равномерности распределения. Попытка восстановления прежней длины хромосомы производится после формирования всех хромосом генотипа. Дополнение хромосом производится с использованием элементов из «очереди на добавление». Внесение нераспределенных элементов в генотип осуществляется аналогично второму этапу оператора мутации.

Рассмотренные реализации генетических операторов позволяют получать разнообразные генотипы и формировать максимально пригодное распределение элементов.

Л.С. ДАШКИНА

Рязанский государственный радиотехнический университет

## ОБЗОР СПОСОБОВ ОЦЕНКИ КАЧЕСТВА КЛАСТЕРИЗАЦИИ

*Рассматриваются способы оценки качества кластеризации, а также программное обеспечение, реализующее их.*

Кластеризация является важным приложением интеллектуального анализа данных. Ее используют, когда отсутствуют априорные сведения относительно классов, к которым можно отнести объекты исследуемого набора данных, либо, когда число объектов велико, что затрудняет их ручной анализ. Так, кластеризация применяется в информационном поиске, розничной торговле, телекоммуникациях, страховании и сегментации изображений [1].

Результатом работы всех алгоритмов кластеризации является разбиение исходных данных на группы сходных объектов (кластеров). Для того чтобы алгоритм построил этот набор, необходимо знать количество кластеров. Чаще всего априорной информации о количестве кластеров нет, поэтому для наиболее лучшего разбиения в алгоритм кластеризации добавляют адаптивный механизм выбора оптимального решения среди множества возможных. Выбор оптимального решения основывается на понятии качества кластеризации. Качество кластеризации – это степень приближения результата кластеризации к идеальному решению [2].

Трудность оценки качества кластеризации связана с тем, что мы не в состоянии зрительно представить геометрические особенности многомерного пространства. При интерпретации полезно использовать расстояние между центрами кластеров. Такую информацию можно свести к таблице, на пересечении  $i$  строки и  $j$  столбца которой будет указано расстояние между центрами  $i$  и  $j$  кластеров. Чем больше расстояние между центрами кластеров, тем выше качество кластеризации, но таблица расстояний не является достаточной для получения содержательных выводов. Кроме расстояния между кластерами, важно иметь представление об относительном расположении образов внутри кластера. Для этого можно определить рассеяние характеристик кластера относительно средних значений. Эту информацию также можно оформить в виде таблицы, на этот раз таблицы дисперсий, в которой будут указаны значения дисперсий (или ско) по каждому атрибуту [3]. Такие способы удобны при интерпретации результатов, но с их помощью трудно понять, какое число кластеров приводит к лучшему разбиению.

Оценить качество кластеризации можно еще двумя способами — экспертным и формальным. Экспертный выбор наилучшего решения задачи заключается в оценке решения специалистами в данной предметной области. Но экспертная оценка зачастую объективно невозможна из-за большого объема и сложности данных. Поэтому важную роль играют формальные критерии оценки качества кластеризации [2].

Среди основных формальных критериев можно выделить следующие:

1) Индекс Данна (Dunn's index) [4]: этот индекс изначально предложен для использования при идентификации «компактных и хорошо отделенных кластеров». Он вычисляется для каждого кластера  $c$  по формуле (1):

$$DI(c) = \min_{i \in c} \left\{ \min_{j \in c, i \neq j} \left\{ \frac{\min_{x \in C_i, y \in C_j} d(x, y)}{\max_{k \in c} \left\{ \max_{x, y \in C} d(x, y) \right\}} \right\} \right\} \quad (1)$$

где  $d(x, y)$  – расстояние между объектами  $x$  и  $y$ .

Этот индекс использует расстояние между кластерами и диаметр кластера (наибольшее расстояние между элементами кластера). Большие расстояния между кластерами и маленькие диаметры кластеров говорят о том, что полученные кластеры компактны и хорошо разделены.

2) Индекс Девиды-Болдуина (Davies-Bouldin index) [5]. Пусть  $R_{i,j}$  – мера того, насколько хорошо объекты разделены на кластеры. Эта мера по определению должна учитывать меру различия между  $i$ -м и  $j$ -м кластером  $M_{i,j}$ , которое должно быть как можно больше, и меру внутрикластерного рассеяния для кластера  $S_i$ , которое должно быть как можно меньше. Индекс вычисляется по формуле (2):

$$DB = \frac{1}{N} \sum_{i=1}^N \max_{j \neq i} \frac{S_i + S_j}{M_{i,j}} \quad (2)$$

3) Оценка «силуэта» (Silhouette) [6]. Пусть  $a_i$  – среднее расстояние между  $i$  и всеми другими данными внутри одного кластера. Мы можем интерпретировать  $a_i$  как меру того, насколько хорошо  $i$  назначен его кластеру (чем меньше значение, тем лучше назначение). Затем определяется среднее различие точки  $i$  с кластером  $c$ , как среднее расстояние от  $i$  до всех точек  $c$ . Пусть  $b_i$  – наименьшее среднее расстояние  $i$  до всех точек любого другого кластера, из которых  $i$  не является членом. Кластер с этим самым низким средним различием называется «соседним кластером»  $i$ , потому что он является следующим наилучшим подходящим кластером для точки  $i$ . Силуэт  $S_i$  определяется по формуле (3):

$$S_i = \frac{b_i - a_i}{\max\{a_i, b_i\}} \quad (3)$$

Чем выше  $S_i$ , тем объект  $i$  лучше подходит своему кластеру.

Среди программных средств, вычисляющих оценки качества кластеризации, можно выделить следующие:

1) Cluster Validity Analysis Platform [7] – программное средство, реализованное на базе Matlab. Включает реализованные алгоритмы кластеризации: иерархический, k-medoidsPAM, k-means и др., а также следующие индексы оценки их работы: оценка «силуэта», индекс Девида-Болдуина, индекс Данна, индекс Calinski-Harabasz и др.

2) Clustering and Data Analysis Toolbox [4] – набор инструментов для Matlab, включающий алгоритмы: k-means, k-medoids, fuzzy c-means и др. и оценки: индекс Данна, модифицированный индекс Данна, коэффициент разбиения (partition coefficient) и др.

3) ELKI [8] – программное обеспечение для интеллектуального анализа данных с открытым исходным кодом, написанное на Java. В ELKI 0.7.1 реализовано множество модификаций k-means (k-meansSort, k-medoidsPAM, x-means и др.) и ряд других алгоритмов. Представлены следующие оценки: оценка «силуэта», индекс Девида-Болдуина, индекс PBM и др.

Главным недостатком первых двух приложений является ограничение на использование из-за коммерческой лицензии Matlab, а к преимуществам применения ELKI помимо открытого распространения можно отнести наличие большого числа алгоритмов кластеризации и более удобной визуализации результатов работы алгоритмов.

Оценка качества кластеризации является сложной задачей. Тем не менее, при правильном подборе индекса для оценки можно эффективнее решать задачу кластеризации и более взвешенно подходить к оценке ее результатов.

### СПИСОК ИСПОЛЬЗУЕМЫХ ИСТОЧНИКОВ

1. Паклин Н. Б., Орешков В. И. Бизнес-аналитика: от данных к знаниям. 2-е изд. – СПб.: Питер, 2013. – 704 с.

2. Анализ данных и процессов: учеб. пособие / А. А. Барсегян, М. С. Куприянов, И. И. Холод, М. Д. Тесс, С. И. Елизаров. 3-е изд. – СПб.: БХВ-Петербург, 2009. – 512 с.

3. Ту Дж., Гонсалес Р. Принципы распознавания образов. М.: Мир, 1978. – 414 с.

4. Balasko B., Abonyi J., Feil B. Fuzzy Clustering and Data Analysis Toolbox [Электронный ресурс], URL: [https://www.researchgate.net/profile/Janos\\_Abonyi/publication/263697045\\_Fuzzy\\_Clustering\\_and\\_Data\\_An](https://www.researchgate.net/profile/Janos_Abonyi/publication/263697045_Fuzzy_Clustering_and_Data_An)

aly-

sis\_Toolbox\_for\_Use\_With\_Matlab/links/0deec53bae2ee507fc000000/Fuzzy-Clustering-and-Data-Analysis-Toolbox-for-Use-With-Matlab.pdf (дата обращения 24.03.18).

5. Davies–Bouldin index [Электронный ресурс], URL: [https://wikivisually.com/wiki/Davies-Bouldin\\_index](https://wikivisually.com/wiki/Davies-Bouldin_index) (дата обращения 24.03.18).

6. P. J. Rousseeuw Silhouettes: a graphical aid to the interpretation and validation of cluster analysis // Journal of Computational and Applied Mathematics, volume 20, 1987. p. 53-65.

7. CVAP: Cluster Validity Analysis Platform (cluster analysis and validation tool) [Электронный ресурс], URL: <http://www.mathworks.com/matlabcentral/fileexchange/14620-cvap--cluster-validity-analysis-platform--cluster-analysis-and-validation-tool> (дата обращения 24.03.18).

8. ELKI: Environment for Developing KDD-Applications [Электронный ресурс], URL: <https://elki-project.github.io/> (дата обращения 24.03.18).

**И.Н. ДМИТРИЕВА**

Рязанский государственный радиотехнический университет

**ВЫБОР МОДЕЛЕЙ ПРЕДСТАВЛЕНИЯ ЗНАНИЯ ПРИ  
ИССЛЕДОВАНИИ ИНТЕЛЛЕКТУАЛЬНЫХ  
ТРАНСПОРТНЫХ СИСТЕМ**

*Рассматриваются основные модели представления знаний в интеллектуальных системах.*

Интеллектуальные транспортные системы представляют собой системы, совмещающие в себе современные информационные и коммуникационные технологии, а также технологии управления. Данные системы применяются для управления транспортными потоками. Основной технологией в интеллектуальных транспортных системах является сочетание интеллектуальных и информационных технологий [1].

Основной проблемой городов является несоответствие пропускной способности улично-дорожной сети реальному спросу на транспортные услуги.

Одна из самых важных целей транспортных систем является проектирование единой системы, позволяющей управлять и отслеживать транспортные потоки в данный момент времени. Только на этой основе в рамках функционирования интеллектуальных транспортных систем могут и будут решаться задачи обеспечения транспортного баланса между реальной пропускной способностью улично-дорожной сети и спросом на дорожное движение.

Всё больше появляются новые методы и возможности для управления транспортными потоками, формируются и новые способы информационного обеспечения для участников дорожного движения. Повышение эффективности работы транспорта, удовлетворение потребностей населения в перевозках достигается при рациональном управлении системы.

Рациональное управление системой способствует сокращению времени доставки пассажиров и товаров, повышению уровня безопасности дорожного движения и качества транспортных услуг и снижению негативного воздействия транспортных средств на окружающую среду.

Системы мониторинга улично-дорожной сети города позволяют увеличить пропускную способность, информировать о реальной ситуации на дорожной сети [2]. В системах собирается и обрабатывается информация о реальной ситуации на дорогах и улицах и параметры

транспортных потоков. Однако возникает ряд проблем с хранением накопленных знаний.

Модель представления знаний представляет собой способ описания знаний для хранения, взаимодействия и доступа в базе знаний. В зависимости от форм представления знаний могут варьироваться характеристики интеллектуальных систем. Широко известными моделями в представления знаний являются: продукционная, семантическая, логическая, фреймовая, нейронные сети.

Продукционные модели совмещают элементы двух моделей: логических и сетевых. Главной чертой моделей является простота анализа. Благодаря однотипному представлению знаний, использование продукционных моделей в реализации очень простое, но при увеличении объемов информации значительно увеличиваются и временные затраты.

Фреймовые модели в представлении используют структурную единицу – фрейм, которая содержит информацию о событии. Достоинством данной модели является описание знаний в любой форме, что делает её одной из самых универсальных, но нахождение решения является иногда невозможным, так как связи между фреймами неоднозначны. Фреймовые модели не приспособлены к адаптации и изменения, которые могут повлечь за собой ряд противоречий и заикливания.

Модель логики предикатов – один из самых известных способов описания знаний, основу данной модели составляет математическая логика. Недостаток данной модели заключается в отсутствии полной выразимости взаимосвязей. Логические модели являются наиболее понятными для использования их в базах знаний.

Семантическая модели описывает отношения и различные понятия в виде дуг и узлов. Преимуществами данной модели являются наглядность представления знаний и соответствие современным представлениям об организации долговременной памяти человека. Но в данных моделях главными недостатками являются дублирование информации и смешение знаний, которые относятся к разным ситуациям.

Задачей нейронной сети является обучение модели реального мира. Сеть обучается множество раз на одном и том же объекте. Когда нейронная сеть обучена и способна отличать разные точки взглядов на объект, то она будет готова распознавать точки зрения, которые не использовались при обучении. Чем больше необходимых параметров, которые нужно задать, тем больше приходится использовать нейронов, происходит увеличение времени обучения сети.

Основными требованиями к моделям представления знаний является универсальность, использование многоуровневых представлений, однородность, наглядность.

В настоящее время моделей, которые бы удовлетворяли таким условиям, не существует. В связи с чем и появляется многообразие и активное развитие представления знаний. В случаях представления знаний в технических системах такой универсальной модели для каждой области просто не существует. Для модели потребуется реализация программных и аппаратных функций интеллектуальных систем, которые в настоящее время просто невозможна.

В качестве исходных моделей представления знаний в транспортных интеллектуальных системах могут быть использованы семантические модели и нейронные сети, также возможна их совместная реализация.

#### **СПИСОК ИСПОЛЬЗУЕМЫХ ИСТОЧНИКОВ**

1. Организация транспортных услуг и безопасность транспортного процесса, Н. В. Пеньшин Тамбов: Издательство ФГБОУ ВПО «ТГТУ», 2014. – 121 с.

2. Основы теории транспортных систем, А.Э. Горев: Учебное пособие, Санкт-Петербург, 2010. – 115 с.



**Д.А. ЕВСТАФЬЕВ**

Рязанский государственный радиотехнический университет

**СХЕМА ИНФОРМАЦИОННОЙ БЕЗОПАСНОСТИ НА ОСНОВЕ  
ВЫЧИСЛИТЕЛЬНОЙ ВРЕМЕННОЙ ВИЗУАЛИЗАЦИИ  
ИЗОБРАЖЕНИЯ**

*В статье предложена схема информационной безопасности, основанная на вычислительной временной визуализации изображения с описанием её работы на примере.*

Информационная безопасность является важным фактором в современных методах коммуникации и распространения. Среди тех областей исследований информационной безопасности, оптическая защита информации становится все более популярной в последние годы. Призрачная визуализация (Ghost imaging (GI)) - это метод, который создает изображение объекта, сопоставляя интенсивность двух световых пучков, ни один из которых независимо не несет информацию о форме объекта. Вычислительная призрачная визуализация (Computational ghost imaging (CGI)) в качестве вычислительной версии GI использует программируемую подсветку и, в значительной степени, упрощает реализацию GI.

Новая схема шифрования информации на основе концепций CTGI (computational temporal ghost imaging) сохраняет преимущества CGI и тем временем преодолевает ее слабость из-за того, что CTGI может восстановить сверхбыстрый сигнал с медленным детектором после длительного времени экспозиции. Исходная информация, которая должна быть зашифрована, преобразуется в одномерный (1D) поток цифровых данных. Ключ шифрования / дешифрования представляет собой стек независимых двумерных (2D) случайных двоичных шаблонов с большим количеством пикселей. Новый 2D-шаблон генерируется путем умножения каждого паттерна стека на соответствующий элемент 1D потока данных в последовательности и суммирования всех взвешенных 2D случайных двоичных паттернов. Этот отображаемый 2D-образ используется как зашифрованная информация, которая рассматривается как зашифрованный текст. Процесс дешифрования может быть реализован посредством использования корреляционного измерения или расчета между зашифрованной информацией и секретными ключами. В этом процессе поток данных 1D и секретный ключ аналогичны временному сигналу, подлежащему восстановлению. Процесс суммирования соответствует времени интегрированного изображения с камерой в течение длительного времени экспозиции.

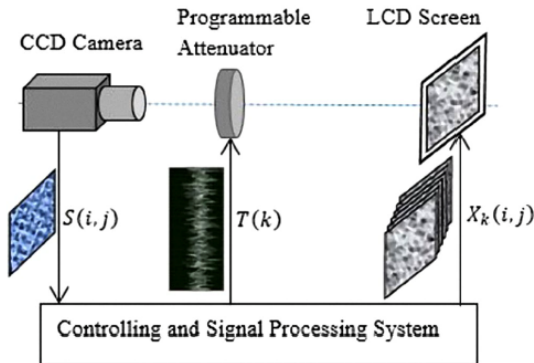


Рис. 1. Схематичное представление CTGI

Принципиальное моделирование, основанное на установке (рис. 1), используется для анализа возможности этой новой схемы шифрования. Настройка включает в себя ЖК-экран, аттенюатор и многопиксельный детектор. В процессе шифрования на ЖК-дисплее отображается последовательность случайных шаблонов, работающих в качестве ключа шифрования. Впоследствии эти модели перемещаются через аттенюатор и изменяют их интенсивности, равномерно зависящие от затухания. Поскольку затухание контролируется 1D-сигнальной последовательностью сигнала, которая преобразуется из 2D исходного изображения, каждое изображение, полученное и наложенное в многопиксельном детекторе, представляет собой модулированный случайный ключ шифрования, взвешенный соответствующим весовым сигналом.

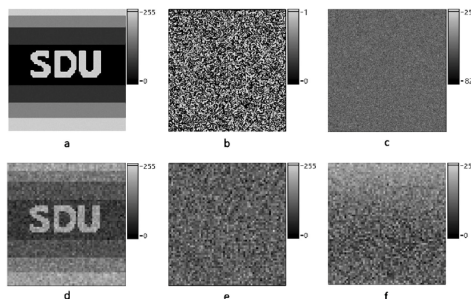


Рис. 2. Экспериментальные результаты предлагаемой схемы безопасности

В данном моделировании исходное изображение  $F_0(m, n)$  представляет собой 8-битное изображение в оттенках серого в разрешении

64 × 64 пикселя, как показано на рис. 2(а). Многоэлементный детектор установлен в динамический диапазон 8 бит, чтобы получить реалистичный результат моделирования. Исходное изображение считывается как открытый текст компьютером и перестраивается на 1D-массив  $T_i$ . Секретный ключ  $K_j(p, q)$  состоит из 4096 независимых 2D случайных двоичных шаблонов с  $256 \times 256$  пикселями. Один из этих секретных ключей изображен на рис. 2(б). Образец взвешенной суммы  $F'(p, q)$  показан на рис. 2(с), который является зашифрованным текстом и может храниться локально или передаваться в приемник. На этапе дешифрования приемник может использовать одну и ту же последовательность секретных ключей для декодирования шифрованного текста, окончательное дешифрованное изображение отображается на Рис. 2(д).

Чтобы проверить безопасность предлагаемой схемы шифрования, в процессе дешифрования вводятся два разных типа неправильных ключевых последовательностей. Рис. 2(е) показывает дешифрованное изображение, полученное после использования первого типа неправильной последовательности. Рис. 2(ф) показывает дешифрованное изображение, полученное после использования второго типа неправильного ключа. Для анализа безопасности системы шифрования нормальная среднеквадратическая ошибка (NRMS) используется как стандарт оценки, и описывается уравнением:

$$NRMS = \frac{\sqrt{\sum_{n=1}^N \sum_{m=1}^M (F_0(m, n) - F_D(m, n))^2}}{\sqrt{\sum_{n=1}^N \sum_{m=1}^M (F_0(m, n))^2}}$$

NRMS восстановленных изображений на рис. 2(е) и (ф) равны 0,02 и 0,03 соответственно. Очевидно, что правильное дешифрование основано на идентичной последовательности секретных ключей, как в пространственном распределении, так и в порядке сортировки.

### СПИСОК ИСПОЛЬЗУЕМЫХ ИСТОЧНИКОВ

1. Information security scheme based on computational temporal ghost imaging [Электронный ресурс], URL: <https://www.nature.com/articles/s41598-017-07816-2.pdf> (дата обращения 24.03.18).
2. 3d computational imaging with single-pixel detectors [Электронный ресурс], URL: <http://science.sciencemag.org/content/340/6134/844> (дата обращения 24.03.18).
3. Ghost imaging lidar via sparsity constraints in real atmosphere [Электронный ресурс], URL: <http://file.scirp.org/Html/34887.html> (дата обращения 24.03.18).

М.А. ЗЕНИНА

Рязанский государственный радиотехнический университет

## ИССЛЕДОВАНИЕ АЛГОРИТМОВ ОБРАБОТКИ ВИДЕОСИГНАЛА

*Рассматриваются алгоритмы для обработки видеосигналов в реальном времени с использованием программируемых логических интегральных схем, с последующей их программной реализацией.*

Программируемые логические интегральные схемы (далее ПЛИС) являются одним из лучших средств обработки зрительной информации. Успешное решение поставленной задачи будет зависеть не только от адекватности выбранного алгоритма, но и от качества исходного видеоизображения [1].

ПЛИС обладают высокой скоростью обработки данных. Данное свойство позволяет получить производительность на порядок выше, чем у цифровых сигнальных процессоров.

Для того, чтобы достичь высокой эффективности алгоритма, программный код создается на языке HDL (язык описания архитектуры, англ. Hardware Description Languages). При помощи программы, написанной на данном языке, появляется возможность изменения структуры интегральной схемы таким образом, что реализация требуемых функций алгебры логики происходит на аппаратном уровне.

Нахождение и отслеживание движущихся объектов в видеопотоке – важная задача для различного класса задач, которым необходимо выделять движущие объекты на сложном фоне. К таким классам можно отнести обнаружение движущегося воздушного объекта, наблюдаемого на фоне ясного или облачного неба.

Идеей интеллектуального видеонаблюдения является автоматическая обработка, поступающих данных с камер в режиме реального времени. Основная мысль предложенного функционала направлена на распознавание аномальных событий.

При решении задачи такого видеонаблюдения стоит выделить несколько подзадач распознавания:

- *Распознавания низкого уровня.* К данным задачам относятся: выделение объекта на изображении, оценка скорости перемещения объекта и т.д;
- *Распознавания высокого уровня.* К ним относятся: распознавания нападения, кражи и т.д.

Существует много способов обнаружения движения. Самыми простыми из них являются *вычитание фона и разность кадров.*

Идеей метода *вычитание фона* является то, что для обнаружения объекта или фиксирования движения в потоке видеоданных используется вычитание.

Математическая представление метода имеет вид:

$$P[f(t)] = P[a(t)] - P[l], \text{ где}$$

$P[f(t)]$  – полученная разность,

$P[a(t)]$  – текущий кадр,

$P[l]$  – референсный кадр.

Референсный кадр снимается при отсутствии движения. Так например, чтобы зафиксировать движение в каком-либо помещении, для начала необходимо сделать и зафиксировать в памяти данного помещения при отсутствии в нем движения. После данной процедуры будет происходить вычитание фона попиксельно. Однако, из-за присутствия цифрового шума, вносимого видеокамерой, а также других факторов, влияющих на изображение, необходимо применить порог срабатывания чувствительности детектора, который применяется к разности кадров. Если значение порога меньше разности кадров, то будет обнаружено движение, в противном случае, движение зафиксировано не будет.

Достоинствами предложенного алгоритма является простота реализации и высокая производительность. К недостаткам относятся:

- Ухудшение качества обнаружения при снижении контрастности объекта с фоном;
- Низкое качество выделенного изображения объекта;
- Высокие требования к стабильности фона относительно сдвигам «система-фон» [2].

К нестабильности фона относятся:

- Изменение освещения;
- Динамический задний план;
- Изменения заднего плана.

Таким образом, любое изменения внешних факторов приведет к ложному срабатыванию детектора.

Метод *разности кадров* имеет схожую реализацию с предыдущим методом. Различие заключается в том, что вместо заднего фона из текущего кадра вычитается кадр, заснятый ранее, после чего эта разность сравнивается с порогом детектора.

Математическое представление метода имеет следующий вид:

$$P[f(t)] = P[l(t)] - P[l(t-1)], \text{ где}$$

$P[f(t)]$  – полученная разность,

$P[l(t)]$  – текущий кадр,

$P[l(t-1)]$  – предыдущий кадр.

Достоинством предложенного метода является относительная устойчивость к внешним факторам. Это означает, что при изменении положения камеры или освещения не будет вызвано долговременного ложного срабатывания детектора, а только кратковременное в пределах двух последовательных кадров.

К недостаткам относятся:

- Зависимость от частоты кадров;
- Отсутствует возможность обнаружения неподвижных объектов;
- Слабое обнаружение объектов, двигающихся с малой скоростью.

Предложенные алгоритмы детектирования объектов могут показаться простыми, однако, при их реализации может возникнуть ряд трудностей. Без модернизации описанные подходы можно применить только тогда, когда изображение содержит минимум шума и есть возможность получить изображения фона без присутствия объектов.

Для реализации методов обработки изображения в реальном времени предлагается использовать ПЛИС фирмы Altera с архитектурой ППВМ. Основным критерием выбора ПЛИС данной фирмы являются: быстрдействие, схемотехнические и конструктивные параметры, наличие развитых бесплатных САПР, возможность программирования в системе непосредственно на плате [3]. Одним из основных преимуществ является создание системы на кристалле, что позволяет реализовать на ПЛИС процессор под управлением операционной системы в режиме реального времени.

### **СПИСОК ИСПОЛЬЗУЕМЫХ ИСТОЧНИКОВ**

1. Держановский А.С., Соколов С.М. Обработка зрительных данных в системе технического зрения реального времени с использованием ПЛИС. – Москва : ИПМ им. М. В. Келдыша РАН, 2016. – 16 с.
2. Гаганов В., Конушин А. Сегментация движущихся объектов в видеопотоке // Компьютерная графика и мультимедиа. Выпуск №3. 2004. с. 45-47.
3. Стешенко В.Б. ПЛИС фирмы ALTERA: Проектирование устройств обработки сигналов / В. Б. Стешенко. – М.:ДОДЭКА, 2000. – 128 с.

**М.Г. КОСТИКОВ, В.В. МИРОНОВ**

Рязанский государственный радиотехнический университет

**АВТОМАТИЗАЦИЯ ПРОЕКТИРОВАНИЯ НОРМАТИВОВ  
ПОТРЕБЛЕНИЯ КОММУНАЛЬНЫХ УСЛУГ**

*В статье описывается подход, позволяющий автоматизировать проектирование (определение) нормативов потребления коммунальных услуг на основе анализа фактического потребления.*

*Введение.* В настоящее время в целях оптимизации и более рационального использования коммунальных ресурсов все чаще используются приборы учета (ПУ). Данная тенденция стимулируется, в том числе, и на законодательном уровне, так, Постановление Правительства предусматривает наличие повышающих коэффициентов, применяемых для расчета с абонентами, не имеющих ПУ и потребляющих ресурс по нормативам. Однако значительная часть населения не имеет ПУ, в некоторых случаях это обусловлено отсутствием возможности установки. Кроме того, нормативный расчет может применяться и к абонентам, имеющим ПУ, в случае отсутствия проверок, несвоевременной передачи показания и т.д. Таким образом, можно утверждать, что, не смотря на постепенное увеличение доли абонентов, которым производится фактический расчет (по приборам учета), расчет по нормативам не теряет своей актуальности.

Нормативы потребления устанавливаются не на федеральном, а на региональном уровне, поскольку при их проектировании необходимо учитывать географические и климатические условий в районе проживания абонента. В течение времени нормативы уточняются и могут меняться. При расчете нормативов основной задачей является как можно более точный прогноз потребления абонента, не имеющего ПУ, учитывая все условия его проживания. В данной статье рассмотрен подход, позволяющий автоматизировать проектирование нормативов потребления.

**1. Основные термины и принципы автоматизации расчета.**

Под абонентом понимается любой физический объект, потребляющий (использующий) коммунальные ресурсы для личных, семейных, домашних и иных нужд. Под коммунальным ресурсом (КР) понимается холодная вода, горячая вода, электрическая энергия, газ, бытовой газ в баллонах, тепловая энергия, твердое топливо, используемое для предоставления услуг [1]. Коммунальная услуга (КУ) - деятельность исполнителя по предоставлению абоненту КР. Расчет включает в себя выполнение ежемесячных начислений и разовых перерас-

четов по различным КУ в натуральном и денежном выражении. Расчетные операции можно разделить на две крупные группы: фактический расчет и нормативный расчет. Под фактическим расчетом понимается расчет начислений на основе потребленного абонентом объема. Объем определяется исходя из показаний ПУ. Нормативный расчет является прогнозным расчетом, начисления производятся исходя из установленных нормативов потребления. В этом случае сначала определяется объем, который предположительно потребляет абонент, а потом на этот объем производится начисление в денежном выражении.

Каждому абоненту соответствует набор характеристик и параметров потребления. Под характеристиками понимаются количественные характеристики абонента, непосредственно влияющие на размер начисления. К ним относятся число зарегистрированных граждан и животных, размер занимаемой жилой и нежилой площади и т.д. Параметр потребления – величина, определяющая характер потребления КУ. К параметрам потребления относятся такие логические (номинальные) характеристики, как режимы работы оборудования, признак наличия или отсутствия горячего водоснабжения и т.д. Под режимом понимается характер потребления абонентом того или иного ресурса, в виде совокупности соответствующих параметров потребления [2]. Каждый режим характеризуется тарифом и нормативом потребления. Норматив потребления КУ – месячный объем (количество) потребления КУ потребителем, используемый при определении размера платы за КУ при отсутствии индивидуальных, общих (квартирных) приборов учета [1].

2. Математические основы расчета объема и стоимости КУ без прибора учета (нормативное потребление).

Расчет объема КУ, израсходованного абонентом за месяц, рассчитывается по формуле

$$V^{норм} = a_1 \cdot b_1 + a_2 \cdot b_2 \dots + a_n \cdot b_n = \sum_{i=1}^n a_i \cdot b_i, \quad (1)$$

где  $a_i$  – значение физической характеристики абонента, (численность, площадь и т.д.);

$n$  – количество физических характеристик абонента,

$b_i$  – значение норматива потребления ресурса.

Величина норматива потребления определяется в зависимости от типа и характеристик установленного у абонента оборудования, а также условий окружающей среды, влияющих на объем потребления ресурса [3].



Режим потребления зависит от комбинации значений набора параметров потребления  $c_j$ , влияющих на объем расхода ресурса:

$$b_i = f_i^V(c_1, c_2, \dots, c_k), \quad (2)$$

где  $k$  – количество параметров.

3. Пример расчета за газ, используемый для пищеприготовления, подогрева воды и отопления.

Коммунальным ресурсом в данном случае является газ.

Физические характеристики  $a_i$  определены следующим образом:  $a_1$  – количество проживающих, чел.,  $a_2$  – размер отапливаемой площади,  $M^2$ .

Логические характеристики  $c_j$  определены как:  $c_1$  – режим работы газовой плиты,  $c_2$  – режим работы водогрейной колонки,  $c_3$  – режим работы отопительного прибора,  $c_4$  – признак наличия или отсутствия центрального горячего водоснабжения (ЦГВС).

Возможные значения характеристики  $c_1$  (режим работы газовой плиты): -1 – газовая плита отключена, 0 – газовая плита отсутствует, 1 – газовая плита включена и используется для пищеприготовления и подогрева горячей воды.

Возможные значения характеристики  $c_2$  (режим работы водогрейной колонки): -1 – водогрейная колонка отключена, 0 – водогрейная колонка отсутствует, 1 – водогрейная колонка включена и используется для подогрева горячей воды.

Возможные значения логической характеристики  $c_3$  (режим работы отопительного прибора): -1 – отопительный прибор отключен, 0 – отопительный прибор отсутствует, 1 – отопительный прибор подключен и используется для отопления, 2 – отопительный прибор подключен и используется для отопления только в отопительный период.

Возможные значения характеристики  $c_4$  (признак наличия или отсутствия ЦГВС): 0 – ЦГВС отсутствует, 1 – есть ЦГВС.

Режимы потребления, а также соответствующие им нормативы, тарифы и параметры потребления для указанных выше услуг приведены в табл. 1 и 2 [2].

**Таблица 1.** Нормативы потребления газа и тарифы на приготовление пищи и подогрев горячей воды

Наименование режима потребления	$D_{f_1^V}, D_{f_1^S}$	$R_{f_2^V}$ , м3	$R_{f_2^S}$ , руб.
Плита без ЦГВС	$(c_1 > 0) \& (c_2 \leq 0) \& (c_4 = 0)$	16.5	14.42
Плита с ЦГВС	$(c_1 > 0) \& (c_2 \leq 0) \& (c_4 = 1)$	10	8.74
Плита и водогрейная колонка включены	$(c_1 > 0) \& (c_2 > 0)$	25	21.86
Водогрейная колонка	$(c_1 \leq 0) \& (c_2 > 0)$	15	13.12
Плита и водогрейная колонка отсутствуют или отключены	$(c_1 \leq 0) \& (c_2 \leq 0)$	0	0

**Таблица 2.** Нормативы расхода газа и тарифы на отопление

Наименование режима потребления	$D_{f_2^V}, D_{f_2^S}$	$R_{f_2^V}$ , м3	$R_{f_2^S}$ , руб.
Отопительный котел в режиме без отключения на лето	$(c_3 = 1)$	7.2	6.29
Отопительный котел в режиме с отключением на лето	$(c_3 = 2)$	6.2	5.42
Отопительный котел отсутствует или отключен	$(c_3 \leq 0)$	0	0

Пусть, абонент - семья из трех человек, проживающая в индивидуальном доме. Размер отапливаемой площади 50.3 м2. У абонента установлены и подключены газовая плита и отопительный котел, который работает в режиме без отключения на лето. ЦГВС отсутствует.

Расчет объема ежемесячно расходуемого газа, для заданного абонента:

$$V^{норм} = a_1 \cdot b_1 + a_2 \cdot b_2 = a_1 \cdot f_1^V(c_1, c_2, c_4) + a_2 \cdot f_2^V(c_3) = 3 \cdot f_1^V(1, 0, 0) + 50.3 \cdot f_2^V(1) = 3 \cdot 16.5 + 50.3 \cdot 7.2 = 411.66 \text{ м}^3$$

#### 4. Принципы проектирования нормативов потребления.

Если учесть, что каждый режим потребления характеризуется собственными нормативами с одной стороны, а также совокупностью параметров потребления, определяющих его активность, с другой стороны, можно выделить явную связь между параметрами потребления

абонента и применяемыми к нему нормативами. Признак наличия или отсутствия ПУ (применения фактического или нормативного расчета), также является параметром потребления.

Для автоматизированного проектирования нормативов предлагается следующий алгоритм действий.

1) Из всей совокупности абонентов необходимо выделить группу абонентов, по которым осуществлялся фактический расчет в течение заданного исторического периода. По каждому из таких абонентов есть возможность получения информации о начисленных объемах в разрезе месяцев.

2) Выполнить кластеризацию по этой группе абонентов. В данном пункте необходимо разделить абонентов на подгруппы в соответствии с совокупностью значений параметров потребления. Другими словами выполняется кластеризация абонентов с фактическим потреблением по режим потребления.

3) Восстановление отсутствующих данных. Теоретически абоненты могут в течение некоторого периода не передавать информацию о потребленных объемах, передав в последующем суммарный объем за несколько месяцев. Учитывая тот факт, что нормативы устанавливаются за один месяц, подобные случаи могут исказить конечный результат. Таким образом, необходимо провести интерполяцию и восстановить промежуточные данные. С точки зрения Постановления Правительства расчетным периодом является месяц, поэтому для корректного расчета необходимы потребленные объемы за каждый месяц рассматриваемого периода.

4) Отсевание сбойных результатов. Подавляющая часть абонентов будут иметь потребление в некотором интервале. Однако могут возникать относительно редкие случаи, когда абоненты потребляют либо чрезвычайно малый, либо чрезвычайно большой объем. Эти случаи, не смотря на свою редкость, могут также довольно сильно исказить результат проектирования [4]. Особенно это касается чрезвычайно большого потребления. Таким образом, на данном этапе необходима разработка критерия позволяющего отсеивать сбойные результаты.

5) Проведение статистического анализа потребления. На данном этапе производится анализ результатов. Необходима разработка метода анализа. Простейший подход – нахождение среднего объема за некоторый исторический период даст довольно поверхностную оценку, поскольку не учитывает сезонный тренд, который может оказывать влияние на большинство КР [4]. По сути, на этом этапе получают новые значения нормативов потребления, для каждой группы абонентов, а значит и для каждого режима.

Модули управления разработаны на основе общих идей теории управления, развитые в работах [5, 6, 7].

*Заключение.* Рассмотренный подход позволяет автоматизировать задачу проектирования нормативов потребления. Задача актуальна и, как правило, раз в несколько лет возникает перед энергетическими комиссиями регионов. Стоит отметить, что задача является довольно сложной, в статье приведен лишь поверхностный обзор алгоритма определения нормативов. Почти каждый пункт этого алгоритма, например, кластеризация, интерполяция, отсеивание сбойных результатов, представляет собой отдельную задачу, нуждающуюся в детализации. Однако использование предлагаемого подхода позволит, во-первых автоматизировать процесс проектирования, а во-вторых наиболее точно спрогнозировать нормативы потребления, так как расчет производится на основе анализа фактического потребления, полученного в этом же регионе по абонентам с идентичными параметрами.

### СПИСОК ИСПОЛЬЗУЕМЫХ ИСТОЧНИКОВ

1. Постановление правительства Российской Федерации о порядке предоставления коммунальных услуг гражданам от 6 мая 2011г. №354.
2. Аникеев С.В., Костиков М.Г., Маркин А.В. Автоматизация настройки расчета за коммунальные услуги в системе «Абонент». // Сборник "Математическое и программное обеспечение вычислительных систем". – Рязань, РГРТУ, 2011.
3. Аникеев С.В., Костиков М.Г., Маркин А.В. Математическая модель определения активных услуг на основе матричной алгебры. // Вестник Рязанского государственного радиотехнического университета. 2016. №6. С. 54-62.
4. Аникеев С.В., Борзых В.Е., Маркин А.В., Шубенков Е.Е. Задача статистической оценки норм потребления газа населением // Информатика и прикладная математика: Межвузовский сборник статей РГПУ имени С.А. Есенина. Рязань, 2005.
5. Миронов В.В., Северцев Н.А. Методы анализа устойчивости систем и управляемости движением. / ВЦ РАН.- М.: Изд-во РУДН, 2002. 165 с.
6. Миронов В.В., Митрохин Ю.С. Задача об управляемости систем оптимального регулирования и методы ее решения // Вестник РГРТА: Рязань, 2000. Вып. 7. С. 57-64.
7. Миронов В.В. Фундаментальные математические проблемы. Комментарии / XXVIII межд. науч. конф. «Математические методы в технике и технологиях». – Рязань: РГРТУ, 2015. С. 3-12.

**А.А. КУЗНЕЦОВ**

Рязанский государственный радиотехнический университет

## **АЛГОРИТМЫ ВИЗУАЛИЗАЦИИ ДЕРЕВЬЕВ БОЛЬШОЙ РАЗМЕРНОСТИ**

*Рассматриваются алгоритмы визуализации деревьев, содержащих большое количество вершин.*

При помощи графов можно представить любую информацию, которую можно описать с помощью объектов (вершин) и связей между ними (ребра, дуги). Многие системы визуализации информации нуждаются в таком отображении графов, при котором получаемые изображения просты для рассмотрения и понимания. К примерам графов таких систем можно отнести:

- граф цитирований публикаций, в котором вершинами отображаются публикации разных годов, а ребрами указаны ссылки одних публикаций на другие. Число публикаций может быть большим;

- концептуальные карты. Они представляют собой графические инструменты для представления и поиска знаний и визуализируются в виде графов, в узлах которых располагаются понятия предметной области, а в качестве связей выступают отношения между понятиями [1].

- модель учебной дисциплины, рассмотренная в статье [2]. Она представляет собой дерево, корнем которого является сама дисциплина, узлами – подразделы дисциплины, а листьями – конечные понятия.

Целью данной статьи является обзор алгоритмов, предназначенных для визуализации особого вида графов, деревьев, в виде, удобном для чтения и анализа. Такое представление должно раскрывать структуру дерева наиболее лучшим образом, чтобы человеку было легче в нем ориентироваться.

Изображение графа на плоскости – это отображение вершин графа в множество точек плоскости [3]. В данной статье будут рассмотрены изображения корневых деревьев, но их можно применить и к свободным деревьям (без выделенного корня), обозначив определенную вершину (например, одну из центральных) в качестве «корня». Среди основных способов изображения деревьев можно выделить следующие:

- послойный (layered);
- радиальный (radial) [3];
- шаровой (balloon);
- *hv*-изображение (*hv-drawing*) [3].

В книге [3] описано поуровневое изображение бинарного дерева, которое можно обобщить на случай произвольного корневого дерева. Такое размещение использует нисходящий метод (от корня к листьям), при котором у вершин фиксируется координата по оси ординат ( $y$ -координата), а вычисление  $x$ -координат вершин отличается в различных алгоритмах, использующих эту схему. Ниже будет предложен послыйный способ рисования дерева, схожий с поуровневым, но использующий восходящий метод (от листьев к корню).

Послойный метод, как и поуровневый, предполагает размещение вершин на определенных уровнях (слоях). Авторы книг [4, 5] по-разному трактуют понятие «уровень вершины», поэтому в дальнейшем будем использовать понятия высоты и глубины.

Высота узла в дереве (*height*) – это длина (самого длинного) пути из этой вершины в самый удаленный лист [3, 5]. Высота листа равна нулю. Глубина вершины дерева (*depth*) – это длина пути от корня дерева до этой вершины [3, 5]. Длина пути между двумя вершинами в обоих случаях – это число ребер в цепи, соединяющей эти вершины.

Первый вариант послыйного изображения задействует понятие высоты узла и представлен на рис. 1. Такое размещение требует одного обхода дерева в обратном порядке (постфиксный порядок, *LRT*-обход, *post-order*), что занимает линейное время. Алгоритм начинается свою работу с инициализации количества листьев  $leaves = 0$  и вызова процедуры *обратный\_обход*, в которую в качестве аргумента передается корень дерева. Процедура имеет вид:

проц *обратный\_обход* (*вершина v*)

1. если  $v$  не лист

2. [вызов] для каждого ее потомка  $q$

*обратный\_обход* ( $q$ )

3. [посещение]  $y = height(v) \cdot R_h$

$$x = x_{\min} + \frac{x_{\max} - x_{\min}}{2}$$

4. иначе

5. [посещение]  $y = 0$

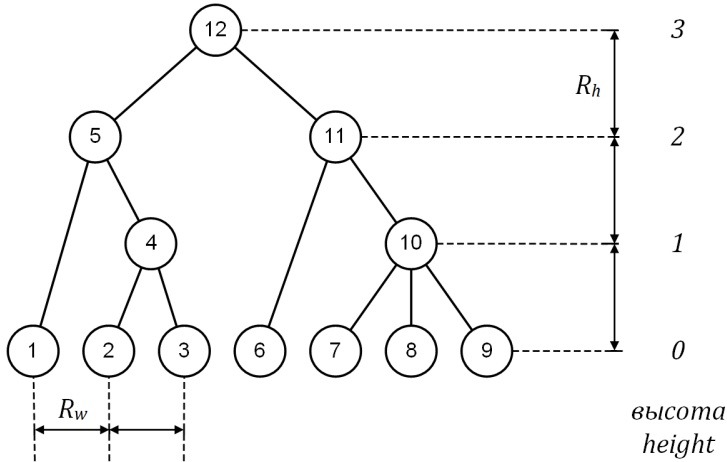
$$x = leaves \cdot R_w$$

$$leaves = leaves + 1$$

На 3 шаге  $x_{\min}$  – минимальная  $x$ -координата среди потомков вершины, а  $x_{\max}$  – максимальная. Высота узла  $v$  (шаг 3) определяется по формуле (1):

$$height(v) = \max_{q \in Q} height(q) + 1 \quad (1)$$

где  $Q$  – множество потомков вершины  $v$ .



**Рис. 1.** Пример послойного изображения дерева (по высоте)

Второй вариант послойного изображения задействует понятие глубины узла и представлен на рис. 2. Алгоритм также начинается с инициализации количества листьев и вызова процедуры обратный\_обход, в которую в качестве первого аргумента передается корень, а в качестве второго – глубина корня  $depth = 0$ . Процедура имеет вид:

проц *обратный\_обход* (*вершина v*, *глубина depth*)

1. если  $v$  не лист

2. [вызов] для каждого ее потомка  $q$

*обратный\_обход* ( $q$ ,  $depth + 1$ )

3. [посещение]  $y = -depth \cdot R_h$

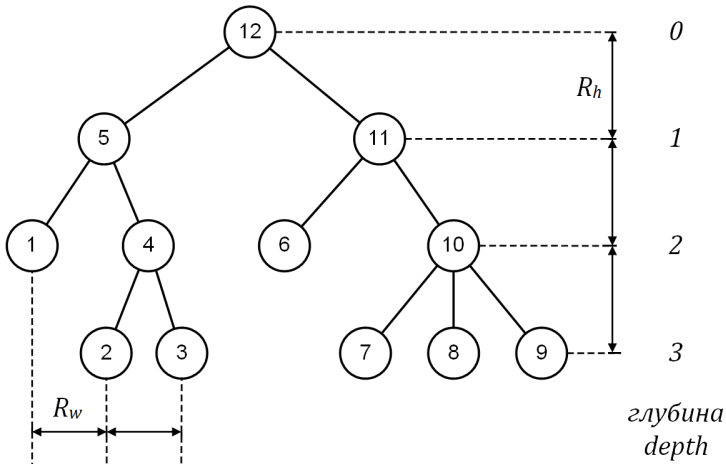
$$x = x_{\min} + \frac{x_{\max} - x_{\min}}{2}$$

4. иначе

5. [посещение]  $y = -depth \cdot R_h$

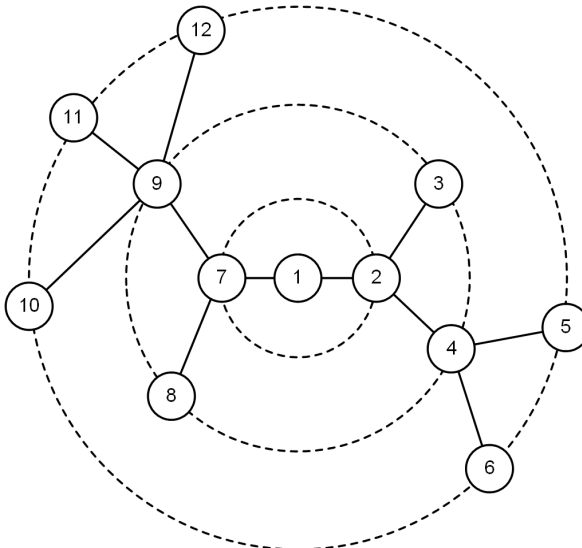
$$x = leaves \cdot R_w$$

$$leaves = leaves + 1$$



**Рис. 2.** Пример послойного изображения дерева (по глубине)

Радиальное изображение, приведенное в книге [3], представлено на рис. 3. Это размещение является вариацией поуровневого размещения, только уровни имеют вид концентрических окружностей, а поддерева занимают секторные сегменты.



**Рис. 3.** Пример радиального изображения дерева



Выбор угла  $\beta_p$  секторного сегмента  $W_p$  для поддерева с корнем  $p$  определяется с учетом числа вершин  $l(p)$  в поддереве следующим образом: пусть  $p$  лежит на уровне  $C_i$  (рис. 4), тогда для каждого сына  $q$  вершины  $p$  имеем:

$$\beta_q = \min\left(\frac{l(q)\beta_p}{l(p)}, \tau\right) \quad (2)$$

где  $\tau$  – угол области  $F_p$ , определяемый точками пересечения  $p(a)$  и  $p(b)$  уровня  $C_{i+1}$  и касательной к уровню  $C_i$ , проведенной через точку  $p$ . Ниже приведен один из возможных вариантов реализации радиального рисования дерева.

Перед вычислением координат вершин, требуется обойти дерево в обратном порядке с целью вычисления числа вершин  $l(p)$  в каждом поддереве  $p$  (в качестве  $l(p)$  можно использовать число листьев поддерева) и вычисления высоты дерева  $height$ , если нужно заранее знать размер изображения дерева. Само вычисление координат  $x$  и  $y$ , при заданном расстоянии между уровнями  $levelSpace$ , происходит следующим образом:

[инициализация] установить корню дерева  $p$ :

$$p(a) = 0$$

$$p(b) = 2\pi$$

$$x = y = levelSpace \cdot height$$

Выполнить прямой обход (префиксный, TLR-обход, pre-order):  
 проц *прямой\_обход* (*вершина p*, *глубина depth*)

1. если  $p$  не лист

2. [посещение] принять угол  $a = p(a)$ ,

$$\text{сегмент } \tau = |b - a|$$

для каждого потомка  $q$

$$\text{угол начала сегмента } q(a) = a$$

$$\text{угол приращения } ha = \frac{1}{2} \cdot \tau \frac{l(q)}{l(p)}$$

$$\text{приращение угла } a = a + ha$$

$$x\text{-координата } q(x) = p(x) + \cos a \cdot (depth + 1) \cdot levelSpace$$

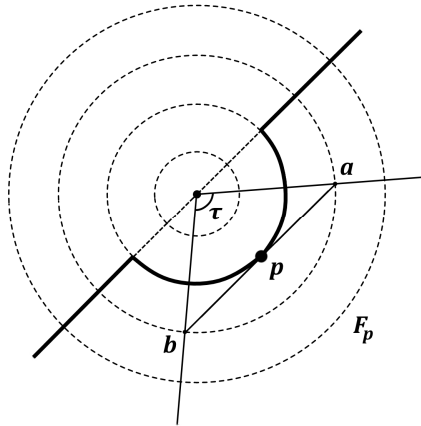
$$y\text{-координата } q(y) = p(y) + \sin a \cdot (depth + 1) \cdot levelSpace$$

$$\text{приращение угла } a = a + ha$$

$$\text{угол конца сегмента } q(b) = a$$

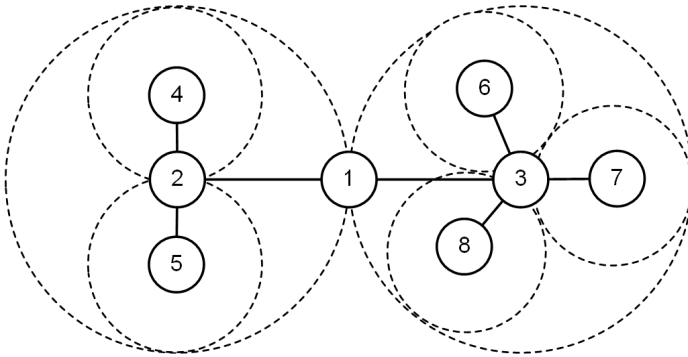
3. [вызов] для каждого потомка  $q$

$$\text{прямой\_обход}(q, depth + 1)$$



**Рис. 4.** Область размещения при радиальном размещении [3]

Шаровое изображение представлено на рис. 5. При таком размещении все вершины выстраиваются вокруг родительского узла, а не вокруг корневого, как при радиальном размещении.



**Рис. 5.** Пример шарового изображения дерева

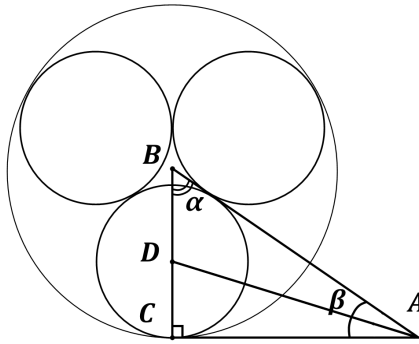
При определении области размещения можно использовать теорему о биссектрисе [6] (рис. 6):

$$\frac{DB}{CD} = \frac{AB}{AC} \quad (3)$$

Зная, что  $\alpha = \frac{\pi}{n}$ , где  $n$  – число потомков корневой вершины и  $\beta = 90 - \alpha$ , радиус  $r$  окружности вокруг одного из потомков ( $CD$  на рис. 6) равен:

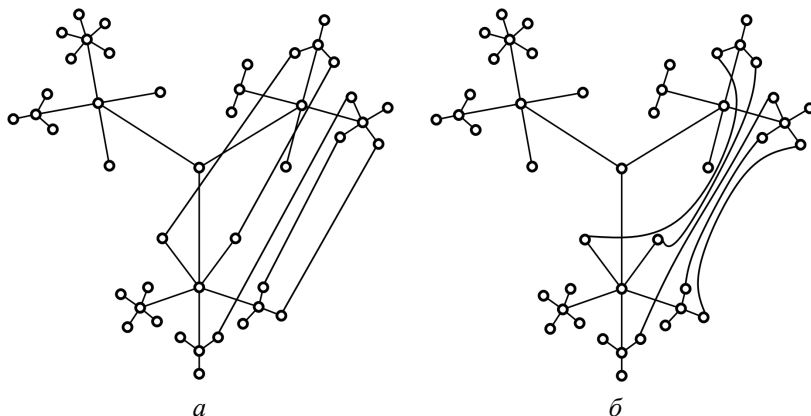
$$r = \frac{R}{1 + \frac{1}{\cos \beta}} \quad (4)$$

где  $R$  – радиус окружности, описанной вокруг корневой вершины ( $BC$  на рис. 6).



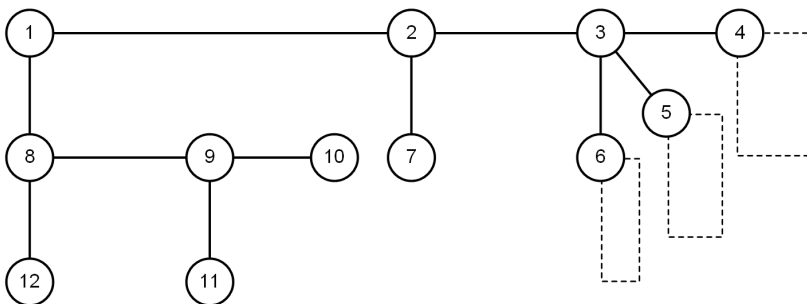
**Рис. 6.** Область размещения при шаровом изображении

Если в большом графе содержится большое количество ребер (рис. 7а), его удобно изобразить с помощью шарового размещения, применив метод связывания ребер [7]. Такое представление показано на рис. 7б. Метод связывания ребер использует принцип прокладки кабельных сетей: объединение проводов в жгуты. Чтобы ребра графа объединить в «жгуты», нужно их отображать кривыми, причем в качестве таких кривых подходят кусочно-заданные кубические  $B$ -сплайны (в частности кривые Безье), двумя опорными точками которых могут выступать координаты одного из предков первой вершины и одного из предков второй вершины. Кривые Безье удобны при написании приложений, так как их отрисовка уже реализована в библиотеках разных языков программирования (например `QPainterPath::cubicTo(c1, c2, endPoint)` в библиотеке `Qt` [8]).



**Рис. 7.** Представление графа с большим числом ребер:  
*a* – без применения связывания ребер, *б* – с применением

*Hv*-изображение, приведенное в книге [3], представлено на рис. 8. Общая схема построения *hv*-изображения состоит в рекурсивном построении поддеревьев с их объединением с помощью горизонтальной (или вертикальной) комбинации, откуда и название этого способа (*horizontal vertical*). Данный алгоритм также характеризуется линейной сложностью.



**Рис. 8.** Пример *hv*-изображения дерева

Каждый из вышеуказанных способов изображения деревьев обладает своими преимуществами и может быть применен в различных областях.

Послойное рисование уместно использовать при визуализации любых иерархических структур (например, классификаторов). Так, просматривая дерево от корня к листьям, человеку удобнее постепенно

изучать структуру, уровень за уровнем, углубляясь в содержание исследуемого объекта.

Радиальное изображение также можно применить при отображении иерархических данных. Но такой способ размещает листья дерева по окружности. Это свойство дает возможность визуально увеличить расстояние между вершинами одного уровня по мере их отдаления от корня.

Шаровое размещение позволяет визуально сконцентрировать листья подеревьев в группы, чтобы можно было детальнее рассмотреть структуру конкретного узла. Эту особенность можно использовать, например, при отображении графов цитирований, так как вершины (публикации) удобно собирать в группы по годам, а ребра (ссылки одних публикаций на другие) можно «стягивать в жгуты» (метод связывания ребер).

Главное преимущество *h<sub>v</sub>*-изображения заключается в наименьшем занимаемом пространстве. Его применение делает возможным просмотр всей структуры дерева большого размера без сливания смежных вершин.

## СПИСОК ИСПОЛЬЗУЕМЫХ ИСТОЧНИКОВ

1. Корячко В. П., Бакулева М. А., Орешков В. И. Интеллектуальные системы и нечеткая логика. – М.: КУРС, 2017. – 352 с.
2. Кузнецов А. А. Изучение процессов освоения материала учебных курсов // Современные технологии в науке и образовании – СТНО-2018. – Рязань: РГРТУ, 2018.
3. Касьянов В. Н., Евстигнеев В. А. Графы в программировании: обработка, визуализация и применение. – СПб.: БХВ-Петербург, 2003. – 1104 с.
4. Евстигнеев В. А., Касьянов В. Н. Теория графов: алгоритмы обработки деревьев. – Новосибирск: ВО «Наука». Сибирская издательская фирма, 1994. – 360 с.
5. Ахо А., Хопкрофт Дж., Ульман Дж. Построение и анализ вычислительных алгоритмов. – М.: Мир, 1979. – 536 с.
6. Понарин Я. П. Элементарная геометрия. Т. 1. — М.: МЦНМО, 2004. — 312 с.
7. Визуализация графов. Метод связывания ребер [Электронный ресурс], URL: <https://habrahabr.ru/post/116758/> (дата обращения 26.03.18).
8. Qt Documentation. QPainterPath Class [Электронный ресурс], URL: <http://doc.qt.io/qt-5/qpainterpath.html> (дата обращения 26.03.18).

**В.Ю. ЛИКУЧЕВ**

Рязанский государственный радиотехнический университет

## **СПЕЦИАЛЬНАЯ ЗАДАЧА ДЛЯ ГРУППЫ БЕСПИЛОТНЫХ ЛЕТАТЕЛЬНЫХ АППАРАТОВ**

*В рамках статьи обосновывается переход к глубокому рассмотрению принципов организации и функционирования групп беспилотных летательных аппаратов, а также формулируется специальная задача для такой группы с последующим рассмотрением возможных путей ее решения.*

Сегодня беспилотные летательные аппараты (далее – БПЛА) представляют собой интенсивно развивающийся сегмент авиатехники, а их разработка является перспективным направлением. Возможности БПЛА позволяют выполнять самый широкий круг задач, начиная от мониторинга местности и доставки грузов, заканчивая проведением военных операций, таких как боевая разведка и нанесение огневых ударов.

Расширяя спектр применений беспилотных летательных аппаратов и беспилотных систем вообще, необходимо перейти к детальному рассмотрению принципов организации и функционирования целых групп, состоящих из объектов данного типа, ведь с увеличением количества подобных технических единиц возрастает сложность решаемых задач. С другой стороны, системы управления обширными техническими группами требуют дополнительных мер повышения и сохранения эффективности, особенно в тех условиях, где на работу управляющих и подконтрольных органов накладываются многочисленные ограничения. В случаях, когда беспилотные группы используются в военно-тактических целях, основной проблемой становится именно организация эффективного автономного управления, так как средства радиоэлектронной борьбы противника нередко дестабилизируют процессы технического функционирования.

Увеличение степени автономности БПЛА возможно за счет создания системы управления, опирающуюся на совокупность известных принципов определения положения аппарата в пространстве и параметров его движения, то есть на совокупность различных методов навигации. Для БПЛА применимы инерциальный, спутниковый и видеоинформационный методы. Первый оперирует показаниями инерциальных устройств, расположенных на борту аппарата (гироскопы, акселерометры, высотомеры). Второй – данными со спутников, а третий – информацией, получаемой приборами внешнего видеонаблюдения. Комбинация перечисленных методов позволяет центральной системе

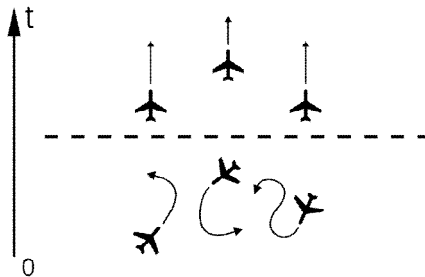
формировать точное представление о движении аппарата для последующей выдачи управляющих команд на подконтрольные органы. Данные от разных навигационных компонентов взаимодополняют и компенсируют друг друга, благодаря чему корректировка движения осуществляется даже если в одном из устройств произошло накопление погрешности.

Особую роль в построении управляющих систем для БПЛА играет условие миниатюризации электронных средств. Появление микроэлектромеханических систем, таких как электронные гироскопы и акселерометры, позволило во много раз сократить размеры навигационных платформ, а значит конструкций самих аппаратов. Последнее является важным условием для выполнения БПЛА обширного комплекса задач.

Таким образом, создание бортовой системы управления на основе комбинированного метода навигации, а также с использованием современной компонентной базы позволяет, во-первых, увеличить степень автономности БПЛА, а, во-вторых, сократить их габариты.

Теперь рассмотрим специальную задачу для БПЛА. Ее сущность заключается в следующем: по заданному маршруту подлететь к определенной точке и занять свое положение в строю идентичных БПЛА; далее – в определенный момент времени выполнить некоторое действие, не связанное с изменением положения в пространстве; далее – возвратиться на место старта. Точки, откуда стартуют БПЛА могут быть разнесены так, что масштабированная топология их расположения не совпадает с топологией составляемой формации.

В общем виде задача для одного БПЛА может быть сформулирована следующим образом: требуется занять строго определенное место в воздушном пространстве к строго определенному моменту времени. Групповая формулировка: множество изначально разнесенных в пространстве БПЛА (старт из разных точек) должно быть упорядочено в четкую по структуре формацию для совместного выполнения некоторого действия, либо функции. Расстояние между элементами структуры при выполнении основной групповой функции не должно выходить за пределы определенной погрешности. Условие для множества БПЛА проиллюстрировано на рис. 1.



**Рис. 1.** Условие задачи о формации

Для выполнения задачи необходимо снабдить БПЛА управляющей системой с возможностью занесения в нее данных о маршруте движения. Сравнивая эти данные с показаниями устройств навигации, система должна будет путем выработки управляющих команд, полученных на основе оценки сравнения, корректировать движение аппарата согласно заданному маршруту.

Однако, всякого рода помехи могут усложнить процесс выделения полезного сигнала с навигационных систем относительного принципа действия, а также с наземных пунктов управления, тем самым дестабилизировать работу подконтрольных объектов. Таким образом, БПЛА, оснащенные системой управления, которая опирается только на комбинирование навигационных методов, способны выполнять лишь первую часть задачи, то есть подлет к месту формации. На последующих этапах (построение формации, функционирование в составе формации) система управления должна обеспечивать еще более высокую и постоянную во времени точность регулирования пространственных состояний единиц (их координат, углов наклона), и, следовательно, устойчивость всей группы, а для этого необходимо по возможности минимизировать долю внешнего регулирования, или, другими словами, приблизить систему к самоорганизующейся.

Точность взаимного регулирования единиц группы возрастает с их пространственным сближением. Находясь на относительно близком расстоянии друг от друга и пользуясь специальными техническими средствами, объекты становятся чувствительными к изменению состояний своих «соседей». Так при отрыве одного БПЛА от общей формации, остальные с помощью радиоэлектронных сигналов произ-



водят корректировку его пространственного положения, стремясь вернуть его в состав коллектива.

Для построения такого рода системы необходимо разработать формальную модель ее функционирования, а также варианты аппаратной реализации средств взаимодействия элементов внутри группы. Эти вопросы призвана решать вводимая *концепция искусственных коллективов*. Данная концепция должна объединить многие отрасли науки и техники, связанные с построением самоорганизующихся систем, действующих с высокой точностью и мало подверженных мешающим воздействиям. Разработка концепции активно ведется в настоящее время: рассматриваются всевозможные структуры групп объектов, алгоритмы взаимодействия внутри групп, законы регулирования пространственных состояний.

Искусственный коллектив – это совокупность подобных технических единиц, совместно выполняющих определенную задачу при условии полной автономности. Основное отличие таких групп от роевых структур – малые погрешности состояний элементов при непосредственном функционировании. Пример искусственного коллектива – формация беспилотных летательных аппаратов, реализующая функцию фазированной антенной решетки для подавления радиоэлектронных сигналов объекта противника. Расстояние между элементами решетки, очевидно, должно быть фиксированным.

Развитие управляющих систем в аспектах автономности и малогабаритности позволит вывести БПЛА самых примитивных конфигураций на новый уровень тактического использования, обеспечит возможность их многоцелевого применения не только в гражданских, но и в боевых приложениях, что может повысить военный потенциал страны.

**Е.С. МАХНАЧ**

Рязанский государственный радиотехнический университет

## **ИССЛЕДОВАНИЕ МЕТОДОВ И СРЕДСТВ РЕАЛИЗАЦИИ СИСТЕМЫ «УМНЫЙ ДОМ»**

*Рассматриваются методы реализации системы автоматизированного управления «Умный дом».*

Реализация систем управления зданиями – это новейшее направление в области автоматизации, которое называется Smart House – от английского «умный дом». Это может быть не только жилой дом, но также и государственное учреждение, стадион и даже аэропорт. Это набор совокупных технических средств и программного обеспечения, которые используются для контроля всей системы автоматизации инженерных подсистем в целом. К таким подсистемам относятся отопление, водоснабжение, кондиционирование, освещение, подсистемы доступа, охраны и безопасности, аудио-видеотехники и ряд других. Чем сложнее эти подсистемы и чем больше функций они выполняют, тем управление ими становится все сложнее, при этом значительно растут расходы на содержание этих подсистем.

Автоматизация инженерных систем имеет ряд неоспоримых достоинств на всех уровнях, не зависимо от глубины внедрения. Это уменьшение расходов на электроэнергию, эксплуатационных затрат, повышение безопасности, контроль износа оборудования, упрощение управления системой в целом, и, как следствие, предотвращение аварийных ситуаций, возможность создания индивидуальной программы работы для каждой подсистемы и многое другое.

Структура системы «умный дом» состоит из:

1. Контроллер – регулирует работу всей системы, выполняя команды самостоятельно или исполняя команды, которые выбрал человек заранее. Именно контроллер обеспечивает правильную работу всей системы в целом так как, с него подаются сигналы на устройства, которые оказывают воздействие на объект.

2. Информационная шина – служит коммуникацией между управляющими устройствами и исполнительными.

3. Исполнительные устройства – реализуют функцию, определенную контроллером.

Существуют два типа реализации системы управления “умным домом”: централизованная и распределенная.

Основу централизованной системы составляет центральный контроллер с подключенными к нему модулями. Такой тип построения позволяет реализовывать сложные системы с большим количеством исполнительных устройств, однако эта система является уязвимой, так как при выходе из строя контроллера вся система становится не работоспособной.

В распределенных системах управления «умным домом» каждое устройство является самостоятельным. Такой тип систем отличается надежностью, так как каждое устройство является автономным. Но из-за того, что каждый модуль имеет собственный контроллер, распределенный тип систем является более громоздким, медленным и дорогостоящим.

Управление технологией «умного дома» ведется через:

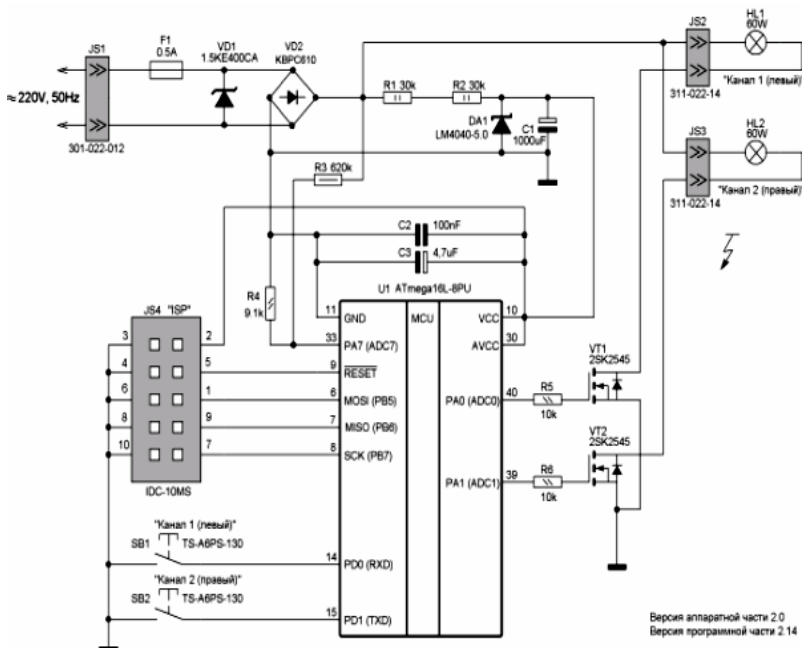
- пульт дистанционного управления
- компьютер
- смартфон.

Система управления освещением является одной из составных частей системы «Умный дом». Устройство, которое предназначено для регулирования яркости, называют производным словом «диммер» от английского «dimmer», или светорегулятор. Управление освещенностью с их помощью позволяет не только повысить экономический эффект, но и предоставляет возможность реализовать множество дополнительных функций в регулировании освещением. Энергосберегающие люминесцентные лампы, которые довольно широко используются, при всех своих преимуществах имеют, во-первых, высокую цену, а во-вторых, для них необходима сложная схема регулирования яркости.

Существующие на данный момент устройства регулирования яркости ламп имеют следующие недостатки:

- возможность регулирования только одного канала;
- отсутствие стабилизации яркости;
- влияние радиопомех.

Поэтому необходимо создать устройство, регулирующее яркость ламп, в которых отсутствуют перечисленные недостатки, для установки в двухканальный светильник. Основой устройства, функциональная схема которого изображена на рис. 1, является микроконтроллер ATmega16L корпорации ATMEL. Это 8-разрядный высокопроизводительный AVR микроконтроллер с малым потреблением, относительно недорогой, простой и широко используемый.



**Рис.1.** Схема принципиальная электрическая устройства автоматического регулирования света

На сегодняшний день в мире созданы и используются различные системы управления «умным домом». Но, к сожалению, они практически не используются в России без адаптации, так как разработаны с учетом иностранных стандартов. Также фактически отсутствует поддержка подобных отечественных решений. При этом основная трудность состоит в частном, а не в общем подходе к реализации данных систем. Именно поэтому цель разработки единых принципов и универсальных моделей и технологий объединения инженерного оборудования является актуальной.

### СПИСОК ИСПОЛЬЗУЕМЫХ ИСТОЧНИКОВ

1. Концепция системы «Умный Дом» — [Электронный Ресурс]. URL: <http://www.ascentis.ru/smart/smtheory/39-smtheorycon>.
2. Мологонцева О.Ф. Проектирование цифровых устройств на однокристалльных микроконтроллерах / В.В. Сташин, А.В. Урусов, О.Ф. Мологонцева. - М. :Энергоатомиздат, 1990. – 224с.

**М.С. МИХЕЕВА**

Рязанский государственный радиотехнический университет

## **ИССЛЕДОВАНИЕ МЕТОДОВ И СРЕДСТВ ТЕХНОЛОГИЧЕСКОЙ ТРЕНИРОВКИ ИЗДЕЛИЙ РЭС**

*Рассматриваются теоретические вопросы применения технологической тренировки изделий радиоэлектронных средств.*

Технологическая тренировка представляет собой испытания аппаратуры с целью выявления и устранения приработочных отказов.

Тренировка является производственным способом обеспечения надежности и предназначена для:

- выявления и устранения дефектов, допущенных в процессе изготовления изделия;
- выявления и устранения дефектных комплектующих изделий;
- стабилизации электрических параметров изделия;
- контроля технологического процесса изготовления изделия.

Длительность тренировки изделия согласно ОСТ4.052.000-86  $T_{тр}, ч$  определяется по формуле:

$$T_{тр} = T_H \cdot \ln \frac{1}{1 - \rho \frac{1}{mQ}} \quad (1)$$

где  $T_H$  – среднее время выявления элементов пониженной надежности при тренировке, ч;

$\rho$  – требуемая вероятность обнаружения всех элементов пониженной надежности за время тренировки;

$m$  – коэффициент пропорциональности, определяющий количество элементов пониженной надежности в изделии в зависимости от его сложности;

$Q$  – количество комплектующих электрорадиоэлементов в изделии, шт.

При разработке технологической тренировки значения  $m, \rho, T_H$  выбирают исходя из результатов тренировки изделий, сходных с разрабатываемым изделием областью применения, конструктивно-техническими решениями, основной элементной базой.

При отсутствии указанных данных в соответствии с ОСТ4.052.000-86 рекомендуются следующие значения  $m, \rho, T_H$ :

$$m = 10^{-3}; \rho = 0,99; T_H = 32ч$$

Технологическая тренировка проводится на заключительном этапе изготовления радиоэлектронных средств, перед предъявлением опытных образцов на предварительные испытания.

Наиболее отработанной в настоящее время является методика отбраковочных испытаний, предусмотренная стандартом США MIL-STD-883.

Технологическая тренировка проводится на заключительном этапе изготовления РЭС перед предъявлением опытных образцов на предварительные испытания.

Объективной тенденцией улучшения конструкций РЭС является непрерывный рост её сложности, что можно объяснить расширением области разрешаемых проблем при одновременном повышении требований к производительности и надежности работы. В настоящее время довольно остро стоит вопрос о качестве изделий, выпускаемых на предприятиях. Одним из методов отбраковки изделия является технологическая тренировка изделий, режимы которой регламентируются ОСТ 4.052.000 – 86.

Известно, что отказы возможны даже в хорошо освоенном производстве. По этой причине распространенными способами повышения качества и надежности изделий является технологическая тренировка их на выходном контроле завода-изготовителя. Теория и опыт показывают, что воздействие температуры, напряжения и мощности при тренировке ускоряют ранние отказы, тем самым сокращая их число при эксплуатации радиоэлектронных средств. В условиях непрерывного усложнения современных РЭС многие потребители считают применение тренировок на входном контроле важнейшим средством повышения качества и надежности РЭС.

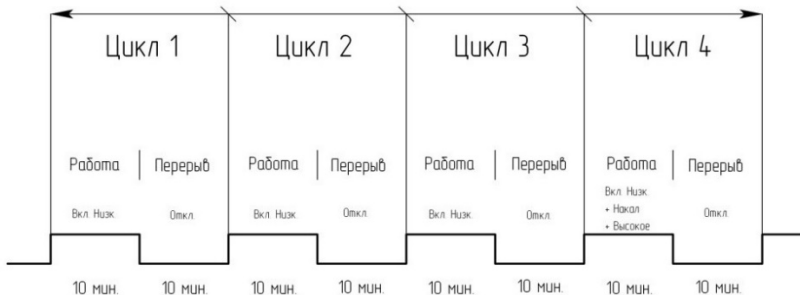
В России в реальных условиях эксплуатации изделия обычно работают периодически, а это значит, что они подвергаются неожиданной подаче электрического напряжения, а затем внезапному его выключению. Зачастую изделия надежно работают в постоянно включенном режиме, но их надежность уменьшается в зависимости от включения и выключения режима. Поэтому при электрической тренировке в последние годы стал широко применяться прерывистый режим – циклическое изменение электрических воздействий "включено-выключено", получивший название – режим энергоциклирования.

Включение-выключение напряжения питания и изменение напряжения питания от минимального до максимального значения применяются как дополнительные воздействия, ускоряющее процесс технологической тренировки.

Условный пример энергоциклирования изображен на рис. 1.

В течение 1 – 3 циклов одновременно на изделие подаются низкие напряжения: +5В, +15В, -15В, +28,5В.

В течение 4 цикла сначала подаются все низкие напряжения одновременно, затем в течение 40 с подается напряжение «Накал» ~115В (1000Гц) и после паузы 12 секунд производится включение напряжения «Высокое» ~115В (1000Гц). Отключение производится в обратном порядке: отключение «Высокое», отключение «Накал», одновременное отключение низких напряжений.



**Рис. 1.** Период подачи питающих напряжений на изделие

Большое распространение данный режим получил при изготовлении мощных транзисторов в пластмассовых корпусах. Если мощный транзистор испытывает многократные включения и выключения, то при низкокачественной пайке кристалла этот дефект будет выявлен при тренировке. При этом перепад температуры перехода, который соответствует режимам включения и выключения, будет более 100°C.

Для транзисторов, имеющих меньшую мощность, данный перепад температуры, соответствующий режиму "Включено-выключено", не имеет большой значимости, следовательно, до первого отказа они выдерживают существенно большее (на два и более порядка) число циклов, чем транзисторы, имеющие большую мощность.

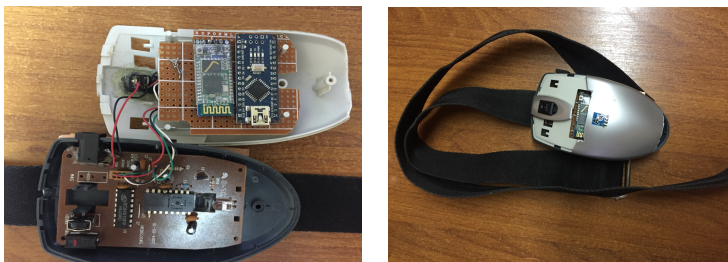
**В.Л. НГУЕН**

Рязанский государственный радиотехнический университет

## **ПОЛУЧЕНИЕ ЭКСПЕРИМЕНТАЛЬНЫХ ДАННЫХ С ПОМОЩЬЮ СИСТЕМЫ ИЗМЕРЕНИЯ И КОНТРОЛЯ ПАРАМЕТРОВ ДЫХАНИЯ**

*В статье представлены экспериментальные результаты, полученные с помощью системы измерения и контроля параметров дыхания, которые полезны в медицине при проведении процедуры физиотерапии.*

Для проведения эксперимента используется датчик, который состоит из части компьютерной мыши, в котором есть оптический элемент слежения: оптический датчик, держатель, светодиод, пластиковые элементы; модуль преобразования сигнала на основе Arduino-совместимой микроконтроллерной платы и модуль дистанционной передачи данных Bluetooth HC-05[1].



**Рис. 1.** Макет рабочего устройства

Живот испытуемого обхватывается нерастяжимым ремнем, который, в зависимости от дыхания пациента, будет смещаться на вдохе в одну сторону, на выдохе в противоположную. Для того чтобы ремень смог смещаться он крепится к эластичной ткани. Датчик расположен над поверхностью ремня и измеряет его смещение. Эти данные передаются в компьютер с помощью модуля Bluetooth, затем с модуля преобразования данных поступают на определенный порт персонального компьютера для анализа параметров дыхания.

Программа для обработки данных написана в среде Labview[2]. В программе можно получить важные параметры дыхания такие как: частота, фаза, скорость вдоха и выдоха, период, амплитуды, графическое отображение процесса дыхания в реальном времени. Программа сохраняет данные в файле формата .txt, чтобы потом использовать их для проведения исследования пациента.





124.597	-9.989	5.220	11.000	53.515	-37.010
118.453	3.958	4.799	13.000	47.600	-30.573
118.205	-6.303	5.220	11.000	51.634	-26.980
108.332	-11.198	4.680	13.000	37.085	-24.374
68.910	-17.061	4.920	12.000	34.343	-28.080
63.272	-24.149	5.219	11.000	37.885	-23.254
53.140	-26.926	4.740	13.000	36.385	-30.111
51.671	-33.633	4.680	13.000	27.212	-28.262
31.678	-25.816	4.679	13.000	25.985	-21.817
34.993	-8.573	4.440	14.000	18.468	-43.391
33.537	-38.053	4.620	13.000	31.956	-52.460
42.481	-40.205	4.620	13.000	31.736	-50.740
53.073	-52.690	5.008	12.000	53.081	-18.206
68.342	-4.190	4.281	14.000	33.996	-45.420
f-mean	T-mean	meanv	meanvu		
12.385	4.895	43.780	-36.321		

где max, min – максимальная и минимальная амплитуда дыхательного сигнала, T(s) – период, f – частота дыхания, svдох – скорость вдоха, svuidox – скорость выдоха, f-mean – среднее значение частот, T-mean – среднее значение периодов, meanv – среднее значение скоростей вдоха, meanvu – среднее значение скоростей выдоха.

По официальным данным дети совершают больше количество вдохов, чем взрослые, также как и женщины дышат быстрее, чем мужчины. Тем не менее, существуют средние показатели частоты дыхания, характерные для разных возрастных групп. Новорожденные в возрасте от 1 до 12 месяцев совершают 30-60 вдохов в минуту, дети 1-2 лет - 24-40 вдохов, дети дошкольного возраста (3-5 лет) - 22-34 вдоха, школьники (6-12 лет) - 18-30 вдохов. Для подростков от 13 до 17 лет нормой частоты дыхания считается 12-16 вдохов в минуту, а для взрослых - 12-18 вдохов. Наши результаты соответствуют норме для взрослого человека.

Проведенный эксперимент доказывает то, что метод измерения [3] простой, легко реализуемый, дает высокую точность результата измерения, полученные экспериментальные измерения является важным для медицинских исследований.

### СПИСОК ИСПОЛЬЗУЕМЫХ ИСТОЧНИКОВ

1. Настройка bluetooth-модулей HC-05/06 [Электронный ресурс], URL: [robotclass.ru/articles/bluetooth-hc-05-06/](http://robotclass.ru/articles/bluetooth-hc-05-06/)

2. Биотехнические, медицинские, экологические системы и робототехнические комплексы – Биомедсистемы 2017: В.Л. Нгуен, М.Б. Каплан. ст. 547

3. Биотехнические, медицинские, экологические системы и робототехнические комплексы – Биомедсистемы 2016: В.Л. Нгуен, М.Б. Каплан. ст. 356

**Р.И. НИКИТИН**

Рязанский государственный радиотехнический университет

## **УСКОРЕННЫЕ ИСПЫТАНИЯ**

*Рассматривается ускорение климатических и механических испытаний.*

Испытания на надежность – обязательный вид испытаний [1]. Общая их цель – установить, какой надежностью обладают изделия в условиях, предписываемых методикой испытаний.

Главная проблема испытаний современных изделий электронной техники – ограниченность испытаний по времени и объему, поскольку для определения соответствия изделий высоким требованиям надежности необходимо проведение длительных испытаний для больших объемов выборок. Для подавляющего большинства изделий длительность испытаний на надежность составляет несколько тысяч часов. Поэтому важным направлением современных исследований в области надежности является разработка методов ускорения испытаний. Наиболее универсальный вид ускоренных испытаний — испытания в форсированных режимах, когда за счет увеличения интенсивности нагрузки удается оценить надежность изделий за более короткое время.

Проведение форсированных ускоренных испытаний связано с необходимостью выбора критерия, с помощью которого можно было бы оценить характеристики надежности в нормальном режиме по характеристикам, полученным при повышенных нагрузках. В качестве такого критерия используют, например, коэффициент подобия

$$K = T_n / T_y \quad (1)$$

где  $T_n$ ,  $T_y$  - средние наработки до отказа (на отказ для восстанавливаемых устройств) в условиях нормальных и форсированных испытаний.

С увеличением коэффициента подобия  $K$  время испытаний уменьшается, однако пределы изменения этого коэффициента ограничены. Безоглядное повышение электрических нагрузок, нагрева, охлаждения и других воздействий с целью сокращения времени испытаний может исказить исходную модель отказов и сделать некорректным перенос результатов ускоренных испытаний на нормальные условия эксплуатации.

В качестве основных ускоряющих факторов при проведении испытаний РЭС в настоящее время применяют повышенные тепловые и механические нагрузки.

Для большинства механизмов отказов интегральных микросхем связь между долговечностью и рабочей температурой описывается уравнением Аррениуса [2]:

$$V_R = V_{R0} \exp\left(-\frac{E_A}{kT}\right) \quad (2)$$

где  $E_A$  – энергия активации;  $k$  – постоянная Больцмана;  $T$  – температуры по Кельвину.

Из ускоряющей температурной модели Аррениуса следует зависимость интенсивностей отказов изделий от температуры в ускоренных испытаниях и при эксплуатации; для экспоненциального распределения времени работы до отказа эта зависимость выражается в виде

$$\lambda_{И,Э} = \Lambda \exp\left(-\frac{E_A}{kT_{И,Э}}\right) \quad (3)$$

где  $\lambda_{И}$ ,  $\lambda_{Э}$  – интенсивности отказов изделий с полупроводниковыми элементами в условиях ускоренных испытаний и в условиях эксплуатации;  $T_{И}$ ,  $T_{Э}$  – температуры по Кельвину ускоренных испытаний и эксплуатации соответственно.

Согласно (3) отношение этих интенсивностей  $\lambda_{И}$ ,  $\lambda_{Э}$  имеет вид

$$\frac{\lambda_{И}}{\lambda_{Э}} = \exp\left(\frac{-E_A}{k} \left(\frac{1}{T_{И}} - \frac{1}{T_{Э}}\right)\right) \quad (4)$$

Располагая значением  $\lambda_{И}$ , из уравнения (4) при известных значениях  $T_{И}$ ,  $T_{Э}$ ,  $E_A$  и  $k$  (см. выше) можно определить интенсивность отказов  $\lambda_{Э}$  в условиях эксплуатации.

Этот наиболее простой случай, который предлагается смоделировать в среде *LabVIEW*, не всегда имеющий место в реальности. Даже при корректном принятии модели (2) и большой вероятности распределения наработки по экспоненциальному закону значение энергии активации  $E_A$  бывает не известно с достаточной точностью. В таком случае нахождение значения  $\lambda_{Э}$  требует проведения ускоренных испытаний при двух температурах:  $T_{И1}$ ,  $T_{И2}$  (упрощенно обозначим  $T_1$  и  $T_2$ ).

Тогда получается два уравнения:

$$\lambda_{И1} = \Lambda \exp\left(-\frac{E_A}{kT_1}\right), \quad \lambda_{И2} = \Lambda \exp\left(-\frac{E_A}{kT_2}\right) \quad (5)$$

из которых становится возможным найти неизвестные  $\Lambda$  и  $E_A$ . После этого модель (2) позволяет вычислить искомое значение интенсивности отказов при эксплуатации

$$\lambda_3 = \Lambda \exp\left(-\frac{E_A}{kT_3}\right) \quad (6)$$

Следующая часть данной работы посвящена моделированию ускоренных механических испытаний. Цель механических испытаний – проверка работоспособности изделий в условиях воздействия механических факторов: вибрации, ударных и линейных нагрузок, звукового давления. Надо сразу отметить, что методы проведения ускоренных механических испытаний и пересчета их результатов на реальные условия эксплуатации до сих пор разработаны недостаточно [3].

Одним из самых опасных и наиболее часто встречающихся на практике видов механических воздействий, влекущим наибольшее количество отказов РЭС, является вибрация. Наиболее распространенный её вид - стационарная случайная вибрация, описывается с помощью функции спектральной плотности  $S_{GB}(\omega)$ .

Спектральная плотность реакции механической системы  $S_{GA}(\omega)$  при воздействии на нее случайного воздействия  $S_{GB}(\omega)$  находится по формуле [4]

$$S_{XA}(\omega) = |l_{BA}(j\omega)| \cdot S_{GB}(\omega) \quad (7)$$

где  $|l_{BA}(j\omega)|$  - функция динамической податливости механической системы.

При изучении действия вибрации на объект, имеющий в реальности множество степеней свободы, во многих случаях допустимо пренебрежение всеми формами колебаний, за исключением одной преобладающей [4]. Тогда объект моделируется колебательной системой с одной степенью свободы, имеющей массу  $m$ , коэффициент упругости  $c$ , коэффициент вязкого трения  $b$ . Тогда выражение динамической податливости принимает вид

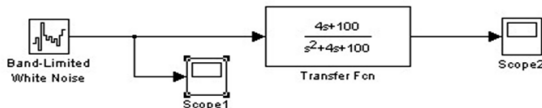
$$|l_{BA}(j\omega)| = m^{-1} \left[ (\omega_0^2 - \omega^2)^2 + 4\beta^2 \omega_0^2 \omega^2 \right]^{-1/2}, \quad \beta = \frac{b}{2m\omega_0} \quad (8)$$

При воздействии на такую систему случайной вибрации дисперсия напряжений в упругом элементе представляется выражением вида

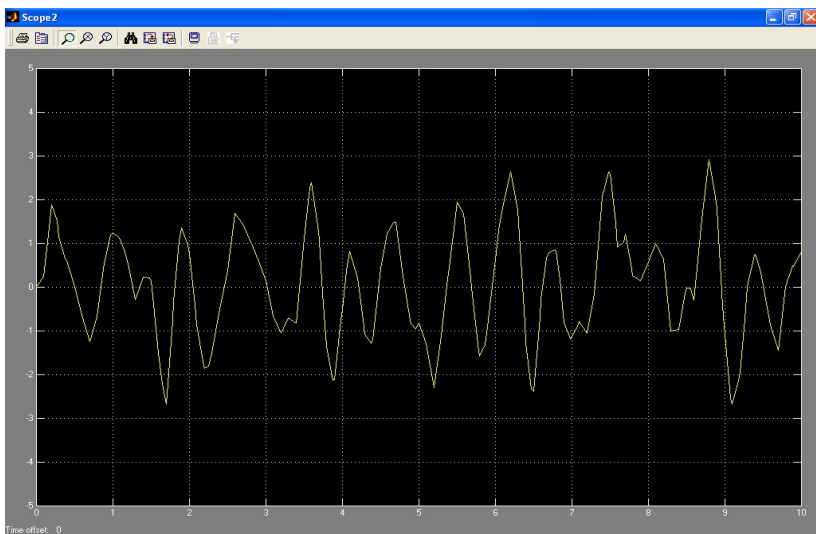
$$\sigma_A^2 = \frac{\pi \cdot c^2 S_{GA}(\omega) \omega_0}{4\beta m^2} \quad (9)$$

Для построения методики ускоренных механических испытаний практический интерес представляет замена действия случайной вибрации эквивалентной ей гармонической вибрацией. Корректность идеи такой замены основывается на фильтрующих свойствах колебательной системы, благодаря которым ее реакция на случайную вибрацию с

достаточно равномерным спектром будет содержать явно выраженную гармоническую составляющую. Эквивалентность подразумевает равенство максимальных реакций (виброизолированного блока). На рис. 1 и 2 представлены результаты моделирования действия широкополосной случайной вибрации на систему виброзащиты в среде Matlab.



**Рис. 1.** Схема моделирования действия широкополосной случайной вибрации на систему виброзащиты



**Рис. 2.** Сигнала ускорения виброизолированного блока; отчетливо наблюдается гармоническая составляющая с периодом  $T_0 \approx 0,63$  с в соответствии с заданной собственной частотой испытываемой системы ( $\omega_0 = 2 \cdot \pi / T_0 = 10$  1/с)

Амплитуда напряжений  $\sigma_N$  в упругом элементе, возникающих в результате действия эквивалентной гармонической вибрации, определяется по формуле

$$\sigma_N = (\sqrt{2\pi}/2)^{1/a} \sqrt{\frac{\alpha}{e} \sigma_A^2} \quad (10)$$

Располагая величиной  $\sigma_N$ , можно по кривой усталости найти число циклов  $N$ , соответствующее разрушению образца.

Процесс проведения ускоренных вибрационных испытаний представляется следующим.

Сначала для заданного уровня вибраций, действующей на объект в реальных условиях, устанавливается величина дисперсии напряжений в конструкционном элементе - см. формулу (9). Эта величина  $\frac{\sigma_A^2}{\sigma_A}$  пересчитывается по формуле (10) в напряжение  $\sigma_{N_3}$ . Полученному значению  $\sigma_{N_3}$  на кривой Веллера соответствует числу циклов  $N_3$ . Наша задача заключается в том, чтобы уменьшить это число циклов, увеличив гармоническую нагрузку на элемент. Напряжение при такой увеличенной нагрузке обозначим этому напряжению будет соответствовать число циклов  $N_u$ .

Время выдержки испытуемого изделия под действием гармонической вибрации связано с её частотой, заданной числом колебаний в секунду (Гц) и числом циклов простым соотношением  $t = N/f$ , с.

Сказанное выше можно проиллюстрировать следующим численным примером. Рассмотрим модель конструкционного элемента, представляющую собой упругую консоль с закрепленной на её конце массой при следующих численных значениях параметров:  $m = 20\text{ г}$ ;  $l = 10\text{ мм}$ ; ширина консоли  $b = 5\text{ мм}$ ; толщина  $h = 0,5\text{ мм}$ .

Параметры действующего виброускорения: амплитуда  $A = 0,1\text{ мм}$ ; частота  $f = 100\text{ Гц}$ .

Амплитуда виброускорения определяется как

$$W_m = A \cdot 4\pi^2 \cdot f^2 \quad (11)$$

где  $\omega = 2\pi f$  есть круговая частота. Обусловленная этим ускорением сила инерции действует с амплитудой

$$F_{\max} = m \cdot W_m \quad (12)$$

Максимальное напряжение изгиба возникает в месте заделки консоли; это напряжение подсчитывается по формуле

$$\sigma_{\max} = 6F_{\max}l/(bh^2) \quad (13)$$

При заданных численных значениях величин, входящих в формулы (1) - (3) получаем.

$$W_m = 40\text{ м/с}^2; \quad F_{\max} = 2\text{ Н.}$$

$$\sigma_{\max} = 6 \cdot 2\text{ Н} \cdot 10\text{ мм} / (5\text{ мм} \cdot 0,25\text{ мм}^2) = 100\text{ Н/мм}^2 \approx 40\text{ МПа.}$$

Этому значению изгибного напряжения на кривой Вёллера соответствует число циклов  $N_3$ , до разрушения  $5 \cdot 10^7$ . Нетрудно подсчитать время до разрушения: оно составляет примерно 150 часов. Увеличением испытательного виброускорения на той же частоте всего в 2 раза можно согласно кривой Веллера для данного материала уменьшить число циклов и соответственно время испытания в 500 раз.



На сегодняшний день существует технология, которые успешно комбинируют механические и климатические испытания, она называется HALT (Highly Accelerated Life Test — ускоренные ресурсные испытания) & HASS (Highly Accelerated Stress Screen — ускоренный выборочный контроль) использует комбинацию нагрузок для того, чтобы как можно раньше выявить недостатки, допущенные на этапах проектирования производства (часто на уровне изготовления плат). Применение данных технологий позволяет повысить надежность конечных изделий.

### **СПИСОК ИСПОЛЬЗУЕМЫХ ИСТОЧНИКОВ**

1. ГОСТ Р 27.607-2013. Надежность в технике. Управление надежностью. Условия проведения испытаний на безотказность и статистические критерии и методы оценки их результатов.
2. Ускоренные вибрационные испытания [Электронный ресурс], URL: [http://tech-e.ru/2010\\_3\\_50.php](http://tech-e.ru/2010_3_50.php).
3. Шишмарев В.П. Надежность технических систем. – М.: Изд. «Академия», 2010.
4. Вибрации в технике: Справочник. В 6-ти т. Т. 6, Защита от вибраций и ударов / Под ред. К.В.Фролова, 1981. 456 с.

**И.Г. ОСИН**

Рязанский государственный радиотехнический университет

## **ОРГАНИЗАЦИЯ УЧЕТА ТРАНЗАКЦИЙ В РАСПРЕДЕЛЕННОЙ БАЗЕ ДАННЫХ**

*Рассматриваются возможности технологии Blockchain в распределенных системах обмена и обработки больших объемов конфиденциальных данных.*

Обмен информацией между пользователями распределенной вычислительной системы производится транзакциями – логически осмысленными операциями, которые имеют смысл и могут быть совершены только полностью.

Транзакция – это группа логически определенных последовательных операций по работе с данными, обрабатываемая или отменяемая целиком. Транзакция считается подтвержденной, когда проверены ее формат и подписи, и когда сама транзакция объединена в группу с несколькими другими и записана в специальную структуру – блок.

Блок состоит из заголовка и списка транзакций. Заголовок блока включает в себя свой хеш, хеш предыдущего блока, хеши транзакций и дополнительную служебную информацию.

Технология обработки цепочек блоков получила название Blockchain.

Blockchain — это защищенный от несанкционированного доступа цифровой реестр общего пользования, который ведет учет транзакций в публичной или закрытой одноранговой сети. Распределенный между всеми узлами сети реестр непрерывно записывает историю операций с активами между одноранговыми узлами сети в виде блоков информации.

Все утвержденные блоки транзакций соединяются в цепочку — с начального блока до последнего добавленного. Таким образом, Blockchain выступает в качестве единого источника достоверных данных, а участники Blockchain-цепи видят только те транзакции, которые относятся именно к ним.

Узлы Blockchain-сети используют специальный протокол консенсуса для согласования содержимого реестра, а также криптографические алгоритмы хеширования и электронно-цифровые подписи для обеспечения целостности транзакции и передачи ее параметров.

Механизм консенсуса гарантирует, что распределенные реестры являются точными копиями, что снижает риск появления мошеннических транзакций, поскольку постороннее вмешательство может возникнуть во многих местах одновременно. Криптографические алго-

ритмы хеширования, такие как алгоритм вычислений SHA256, гарантируют, что любое изменение входных данных транзакции, даже самое незначительное, приведет к появлению другого значения хеша в результатах расчетов, что указывает на вероятность компрометации входных данных транзакции. Электронно-цифровые подписи гарантируют, что транзакции осуществляются легитимными отправителями (подписаны закрытыми ключами), а не злоумышленниками.

Децентрализованная одноранговая Blockchain-сеть лишает отдельных участников или групп участников возможности контролировать базовую инфраструктуру или дестабилизировать всю систему. Все участники сети равны и подключаются к ней по одним и тем же протоколам. Участниками могут быть физические лица, государственные структуры, организации или объединения всех перечисленных типов участников.

По сути система записывает хронологический порядок проведения транзакций со всеми узлами сети, признавшими действительность транзакций посредством выбранной модели консенсуса. Результатом являются не подлежащие отмене транзакции, согласованные всеми участниками сети децентрализованно.

Преимущества технологии Blockchain: хранение у всех членов системы одновременно, благодаря чему становится невозможным ее взлом и похищение; прозрачность данных совершенных транзакций, благодаря чему любой пользователь может отследить информацию о переводе средств и удостовериться в том, что платеж в системе действительно был отправлен; невозвратность всех транзакций: плательщик не может отозвать или заморозить отправленный денежный перевод «задним числом», обманув таким образом получателя; передача кодов денежных единиц и других виртуальных ценностей от плательщика к получателю напрямую, без участия посредников и без оплаты комиссии.

Недостатки технологии Blockchain: низкая скорость переводов в случае перегруженности базы. Функционирование Blockchain требует очень больших мощностей, поэтому создание и проверка новых блоков может занять значительное время. К примеру, в системе биткоин одна транзакция может совершаться до 4-5 часов и более, неопределенный нормативный статус. Можно сказать, что Blockchain и криптовалюты находятся за пределами законодательного регулирования большинства стран. Использовать криптовалюты при расчетах в сети возможно на свой страх и риск, потому как они не регулируются законом. изначальная большая стоимость внедрения технологии. Ис-

пользуя Blockchain при передаче электронных ценностей, можно существенно сэкономить на оплате услуг посредников и гарантов.

Однако само создание системы и внедрение ее в какую-либо сферу является очень затратным.

#### **СПИСОК ИСПОЛЬЗУЕМЫХ ИСТОЧНИКОВ**

1. Свон М. Блокчейн: Схема новой экономики. – Олимп-Бизнес. 2016 – 240с.
2. Генкин А., Михеев А. Блокчейн. Как это работает и что ждет нас завтра. – М: Альпина Пабlishер, 2017 – 592с.

**Е.Е. ПЕРЕВЕЗЕНЦЕВ, В.В. РОМАШКОВА**

Рязанский государственный радиотехнический университет

## **АСПЕКТЫ РЕАЛИЗАЦИИ БИБЛИОТЕКИ ГРАДИЕНТНОГО БУСТИНГА CATBOOST**

*Рассматриваются механизмы работы библиотеки градиентного бустинга CatBoost, анализируются ее возможности по работе с категориальными данными.*

CatBoost – это библиотека градиентного бустинга от компании Яндекс, которая использует данную библиотеку в своей поисковой системе, а также в других проектах [1]. Ранее в Яндекс разрабатывали и использовали библиотеку Матрикснет. Позднее в Яндекс создали CatBoost и выложили в сеть с исходным кодом и серией экспериментов, которые показывают превосходство CatBoost над H2O, XGBoost, LightGBM от Microsoft, причем как после подбора гиперпараметров, так и с самыми обычными параметрами.

Другие подобные библиотеки работают на искусственных нейронных сетях (ИНС). Главным преимуществом CatBoost является более точный результат при относительно малом количестве данных, в то время как ИНС требуют для работы огромные массивы однородной информации. Одна из существенных проблем многих библиотек – необходимость предоставления данных в числовом виде. Существуют определенные проблемы, связанные с работой с нечисловыми данными, например, с такими как «жанр музыки», «порода собак», «вид двигателя», которые понятны человеку, но не понятны компьютеру.

В CatBoost предусмотрена возможность оценить так называемые категориальные признаки, то есть признаки, которые принимают одно значение из конечного набора значений. Благодаря данной способности в CatBoost не нужно тратить время на перевод этих данных в числовые. В CatBoost для работы с числами используются регрессии типа MSE, MAE, QuantileError, Logquantile, Poisson, а для работы с категориальными признаками – классификации Logloss и CrossEntropy, мультиклассификация SoftMax и метрики AUC, Accuracy, Precision, Recall, F1 (1 vs all).

Градиентный бустинг может быть представлен в виде деревьев [2]. При работе с числовыми данными реализуется выборка по условию (рис. 1 и 2). С категориальными признаками отбор идет более широко, исходя из всех возможных перечислений (рис. 3 и 4). При работе с нечисловыми данными в CatBoost используют несколько способов обработки. Рассмотрим два из них.

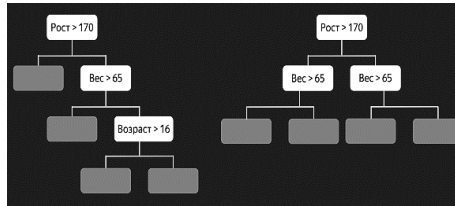


Рис. 1. Бинарные деревья

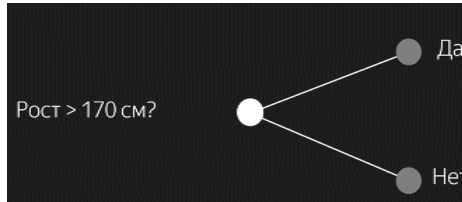


Рис. 2. Числовые данные



Рис. 3. Категориальные данные

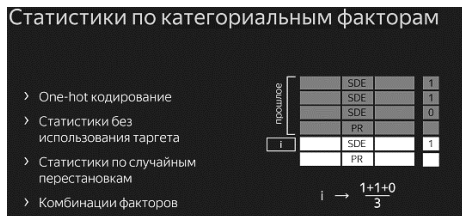


Рис. 4. Статистика по категориальным факторам

Наиболее популярный и понятный способ – one-hot incoding. Пусть есть узел – профессия (рис. 3), из которого выходит пять вариантов – профессий. При этом вместо одного узла появляется пять бинарных факторов, которые отвечают на вопрос: «Является ли человек, например, программистом?». Ответ может быть «Да» или «Нет». Такой вопрос можно задать к любой профессии на рис. 3. Второй способ основан на статистике без использования таргета (рис. 4). Этот способ рассчитывает какую-либо статистику по данной категории, например, количество вхождений (количество программистов, дизайнеров и т.д.).

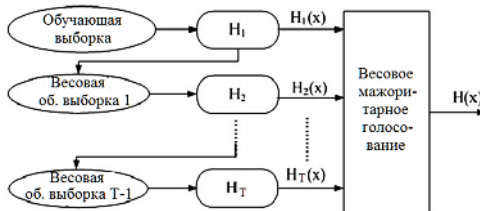


Рис. 5. Общая схема бустинга

В CatBoost обучение полностью работает за счет базового принципа градиентного бустинга (рис.5). Бустинг – итерационный алгоритм, реализующий «сильный» классификатор, который позволяет добиться произвольно малой ошибки обучения (на обучающей выборке) на основе композиции «слабых» классификаторов, каждый из которых лучше, чем просто угадывание, т.е. вероятность правильной классификации больше 0.5.

Особенности бустинга: использование весовой версии одних и тех же обучающих данных вместо случайного выбора их подмножества; слабые классификаторы образуются последовательно, различаясь только весами обучающих данных, которые зависят от точности предыдущих классификаторов; большие веса назначаются “плохим” примерам, что позволяет на каждой итерации сосредоточиться на примерах, неправильно классифицированных; базовые классификаторы должны быть слабыми, из сильных хорошую композицию не построить (“бритва Оккама”).

Укажем причины, почему базовые классификаторы должны быть слабыми: сильный классификатор, давая нулевую ошибку на обучающих данных, не адаптируется и композиция будет состоять из одного классификатора; один, даже сильный, классификатор может дать “плохое” предсказание на данных тестирования, давая “хорошие” результаты на не обучающих данных.

CatBoost обеспечивает возможность работы в Python, который предоставляет пользователю широкие возможности по работе с библиотеками машинного обучения.

## СПИСОК ИСПОЛЬЗУЕМЫХ ИСТОЧНИКОВ

1. Машинное обучение [Электронный ресурс], URL: [https://compscicenter.ru/media/slides/machine\\_learning\\_2\\_2018\\_spring/2018\\_02\\_26\\_machine\\_learning\\_2\\_2018\\_spring.pdf](https://compscicenter.ru/media/slides/machine_learning_2_2018_spring/2018_02_26_machine_learning_2_2018_spring.pdf)

2. Паклин Н.Б., Орешков В.И. Бизнес-аналитика: от данных к знаниям: учебное пособие. 2-е изд. СПб: Питер, 2013. С. 428-472.

**А.С. ПЕХНОВ**

Рязанский государственный радиотехнический университет

**АНАЛИЗ ПРИМЕНЕНИЯ БИБЛИОТЕКИ RHRMORPHY**

*Оценивается целесообразность классификации текстов с использованием библиотеки rhrMorphy*

Классификация научно-технических текстов на основе ключевых слов (терминов) к тем или иным секциям научно-технической конференции поднимает вопрос о поиске данных слов по научной статье с учётом морфологии. Под ключевыми словами будем понимать термины из различных научных областей. Термины сгруппированы в секции по принципу принадлежности соответствующим наукам. Необходимо чтобы поиск ключевых слов заключался не в обыденном их совпадении в тексте, но в сопоставлении текста со словарём, в котором имеются все возможные словоформы терминов. Решением этой проблемы на языке RHR может стать rhrMorphy.

rhrMorphy – библиотека языка RHR, методы которой позволяют определить:

- все словоформы необходимого термина,
- корень термина,
- начальную форму термина,
- изменение формы слова в соответствии с заданными грамматическими характеристиками,
- изменение формы слова по заданному образцу.

А также осуществляют работу с окончаниями и префиксами. Позволяет получить грамматическую информацию для слова: часть речи, род, число, падеж и т.п.

Вышеперечисленные функции позволяют значительно повысить точность классифицирования обрабатываемой статьи, так как в её тексте будет обнаружены все словоформы из наборов указанных нами терминов. Как уже отмечалось выше, для каждого термина методами библиотеки rhrMorphy будет найден список всех его возможных словоформ. На первый взгляд это является излишней задачей, поскольку у рассматриваемой библиотеки существует метод получения корня слова, и достаточно будет подсчитать число вхождения корня слова в текст статьи. Однако в данном случае точность классифицирования текстов упадёт, по причине того, что однокоренные слова не всегда имеют одно семантическое значение. Также этот корень нашего термина для другого слова может оказаться и не корнем вовсе.



По своей сути `phpMorphy` является простым, удобным и эффективным инструментом.

К преимуществам использования данной библиотеки можно отнести:

- открытость, библиотека находится в открытом доступе,
- простота подключения и использования, поскольку `phpMorphy` является библиотекой языка PHP,
- поддержка русского, английского и немецкого языков.

По своей сути `phpMorphy` является готовым модулем, который не требует доработки. То есть его методы могут быть использованы без каких-либо затруднений для повышения эффективности классифицировании текстов.

Для своей работы `phpMorphy` использует словари в двух формах:

- исходной, `xml` файл, который хранит в себе основы слов, правила изменения, грамматическая информация,
- бинарной, компилируется из исходной формы, преимуществом бинарной формы является совместимость с множеством платформ

Из недостатков данной библиотеки наиболее критичными являются:

- возможность использовать одновременно только 1 словарь,
- все буквенные символы должны быть в верхнем регистре.

Однако данные недостатки являются незначительными применительно к решаемой задаче классифицирования научно-технических статей, поскольку для её решения требуется только русский язык, а перевод всех буквенных символов текста в верхний регистр на языке PHP производится достаточно просто.

Аналогами данной библиотеки с поддержкой русского языка, применительно к классифицированию статей научно-технической конференции, являются `ruMorphy2`, `MyStem`, скрипт проекта «Вавилонская башня», `Rumog`. Сравним каждый из аналогов с `phpMorphy` с учётом особенностей решаемой задачи.

`ruMorphy2` – наиболее новый продукт по сравнению с `phpMorphy`, от тех же разработчиков. Написан на языке Python, работает быстрее, однако, всё ещё дорабатывается и не имеет такого большого функционала

`MyStem` – морфологический анализатор от разработчиков компании Яндекс. Является готовой программой с консольным интерфейсом, имеется в открытом доступе, однако не имеет открытого исходного кода. Не имеет функции получения всех словоформ слова по имеющемуся термину, что делает его неприменимым к текущей задаче.

Скрипт проекта «Вавилонская башня» аналогично MyStem не имеет на столько подходящих методов как phpMorphy.

Скрипт Rumog написан на языке PERL, последнее обновление получил в 2002 году, а также не является бесплатным для использования.

Таким образом, на данный момент из этих пяти продуктов наиболее подходящим для классифицирования научно-технических статей на русском языке является phpMorphy, поскольку ruMorphy2 требует доработки, а MyStem и «Вавилонская башня» не в полной мере являются аналогами этих библиотек, а Rumog является весьма устаревшим по сравнению со всеми вышеперечисленными аналогами.

### **СПИСОК ИСПОЛЬЗУЕМЫХ ИСТОЧНИКОВ:**

1. Поиск с учетом русской морфологии [Электронный ресурс] // URL: [https://valera.ws/2007.09.05~morpho\\_search\\_in\\_mysql/](https://valera.ws/2007.09.05~morpho_search_in_mysql/)
2. Документация по phpMorphy [Электронный ресурс] // URL: <http://phpmorphy.sourceforge.net/dokuwiki/manual>
3. Морфологическая библиотека PHPMorphy [Электронный ресурс] // URL: <https://programming086.blogspot.ru/2010/02/phpmorphy.html>
4. Реализация морфологического поиска на Kohana (библиотека phpMorphy) [Электронный ресурс] // URL: <https://habrahabr.ru/post/165715/>

П.С. ПИСАКА

Рязанский государственный радиотехнический университет

**СРАВНИТЕЛЬНЫЙ АНАЛИЗ ОПТИМАЛЬНЫХ  
КОГЕРЕНТНЫХ И НЕКОГЕРЕНТНЫХ АЛГОРИТМОВ  
ОБРАБОТКИ ИНФОРМАЦИИ БОРТОВОЙ СИСТЕМЫ  
ТЕЛЕМЕТРИЧЕСКИХ ИЗМЕРЕНИЙ ПРИ ТЕРРИТОРИАЛЬНО-  
РАСПРЕДЕЛЕННЫХ ПРИЕМНЫХ СТАНЦИЯХ**

*Рассмотрены особенности разработки оптимальных алгоритмов обработки телеметрической информации изделий РКТ при территориально-распределенных приемных станциях. Проведено сравнение оптимальных алгоритмов обработки когерентных и некогерентных сигналов ТМИ. На основе анализа натурной ТМИ получены значения показателя вероятности правильного обнаружения для разработанных алгоритмов при заданной вероятности ложной тревоги.*

В процессе проведения летных испытаний изделий РКТ, бортовая система телеметрических измерений осуществляет формирование и передачу телеметрической информации (ТМИ), содержащей сведения о параметрах состояния бортовых систем и агрегатов. Для оценки состояния объекта необходимо осуществление обработки и анализа ТМИ. С целью выполнения указанных работ, на протяжении всей трассы полета изделия РКТ установлены территориально-распределенные приемные станции, осуществляющие прием и регистрацию переданных сигналов ТМИ. Поскольку антенны приемных станций имеют ограниченную зону радиовидимости, а кроме того, в канале передачи присутствуют различные мешающие факторы, информация, полученная от одной приемной станции, не характеризует, с высокой степенью достоверности, состояние изделия РКТ на всей трассе полета.

Для устранения указанной особенности необходима разработка алгоритмов обработки сигналов ТМИ одновременно от нескольких территориально-распределенных приемных станций с целью улучшения качества. Применение подобных алгоритмов позволит сформировать единый сигнал ТМИ на всей трассе полета изделия РКТ, а кроме того повысит качество информации на отдельных участках трассы полета изделия РКТ. В качестве показателей качества выбраны показатели вероятности правильного обнаружения  $P_0$  сигнала ТМИ и вероятности ложной тревоги  $P_{лт}$ .

Алгоритмы обработки сигналов ТМИ от нескольких территориально-распределенных приемных станций должны быть построены на основе модели сигнала системы передачи ТМИ. В работе [1] на основе

анализа натурной ТМИ, полученной при проведении летных испытаний изделий РКТ разработана статистическая модель сигнала системы передачи ТМИ в виде двухкомпонентного гамма-распределения для областей независимых и дружных флуктуаций амплитуд сигнала соответственно. Таким образом, модель учитывает наличие корреляционных связей между сигналами ТМИ, принятых на разных территориально-распределенных приемных станциях. В работе [2] показано, что при выбранном уровне значимости  $\alpha = 0,01$  теоретическая модель сигнала системы передачи ТМИ соответствует экспериментальным данным по критериям согласия А.В. Колмогорова.

Алгоритмы обработки сигналов ТМИ могут быть разделены на две группы – оптимальные и квазиоптимальные алгоритмы соответственно. Оптимальные алгоритмы основаны на обработке «чистого» сигнала ТМИ, т.е. сигнала без изменения структуры. К подобным алгоритмам относятся алгоритмы обработки когерентных и некогерентных сигналов ТМИ соответственно. Квазиоптимальные алгоритмы предполагают осуществление над сигналами ТМИ предварительных операций фильтрации, детектирования и т.п.

В процессе полета изделия РКТ, радиопередатчик производит передачу частотно-манипулированного группового телеметрического сигнала  $S(t)$ , который может быть представлен в комплексной форме как

$$S(t) = as_0(t - t_z) \exp[j(\omega_0 + \Omega)(t - t_z) - j\varphi], \quad (1)$$

где  $a$  – эффективное значение огибающей сигнала ТМИ,  $s_0(t - t_z)$  – комплексная огибающая,  $\varphi$  – начальная фаза сигнала,  $t_z$  – запаздывание сигнала,  $\omega_0$  – частота сигнала,  $\Omega = e(t)$  – модулирующая частота.

Таким образом, телеметрический сигнал  $S_i(t)$  в зоне радиовидимости  $i$ -ой,  $i = 1, \dots, m$  приемной станции, под воздействием шумов и помех принимает вид

$$X_i(t) = \sum_{i_1} X_{i i_1}(t) = \sum_{i_1} (S_{i i_1}(t) + \Pi_{i_1}(t) + \text{Ш}_i(t)), \quad (2)$$

где  $i$  – номер приемной станции,  $i = 1, \dots, m$ ,  $X_i(t)$  – результирующий сигнал на входе  $i$ -ой,  $i = 1, \dots, m$  приемной станции,  $i_1$  – число лучей сигналов в зоне радиовидимости  $i$ -ой,  $i = 1, \dots, m$  приемной станции,  $X_{i i_1}(t)$  – сигнал ТМИ в  $i_1$ -ом луче в зоне радиовидимости  $i$ -ой,  $i = 1, \dots, m$  приемной станции искаженный действием шумов и помех,

$S_{ii_1}(t)$  – исходный сигнал ТМИ, заданный выражением вида (1) в  $i_1$ -ом луче в зоне радиовидимости  $i$ -ой,  $i=1, \dots, m$  приемной станции,  $\Pi_i(t)$  – совокупность воздействующих помех на сигнал в зоне радиовидимости  $i$ -ой,  $i=1, \dots, m$  приемной станции,  $\Pi_i(t)$  – совокупность воздействующих шумов на сигнал в зоне радиовидимости  $i$ -ой,  $i=1, \dots, m$  приемной станции.

Можно показать, что для сигнала ТМИ, заданного выражениями (1,2) с учетом модели сигнала в виде двухкомпонентного гамма-распределения, логарифм отношения правдоподобия представляется как

$$\bar{\Lambda} = \int \int_{\mathbf{a} \varphi} f(\mathbf{a})f(\varphi) \exp[\mathbf{a} \operatorname{Re}[\bar{\mathbf{E}}(\varphi)\mathbf{G}]] - \frac{1}{2} \mathbf{a} \bar{\mathbf{E}}^*(\varphi) \mathbf{C} \bar{\mathbf{E}}(\varphi) \mathbf{a}] d\mathbf{a} d\varphi,$$

$$\mathbf{G} = \frac{1}{2\pi} \int_{-\infty}^{\infty} \bar{\Psi}^*(\omega) \mathbf{v}(\omega) [(1-\kappa) \mathbf{X}_{\text{нф}}(\omega, \varphi) + \kappa \mathbf{X}_{\text{дф}}(\omega, \varphi)] d\omega, \quad (3)$$

$$\mathbf{C} = \frac{1}{2\pi} \int_{-\infty}^{\infty} \bar{\Psi}^*(\omega) \mathbf{v}(\omega) \bar{\Psi}(\omega) d\omega,$$

где  $\bar{\Lambda}$  – безусловное отношение правдоподобия;  $f(\varphi)$  – функции плотности вероятности распределения элементов вектора  $\varphi$ ,  $\bar{\Psi}(\omega)$  – диагональная матрица спектров  $\Psi(\omega)$  сигналов, не содержащих случайные параметры;  $\bar{\mathbf{E}}(\varphi)$  – диагональная матрица случайных начальных фаз  $\varphi$ ;  $\mathbf{a}$  – вектор случайных эффективных значений огибающей сигнала  $\mathbf{a}$ ,  $f(\mathbf{a})$  – функции плотности вероятности распределения элементов вектора  $\mathbf{a}$ ,  $\mathbf{G}$  – вектор результатов согласованной фильтрации для результирующих сигналов  $X_i(t)$ ,  $i=1, \dots, m$  в зоне радиовидимости территориально-распределенных приемных станций,  $\mathbf{C}$  – вектор констант,  $\bar{\mathbf{E}}(\varphi)$  – диагональная матрица случайных начальных фаз  $\varphi$ ,  $\mathbf{X}_{\text{нф}}(\omega)$  – вектор преобразований Фурье элементов вектора результирующих сигналов для области независимых флуктуаций амплитуд сигналов,  $\mathbf{X}_{\text{дф}}(\omega)$  – вектор преобразований Фурье элементов вектора результирующих сигналов для области дружных флуктуаций амплитуд сигналов,  $\mathbf{v}(\omega)$  – преобразования Фурье корреляционной матрицы сигналов ТМИ.

Для когерентных сигналов ТМИ  $S_i(t)$ ,  $i=1, \dots, m$  в зоне радиовидимости приемных станций, фазы  $\varphi_i$  когерентны, а эффективные зна-

чения огибающей сигналов  $a_i$  связаны между собой. Можно показать, что оптимальный алгоритм  $L_0$  обработки когерентных сигналов ТМИ может быть представлен в виде

$$L_0 = \ln(\bar{L}_0) = \left| \mathbf{A}^T \mathbf{E} \mathbf{G} \right| = \left| (1 - \kappa) \sum_{i=1}^m a_i e^{-j\Delta\varphi_{i1}} G_i + \kappa \sum_{i=1}^m A_{i1} e^{-j\Delta\varphi_{i1}} G_i \right| =$$

$$= \left| (1 - \kappa) \sum_{i=1}^m a_i e^{-j\Delta\varphi_{i1}} \frac{1}{2} \int_{-\infty}^{\infty} \Psi_0(\omega - \omega_0 - \Omega_i) e^{j\omega t_i} v_i(\omega) X_{\text{нфи}}(\omega) d\omega + \right. \quad (4)$$

$$\left. + \kappa \sum_{i=1}^m A_{i1} e^{-j\Delta\varphi_{i1}} \frac{1}{2} \int_{-\infty}^{\infty} \Psi_0(\omega - \omega_0 - \Omega_i) e^{j\omega t_i} v_i(\omega) X_{\text{дфи}}(\omega) d\omega \right|.$$

Значение оптимального алгоритма обработки когерентных сигналов ТМИ  $L$ , заданного выражением вида (4) сравнивается со значением порога, определяемого заданной вероятностью ложной тревоги  $P_{\text{лт}}$ .

Для некогерентных сигналов ТМИ  $S_i(t)$ ,  $i = 1, \dots, m$  в зоне радиовидимости приемных станций, информация о фазах  $\varphi_i$  недоступна. Можно показать, что оптимальный алгоритм  $L_1$  обработки некогерентных сигналов ТМИ может быть представить в виде

$$L_1 = \ln(\bar{L}_1) = \dots = \sum_{i=1}^m \frac{1 + \kappa(A_{i1}^2 - 1)}{(1 + q_i^2)} G_i =$$

$$= (1 - \kappa) \sum_{i=1}^m \frac{1}{(1 + q_i^2)} \frac{1}{2} \int_{-\infty}^{\infty} \Psi_0(\omega - \omega_0 - \Omega_i) e^{j\omega t_i} v_i(\omega) X_{\text{нфи}}(\omega) d\omega + \quad (5)$$

$$+ \kappa \sum_{i=1}^m \frac{A_{i1}^2}{(1 + q_i^2)} \frac{1}{2} \int_{-\infty}^{\infty} \Psi_0(\omega - \omega_0 - \Omega_i) e^{j\omega t_i} v_i(\omega) X_{\text{дфи}}(\omega) d\omega.$$

Значение оптимального алгоритма обработки некогерентных сигналов ТМИ  $L$ , заданного выражением вида (5) сравнивается со значением порога, определяемого заданной вероятностью ложной тревоги  $P_{\text{лт}}$ .

Экспериментальные исследования, проведенные на основе анализа натурной ТМИ, полученной по результатам серии летных испытаний изделий РКТ показали, что применение разработанного оптимального алгоритма обработки когерентных сигналов ТМИ обеспечивает среднюю вероятность обнаружения  $\bar{P}_0 = 0,9826$  при вероятности ложной тревоги  $P_{\text{лт}} = 10^{-3}$ . В случае применения разработанного оп-

тимального алгоритма обработки некогерентных сигналов ТМИ обеспечивается средняя вероятность обнаружения  $\bar{P}_0 = 0,9605$  при вероятности ложной тревоги  $P_{лт} = 10^{-3}$ . Таким образом, применение реализуемого оптимального алгоритма обработки некогерентных сигналов ТМИ приводит к уменьшения средней вероятности обнаружения  $\bar{P}_0$  на  $\Delta = 0,0221$ .

### **СПИСОК ИСПОЛЬЗУЕМЫХ ИСТОЧНИКОВ**

1. Кириллов С. Н., Писака П. С. Имитационная модель канала связи системы передачи телеметрической информации космических ракет-носителей // 7-я международная научно-техническая конференция «К.Э. Циолковский – 160 лет со дня рождения. Космонавтика, Радиоэлектроника. Геоинформатика»: Тез. докл. Рязань: Рязанский государственный радиотехнический университет, 2017. С. 291-293.

2. Писака П. С. Обоснование модели канала передачи телеметрической информации // Материалы XXII Всероссийской научно-технической конференции студентов, молодых ученых и специалистов «Новые информационные технологии в научных исследованиях». Рязань: Рязанский государственный радиотехнический университет, 2017. С. 118-119.

**Г.А. ПОПОВ, С.В. СКВОРЦОВ**

Рязанский государственный радиотехнический университет

## **ТЕГИРОВАНИЕ ТЕКСТОВ НА ОСНОВЕ ТЕХНОЛОГИИ WORD2VEC**

*Рассматривается задача автоматизации процесса тегирования текстов на основе технологии word2vec при использовании списка заранее заданных тегов.*

В настоящее время информационным системам приходится работать с большим объемом текстовой информации, при этом предоставляя пользователям только ту часть информации, которая им интересна. В связи с этим эти блоки информации помечаются специальными метками, которые кратко описывают данные и называются тегами.

В информационных системах теги представляются как неструктурированные ключевые слова, описывающие части информации. Такие метаданные помогают находить нужную информацию в процессе их просмотра или через поисковый запрос.

В статье рассматривается задача автоматизации процесса присвоения тегов блокам информации на основе технологии Word2Vec [1], что позволит повысить эффективность работы с тегами.

Word2vec - это группа связанных моделей, которые используются для создания векторов слов. Такие модели представляют собой мелкие двухслойные нейронные сети, которые обучаются реконструировать языковые контексты слов. Word2vec принимает в качестве своего ввода большой корпус текста и создает векторное пространство, обычно в несколько сотен измерений, причем каждому уникальному слову в корпусе присваивается соответствующий вектор в пространстве. Для слов, которые имеют общие контексты в корпусе, соответствующие векторы расположены в непосредственной близости друг от друга в полученном векторном пространстве [1].

Word2vec является практически единственным алгоритмом глубокого обучения (deep learning), который сравнительно легко можно запустить на обычном персональном компьютере (ПК) и который строит распределенное представление слов за приемлемое время [2].

Данный подход был применен для автоматизированного тегирования большого набора статей по IT-тематике с заданным набором тегов, включающем, например, такие как «Java», «QA», «.Net», «Agile», «Database», «Blockchain» и т.п.

Первый стадией для создания программы тегирования является построение модели. Данный этап является основополагающим, поскольку от его успеха сильно зависит точность присвоения тега тексту.



Для этого может быть использовано несколько алгоритмов, в основе которых лежат известные подходы CBOW и Skip-gram [3].

CBOW и Skip-gram — это нейросетевые архитектуры, которые описывают каким образом нейросеть обучается на контрольных данных и запоминает представления слов. В основе указанных архитектур лежат разные принципы. Для CBOW принцип работы заключается в предсказывании слова при данном контексте, а для skip-gram наоборот — предсказывается контекст при данном слове.

Continuous Bag of Words (CBOW) – обычная модель «мешка слов» с учётом четырёх ближайших соседей термина (два предыдущих и два последующих слова) без учёта порядка следования, k-skip-n-gram — это последовательность длиной n, где элементы находятся на расстоянии не более, чем k друг от друга.

Качество векторов слов значительно увеличивается с ростом объемов контрольных данных, используемых при обучении нейросети. Для исследовательских целей можно рассмотреть возможность использования наборов данных, доступных в режиме онлайн:

- первый миллиард символов из Wikipedia;
- сайт WMT11, содержащий текстовые данные для нескольких языков (перед началом обучения модели необходимо удалить дублирующиеся предложения);
- набор данных из «One Billion Word Language Modeling Benchmark», который включает уже предварительно обработанный текст объемом почти миллиард слов;
- дампы базы данных stack overflow.

Результирующий векторный файл слов может использоваться во многих приложениях для обработки естественного языка и машинного обучения.

Простым способом исследования полученных представлений является поиск ближайших слов для заданного пользователем слова. Для этого служит дистанционный инструмент [4]. Например, если ввести слово «france», то будут отображаться самые похожие слова и их расстояния до «france», которые должны выглядеть примерно так: spain (0.678515) belgium (0.665923) netherlands (0.652428) italy (0.633130) switzerland (0.622323).

Недавно было показано, что векторы слов учитывают многие другие языковые закономерности. В частности, пусть  $\text{vec}(\text{Париж})$  обозначает вектор слова Париж, тогда  $\text{vec}(\text{Париж}) - \text{vec}(\text{Франция}) + \text{vec}(\text{Италия})$  приводит к вектору, который очень близок к  $\text{vec}(\text{Рим})$ , а  $\text{vec}(\text{король}) - \text{vec}(\text{человек}) + \text{vec}(\text{женщина})$  близок к  $\text{vec}(\text{королева})$ .

Чтобы наблюдать сильные закономерности в векторном пространстве слов, необходимо подготовить модели на большом наборе данных с достаточной векторной размерностью. Используя инструмент `word2vec`, можно готовить модели на огромных наборах данных (до сотен миллиардов слов).

При обучении модели следует учитывать, что на итоговый результат сильно влияет тематика входных текстов. Так для модели, обученной на данных из Wikipedia, ближайшим вектором к `vec(Java)` является `vec(Coffee)`, а для модели, обученной на `stack overflow - vec(language)`. Для решаемой задачи идеально подходит модель, построенная на дампе данных `stack overflow` (60 гигабайт текстовых данных).

Вторым этапом создания программы тегирования является реализация следующего алгоритма.

1. Удаление самых часто повторяющихся слов языка из текста.
2. Определение квадратичного расстояния от выбранного тега до каждого слова в тексте с помощью `word2vec`.
3. Удаление слов, расстояние которых от текущего тега меньше чем 0.01.
4. Определение среднего расстояния от оставшихся слов до тега.
5. Если среднее расстояние больше чем 0.7, то присвоить данный тег данному тексту.

Разработанная программа успешно справляется с поставленной задачей и в 96% правильно определяет принадлежность тегов заданному тексту. Ошибочные результаты наблюдаются только для 1% случаев.

### СПИСОК ИСПОЛЬЗУЕМЫХ ИСТОЧНИКОВ

1. Google code [электронный ресурс] // Word2vec introduction // URL: <https://code.google.com/archive/p/word2vec/>
2. Google drive [электронный ресурс] // Word2vec model // URL: <https://drive.google.com/file/d/0B7XkCwpI5KDYNINUTTISS21pQmM/edit>
3. Gensim site [электронный ресурс] // Deep learning with word2vec // URL: <https://radimrehurek.com/gensim/models/word2vec.html>
4. Tensor Flow [электронный ресурс] // Vector Representations of Words // URL: <https://www.tensorflow.org/tutorials/word2vec>

**С.В. ПОПОВ**

Рязанский государственный радиотехнический университет

**ПРИМЕНЕНИЕ ВИДЕОСЕНСОРНЫХ СИСТЕМ В БПЛА**

*В докладе рассматривается принцип работы видеосенсорных систем, применение данных систем в беспилотных летательных аппаратах (БПЛА), а также в самоорганизующихся формациях БПЛА.*

Сегодня БПЛА являются распространённым средством решения большого спектра задач. Так, применяясь в сельскохозяйственной промышленности, они производят оценку состояния полей, удобряют их и дезинфицируют, при этом, человек является лишь контролирующим звеном в этой автоматизированной системе.

Для полноценного функционирования БПЛА используют множество различных сенсоров, позволяющих беспилотному аппарату стабилизировать полет и оценивать обстановку с высоким уровнем достоверности. К ним можно отнести:

- Сенсоры стабилизации;
- Сенсоры параметров окружающей среды;
- Сенсоры приближения;
- Сенсоры обстановки;
- Беспроводные передатчики;
- МЭМС-сенсоры (микроэлектромеханические системы).

Несмотря на высокий уровень самостоятельности БПЛА, чаще всего говорят о имитации искусственного интеллекта. Поэтому, всегда должен присутствовать оператор-контролер максимально осведомленный об основных параметрах и состоянии беспилотного аппарата, при выполнении им определенной задачи. Одним из распространенных способов контроля является, использование оптических сенсоров, которые обычно относят к сенсорам обстановки.

БПЛА не ограничиваются использованием в коммерческих целях. Сегодня вооруженные силы большинства стран уже приняли в свои ряды беспилотные аппараты, для выполнения различного рода задач, от разведки труднодоступной местности, до применения их в качестве носителей специального вооружения.

Чаще всего БПЛА требует постоянного, или временного, соединения с управляющим центром. А это говорит о необходимости наличия беспроводной связи. В случае разрыва соединения (зашумление, повреждение аппаратуры связи), беспилотный аппарат будет не способен выполнить поставленную перед ним задачу. Именно на это рассчитаны средства для борьбы с БПЛА.

Основной информацией передаваемой беспилотному аппарату является координаты собственного местоположения, координаты цели (иногда движущейся), наличие и координаты наземных и воздушных средств подавления.

Видеосенсор представляет собой пассивный оптический сенсор – чувствительный элемент, который принимает определенный спектр света. Его выходными данными являются изображения. Обработка изображения должна осуществляться специальной системой, которая обеспечит разгрузку центрального управляющего звена летательного аппарата. Эффективнее всего данную задачу выполняет ПЛИС [1]. Видеосенсорная система – это набор специальных электромеханических устройств, обеспечивающих высокий уровень вариации основных параметров видеосенсора. К данным параметрам можно отнести: фокусное расстояние, размер диафрагмы, чувствительность, направление визирования и т.д.

При использовании видеосенсорных систем в БПЛА, возникают задачи локализации беспилотного аппарата, а так же поиска и выявления параметров других воздушных объектов. Первая задача разрешается с помощью сверки текущего визируемого пространства земной поверхности с заранее загруженными картами местности, совместно с другими сенсорными системами [2]. Вторая задача является более сложной и требует построения специальной системы.

Для описания движения визируемого объекта необходимы параметры относительного углового перемещения, а так же относительной дальности. Если определение углового перемещения, как и обнаружение воздушного объекта, можно осуществить с помощью традиционных алгоритмов выделения движущихся объектов [3], то определение дальности становится нетривиальной задачей. Самым простым способом является сравнение визируемого объекта с некоторыми известными (эталонными) имеющимися базе данных БПЛА, и на основе масштабного коэффициента рассчитать дальность. В случае невозможности распознавания объекта, или отсутствия его в базе эта задача решается с помощью варьирования некоторых параметров видеосенсорной системы. Так, увеличение фокусного расстояния позволяет уменьшить угол обзора и искусственно приблизить визируемый объект, а увеличение диафрагмы будет уменьшать дальность резкости. Таким образом, появление размытости на изображении, позволяет оценить расстояние до объекта.

Использование автономной лазерной оптической навигации в БПЛА позволяет с достаточной точностью ориентирования на местности, при полете на предельно низкой высоте [4]. Однако при полете на

бóльших высотах применяется ориентирование с помощью камеры. При использовании видеосенсора необходимы специальные устройства обработки визуальных данных. Такой способ определения местоположения в воздухе имеет следующие преимущества:

- Обеспечение скрытности при исследовании местности;
- Автономная локализация;
- Возможность воспроизведения процесса выполнения отработанной задачи, без установки дополнительного оборудования.

Для формации БПЛА, обеспечивается распределение задач исследования рельефа местности и поиска других летательных аппаратов, что заметно повысит общую эффективность формации. Еще один положительным эффектом является повышение точности определения дальности до нераспознанного объекта, при его визировании несколькими БПЛА.

Передача информации внутри формации возможно с применением беспроводных соединений малой дальности, таких как Bluetooth. В этом случае формация будет более устойчива к средствам подавления, в отличие от удаленного управления.

Таким образом, применяя видеосенсоры в БПЛА, можно повысить уровень автономности, а также уменьшить вероятность подавления, при этом не теряя в функциональности.

### **СПИСОК ИСПОЛЬЗУЕМЫХ ИСТОЧНИКОВ**

1. Держановский А.С., Соколов С.М. Обработка зрительных данных в системе технического зрения реального времени с использованием ПЛИС // Препринты ИПМ им. М.В.Келдыша. 2016. № 126. 16 с.
2. Фетисов В.С., Артемьев А.Е. Комплексирование радиотехнических и видеотехнических средств в системах локального позиционирования малых беспилотных летательных аппаратов. Альманах современной науки и образования. Тамбов, 2016, № 10, с. 91-95.
3. Козлов В.А., Потапов А.С. Анализ методов выделения движущихся объектов на видеопоследовательностях с шумами. Научно-технический вестник Санкт-Петербургского государственного университета информационных технологий, механики и оптики. Санкт-Петербург, 2011г. №3 (73), с.39-43.
4. Буянов И.А., Исламов В.К. Автономная система ориентирования беспилотного летательного аппарата: состав и схема функционирования в формате 3D. Молодой ученый. Казань, 2017, №50 (184) с. 24-30.

**Р. Р. САЛИМГАРЕЕВ, С. В. СКВОРЦОВ**

Рязанский государственный радиотехнический университет

## **АВТОМАТИЗИРОВАННОЕ ИЗМЕНЕНИЕ ПАРАМЕТРОВ ТОЧКИ ДОСТУПА ЧЕРЕЗ БЕСПРОВОДНОЕ ПОДКЛЮЧЕНИЕ**

*В статье рассматривается автоматизированное изменение параметров точки доступа через беспроводное подключение.*

В последние годы постоянно уменьшаются продажи компьютеров и одновременно увеличиваются продажи смартфонов. Таким образом, всё большее количество пользователей выходит в сеть интернет, используя беспроводные технологии передачи данных, в том числе и Wi-Fi. Для большинства сценариев использования такой способ выхода в интернет ничем не уступает проводному подключению, а зачастую и превосходит его. Однако такой способ подключения имеет свои недостатки, один из которых – меньшая безопасность.

Первым протоколом для защиты данных Wi-Fi сетей был WEP, однако в нём была найдена уязвимость, позволяющая довольно быстро подобрать пароль [1]. На смену WEP пришёл протокол WPA, а затем WPA2. Данные протоколы позволили существенно повысить безопасность Wi-Fi сетей, однако и они уязвимы к атакам, например, с помощью словарей. Короткие словарные пароли могут быть подобраны за относительно небольшое время, если были перехвачены пакеты, посылаемые во время установления соединения [2].

В январе 2018 года был анонсирован протокол WPA3, реализацию которого обещают раскрыть к концу года. Среди нововведений имеется: настройка сетей с клиентского устройства, персонализированное шифрование данных, а также новый способ шифрования, обеспечивающий защиту даже при использовании слабых паролей [3]. Однако для его внедрения потребуется время, поэтому на данный момент самым защищённым протоколом является WPA2.

Использование WPA2 само по себе не гарантирует защиты от взлома. Однако использование сложных паролей, а также их регулярная смена значительно повышают безопасность. Дополнительно повысить безопасность может смена SSID, так как он также участвует в шифровании. С другой стороны, эти меры требуют от пользователя большого количества действий, которые занимают довольно много времени, но могут быть легко автоматизированы. Среди этих действий: открытие интерфейса настроек Wi-Fi на маршрутизаторе, подбор и ввод сложных SSID и пароля, добавление точки доступа с новыми параметрами на пользовательском устройстве.

Рассмотрим один из способов автоматизации изменения параметров точки доступа для схемы подключения, показанной на рис. 1.



**Рис. 1.** Схема взаимодействия маршрутизатора и клиента

Пользователь запускает на смартфоне специальное приложение, которое загружает стандартный веб-интерфейс маршрутизатора. Приложение необходимо для вызова функций системы, недоступных из браузера, таких как сохранение имени и пароля точки доступа, принудительное подключение к определённой точке доступа. Через веб-интерфейс пользователь выбирает функцию изменения текущего пароля, после чего отправляется запрос на сервер управления маршрутизатором и одновременно с этим начинается ожидание подключения телефона к новой точке доступа. Если веб-интерфейс открыт не через приложение или автоматическое подключение к устройству не произошло, то показывается специальная страница (рис. 2).



**Рис. 2.** Страница неудачной попытки подключения к маршрутизатору

Автоматическое подключение обычно занимает не более 30 секунд и не требует от пользователя никаких дополнительных действий. Если подключения не произошло, то пользователь может самостоятельно добавить точку доступа и подключиться к ней, используя страницу, показанную на рисунке 2.

Помимо рассмотренного способа определения настроек точки доступа возможно изменение параметров без участия пользователя, например, по расписанию. Для этого требуется модификация серверной части маршрутизатора: после получения запроса от приложения сервер должен запускать специальный таймер. Если до истечения за-

данного времени смартфон подключился к новой точке доступа, то она сохраняется в качестве основной. Если время истекло, то маршрутизатор восстанавливает настройки текущей точки доступа.

К сожалению такой способ плохо подходит, если точка доступа используется несколькими клиентами. Однако современные маршрутизаторы позволяют одновременно создавать несколько точек доступа с разными SSID. Это позволяет выделить каждому клиенту свою беспроводную сеть, которую можно менять независимо от других, что затруднит осуществление некоторых видов атак.

Таким образом, при регулярном изменении пароля точки доступа можно сделать беспроводную сеть практически такой же безопасной, как и проводную. А возможная автоматизация данного процесса позволяет повысить безопасность, не требуя дополнительных усилий со стороны пользователя.

### **СПИСОК ИСПОЛЬЗУЕМЫХ ИСТОЧНИКОВ**

1. Breaking 104 bit WEP in less than 60 seconds / Erik Tews, Ralf-Philipp Weinmann, Andrei Pyshkin // WISA'07 Proceedings of the 8th international conference on Information security applications, 2007
2. Exposing WPA2 security protocol vulnerabilities / Achilleas Tsitroulis, Dimitris Lampoudis, Emmanuel Tsekleves // International Journal of Information and Computer Security 6(1):93-107, 2014
3. Wi-Fi Alliance® introduces security enhancements [Электронный ресурс]. – URL: <https://www.wi-fi.org/news-events/newsroom/wi-fi-alliance-introduces-security-enhancements> (дата обращения 13.03.2018)



**А.Н. САПРЫКИН, Е.Н. САПРЫКИНА**

Рязанский государственный радиотехнический университет

**СТРУКТУРА ХРОМОСОМЫ ГЕНЕТИЧЕСКОГО  
АЛГОРИТМА БАЛАНСИРОВКИ ПОТОКОВ  
ДАНЫХ В СЕТЯХ ДАТА-ЦЕНТРОВ**

*Рассматривается способ кодирования хромосом генетического алгоритма, использующегося для балансировки потоков данных в сетях дата-центров. Описывается принцип организации структуры особи, ее генотипа и фенотипа. Проводится анализ структуры хромосомы алгоритма для дальнейшей модификации.*

В работе [1] описывается возможность применения двухфазной многопутевой маршрутизации в сетях дата-центров. Для балансировки их сетевой нагрузки авторами был предложен модифицированный генетический алгоритм [2], распределяющий каждый входной поток на несколько частей, идущих от источника в приемник по различным маршрутам, что позволяет снизить загрузку сетевых каналов и организовать в них резервные полосы.

Структура хромосомы разработанного алгоритма зависит от входных параметров: числа узлов в сети и количества параллельных маршрутов, по которым будет распределен входной поток. При делении входного потока на две части, каждая из которых может проходить через один промежуточный узел, хромосома представляет собой следующий набор элементов, записанных последовательно друг за другом: балансировочный параметр, промежуточный узел для первой части потока и промежуточный узел для второй части потока.

Балансировочный параметр  $q$  показывает, какая доля потока будет проходить через первый промежуточный узел. Остальная часть потока определяется автоматически, как  $100 - q$  и проходит через второй промежуточный узел. На кодирование балансировочного параметра отводится 7 бит.

Для предотвращения появления запрещенных хромосом, содержащих несуществующие номера промежуточных узлов, число бит, отводимых для их кодирования, варьирует в зависимости от числа узлов в сети и принимает значения от 3 до 7 бит.

Пример кодирования хромосомы представлен на рис. 1.



**Рис. 1.** Структуры хромосомы

В представленной хромосоме значение балансирующего параметра равно 74, что означает, что 74% потока будет проходить через первый промежуточный узел (узел №109), а 26% потока будет проходить через второй промежуточный узел (узел №65).

Генотип – это набор хромосом для выбранной особи. Количество хромосом в генотипе определяется количеством потоков в сети.

Пример жизнеспособной особи, состоящих из 3 хромосом (на кодирование значения каждого промежуточного узла отводится 3 бита), и ее фенотипа (декодированная структура) приведен в табл. 1.

**Таблица 1.** Генотип и фенотип особи генетического алгоритма

Генотип	Фенотип		
	Балансировочный параметр	Промежуточный узел для первой части потока	Промежуточный узел для второй части потока
0101100000010	44	0	2
0110110011000	54	3	0
1001100010011	76	2	3

При увеличении числа параллельных подпотоков описанную структуру хромосомы можно легко модифицировать, добавив нужное количество балансирующих параметров и промежуточных узлов. Осуществляя подобную модификацию следует учитывать, что чрезмерное удлинение хромосомы может негативно сказаться на скорости работы и сходимости генетического алгоритма.

## СПИСОК ИСПОЛЬЗУЕМЫХ ИСТОЧНИКОВ

1. Сапрыкина, Е.Н. Использование двухфазной многоканальной маршрутизации в сетях дата-центров / Е.Н. Сапрыкина, А.Н. Сапрыкин // Новые информационные технологии в научных исследованиях: материалы XXI Всероссийской научно-технической конференции студентов, молодых ученых и специалистов. – Рязань: РГРТУ, 2016. – С. 337–338.

2. Сапрыкина, Е.Н. Генетический алгоритм балансировки нагрузки в сетях центров обработки данных, использующих двухфазную многоканальную маршрутизацию / А.Н. Сапрыкин, Е.Н. Сапрыкина // Современные технологии в науке и образовании - СТНО-2017: сб. тр. междун. науч.-практич. И научн.-практ. конф.: в 8 т.; под ред. О.В. Миловзорова. – Рязань:РГРТУ, 2017. – Т.4 – С. 253-255.

**А.И. СОБОЛЕВА, А.Ю. ГРОМОВ**

Рязанский государственный радиотехнический университет

## **ИСПОЛЬЗОВАНИЕ БИБЛИОТЕКИ KERAS ДЛЯ ПРОГНОЗИРОВАНИЯ УХОДА РАБОЧИХ КАДРОВ**

*В данной статье затрагивается задача прогнозирования контроля ухода рабочих кадров, а именно возможности обработки и анализа данных с помощью библиотек языка Python.*

Успех любой современной компании зависит не только от стратегии достижения поставленных целей, но и от компетентности сотрудников, поэтому текучесть персонала является актуальной проблемой бизнеса.

Источником репрезентативной выборки для обучения в контексте рассматриваемой задачи является платформа для соревнований и научных исследований данных по прогнозированию и аналитике Kaggle.

Поля в полученном наборе включают:

- 1) уровень удовлетворенности сотрудника своей работой;
- 2) количество проектов;
- 3) средние месячные часы;
- 4) время, проведенное в компании;
- 5) были ли несчастные случаи на производстве;
- 6) прошли ли эти случаи за последние 5 лет;
- 7) отделы;
- 8) зарплата;
- 9) остался ли сотрудник.

Для работы с набором данных используются следующие основные библиотеки:

1. Numpy. Осуществляет поддержку больших многомерных массивов и матриц, а также математических функций для операций с ними.
2. Pandas. Высокоуровневая библиотека поверх Numpy для обработки и анализа данных. Производится работа с табличной структурой данных – DataFrame.
3. Seaborn, matplotlib. Библиотеки для отображения и визуализации данных.
4. Keras. Предназначена для глубокого обучения.

Разработка осуществлялась в Microsoft Azure Notebooks, что представляет собой облачное решение, предоставляющее работу с Jupyter Notebook для создания красивых аналитических отчетов, по-

звolyющих хранить вместе код, изображения, комментарии, формулы и графики.

Python предоставляет большие возможности для проведения разведочного анализа данных. Исходные данные, хранящиеся в csv-файле, загружаются в DataFrame.

Можно узнать общую информацию о DataFrame, наименование столбцов, типы переменных, количество наблюдений, общее использование памяти. После выполнения метода `info()`, видно, что для лучшего понимания необходимо переименовать столбец с информацией об отелах.

Результат метода `description()` показывает основные характеристика набора. Можно сделать вывод, что примерно 24% сотрудников покинули компанию. Уровень удовлетворенности составляет около 62%, а производительность - около 72%. Также есть возможность проанализировать, какие значения наблюдались в 25, 50 и 75 процентах случаев.

Средства языка позволяют строить сводные таблицы (используя `pivot_table`), принимающие на входе данные о столбцах, группирующие их, формируя двумерную таблицу реализующую многомерное обобщение данных. Это своеобразный аналог OLAP.

Также в предварительном анализе широко используются круговые диаграммы, диаграммы размаха (представлено на рис. 1). В данных по времени, проведенном в компании, наблюдаются выбросы – значения 7, 8, 9, 10 расположены в аномальном расстоянии от других значений.

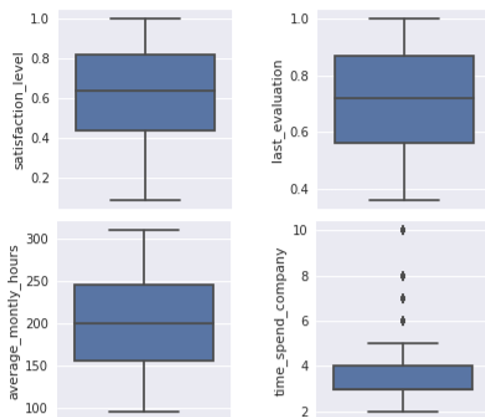


Рис. 1. Диаграммы размаха

Существуют возможности анализировать коэффициент корреляции (рис. 2).

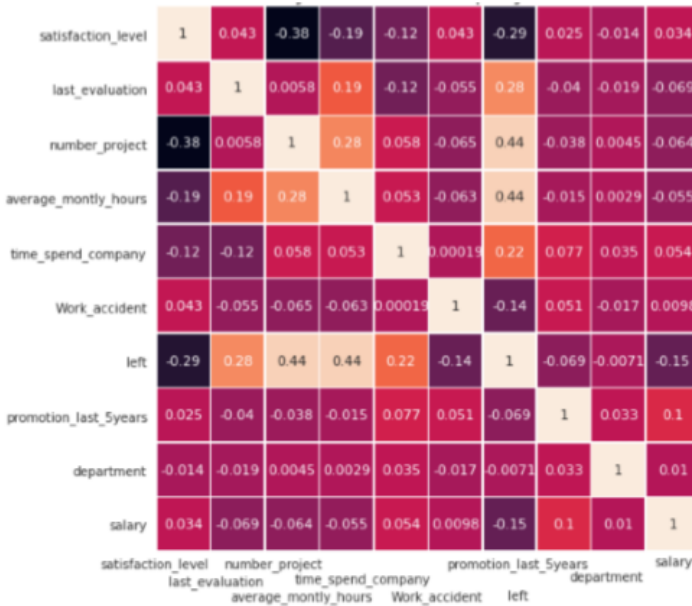


Рис. 2. Коэффициент корреляции

Проведенный разведочный анализ помогает сформировать предварительные суждения о наборе.

Перед непосредственным построением модели нейронной сети[1] происходит подготовка данных. Необходимо закодировать все текстовые значения в ручную или воспользовавшись возможностями модуля `sklearn.preprocessing`. Предпочтительнее последний вариант. С помощью объекта `LabelEncoder` значения входного списка, нуждающиеся в кодировании, однозначно отображаются в уникальные числовые, при этом теряется смысл категориальности и возникают ложные интерпретации, основывающиеся на свойствах выбранной кодировки. Чтобы этого избежать, надо воспользоваться `one-hot`-кодированием. Для вектора с размерностью равной количеству возможных категорий, описывающего кодируемый категориальный признак, все координаты обнуляются, а значение координаты кодируемой категории берется за единицу. Для анализируемой выборки осуществляется кодирование таких качественных данных, как заработная плата и наименование отдела.

Объект `model` библиотеки `Keras`[2] описывает в каком порядке и какие именно слои содержит нейронная сеть. Само описание происходит последовательно, через `Sequential API`:

```
classifier = Sequential()
classifier.add(Dense(input_dim = 20, activation = 'relu'))
classifier.add(Dropout(0.1))
classifier.add(Dense(units = 21, activation = 'relu'))
classifier.add(Dense(units = 1, activation = 'sigmoid'))
```

В полносвязных `Dense` слоях используется функция активации – полулинейная (выпрямитель). Слой `Dropout`, который нужен для снижения переобучения. При выборе очередного объекта происходит изменение в структуре, а именно каждая вершина выбрасывается с некоторой вероятностью, далее выполняется обратное распространение ошибки, в завершение выброшенные вершины возвращаются. Выход выходного слоя можно рассматривать как функцию распределения вероятности ухода сотрудника для подаваемого примера. Слой содержит один нейрон, активационная функция - сигмоидная.

При компиляции задаются параметры для решения задачи бинарной классификации, а именно функция потерь - `binary_crossentropy`, которая характеризует потери при неправильном принятии решений на основе наблюдаемых данных; список метрик оптимизации - `accuracy`:

```
classifier.compile(loss='binary_crossentropy', metrics=['accuracy'])
```

Для обучения нейронной сети устанавливаются количество эпох - проходов нейронной сети по всем записям (выбирается исходя из того, насколько быстро модель с каждым новым проходом приближается к желаемой предсказательной точности), а также количество объектов выборки, берущихся за один шаг:

```
classifier.fit(X_train, y_train, batch_size = 10, epochs = 100)
```

В процессе обучения `API` будет выводить соответствующие строки с величинами функции потерь и метрики для каждой из эпох. В результате была достигнута точность прогноза 96,84%.

В работе рассмотрены возможности библиотеки для быстрого составления и тестирования нейронных сетей `Keras`. Библиотека предоставляет собой «синтаксический сахар», надстройку над более низкоуровневой, обладающей большими возможностями настройки библиотеки `Theano`, изучение которой планируется в дальнейшем. Другим направлением развития работы является построение других моделей нейросети и определение наилучшей модели обучения в рамках решаемой задачи.

**СПИСОК ИСПОЛЬЗУЕМЫХ ИСТОЧНИКОВ**

1. Хайкин С. Нейронные сети. Полный курс. – Вильямс, 2016. – 1104 с.
2. Документация библиотеки глубокого обучения Python Keras – Режим доступа: <https://keras.io/>, свободный (дата обращения 15.04.2018).



**Ю.А. СТРОГОВА, А.Ю. ГРОМОВ**

Рязанский государственный радиотехнический университет

**ПРЕОБРАЗОВАНИЕ НЕЯВНЫХ ЗНАНИЙ ЭКСПЕРТА  
В СФЕРЕ БАНКОВСКИХ УСЛУГ ДЛЯ СИСТЕМ  
ИСКУССТВЕННОГО ИНТЕЛЛЕКТА**

*В статье рассматриваются способы получения знаний, методы выявления неявных знаний, а так же предложены критерии и ограничения алгоритма по извлечению неявных знаний в выбранной области исследования – сфере услуг (банки).*

Стремительное развитие информационных технологий параллельно способствует повышению роли знания, как важнейшего нематериального актива. Умение качественно извлечь собственные знания и в дальнейшем эффективно ими управлять позволяет увеличить конкурентоспособность организации на рынке. С целью осуществления этой задачи требуется создание исчерпывающей системы по управлению знаниями предприятия. Важное место в проведении этого процесса занимает неявное знание (накопленный опыт, мудрость, культура мышления, интуиция), которое ввиду особенностей формализации сложнее поддается представлению. Вопросы и проблематика представления знаний исследуются в разделе искусственного интеллекта – инженерии знаний. В настоящее время развитие данной области искусственного интеллекта продолжается.

В качестве области исследования и рассмотрения поставленной проблемы – извлечение неявных знаний выбрана сфера услуг. В настоящее время роль данной области человеческой деятельности заметно возрастает, ее развитие перспективно для повышения экономического роста. Как научный термин услуга предполагает экономическую деятельность, направленную на удовлетворение потребностей, как индивидуальных, так и общества в целом. В отличие от материального производства, услуга не носит материально-вещественной формы.

Особый интерес в сфере услуг представляют банковские услуги или сфера банков. Эффективность работы банка сопряжена с множеством условий, одним из которых является способность обеспечить лицензирование своей деятельности, и умение поддерживать качество предоставляемых услуг на необходимом качественном уровне. Для осуществления представленного выше условия требуется особый «интеллектуальный сервис по управлению знаниями», в котором будут отражены всеобъемлющие знания банка, как организации с собственной, присущей только этому банку их уникальностью.

Получение знаний для системы по управлению знаниями представляет собой сложный процесс. В научной теории для рассматриваемого явления используются следующие термины, характеризующие способы получения знаний: извлечение, приобретение, формирование знаний.

Извлечение знаний – процесс взаимодействия инженера с источником знаний (эксперт, специализированная литература и др.), в ходе которого выясняется представление эксперта об области исследования.

Приобретение знаний – процесс наполнения базы знаний при помощи программных средств.

Формирование знаний – анализ полученных данных и выявление закономерностей с помощью математического и программного обеспечения [1].

Представленные выше способы по-разному подходят к решению задачи по получению знаний. Извлечение и приобретение знаний являются более изученными процессами, в то время как формирование знаний является более сложной задачей, так как сам способ основывается на описании закономерностей предметной области и требует участия эксперта, которое состоит в изложении накопленного опыта в рамках заданной модели. Таким образом, сам процесс сталкивается с рядом проблем, к примеру, трудности, а порой и невозможности экспертом, ввиду наличия ограничений, представить свой накопленный опыт.

Первостепенным этапом при разработке системы по управлению знаниями является извлечение имеющихся явных и неявных знаний. Для решения поставленной задачи в выбранной предметной области используется комбинированный подход. Среди существующих многочисленных методов, которые подразделяются на коммуникативные и текстологические выбраны наиболее отвечающие поставленной задаче: наблюдение, анкетирование, интервью, анализ документов. На их основе будет разработан новый алгоритм извлечения неявных знаний, отвечающий особенностям исследуемой предметной области. По своей сути, алгоритм представляет процесс социализации (обобществления) в четырехфазной модели трансформации знания между его явной и неявной формами. Данная модель основана на трудах японских исследователей.

Процесс создания алгоритма связан с ограничениями, которые накладываются при его разработке.

Понятие «извлечение знаний» неразрывно связано с получением знаний от эксперта. Данный процесс является трудоемким и сложным, поскольку исследователю (аналитику, инженеру) необходимо предста-

вить знания эксперта (в том числе неявные) в виде модели предметной области. Таким образом, процесс дальнейшей качественной формализации знаний для интеллектуальной системы коррелируется с опытом и квалификацией самого исследователя.

С целью уменьшения описанного выше ограничения предполагается использовать диалог между исследователем и экспертом. Процесс прямого общения позволит в режиме реального времени отслеживать проявление неявных знаний, что поспособствует определению их закономерностей.

Важным критерием при разработке алгоритма является выбор самого эксперта, как источника требуемого объема знаний, которые влияют на основные процессы, происходящие в банке. Как правило, экспертами могут послужить аналитики/специалисты, выполняющие на ежедневной основе функции, отражающие непосредственную работу банка, менеджеры и руководители, которые управляют этими процессами. Наиболее подходящий вариант – эксперты, прошедшие несколько уровней одного процесса и которые постоянно находятся на разных уровнях одного процесса. Это могут быть старшие аналитики/специалисты, которые способны выполнять и постоянно выполняют функции различных уровней. Наличие последнего условия связано с тем, что процессы, происходящие в банках, постоянно меняются в связи с развитием информационных технологий и с течением небольшого времени знания могут стать неактуальными, а обучение проводится только для тех сотрудников, на выполнение функций которых и влияют произошедшие изменения.

Таким образом, разработка алгоритма извлечения и преобразования знаний, включающих как явные, так и неявные, в рассматриваемой предметной области сопряжено с особенностями самой сферы банковских услуг, так и сложностями выявления и преобразования неявных знаний.

### **СПИСОК ИСПОЛЬЗУЕМЫХ ИСТОЧНИКОВ**

1. Гаврилова Т.А., Хорошевский В.Ф. Базы знаний интеллектуальных систем. – СПб.: Питер, 2001. – 384 с.: ил.

**А.А. ТРИШИН, А.П. БУРЯКОВ**

Рязанский государственный радиотехнический университет

## **ПОВЫШЕНИЕ БЕЗОПАСНОСТИ СИСТЕМЫ ОБМЕНА ИНФОРМАЦИЕЙ СПЕЦИАЛЬНОГО НАЗНАЧЕНИЯ**

*Рассматриваются основные способы, перспективные средства и методы повышения безопасности системы обмена данными специального назначения.*

Развитие телекоммуникационных технологий привело в последние годы к серьезным изменениям в современных инфокоммуникационных сетях.

Для обеспечения оперативности, непрерывности, устойчивости и скрытности автоматизированного управления войсками и связью объединений (соединений) создаются СОД объединения (соединений).

Сложность телекоммуникационных систем и многообразие процессов обработки и передачи информации различного вида привели к необходимости создания многоуровневой модели взаимосвязи средств, в которых осуществляются эти процессы[1].

ЭМВОС разработана в 1984 г., как стандарт, на основе которого пользователи телекоммуникационного оборудования могли осуществлять взаимодействие друг с другом. Модель OSI (рис.1) разделяет сложность передачи информации на 7 менее крупных уровней или слоев, 2 низших уровня реализуются аппаратно и программно, остальные – программно.



**Рис. 1.** Уровни ЭМВОС

Современное развитие беспроводных инфокоммуникационных систем выявило серьезную проблему. Почти весь частотный диапазон на данный момент времени лицензирован.

Значительно повысить эффективность использования полосы частот предоставляет механизм динамического управления спектром. Он очень сложен технически, и может использоваться только в интеллектуальных радиосистемах. Для описания интеллектуальных систем был введен термин – когнитивное радио.

При передаче сигналов данных через физическую среду возможно воздействие на них мультипликативных и аддитивных помех, которые могут вызвать искажения сигналов.

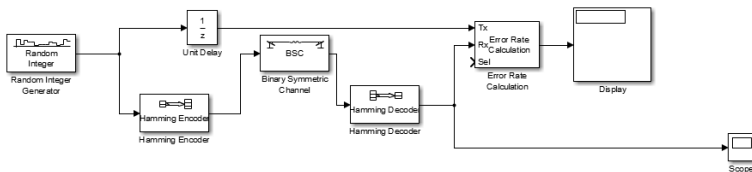
В каналах ТЧ и ШК передача осуществляется аналоговыми сигналами, образованными, как правило, на основе методов амплитудной (АМ), частотной (ЧМ), фазовой (ФМ) модуляции, а также их комбинации. Для увеличения скорости обмена данными применяют комбинированные методы модуляции. Наиболее известными являются методы QAM модуляции (Quadrature Amplitude Modulation).

Применение многопозиционной системы сигналов, кроме повышения скорости передачи в дискретном канале, дополнительно позволяет повысить достоверность принимаемых дискретных сигналов при введении ограничения на скорость передачи данных.

При передаче данных наряду с традиционными методами амплитудной, частотной и фазовой модуляции в современных телекоммуникационных системах используются методы цифрового кодирования.

В качестве формирования СКК участвуют операции отображения информационной последовательности в кодовую путем внесения избыточности. Помехоустойчивое кодирование, повышающее энергетическую эффективность СЭС, является одной из важнейших операций формирования СКК.

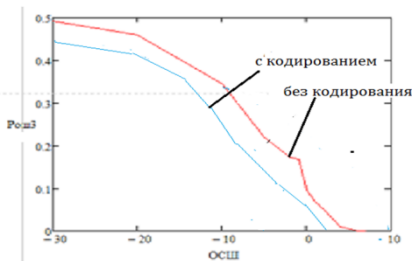
Проведем исследование наиболее распространенных видов помехоустойчивого кодирования, использующихся в современных СОД, вероятностей ошибок декодирования в различных условиях работы цифровых систем связи. Для исследования будем использовать программу расчета *MatLab*.



**Рис. 2.** Обобщенная структура для кода Хемминга

Анализ графических зависимостей показывает, что чем меньше ОСШ, тем больше вероятность ошибки перепутывания битов.

Кодирование можно изучать как инструмент, осуществляющий разные компромиссы системы. На рис.3 приведен сопоставительный вид двух кривых, описывающих зависимость достоверности передачи от ОСШ. Одна кривая определяет обычную схему модуляции без кодирования (синяя кривая), а вторая с использованием кодирования (красная).



**Рис. 3.** Сравнение ошибок передачи при использовании схемы с кодированием и схемы без кодирования

В современных телекоммуникационных системах одной из проблем является обеспечение требуемой достоверности передаваемой информации различного вида. Эта проблема особенно важна при передаче данных по реальным каналам и линиям связи с низкой помехозащищенностью.

В большинстве применяемых способов обеспечения достоверности используется введение избыточности в передаваемое сообщение.

Проведенные исследования показали, что преобразование последовательности данных в новую “улучшенную”, обладающую избыточностью, позволяет существенно уменьшить вероятности ошибок, тем самым повысить безопасность передаваемой информации, что в свою очередь приводит к существенному снижению сроков на принятия решений.

### СПИСОК ИСПОЛЬЗУЕМЫХ ИСТОЧНИКОВ

1. Титов В. С., Аверьянов Е. Г., Ащеулов С. В., Бабич Б. И., Дормидонтов А. А., Зубакин В. В., Калюка В. И., Мишин А. И., Молокович И. А., Макаров М. И., Сазонов В. В. Теоретические основы передачи данных: Уч. пособие. / Под ред. В. С. Титова – СПб.: ВАС, 2013. – 212 с.
2. Бергсекас Д., Галлагер Р. Сети передачи данных. М.: Мир, 1989. 544 с.
3. Блэк Ю. Сети ЭВМ: Протоколы, стандарты, интерфейсы. М.: Мир, 1990. 506 с.

К.Н. ТУРБИН, А.А. ПЕТУХОВ

Рязанский государственный радиотехнический университет

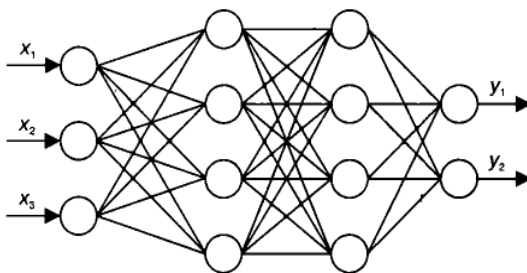
## МЕТОДЫ ОПТИМИЗАЦИИ РАБОТЫ НЕЙРОСЕТЕЙ

*Рассматриваются методы и алгоритмы, предназначенные для улучшения производительности и эффективности работы нейронных сетей.*

Нейронные сети являются одним из самых популярных методов машинного обучения. Они приобрели широкую популярность ввиду их высокой эффективности для решения широкого класса трудно формализуемых задач, таких как машинное зрение, распознавание образов, поиск скрытых зависимостей, кластеризация и других.

Существует несколько критериев эффективности работы нейронной сети. При одинаковой точности вычислений более эффективной будет считаться нейросеть, процесс обучения и вычисления выходных значений которой требует меньшее количество вычислительных ресурсов. Соответственно, среди нейронных сетей, обучающихся с одинаковой скоростью и одинаковыми затратами вычислительных ресурсов, более эффективной считается та, которая позволяет добиться более точного выходного значения.

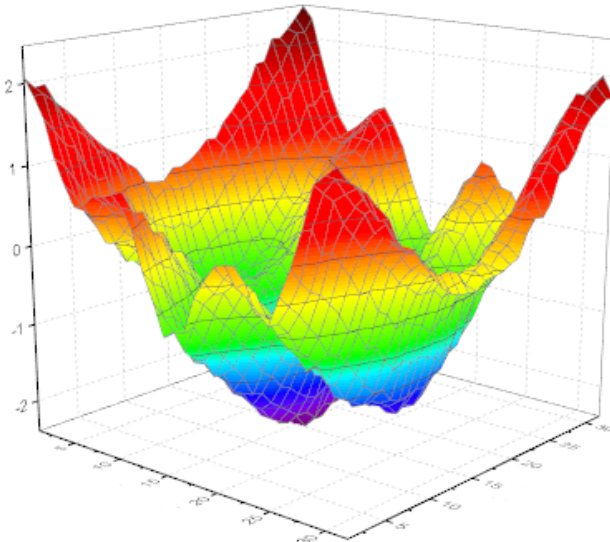
Классической реализацией нейронной сети принято считать многослойный персептрон, представляющий собой многослойную сеть, содержащую входной слой, выходной слой и произвольное количество скрытых слоёв.



**Рис. 1.** Пример нейронной сети (многослойный персептрон)

При этом, каждый нейрон каждого слоя имеет связь с каждым нейроном следующего и предыдущего слоя. Значения весов в такой реализации нейросети изменяются с использованием алгоритма обратного распространения ошибки, реализующего градиентный спуск по n-мерной поверхности функции ошибки, формируемой при обучении на

обучающем наборе данных. При этом  $n$  – это размерность вектора выходных данных.



**Рис. 2.** Пример трёхмерной поверхности функции ошибки

Основным фактором, влияющим на точность нейронной сети, является вероятность попадания в локальный минимум гиперповерхности ошибки. При подобном явлении дальнейшее обучение сети не будет увеличивать точность получаемых выходных данных. При этом локальные минимумы могут располагаться сколько угодно далеко от глобального и их может быть сколько угодно много.

Для сокращения влияния данной проблемы существуют модификации градиентных алгоритмов, например, алгоритм «упругого» распространения ошибки. В этом алгоритме добавлено специальное правило – если при следующем шаге спуска у частной производной поверхности ошибок меняется знак (что означает, что алгоритм «перешагнул» через минимум), это действие откатывается, а коэффициент скорости обучения уменьшается, иначе – увеличивается. Это правило позволяет избавиться от необходимости задавать глобальный параметр скорости обучения вообще. В результате, алгоритм упругого распространения ошибки показывает более быструю сходимость, однако требует больше памяти для каждого нейрона и является несколько более сложным для вычисления и реализации.



Применимы и неградиентные методы оптимизации, например, алгоритм имитации отжига. Данный алгоритм является одним из вариантов случайного поиска и имеет физический аналог – процесс, происходящий при остывании расплавленного металла. При формировании кристаллической решетки атомы металла движутся хаотично, но при дальнейшем снижении температуры энергия атомов уменьшается и их колебания уменьшаются до тех пор, пока они не займут положение с минимальной энергией.

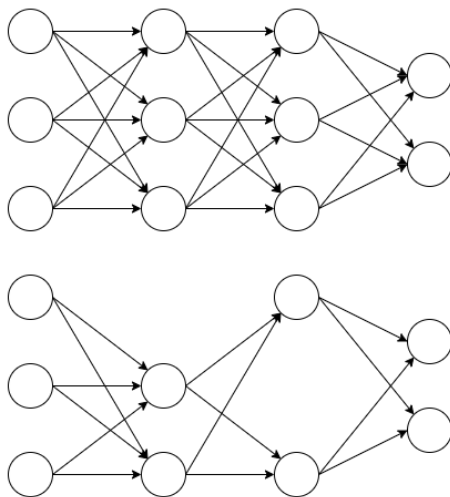
В математической модели начало процесса описывается некоторым случайным начальной точкой и далее на каждом шаге алгоритма случайным образом определяется следующая точка в окрестности текущей точки, которая уменьшается при “остывании” модели. При поиске глобального минимума вероятность перехода равна единице, если значение функции в новой точке меньше текущего значения функции. Если выбрана точка со значением функции большим, чем в текущей точке, то вероятность перехода зависит от текущей температуры. С каждым шагом температура системы уменьшается. Из-за своей вероятностной природы алгоритм не гарантирует нахождение глобального экстремума, но быстро сходится к решению.

Для предотвращения переобучения нейронной сети применяются алгоритмы регуляризации, такие как  $L-1$  и  $L-2$ .

Алгоритмы  $L-1$  и  $L-2$  добавляют в функцию потерь коэффициент “штрафа”, который регулирует ограничение величины весов нейрона. В результате предотвращается чрезмерный рост величины весов нейронов при обучении и сеть лучше обобщает новые данные, ранее ей не встречавшиеся.

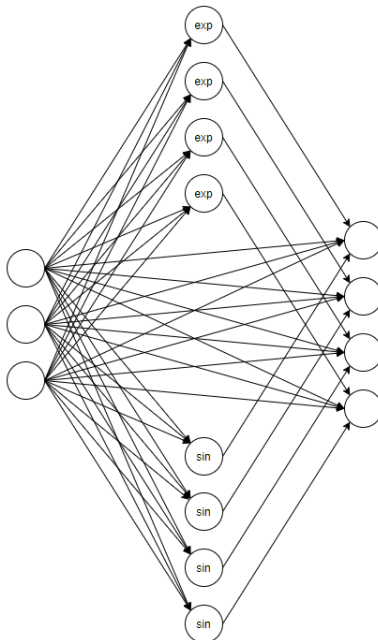
Основным фактором снижения вычислительной нагрузки на аппаратное обеспечение является снижение количества как самих нейронов, так и количества связей между ними – синапсов. При этом важной является задача определения незначимых синапсов и нейронов. При классической реализации алгоритма обратного распространения ошибки отсутствие корреляции между какими-либо входными параметрами может внести излишние шумы в выходной результат.

Метод *Dropout* с заданной вероятностью выключает некоторые нейроны, что в результате упрощает структуру сети. Посредством такого прореживания получают набор нейросетей и проводят их обучение на одинаковом обучающем наборе. Полученные веса синапсов затем усредняются в исходной структуре нейросети. Данный метод предотвращает совместную адаптацию между скрытых слоёв, которая приводит к переобучению сети.



**Рис. 4.** Пример работы метода *Dropout*

Для улучшения результатов, можно также менять и архитектуру сети. Например, нейронные сети Ворда предусматривают наличие в скрытом слое нескольких блоков нейронов, которые используют разные передаточные функции. Это позволяет расширить диапазон решаемых задач, например, добавлением блока с нейронами, имеющими периодическую передаточную функцию возможно обучить сеть выявлять периодические зависимости.



**Рис. 5.** Пример нейронной сети Ворда

Следует отметить, что перечисленные выше методы являются лишь малой частью существующих на данный момент методик и алгоритмов, позволяющих оптимизировать существующие технологии построения, обучения и использования нейросетей и подстроить их под решение конкретных прикладных задач с минимальными издержками.

### **СПИСОК ИСПОЛЬЗУЕМЫХ ИСТОЧНИКОВ**

1. Круглов В.В., Борисов В. В. Искусственные нейронные сети. Теория и практика. — 2-е изд., стереотип. — М.: Горячая линия-Телеком, 2002. — 382 с: ил.
2. Паклин Н.Б., Орешков В.И. Бизнес-аналитика: от данных к знаниям: учеб. пособие. — 2-е изд., испр. — СПб.: Питер, 2013. — 704 с.: ил.
3. Савин А. Н. Применение алгоритма оптимизации методом имитации отжига на системах параллельных и распределённых вычислений / Савин А. Н., Н. Е. Тимофеева // Изв. Саратов. ун-та. Нов. сер. Сер. Математика. Механика. Информатика, 2012, 12:1, 110–116

**А.М. ФАМ, А.П. ШИБАНОВ**

Рязанский государственный радиотехнический университет

## **РАЗРАБОТКА ПРОГРАММЫ КОМБИНИРОВАННОЙ СИСТЕМЫ МОДЕЛИРОВАНИЯ НА ОСНОВЕ СЕТЕЙ ПЕТРИ И GERT-СЕТЕЙ**

*В статье рассматриваются разработка программного обеспечения комбинированной системы моделирования на основе сетей Петри и GERT-сетей на языке C#. В сочетании с набором библиотек React.NET она представляет собой библиотеку с открытым исходным кодом для программирования дискретно-событийного моделирования, выполняемого поверх Microsoft®. NET Framework, с интерфейсом GraphX.*

*Введение.* Математические сети и их исследование играют все более важную роль в описании и совершенствовании системного анализа из-за легкости доступа и визуализации, с помощью которой системы могут быть представлены и смоделированы в сетевой форме.

Использование математических сетей может принести преимущества: способность моделировать сложные системы путем объединения простых систем; формальная процедура получения системных функций из элементарных сетей; средства для определения требований к данным для анализа системы и планирования системы управления.

*Целью работы* является разработка программного обеспечения с применением аппарата сетей Петри и GERT-сетей [1,2] для исследования простых и сложных систем, что вносит определенный вклад в совершенствование социальных и технических процессов.

*Теоретическая часть.*

Сетевой анализ – метод анализа, основанный на определении стохастических моментов времени начала и окончания нереализованных частей проекта. Он позволяет связывать выполнение различных работ и процессов во времени, получая прогноз общей продолжительности реализации всего проекта. Сетевые операции могут быть как вероятностными, так и детерминированными. Вероятностные сети могут ограничиваться выполнением лишь части условий и событий, заданных предшествующими элементами. Для системного моделирования важную роль играют GERT-сети (Graphic Evaluation and Review Technique).

GERT – это метод анализа сетей, который описывается следующими характеристиками: определена вероятность того, что ветвь сети действительно выполняется в сети; а также случайное время, связанное с ветвью, если данная ветвь выполняется. Такие сети будут называться стохастическими сетями и состоять из набора ветвей и узлов.

Реализация сети – это определенная последовательность ветвей и узлов, которые проходятся по всевозможным траекториям из единственного входного узла до единственного выходного узла. Они описывают сеть для одного эксперимента. Можно упомянуть преимущества GERT: они позволяют выполнять циклы, реализовывать вероятностный выбор следующей операции и позволяют получить несколько конечных результатов проекта. Можно отметить, как масштабность и разнообразие исследуемых систем, так и сложность и математического анализа GERT-сетей.

Сеть Петри представляет собой двудольный ориентированный граф, в котором позициям соответствуют вершины, а переходам – вершины. Распределение маркеров по позициям называют маркировкой. Маркеры могут перемещаться в сети. Каждое изменение маркировки называют событием, причем каждое событие связано с определенным переходом. Каждому условию в сети Петри соответствует определенная позиция. Совершению события соответствует срабатывание (возбуждение или запуск) перехода, при котором маркеры из входных позиций этого перехода перемещаются в выходные позиции. Последовательность событий образует моделируемый процесс. Использование сети Петри приносит выгоды: 1) пользователи описывают системы с графическим обозначением; они непосредственно не связаны с математическими основами; 2) они могут использоваться на ранних этапах жизненного цикла разработки, что позволяет внести изменения в проект относительно недорого; 3) сеть может использоваться на различных уровнях абстракции; 4) может вмещать информацию о времени; 5) используется язык, который позволяет автоматически генерировать конкретную симуляцию. Чтобы выполнить моделирование, нужно использовать концепцию дискретно-событийного моделирования (англ. discrete-event simulation, DES). Каждое событие происходит в определенный момент времени и изменяет состояние системы. Между последовательными событиями никаких изменений в системе не предполагается; таким образом, процесс моделирования может заключаться непосредственно в переходе во времени от одного события к другому.

#### *Имитационное моделирование.*

Для представления интерфейса программного обеспечения используется GraphX for .NET. Это расширенная библиотека раскладки и визуализации с открытым исходным кодом, которая поддерживает различные алгоритмы компоновки и предоставляет множество инструментов для визуальных настроек. Он способен отображать большое



узлы и дуги с параметрами, которые представляют характеристики моделируемой системы. Кроме узлов Генератор, Очереди и Терминатор, которые должны иметь большинство программ моделируемой системы, узел Конъюнкция создается для имитации перехода в сети Петри. Параметры узла содержат условия, при котором может быть срабатывание перехода. Определяется параметр, указывающий задержку срабатывания перехода. То есть, когда требуемые условия выполняются, срабатывание может произойти немедленно или после определенной задержки. Задержка может характеризоваться распределением некоторой выходной случайной величины GERT-сети.

На рис. 3 приведено обозначение узла Конъюнкция.

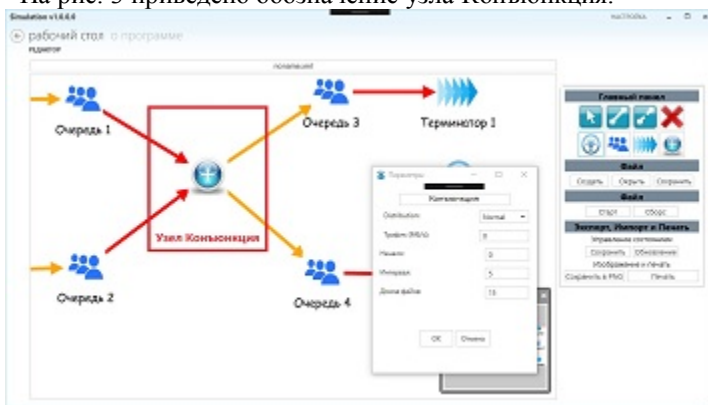


Рис. 3. Узел Конъюнкция

В некоторых случаях можно использовать одну из вышеприведенных сетей для имитации систем. Объединение этих двух сетей позволяет использовать преимущества каждой сети. Это помогает сделать процесс моделирования более удобным, помогает экономить средства и время на проектирование.

### СПИСОК ИСПОЛЬЗУЕМЫХ ИСТОЧНИКОВ

1. Кудрявцев Е.М., Microsoft Project. Методы сетевого планирования и управления проектом. – М.: ДМК Пресс, 2011. – 240 с.
2. Питерсон Дж. Теория сетей Петри и моделирование систем. – М: Мир, 1984. – 264 с.

**В.Т. ФАМ, В.В. СУСКИН**

Рязанский государственный радиотехнический университет

## **ИСПОЛЬЗОВАНИЕ МЕТОДА ГРАДИЕНТНОГО СПУСКА ДЛЯ ОПРЕДЕЛЕНИЯ ПАРАМЕТРОВ ФОТОЭЛЕКТРИЧЕСКОГО МОДУЛЯ С ОХЛАЖДЕНИЕМ**

*Предлагается метод определения параметров фотоэлектрического теплового модуля, функционирующего в области высокой температуры окружающей среды с помощью метода градиентного спуска.*

В настоящее время в регионах с наиболее высокой солнечной инсоляцией 1800 кВт час в год на  $1\text{ м}^2$  и более практически отсутствуют солнечные установки прямого преобразования солнечной энергии в электрическую.

Такое положение дел вызвано двумя причинами, связанными с высокой температурой окружающей среды в этих регионах:

– При температуре окружающей среды около  $+45^\circ\text{C}$  собственная температура фотоэлектрического модуля (ФЭМ) становится выше температуры  $+85^\circ\text{C}$  – максимально допустимой гарантийной температуры работы модуля;

– при высокой температуре окружающей среды резко, на 30-40 %, падает мощность ФЭМ, что делает проекты по созданию ФЭМ малоэффективными.

Для решения этой задачи наиболее в конструктивном отношении применяются фотоэлектрические модули с охлаждением (далее – ФЭМО), использующие радиатор с проточной жидкостью.

Однако, в настоящее время использование фотоэлектрические модули с охлаждением крайне ограничено, что связано с увеличением стоимости модуля. Поэтому разработка методов определения параметров ФЭМО, которые имеют низкие затраты изготовления ФЭМО, функционирующие в областях с высокой температурой окружающей среды, является актуальной задачей.

В данной статье рассматривается метод определения параметров двух конструкции фотоэлектрического модуля с охлаждением, которые изображены на рис. 1 [1, 2].



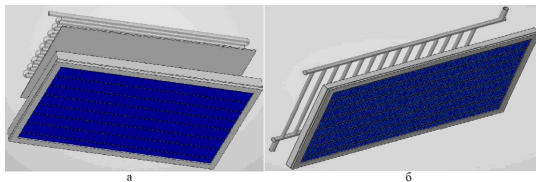


Рис. 1. конструкции ФЭМО

а – ФЭМО типа «лист – труба», б – ФЭМО канальным радиатором

В работах [1, 2] также получены математические модели температурного режима двух конструкции ФЭМО, которые определяется в виде:

$$T_{\text{сред}} = \frac{T_{\text{вых}} - T_0 - \frac{K}{\lambda}}{m \cdot L} \cdot \tanh(m \cdot L) + T_0 + \frac{K}{\lambda} \quad (1)$$

$$\text{где } m^2 = \frac{\lambda}{\lambda_4 \cdot s \cdot \frac{\beta_1}{\beta_2}}; \quad \beta_1 = \frac{1}{\frac{\delta_1}{\lambda_1} + \frac{\delta_2}{\lambda_2}}; \quad \beta_2 = \frac{2 \cdot \pi}{\frac{1}{\lambda_3} \cdot \ln \frac{R_2}{R_1}}; \quad K - \text{коэффициент}$$

поглощения;  $T_0$  – температура окружающей среды;  $\lambda, \lambda_1, \lambda_2, \lambda_3, \lambda_4$  – коэффициент теплопроводности окружающей среды, ФЭМ, поглотителя, радиатора и жидкости;  $\delta_1, \delta_2$  – толщина ФЭМ и поглотителя;  $R_1, R_2$  – внешний и внутренний радиус трубы радиатора;  $L$  – длина единицы трубы.

$$T_{\text{сред}} = (T_{a,0} - \frac{K}{\lambda} - T_0) \cdot \frac{(\beta + \lambda) \cdot \dot{m} \cdot C_p}{\beta \cdot \lambda} \cdot \left[ e^{\frac{\beta \cdot \lambda}{(\beta + \lambda) \cdot \dot{m} \cdot C_p} \cdot L} - 1 \right] \cdot \frac{\tanh[m \cdot a]}{m} + T_0 + \frac{K}{\lambda} \quad (2)$$

$$\text{где } \beta = \frac{2 \cdot \pi}{\frac{1}{\lambda_2} \cdot \ln \frac{D}{d}}; \quad m^2 = \frac{\lambda}{\lambda_1 \cdot \delta \cdot L}; \quad K - \text{коэффициент поглощения}$$

ФЭМ,  $\delta$  – толщина ФЭМ;  $\lambda, \lambda_1, \lambda_2$  – коэффициент теплопроводности окружающей среды, ФЭМ и радиатора;  $d, D$  – внутренний и внешний диаметр трубы радиатора;  $T_0$  – температура окружающей среды;  $L$  – длина единицы трубы;  $\dot{m}, C_p$  – массовой расход и теплоемкость теплоносителя.

Формула (1) - математическая модель температурного режима ФЭМО типа «лист – труба».

Формула (2) – математическая модель температурного режима ФЭМО канальным радиатором.

Для увеличения КПД ФЭМ, необходимо определить параметры ФЭМО, чтобы получить минимальное значение  $T_{сред}$ . Следовательно, получить минимальную стоимость ФЭМО.

Из формул (1) и (2) видно, что функцией  $T_{сред} = f(x)$  является многомерная функция с переменными  $x = (T_0, \lambda, \delta, \dots)$ . Для нахождения минимального значения  $T_{сред} = f(x)$  мы предлагаем использовать метод градиентного спуска.

Известно, что градиент функции  $T_{сред} = f(x)$  в каждой точке направлен в сторону наискорейшего локального возрастания этой функции. Следовательно, для поиска минимума необходимо спускаться в противоположном направлении. Если минимизируемая функция дифференцируема и ограничена снизу, а ее градиент удовлетворяет условию Липшица, то интегральный процесс [3]:

$$x_i^{k+1} = x_i^k - t_k \cdot \text{grad}_{x_i} f(x) \quad (3)$$

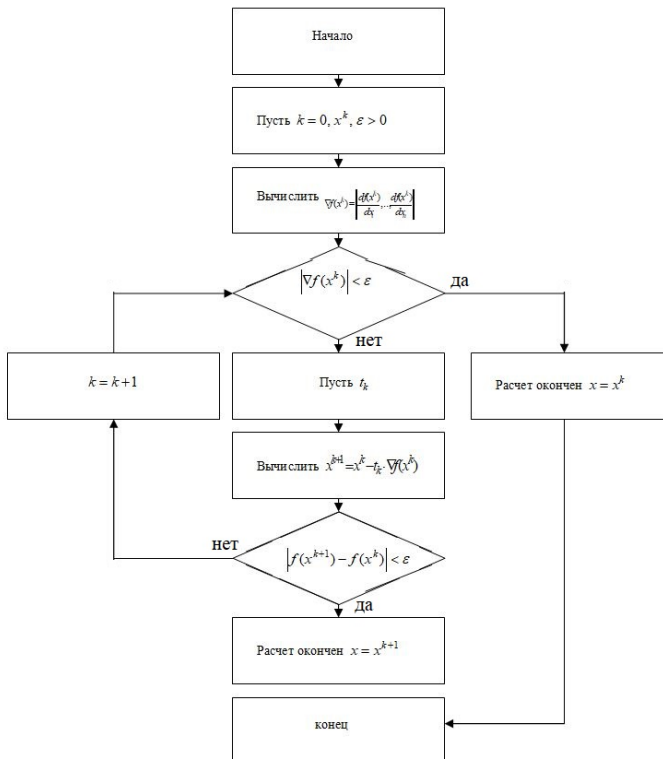
Будет сходиться к минимуму функция  $f(x)$  из произвольной начальной точки с координатами  $x_1^0, x_2^0, \dots, x_n^0$ .

Параметр  $t_k$  в формуле (3) определяет длину шага в направлении спуска. Длину шаг  $t_k$  можно выбирать из условия минимизации функции вдоль направления, противоположного градиенту. В последнем варианте итерации по формуле (3) проводятся с начальным шагом  $t_k$  до тех пор, пока функция  $f(x)$  убывает, т.е. выполняется условие:

$$f(x_1^{k+1}, \dots, x_n^{k+1}) - f(x_1^k, \dots, x_n^k) < \varepsilon \quad (4)$$

При невыполнении этого условия шаг  $t_k$  уменьшается вдвое, вычисляются координаты  $x_i^{k+1}$  с новым шагом и вновь проверяется условие (4). Дробление шага продолжается до тех пор, пока не будет выполнено условие(4).

С помощью градиентного спуска минимум гладких функций находится значительно быстрее, чем при использовании других методов. Однако наряду с вычислением функции  $f(x)$  на каждой итерации градиентного метода приходится вычислять составляющие градиента этой функции.



**Рис. 2.** Блок – схема алгоритма минимизации функции многих переменных методом градиентного спуска

Программа для оптимизации параметров ФЭМО реализована на Matlab с помощью GUI. Такой подход выгоден тем, что в то время как процессы, обрабатываемые системой, довольно ресурсоемки, потоки, создаваемые нашей программой, могут быть созданы и обработаны за 200 мкс.

Разработанная программа (рис. 3) позволяет определить параметры ФЭМО, оптимизируя значение конструктивно-технологических и эксплуатационных параметров ФЭМО и вычисляя минимальное значение температуры корпуса ФЭМО. Результатам моделирования является оптимальная характеристика ФЭМО, которая позволяет оптимизировать пропускную способность конструкции ФЭМО, и, следовательно, минимизировать затраты на оборудование. Пользователь задает или выбирает материал ФЭМО, начальные параметры и допустимые погрешности.

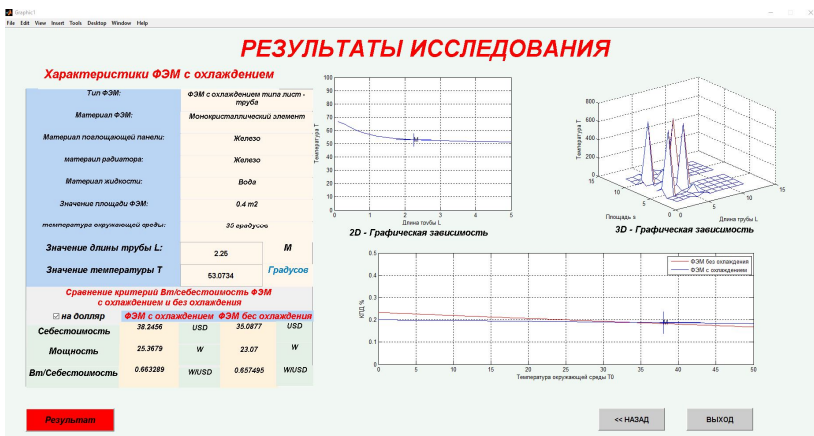


Рис. 3. Главное окно разработанной программы

Результатами работы программы являются:

- Оптимальные конструктивно-технологические и эксплуатационные параметры ФЭМО, функционирующего в повышенной температуре окружающей среды.
- Сравнение критерий Вт/себестоимость ФЭМ с охлаждением и без охлаждения.
- График 2D и 3D зависимости температуры корпуса и КПД ФЭМО от параметров.

### СПИСОК ИСПОЛЬЗУЕМЫХ ИСТОЧНИКОВ

1. Сускин В. В., Фам В. Т. Математическая модель конструкции фотоэлектрической панели с охлаждением // Вестник РГРТУ. 2017. – № 61. С. 119 – 123.
2. Сускин В. В., Фам В. Т. Математическая модель температурного режима фотоэлектрической панели с охлаждением тепловым канальным радиатором // Вестник РГРТУ. 2017. – № 62. С. 143 – 148.
3. Сухарев А.Г., Тимохов А.В., Федоров В.В. Курс методов оптимизации. – М.: наука, 1986. – 328 с.

**К.Ю. ФИЛАТОВ**

Рязанский государственный радиотехнический университет

**АНАЛИЗ СВОЙСТВ ЮЗАБИЛИТИ***Рассматриваются основные принципы построения пользовательских интерфейсов, свойства юзабилити.*

При создании информационных систем различного назначения, всегда встает вопрос о моделировании интерфейса пользователя. Пользовательский интерфейс (англ. User Interface, UI) является коммуникационным каналом, по которому осуществляется взаимодействие пользователя и компьютера. Выделяют две основные характеристики пользовательского интерфейса: функциональность и юзабилити. Функциональность показывает, насколько реализованные функции соответствуют требованиям. Согласно стандарту ISO 9241-11: «Юзабилити — степень, с которой продукт может быть использован определёнными пользователями при определённом контексте использования для достижения определённых целей с должной эффективностью, продуктивностью и удовлетворённостью».

Говоря другими словами, юзабилити — это то, в какой степени готовый продукт удобен и прост в понимании для конечных пользователей. Сегодня данный термин обычно используется в отношении систем с графическим интерфейсом: сайты, программы, операционные системы, мобильные приложения, интерфейсы различных терминалов и т.п.

Определение содержит в себе три важных понятия: определенные пользователи, определенный контекст использования и определенные цели, что означает: юзабилити для все быть не может. Что отлично подходит одному человеку, может оказаться совсем не понятным другому. Очень важно в начале проектирования интерфейса провести исследования и проанализировать интересы и возможности целевой аудитории продукта: какие особенности имеют возможные пользователи системы, работали ли они с подобными продуктами, в каком контексте эти люди будут использовать разрабатываемый продукт, для чего люди будут использовать именно этот продукт, какие цели они преследуют.

Продуктивность, эффективность и удовлетворенность — это главные понятия юзабилити, с помощью которых можно измерить юзабилити интерфейса. Используя эти термины можно оценить удобство разрабатываемого продукта. Чтобы оценить удобство разрабатываемого продукта используют две основных оценки:

- прямая оценка, которая рассчитывается на основе анализа результативности, эффективности и удовлетворённости, которые достигаются в процессе использования продукта в конкретных целевых задачах: если одна система будет более привлекательна с точки зрения эргономики, то прямая оценка должна четко это показывать.

- косвенная оценка, опирается на анализ отдельных конкретных подхарактеристик, которые фиксирует некоторые свойства разрабатываемой системы при определенных эксплуатационных условиях.

Выделяют следующие основные компоненты (факторы) юзабилити:

- легкость обучения - быстро ли пользователь обучается работать с системой;

- эффективность обучения - быстро ли пользователь работает после обучения;

- запоминаемость обучения - насколько легко пользователь восстанавливает эффективность использования системы после перерыва в работе;

- ошибки - часто ли пользователь допускает ошибки в работе и насколько они серьезные, как быстро/просто осуществляет корректировку ошибок;

В рамках данной темы создается информационное табло, призванное помочь пользователю самостоятельно найти необходимую информацию без привлечения специалистов, что значительно экономит ресурсы компании. В текущей интерпретации это некий информационный стенд, который помогает пользователю сориентироваться куда пойти (какой этаж, какой отдел, часы работы), чтобы получить необходимые услуги. Система полностью автоматизирована: имеется отдельное приложение, призванное администрировать информационное табло в части его наполнения.

Чтобы отвечать требованиям юзабилити, система придерживается основных принципов создания подобных интерфейсов. Принцип золотого сечения, используется в определении пропорций информационных блоков (ширины текстовых боксов, разбивки страницы и т.п.). Это делается для достижения "идеальных" пропорций, при которых внимание пользователя устремлено на нужные области интерфейса. Согласно принципу группировки, экран программы разбит на ясно очерченные блоки элементов: по функциям, по результату от действия, по функциям и результату действия (смешанный тип). Применяя принцип «Бритвы Оккамы», удалось добиться простоты и понятности пользования информационным стендом. Все нужные для поиска информации действия, интуитивно понятны. Пользователь всегда пони-

мает, что происходит и видит отклик системы в любых ситуациях, в том числе связанных с возможными сбоями в системе, например из-за разрыва соединения с базой данных. Согласно принципу «Видимость отображает полезность» были разработаны кнопки быстрого доступа, которые отображают наиболее часто встречаемые поисковые запросы пользователей.

Понятие удобства пользования системой не заканчивается степенью простоты ее эксплуатации. Согласно стандарту серии ISO 9241 данная характеристика имеет более широкий смысл. Чтобы оценить удобство разрабатываемой системы нужно учитывать личные цели, которые преследует пользователь, его эмоции и чувства, определяющие восприятие системы, в том числе и удовлетворенность от работы с продуктом. Пригодность использования во многом зависит от конкретных задач и окружающей среды. Пригодность использования — это далеко не абсолютное понятие, это говорит о том, что оно может по разному открываться в различных условиях эксплуатации.

#### **СПИСОК ИСПОЛЬЗУЕМЫХ ИСТОЧНИКОВ:**

1. Юзабилити [Электронный ресурс] // URL: <https://ru.wikipedia.org/wiki/Юзабилити>.
2. Что такое Юзабилити [Электронный ресурс] // URL: <https://usabilitylab.ru/blog/chto-takoe-yuzabiliti/>.
3. Принципы построения интерфейсов [Электронный ресурс] // URL: <http://www.realcoding.net/articles/printsipy-postroeniya-interfeisov.html>

**И.В. ЦУКАНОВ**

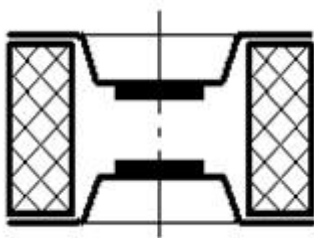
Рязанский государственный радиотехнический университет

## **ПРОБЛЕМА КОНСТРУИРОВАНИЯ ВЫСОКОСТАБИЛЬНОГО НЕУПРАВЛЯЕМОГО РАЗРЯДНИКА**

*В данной статье рассматриваются перспективные направления в области конструирования неуправляемого разрядника. Предлагаются возможные варианты изменения конструкции прибора при стабилизации параметра статического пробоя.*

Проблема конструирования высокостабильного неуправляемого разрядника

Разрядник – это электрический аппарат, искровой промежуток которого пробивается при определенном значении приложенного напряжения, ограничивая тем самым перенапряжения в установке<sup>[1]</sup>. Их сфера применимости очень широка, от бытовых приборов вплоть до самолетов. Существуют различные типы и виды разрядников. В данной статье, мы выделим один из типов – неуправляемые разрядники рис. 1.



**Рис. 1.** Защитный металлокерамический разрядник

Особенность вольт-амперной характеристики защитного разрядника это присутствие порогового напряжения, ниже которого разрядник представляется как изолятор, а выше - как низкоомный проводник.

Данный вид разрядников является защитным. Его переход в проводящее состояние неравносителен замкнутому ключу. В режим низкоомного проводника они переходят тогда, когда увеличение напряжения достигает порогового значения. Из проводящего состояния в непроводящее защитные разрядники возвращаются только после уменьшения напряжения между электродами.

Разработчиками были подняты проблемы: 1) каким образом можно было из сети убрать тысячи или даже десятки тысяч ампер; 2) временной интервал от начала подачи напряжения до момента самого пробоя, устремляя его к нулю; 3) как продлить долговечность прибора



на десятки лет. Решение некоторых проблем осуществили ещё наши предки, но все оставшиеся придется решать нам самим.

Одной из не решённых проблем на современном рынке звучит следующим образом: как добиться увеличения стабилизации параметра статического напряжения пробоя на разряднике.

Одним из вариантов решений данной задачи возможно совершенно новой конструкция разрядника с уже совершенно новыми материалами. Но так же можно воспользоваться менее радикальными вариантами. При создании газовых разрядников применяют различные газы: He, Ar, Ne, а так же их смеси. Зная закон Пашена рис. 2<sup>[2]</sup>, мы можем сказать, что чем выше у нас давление наполнения прибора, тем напряжение пробоя будет больше (используем правую часть ветви закона Пашена). Это один из факторов, которым мы можем регулировать рассматриваемый параметр.

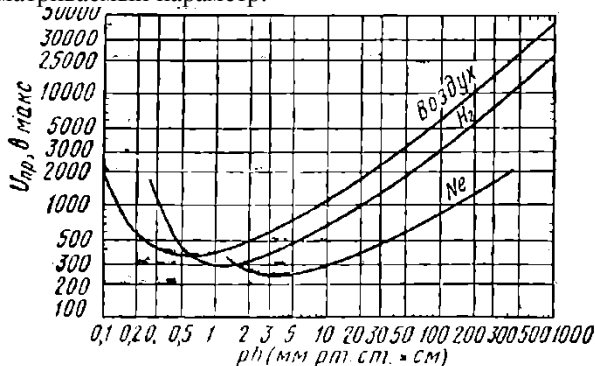


Рис. 2. Закон Пашена

Другой составляющей является добавления в газы различные радиоизотопы или точек (специальные графитовые препараты), которые стабилизируют рассматриваемый нами параметр. Третьем фактором является распыление электрода (эрозия) во время статического и рабочего режима разрядника. Частицы электрода распыляются на керамическое (-ие) кольцо (-а), которое является диэлектриком и вследствие распыления его сопротивление падает и по этому разряд может проходить по поверхности кольца, не смотря на то, что в основных случаях данное расстояние будет в разы больше, по сравнению с межэлектродным промежутком.

Для достижения длительного срока службы разрядника, и стабильности параметра напряжения пробоя статического относительно не рабочего времени, не стоит забывать о таком параметре, как коррозия. Воздействие ветра, воздуха и влажности для металлических дета-

лей сказывается как окислитель, что естественно за собой влечёт коррозию металла и последующую разгерметизацию прибора. Во избежание данного недостатка, применяют различные краски и эмали в ответственных местах, например, в точках сварки и пайки. Другим способом является защита деталей на этапе их приготовления к сборке прибора, таковым может являться никелирование поверхности контактирующей с окружающей средой.

Вывод по проделанной работе можно сделать следующий: идеально точным параметра напряжения пробоя статического не возможно будет сделать, для всего существует погрешность, но устремить существующий показатель к разумному значению в моих силах.

### **СПИСОК ИСПОЛЬЗУЕМЫХ ИСТОЧНИКОВ**

1. Электрические аппараты: Учебник для техникумов. – 3е. изд. Перераб. И доп. – Л.: Энергоиздат. Ленинг. Отд-ние, 1981. – 304 с.
2. Яновский В.П., Луканин А.А. Установка для изучения закона Пашена // Изв.ТПУ. – 2008. – N.2. – P.101-105

**П.Ю. ЧЕСАЛИН, П.В. ШЕЛЕХИН**

Рязанский государственный радиотехнический университет

## **ИСПОЛЬЗОВАНИЕ АЛГОРИТМА RANDOM FOREST ДЛЯ РЕШЕНИЯ ЗАДАЧИ КЛАССИФИКАЦИИ**

В данной статье рассматривается использование алгоритма random forest для решения задачи классификации, рассматриваются его преимущества и недостатки.

Серьёзной проблемой при использовании деревьев решений является проблема переобучения. Одним из вариантов решения данной проблемы является отсечение ветвей (pruning), сначала строится полное дерево, затем производится автоматическое удаление узлов и листьев, имеющих малую значимость (в которые попало небольшое число примеров). Проблема отсечения ветвей заключается в том, что оно снижает точность классификации на обучающем множестве [1]. Другим вариантом борьбы с переобучением является использование алгоритма random forest.

Random forest (с англ. – случайный лес) представляет собой ансамбль множества деревьев решений. Обучение деревьев происходит независимо (на разных подмножествах обучающего множества), что позволяет снизить проблему переобучения, повысить точность и является очень удобным для применения в системах распределенных вычислений.

Алгоритм обучения для random forest использует технику бэггинга (bagging). Пусть  $X = x_1, \dots, x_n$  - обучающий набор, а  $Y = y_1, \dots, y_2$  - метки классов. Бэггинг производится  $B$  раз.

Для  $b = 1, \dots, B$ :

1. Генерируется случайная подвыборка с повторением размером  $n$  из обучающего множества.

2. Строится дерево решений на основе полученной подвыборки. Дерево не подвергается процедуре отсечения ветвей.

Классифицировать новые примеры можно путём проведения голосования: каждое дерево относит пример к одному из классов, результатом будет тот, за который проголосовало наибольшее число деревьев. [3]

Чаще всего число деревьев  $B$  варьируется от нескольких сотен до нескольких тысяч, в зависимости от размера обучающего множества.

Кроме того, важной особенностью алгоритма обучения для random forest является то, что при создании очередного узла дерева используется случайное подмножество всех признаков. Обычно при раз-

мерности пространства признаков  $P$  при создании очередного узла используется  $p = \sqrt{P}$  признаков. [4]

Используя random forest можно произвести оценку важности признаков.

Для оценки важности признака в обучающем множестве первым делом требуется обучение случайного леса на обучающем множестве. Для каждого примера в процессе обучения фиксируется out-of-bag ошибка и усредняется по всем деревьям. При обучении каждого дерева используется не вся выборка и элементы, которые не используются при обучении называются out-of-bag samples. Часть выборки, которая не используется при обучении может использоваться как тестовое множество. Под out-of-bag ошибкой понимается ошибка на описанном выше подмножестве. [5]

Для оценки важности  $j$ -го признака после обучения производится перестановка значений  $j$ -го признака в обучающем множестве и после этого снова рассчитывается out-of-bag ошибка. Важность  $j$ -го признака считается как усредненная разница между out-of-bag ошибкой до и после перестановки для всех деревьев. [4] Признаки, для которых были получены большие значения считаются более важными, а признаки с наименьшим значением оказывают меньшее влияние.

Отдельно выделяются сверхслучайные деревья (Extremely Randomized Trees) [2]. При разбиении в узлах в сверхслучайных деревьях, как и в случайных деревьях, используется случайное подмножество возможных признаков, но вместо поиска наиболее оптимальных порогов разбиения, например, исходя из прироста информации или индекса Джини (Gini index), эти пороговые значения случайно генерируются для каждого признака, и лучший из них выбирается при разбиении узла.

К преимуществам random forest можно отнести следующее:

- Эффективная обработка данных с большим количеством признаков и классов.
- Имеется возможность оценки значимости отдельных признаков.
- Высокая масштабируемость.
- Слабая чувствительность к выбросам.
- Менее подвержен переобучению.

К недостаткам можно отнести следующее:

- результаты сложнее интерпретировать чем в одном дереве решений.
- требовательность к памяти для хранения модели.

Данный алгоритм широко используется в научной работе в связи со своими преимуществами. Например, он может использоваться для проведения оценки качества статей интернет-энциклопедий.

### **СПИСОК ИСПОЛЬЗУЕМЫХ ИСТОЧНИКОВ**

1. Паклин Н.Б., Орешков В.И. Бизнес-аналитика: от данных к знаниям (+ CD): учеб. пособие. – 2-е изд., испр. – СПб.: Питер, 2013. – 704 с.: ил.
2. Т.Н. Хо, A Theory of Multiple Classifier Systems And Its Application to Visual World Recognition, Buffalo, 1992
3. L. Breiman, J.H. Friedman, R.A. Olshen, and C.T. Stone. Classification and Regression Trees. Wadsworth, Belmont, California, January 1984. – p. 368.
4. Quinlan, J. R. 1986. Induction of Decision Trees. Mach. Learn. 1, 1 (Mar. 1986).
5. Quinlan, Ross J. C4.5: Programs for Machine Learning (Morgan Kaufmann Series in Machine Learning). Morgan Kaufmann, January 1993. – p. 302.

**В.С. ЧИКАЛКИН**

Рязанский государственный радиотехнический университет

## **АДАПТИВНЫЙ АЛГОРИТМ РАБОТЫ СИСТЕМЫ СИНХРОННОГО МНОЖЕСТВЕННОГО ДОСТУПА**

Рассматривается адаптивный метод работы синхронного алгоритма множественного доступа.

Сеть космической связи — это система массового обслуживания, в которой заявки представляют собой пакеты данных, которые отправляют станции с орбиты на Землю, а обслуживающим устройством является базовая станция, расположенная на Земле. Многие космические сети работают на основе алгоритма синхронного множественного доступа, однако этот алгоритм имеет существенный недостаток, который сказывается на производительности такой сети.

Связь в системах с множественным доступом можно рассматривать в терминах теории массового обслуживания. Каждый узел имеет очередь пакетов, которые должны быть переданы, а канал с множественным доступом является обслуживающим прибором. В идеальном случае обслуживающий прибор должен рассматривать все ожидающие пакеты как стоящие в одной очереди, которую нужно обслужить, используя соответствующую дисциплину обслуживания очереди.

Недостаток алгоритма заключается в том, что обслуживающий прибор не знает, у каких узлов имеются пакеты для отправки; аналогично в узлах не известно о наличии пакетов в других узлах. Таким образом, интересная особенность проблемы состоит в том, что информация о состоянии очереди является распределенной. Одним из многих методов, которые были разработаны для решения этой проблемы является метод свободного доступа для всех, при котором из узлов обычно посылаются новые пакеты немедленно в надежде на то, что другие узлы передавать не будут.

Основная идея синхронного алгоритма состоит в том, что каждый узел, не имеющий задолженности, просто передает новый поступивший пакет в первом окне после момента прихода (с вероятностью 1), рискуя, таким образом, попасть в случайные конфликты, но это приводит к очень маленькой задержке, если конфликты возникают редко. Таким образом, при синхронной системе АЛОНА пакеты передаются почти немедленно, при этом иногда попадая в конфликты.

Количество поступающих заявок на обслуживание в системе измеряется величиной Пуассоновского входного потока  $\lambda$ . График (рис. 1) показывает скорость ухода заявок в текущем окне и связывает число

переданных заявок в окне (Ось X) и Пуассоновское значение потока уходящих заявок (Ось Y). Видно, что максимальная производительность системы достигается при обращении лишь одной заявки в окне и равняется  $e^{-1} = 0,368$ . При значении входного потока не превосходящем  $0,368$  в системе много оказывается пустых окон, при значении превышающим  $0,368$  в системе появляется слишком много конфликтов и она «сваливается».



**Рис. 1.** Скорость ухода пакетов как функция интенсивности попыток передач для системы синхронная ALOHA

При таком подходе, возникает в сети множество коллизий, в следствии чего производительность такой сети оставляет желать лучшего. Для снижения количества этих коллизий и увеличения производительности сети можно адаптировать алгоритм при помощи псевдобайесовского алгоритма. Идея адаптации состоит в том, что новые и повторные заявки в выбранном станцией окне отправляются в зависимости от нагрузки сети, таким образом лишает алгоритм его главного недостатка, сказывающегося на производительности. Для реализации этой идеи вводится понятие оценки нового «окна» для отправки заявки. В конце каждого окна узлом от базовой станции отправляется значение оценки нового окна  $\hat{n}_{k+1}$ , учитывая которое станция будет отправлять пакет в новом окне не с вероятностью 1, а с вероятностью  $q_r(\hat{n}_{k+1}) = 1/\hat{n}_{k+1}$ .

Оценка нового окна вычисляется исходя из значения оценки предыдущего окна, значения Пуассоновского входного потока, а также в зависимости от того, успешным или конфликтным было окно. Так, если окно было успешным, то к оценке предыдущего окна прибавляется значение Пуассоновского входного потока и отнимается 1, если же окно было конфликтным, то к оценке предыдущего окна прибавляется значение Пуассоновского входного потока и прибавляется значение  $(e - 2)^{-1}$ . В итоге на практике узел при наличии у него пакета для отправки базовой станции отправляет заявку если канал отправки свобо-

ден, в противном случае при загрузке канала заявка не отправляется и узел ожидает «разгрузки» канала.

В итоге алгоритм работы синхронного множественного доступа преобразуется и выглядит следующим образом:

- Вначале каждого окна на основе значения входного потока случайно генерируются заявки узлов.
- Затем узлы, учитывая оценку загрузки сети отправляют свои заявки базовой станции.
- После чего происходит вычисление оценки загрузки сети для нового окна и проверка на предмет коллизии.
- В конце обновляются значения счетчиков различных данных для отображения информации о сети на экране.

### **СПИСОК ИСПОЛЬЗУЕМЫХ ИСТОЧНИКОВ**

1. Д. Бертсекас, Р. Галлагер, Сети передачи данных
2. А. А. Назаров, Ю. Д. Одышев, Исследование сетей связи с протоколами “адаптивная Алоха, 2000



**И.А. ШИРЕНИН**

Рязанский государственный радиотехнический университет

**ШАБЛОН ПРОЕКТИРОВАНИЯ «ФАБРИКА» ПРИ  
РАЗРАБОТКЕ ПРИЛОЖЕНИЙ**

*Рассматривается шаблон проектирования «фабрика», условия необходимости его применения, принципы построения и основные преимущества. Будет рассмотрено применение шаблона при создании внутренней архитектуры и пользовательского интерфейса.*

Шаблон проектирования «Фабрика» относится к порождающим шаблонам проектирования, а значит предназначения для создания новых объектов в системе. Данный шаблон пользуется большой популярностью у разработчиков из-за простоты в реализации и эффективности применения.

Фабрика позволяет классу делегировать создание подклассов. Используется когда:

- Классу заранее неизвестно, объекты каких классов ему нужно создавать.
- Класс спроектирован так, чтобы объекты, которые он создаёт, специфицировались подклассами.
- Класс делегирует свои обязанности одному из нескольких вспомогательных подклассов, и планируется локализовать знание о том, какой класс принимает эти обязанности на себя [1].
- Необходимо скрыть конкретные классы от клиента [2].

Данный шаблон является достаточно распространённым – он применяется даже в стандартных библиотеках *JAVA*, например:

- *java.util.Calendar* - метод *getInstance()*.
- *java.util.ResourceBundle* - метод *getInstance()*.
- *java.text.NumberFormat* - метод *getInstance()*.

Признаки применения паттерна: фабричный метод можно определить по создающим методам, которые возвращают объекты продуктов через абстрактные типы или интерфейсы. Это позволяет переопределять тип создаваемых продуктов в подклассах [3].

Представим ситуацию, в которой у какого-либо элемента интерфейса можно вызвать контекстное меню. При добавлении нового функционала необходимо будет также добавлять пункты контекстного меню.

В *Java* класс *JMenuItem* представляет собой элемент меню. То есть, мы можем создать наследников данного класса, имеющих собственную логику по обработке событий.

Например, меню будет состоять из двух пунктов – загрузить и переименовать. Создаем два класса (*UploadPopupItem* и *RenamePopupItem*),

```
public class UploadPopupItem extends JMenuItem implements
MenuItem{
/* дополнительная логика класса */
@Override
public void perform (JTree tree) {
/* логика обработки события */
}
public class RenamePopupItem extends JMenuItem implements
MenuItem{
/* дополнительная логика класса */
@Override
public void perform (JTree tree) {
/* логика обработки события */
}
}
```

перечисление доступных пунктов меню(*PopupMenuItemTypes*)

```
public enum PopupMenuItemTypes {
    RENAME,
    UPLOAD
}
```

и интерфейс(*MenuItem*), который содержит метод, в котором позже будет логика обработки события.

```
public interface MenuItem {
    void perform(JTree tree);
}
```

Теперь необходимо реализовать класс фабрики. Это будет класс *PopupMenuItemFactory* с методом *createItem(PopupMenuItemTypes itemType)*. Реализация метода:

```
MenuItem item = null;
switch (itemType) {
    case LOAD: item = new RenamePopupItem(); break;
    case UPLOAD: item = new UploadPopupItem(); break;
}
return item;
```

Остается в нужном месте создания интерфейса вызвать метод *createItem*, передать ему в качестве параметра тип создаваемого пункта и получить полноценный элемент, а также добавить его непосредственно в меню, передав в метод *add(JMenuItem menuItem)* класса *JPopupMenu*.

Теперь рассмотрим более общую ситуацию применения шаблона, не привязываясь к интерфейсу. Представим, что необходимо напи-

сать приложение для крупной компании, занимающейся перевозками. Сами перевозки могут осуществляться железнодорожным транспортом, воздушным транспортом или автомобильным. Но в любом случае транспортировка остается транспортировкой.

Применяя шаблон проектирования, можно выделить следующие классы: абстрактный класс *Departure* с методом *send()*,

```
public abstract class Departure {
    public abstract void send();
}
```

два класса *AirDeparture* и *RailwayDeparture*, означающие воздушную и железнодорожную перевозку соответственно,

```
public class AirDeparture extends Departure {
    public void send () { /* логика метода */ }
}
```

```
public class RailwayDeparture extends Departure {
    public void send () { /* логика метода */ }
}
```

перечисление типов перевозок

```
public enum DepartureTypes {
    AIR,
    RAILWAY
}
```

и сам класс фабрики *Factory* аналогично классу *PopupMenuItemFactory*.

В данной статье был описан такой подход к программированию, как применение шаблона проектирования. Он позволяет разработчику однозначно решать часто возникающие однотипные проблемы проектирования.

### СПИСОК ИСПОЛЬЗУЕМЫХ ИСТОЧНИКОВ

1. Фабричный метод (шаблон проектирования) [Электронный ресурс], URL: [https://ru.wikipedia.org/wiki/Фабричный\\_метод\\_\(шаблон\\_проектирования\)](https://ru.wikipedia.org/wiki/Фабричный_метод_(шаблон_проектирования)) (дата обращения 11.03.18).

2. Фабрика – паттерн разработки [Электронный ресурс], URL: <http://technerium.ru/izuchenie-java-na-praktike/fabrika-pattern-razrabotki> (дата обращения 11.03.18).

3. Фабричный метод в Java [Электронный ресурс], URL: <https://refactoring.guru/ru/design-patterns/factory-method/java/example> (дата обращения 11.03.18).

**И.И. ЯКОВЛЕВ, А.И. БАРАНЧИКОВ**

Рязанский государственный радиотехнический университет

## **РАЗРАБОТКА ОБЪЕКТНОЙ МОДЕЛИ ПРЕДСТАВЛЕНИЯ ФУНКЦИОНАЛЬНЫХ ЗАВИСИМОСТЕЙ**

*В статье рассматриваются правила формализации описания предметной области, разрабатываются классы для функциональных зависимостей (сущностей) и предметной области как части алгоритма сравнительного анализа формальных описаний предметных областей.*

В настоящее время ряд практических задач, таких как алгоритм формирования семантического описания предметной области на основе схемы реляционной базы данных и алгоритм сравнительного анализа формальных описаний предметных областей требуют наличия математического и алгоритмического обеспечения механизмов работы с формализованными описаниями предметной области, что говорит об актуальности исследования. Целью разработки является предоставление объектной модели в виде классов, предназначенных для хранения формализованных описаний, необходимого при практическом решении задач консолидации и реинжиниринга реляционных баз данных.

В данной статье рассматривается алгоритмического обеспечения сравнительного анализа формальных описаний предметных областей, общее описание которого было приведено в [1]. Рассматриваются аспекты, обуславливающие взаимосвязь моделей в терминах реляционной алгебры и моделей в виде структур данных, непосредственно используемых в программе.

Для начала оговоримся о правилах формирования корректных входных данных. В качестве входных данных принимается список функциональных зависимостей, составленный экспертом в виде текстового файла, и информацию из базы данных, к которой производится подключение в процессе выполнения приложения. Процесс извлечения функциональных зависимостей на основе семантического анализа данных рассмотрен в [2], поэтому теперь опишем правила составления текстового файла, содержащего экспертное описание:

- каждая функциональная зависимость записывается с новой строки;
- для разделения левых и правых частей функциональных зависимостей используется символ «|»;
- атрибуты разделяются пробелами [5].

Реляционная база данных содержит как структурную, так и семантическую информацию. Структура базы данных определяется чис-

лом и видом включенных в нее отношений, и связями типа "один ко многим", существующими между кортежами этих отношений. Семантическая часть описывает множество функциональных зависимостей, существующих между атрибутами этих отношений [4].

Пусть  $R$  – отношение. Множество атрибутов  $Y$  функционально зависит от множества атрибутов  $X$  ( $X$  функционально определяет  $Y$ ), если в любой момент времени каждому значению  $X$  соответствует ровно одно значение  $Y$ .

Функциональная зависимость обозначается  $X \rightarrow Y$ . Таким образом, если  $X \rightarrow Y$  и для любых кортежей  $t_1, t_2 \in R$   $t_1(X) = t_2(X)$ , то  $t_1(Y) = t_2(Y)$ . Обозначается как  $\Gamma = X \rightarrow Y$  [3].

Приведём элементарный пример составленного по данным правилам формального описания предметной области «Телефонные переговоры». Выделим в ней две функциональные зависимости:

№Абонента | ФИОАбонента АдресАбонента №Договора  
IDСоединения | №Абонента №Абонента ДлительностьСоединения

Отметим, что эксперт, составляющий формальное описание предметной области, может совершить случайные ошибки («промахи»), о наличии которых разрабатываемый алгоритм будет предупреждать работающего с ним аналитика.

При разработке программного обеспечения необходимо создать классы объектов, оптимально и просто отображающие структуру информации, поступающую на вход алгоритма. С этой целью разрабатываются классы `Entity` и `SubjectArea`. Рассмотрим их более подробно.

Класс `Entity` предназначен для представления функциональной зависимости. Поскольку каждая функциональная зависимость состоит из двух частей: левой и правой, а каждая часть, в свою очередь, может состоять нескольких атрибутов, то для реализации такой структуры данных был выбран тип `List<String>`, обладающий в C# широкими функциональными возможностями. Приведён листинг описания разработанного класса:

```
public class Entity{
    public List<String> primaryKeys;
    public List<String> attributes;
    public Entity(List<String> primaryKeys, List<String>
attributes){
        this.primaryKeys = primaryKeys;
        this.attributes = attributes}
    public Entity(){
        this.primaryKeys = new List<String>();
        this.attributes = new List<String>();}}
```

Каждая функциональная зависимость типа Entity включает в себя два списка: primaryKeys и attributes. Таким образом, для формирования Entity необходимо передать два списка строк, в одном из которых будут располагаться атрибуты из левой части функциональной зависимости, а в другом – из правой. Таким образом, поток входных данных из текстового файла должен быть специальным образом подготовлен: каждая строка разбита на левую и правую часть функциональной зависимости по символу «|», левая и правая части разделены на атрибуты.

Класс SubjectArea способен вмещать полноценное формальное описание предметной области в виде списка функциональных зависимостей. Рассмотрим его листинг:

```
public class SubjectArea
{
    public List<Entity> entities; //список ФЗ
    public SubjectArea(List<Entity> entities)
    {
        this.entities = entities;
    }
    public SubjectArea()
    {
        this.entities = new List<Entity>();
    }
}
```

Данный класс представляет собой, как следует из его определения, не что иное, как список разработанных ранее классов Entity, то есть, фактически, список функциональных зависимостей, что и было необходимо разработать.

Поскольку данные классы были реализованы на основе класса List, то разработанные типы можно описать как списки строго типизированных объектов, доступных по индексу.

Алгоритм сравнения и анализа формальных моделей предметной области предоставляет эксперту прикладной инструмент для анализа и коррекции логической структуры БД. Метод позволяет на основе сравнения семантических зависимостей в работающих БД и составленной экспертом формальной модели предметной области, представленной в виде списка функциональных зависимостей (ФЗ) оценить корректность их структуры. При этом метод должен находить различия в формальных моделях и демонстрировать их в удобной форме, производить их анализ и давать рекомендации по их устранению. Необходимо учитывать, что различия могут появиться из-за того, что БД не соответствует актуальному состоянию предметной области или же при составлении формальной модели могли быть допущены ошибки как методом, так и экспертом. После анализа результатов будет получена уточнённая модель данных, на основе которой можно произвести создание или реинжиниринг БД.

На начальных этапах работы алгоритма производится синтез формального описания на основе анализа семантических зависимостей в базе данных, алгоритм работы которого рассмотрен в [2], и чтение множества функциональных зависимостей из файла, содержащего экспертное описание предметной области. Оба входных потока данных преобразуются в объекты класса SubjectArea, содержащие определённое количество объектов класса Entity.

Результаты проделанной работы позволили создать прикладной механизм представления объектной модели формализованных описаний предметных областей в виде функциональных зависимостей и сущностей, необходимый при реализации методов и алгоритмов работы с предметными областями при синтезе описаний, оптимизации, реинжиниринге и консолидации реляционных баз данных.

### **СПИСОК ИСПОЛЬЗУЕМЫХ ИСТОЧНИКОВ**

1. Баранчиков А.И., Яковлев И.И., Громов А.Ю. Алгоритм сравнительного анализа формальных описаний предметных областей // Известия ТулГУ. Технические науки. 2014. №9-2.

2. Баранчиков А.И., Костров Б.В., Громов А.Ю. Метод и алгоритм формирования семантического описания предметной области на основе схемы реляционной базы данных// Современные проблемы науки и образования. – 2015. – № 1-1.;

3. Баранчикова Е.А. Алгоритм автоматической генерации регулярных выражений (РВ) для спам-фильтра на основе обучающей выборки // Информационные и телекоммуникационные технологии. Материалы 34-ой всероссийской научно-технической конференции. Часть 1. Рязань: РВВКУС. - 2009, с. 380-381

4. К.Дж. Дейт, Хью Дарвен Основы будущих систем баз данных. Третий манифест. Перевод: С.Д.Кузнецов, Т.А.Кузнецова, Издательство Янус-К, 2004. - 656 с.

5. Крэнке Д. Теория и практика построения баз данных, 8-е изд. "Питер", 2003. - 800 с.

# **Информационные технологии**

Межвузовский сборник научных трудов

Компьютерная верстка: А.Н. Сапрыкин, А.О. Сапрыкина

Подписано в печать 27.04.18. Формат бумаги 60x84 1/16.  
Бумага офсетная. Печать струйная. Усл. печ. л. 8,5.  
Тираж 500 экз. Заказ № 1423.

Отпечатано в типографии BookJet

390046, г. Рязань, ул. Пушкина, д. 18,  
тел.: +7 (4912) 466-151, сайт <http://bookjet.ru/>