

**МИНИСТЕРСТВО ОБРАЗОВАНИЯ И НАУКИ
ДОНЕЦКОЙ НАРОДНОЙ РЕСПУБЛИКИ**



ИНФОРМАТИКА И КИБЕРНЕТИКА

4 (18)

Донецк – 2019

УДК 004.3+004.9+004.2+51.7+519.6+519.7

ИНФОРМАТИКА И КИБЕРНЕТИКА, № 4 (18), 2019,
Донецк, ДонНТУ.

Представлены материалы по вопросам приоритетных направлений научно-технического обеспечения в области информатики, кибернетики и вычислительной техники.

Материалы предназначены для специалистов народного хозяйства, ученых, преподавателей, аспирантов и студентов высших учебных заведений.

Редакционная коллегия

Главный редактор: Павлыш В. Н., д.т.н., проф.

Зам. глав. ред.: Мальчева Р. В., к.т.н., доц.

Ответственный секретарь: Воронова А. И.

Члены редакционной коллегии: Аверин Г. В., д.т.н., проф.; Аноприенко А. Я., к.т.н., проф.;

Зинченко Ю. Е., к.т.н., доц.; Зори С. А., д.т.н., доц.; Карабчевский В. В., к.т.н., доц.;

Миненко А. С., д.ф-м.н., проф.; Привалов М. В., к.т.н., доц.; Скобцов Ю. А., д.т.н., проф.;

Федяев О. И., к.т.н., доц.; Шелепов В. Ю., д.ф-м.н., проф.

Рекомендовано к печати ученым советом ГОУ ВПО «Донецкий национальный технический университет» Министерства образования и науки ДНР. Протокол № 9 от 27 декабря 2019.

Свидетельство о регистрации СМИ: серия ААА № 000145 от 20.06.2017.

Приказ МОН ДНР № 135 от 01.02.2019 о включении в Перечень рецензируемых научных изданий ВАК ДНР.

Контактный адрес редакции

ДНР, 83001, г. Донецк, ул. Артема, 58, ГОУ ВПО «ДонНТУ»,

4-й учебный корпус, к. 36., ул. Кобозева, 17.

Тел.: +38 (062) 301-07-35, +38 (071) 334-89-11

Эл. почта: infcyb.donntu@yandex.ru

Интернет: <http://infcyb.donntu.org>

© Донецкий национальный технический университет
Министерство образования и науки ДНР, 2019

СОДЕРЖАНИЕ

Информатика и вычислительная техника

Регрессионная модель для прогнозирования температуры вспышки дизельного топлива в закрытом тигле <i>Максимова А. Ю., Иванова А. А., Лозинский Н. С.</i>	5
Исследование устойчивости стеганографических методов для защиты информации при помощи цифровых водяных знаков <i>Завадская Т. В., Крахмаль М. В.</i>	14
Синтез математической модели управления процессом функционирования железнодорожных переездов на основе новых средств формирования извещений <i>Трунаев А. М., Чепцов М. Н., Радковский С. А.</i>	22
Применение генетического алгоритма для решения задач размещения базовых станций в сетях пятого поколения <i>Павловская К. А.</i>	29
Повышение эффективности методов и средств защиты пользовательских мобильных приложений в операционной системе Android <i>Лебедев В. Е., Чернышова А. В.</i>	35
Системный анализ закономерностей мирового развития компьютерных систем <i>Аноприенко А. Я., Сидоров К. А., Максименко Н. С., Койбаши А. А.</i>	41
Основные принципы и подходы при разработке системы управления профессиональными знаниями сотрудников вуза <i>Андреевская Н. К.</i>	49
Задача оптимизации энергопотребления гетерогенной сетью LTE в условиях крупного города <i>Дзюба А. В., Червинский В. В.</i>	57
Альтернативная реализация циклических несистематических кодов на основе регистров конфигурации Галуа и Фибоначчи <i>Дяченко О. Н.</i>	65
<u>Требования к статьям, направляемым в редакцию научного журнала «Информатика и кибернетика»</u>	75

Информатика и вычислительная техника

Регрессионная модель для прогнозирования температуры вспышки дизельного топлива в закрытом тигле

А. Ю. Максимова*, А. А. Иванова*, Н. С. Лозинский **

*ГУ «Институт прикладной математики и механики», г. Донецк

**ГУ «Институт физико-органической химии и углехимии им. Л.М. Литвиненко», г. Донецк
maximova.alexandra@mail.ru, ivanova.iamm@mail.ru, lozinsky58@mail.ru

Аннотация

Рассматривается задача построения функциональной зависимости температуры вспышки дизельного топлива в закрытом тигле от его измеренных нормативных параметров с использованием подходов машинного обучения. Выполнен анализ сырых данных лаборатории контроля качества методами корреляционного анализа. Разработан метод поиска аномальных объектов, подозрительных на выбросы. Применена гребневая регрессионная модель, построены вероятностные оценки параметров линейной модели. Предложенный подход может быть применен при анализе похожих данных в области нефтехимии.

Введение

Регрессионный анализ позволяет определять скрытые зависимости между информативными показателями качества контролируемого продукта. С другой стороны существуют отрасли, в которых массивы данных контроля качества имеют большой объем и ежедневно накапливаются, например, нефтехимия. Поэтому, если применить регрессионный анализ к обработке этих массивов, можно получить зависимости между информативными показателями, которые позволят уменьшить трудозатраты на проведение контроля качества.

Для дизельного топлива (ДТ) информативными показателями являются: температура вспышки в закрытом тигле (ТВЗТ) и фракционный состав (ФС) [1]. Поскольку эти показатели контролируют свойства самой легкой фракции ДТ, между ними должна существовать взаимосвязь [2]. Более того, в [3, 4] построены регрессионные модели зависимости между ТВЗТ и ФС для некоторых видов дизельного топлива, производимых в Индии и Бразилии. В [5] приведено уравнение регрессии для зависимости ТВЗТ от плотности дизельного топлива с содержанием серы до 0,5 % массы.

Упомянутые модели построены для стандартных ДТ промышленного производства. Однако на практике случается контролировать образцы ДТ подвергнутые смешению, как топливами различных производителей, так и горючими материалами различных марок.

Сведения о работах, посвященных обработке данных контроля реальных образцов дизельного топлива, подверженных технологическим и природным воздействиям, не найдены, поэтому авторами проведен анализ таких данных, полученных в лаборатории контроля качества нефтепродуктов г. Донецка.

Постановка задачи

Необходимо построить регрессионную модель для вычисления ТВЗТ дизельного топлива по измеренным показателям (ФС и плотность ρ) и оценить пределы ее применимости и достоверность полученного результата.

Вопросами определения факторов, которые влияют на значение температуры вспышки, занимаются специалисты в предметной области. В литературе приводятся формулы расчета ТВЗТ для известных органических соединений по температуре их кипения и эмпирическим коэффициентам [6, 7]. Для смесей известных органических соединений также можно посчитать температуру вспышки при условии, что точно известно соотношение компонентов смеси. При этом средняя квадратичная погрешность расчета по приведенным формулам достигает 10-13 °С. Дизельные топлива являются сложными смесевыми продуктами, состав которых зависит от особенностей технологического процесса нефтеперерабатывающего предприятия, на котором они произведены, и от используемых для их приготовления компонентов. Вопрос взаимосвязи между фракционным составом дизельного топлива и ТВЗТ рассматривался в работах [3, 4].

В [5] приводятся функциональные зависимости ТВЗТ от плотности и/или вязкости. Целью данной работы является исследование показателей, по которым может быть рассчитано значение ТВЗТ, и восстановление параметров функции регрессии для вычисления значения ТВЗТ по измеренным значениям этих показателей.

Процедура сбора качественных обучающих данных при решении задач анализа

данных занимает значительную часть времени и требует соответствующих ресурсов. В данном исследовании используется обучающая выборка, собранная в течение 10 лет, состоящая из 1157 необработанных наблюдений, содержащих ошибки и пропуски, что связано с тем, что данные были собраны без предварительного плана сбора исходной статистической информации, и в условиях накопления грубых ошибок, обусловленных, в том числе, влиянием человеческого фактора. Данная выборка несбалансированна в том смысле, что содержит значительно больше примеров стандартных образцов с ТВЗТ большей 40 °С. Возможности провести содержательный анализ условий, при которых регистрировались наблюдения, нет. Принимая во внимание особенности выборки, возникла задача поиска аномальных образцов, подозрительных на выбросы.

Разведочный анализ данных

Выборка с сырыми данными содержала как числовые измеренные показатели: ТВЗТ, плотность при 20 °С, ФС с шагом 10 %, кроме последнего значения, измеренного для 96 %, так и нечисловые данные: идентификатор пробы, источник забора образца, название производителя, время забора пробы. При этом у многих образцов есть пропущенные данные, например, известна плотность, но отсутствует ФС, или наоборот.

Введем обозначения для числовых признаков и целевой переменной регрессионной модели: t_{vzt} – ТВЗТ, t_0 – температура кипения,

$t_0, t_{10}, \dots, t_{80}, t_{96}$ – температуры выкипания фракций, ρ – плотность. При этом каждый из этих признаков определяется с нормируемой сходимостью и воспроизводимостью, которая для ТВЗТ составляет 4 °С, для точек ФС – 2-7 °С.

На первом этапе очистим выборку от образов с пропусками и заведомо недопустимыми значениям (например, $t_0 > 500$). После предобработки выборка X^n включает 324 примера ($n = 324$). Оценим тесноту линейной статистической связи между признаками и целевой переменной t_{vzt} , рассчитав значение коэффициента корреляции по формуле Пирсона:

$$r_{xy} = \frac{\sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum_{i=1}^n (x_i - \bar{x})^2 \sum_{i=1}^n (y_i - \bar{y})^2}},$$

где x – информативный показатель,
 y – целевая переменная t_{vzt} ,
 $\bar{x}, \bar{y}$ – выборочные средние x, y .

Анализ полученных коэффициентов корреляции (табл. 1), рассчитанных как для всей выборки, так и для подвыборки образцов, у которых $t_{vzt} > 40$ в соответствии с требованиями ДСТУ 3868-99 [8], показал возможную сильную корреляцию между температурой кипения и ТВЗТ $r_{t_0 t_{vzt}}^{bce} = 0.77$, $r_{t_0 t_{vzt} > 40}^{bce} = 0.85$, а также возможную среднюю корреляцию между температурой выкипания 10 % фракции и 20 % фракции $r_{t_{10} t_{vzt}}^{bce} = 0.69$, $r_{t_{10} t_{vzt} > 40}^{bce} = 0.73$, $r_{t_{20} t_{vzt}}^{bce} = 0.58$, $r_{t_{20} t_{vzt} > 40}^{bce} = 0.62$.

Таблица 1 – Коэффициенты корреляции

	t0	t10	t20	t30	t40	t50	t60	t70	t80	t90	t96	ro
$t_{vzt} > 40$ °С	0.85	0.73	0.62	0.09	0.05	0.39	0.05	0.25	0.16	-0.01	-0.01	0.23
t_{vzt}	0.77	0.69	0.58	0.5	0.43	0.4	0.31	0.25	0.15	0.04	-0.06	0.29

Визуальный анализ корреляционных полей между ТВЗТ и всеми известными показателями (рис. 1) подтверждает наличие достаточно сильной линейной зависимости ТВЗТ от t_0, t_{10}, t_{20} .

В [5] приводится регрессионная модель зависимости t_{vzt} от плотности дизельного топлива с содержанием серы меньше 0,5 % масс

$$t_{vzt} = 117,012\rho_4^{20} - 53,944.$$

Поэтому несколько неожиданным оказывается визуально слабо подтвержденная линейная связь между t_{vzt} и плотностью ρ . Предположительно это может быть связано с достаточно большим числом нестандартных образцов в исследуемой выборке, которые позволяет определить визуальный анализ корреляционных полей.

В результате проведенного разведочного анализа решено строить регрессионную модель с использованием только предикторов t_0, t_{10}, t_{20} .

Выбор регрессионной модели

Множественная линейная регрессия предполагает линейную зависимость исследуемой случайной величины ξ_{tvzt} от предикторов $(x_1, \dots, x_m)$ при условии, что ее дисперсия σ :

$$\xi_{tvzt} = f(x) + \varepsilon,$$

где $f(x) = w_0 + w_1 x_1 + w_2 x_2 + w_3 x_3$;
 x_1, x_2, x_3 – предикторные переменные, соответствующие признакам t_0, t_{10}, t_{20} ;

w_0, w_1, w_2, w_3 – параметры модели;

ε – стохастическая ошибка, которая обусловлена наличием неизвестных факторов, $E(\varepsilon|x) = 0$.

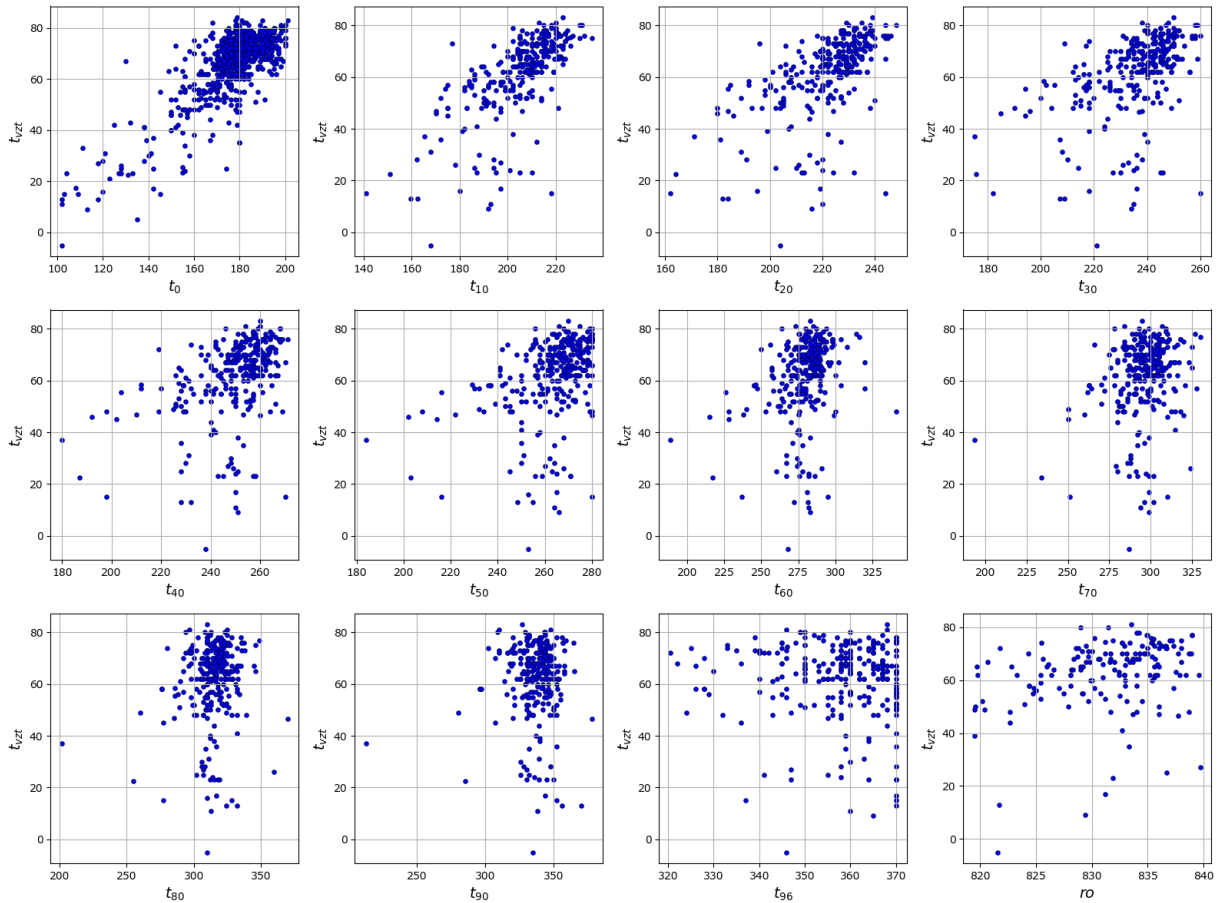


Рисунок 1 – Корреляционные поля полной обучающей выборки

Решим задачу оценки параметров w_0, w_1, w_2, w_3 линейной функции $f(x)$ чтобы вычислить ТВЗТ как

$$y_{tvzt,cp} = w_0 + w_1x_1 + w_2x_2 + w_3x_3.$$

Поскольку было установлено, что

существенно ТВЗТ зависит только от признаков t_0, t_{10}, t_{20} и остальные признаки исключили из рассмотрения, появилась возможность использовать выборку с большим количеством примеров: 324 образца, включая примеры, где данные по плотности отсутствовали в исходном наборе (рис. 2).

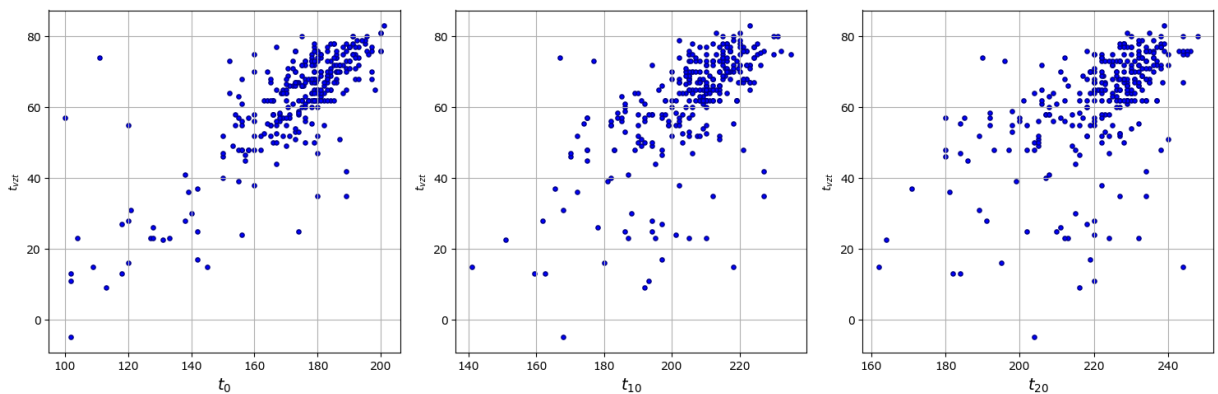


Рисунок 2 – Корреляционные поля

В результате визуального анализа рис. 2 можно предположить, что дисперсия отклонений t_{vzt} от $y_{tvzt,cp}$ одинакова для всех значения предикторов x .

Настройку параметров линейной

регрессионной модели выполним минимизируя квадрат суммы отклонения модельных значений прогнозируемой переменной от реальных значений, содержащихся в обучающей выборке, применив при этом L_2 регуляризацию [9, 10].

Функционал ошибки для данного типа регрессионной модели, которую называют гребневой регрессией, имеет вид:

$$Q(\bar{w}; (\bar{x}_i, y_i)_{i=1}^M) = \sum_{i=1}^M (y_i - f(\bar{x}_i))^2 + \lambda \sum_{j=0}^3 w_j^2 \rightarrow \min_{\bar{w}},$$

где $(\bar{x}_i, y_i)_{i=1}^M$ – обучающая выборка;
 $f(\bar{x}) = w_1 x_1 + w_2 x_2 + w_3 x_3 + w_0$ – функция регрессии;
параметр $\lambda > 0$ – коэффициент регуляризации.

Оценим параметры $\hat{w}_0, \hat{w}_1, \hat{w}_2, \hat{w}_3$ решив уравнение

$$\frac{\partial(Q)}{\partial(w_j)} = 0, j = \overline{0,3}. \quad (1)$$

Полученные в результате решения уравнения (1) параметры $\hat{w}_0, \hat{w}_1, \hat{w}_2, \hat{w}_3$ будут некоторыми оценками настоящих w_0, w_1, w_2, w_3 для исследуемой функциональной зависимости. Далее для удобства оценки коэффициентов будем обозначать просто w_0, w_1, w_2, w_3 .

Чтобы избежать переобучения разобьем имеющуюся в нашем распоряжении выборку X^n на обучающую и контрольную выборки (X^l, X^k) . Контрольную выборку будем использовать только для оценки качества построенной модели. Далее все анализируемые параметры качества модели вычислены по контрольной выборке.

Применим для обучения регрессионной модели метод кросс-валидации по случайным перестановкам, который является базовым инструментом в решении задач машинного обучения. Параметром в данном случае является количество K разбиений выборки на тестовую и обучающую и соотношение количества образцов. Будем использовать разбиение в соотношении 1:3. В качестве результирующей модели выбирается модель с наилучшим значением критерия качества.

Критерием качества модели часто выступает коэффициент детерминации R^2

$$R^2 = 1 - \frac{\sum_{i=1}^M (y_i - \hat{y}_i)^2}{\sum_{i=1}^M (y_i - \bar{y})^2}, \quad (2)$$

где y_i – реальные значения;
 $\hat{y}_i$ – модельные значения;
 $\bar{y}$ – выборочное среднее.

Однако известно, что сравнивать качество модели, опираясь на значение коэффициента детерминации, в случае, если модели построены по разным выборкам, нельзя. Поэтому в некоторых случаях для оценки качества модели будем использовать среднеквадратичную ошибку и визуальный анализ невязок функции

регрессии $\hat{\varepsilon}_i = |y_i - \hat{y}_i|$, а также $E(\hat{\varepsilon})$ и $\sqrt{D(\hat{\varepsilon})}$. Графики для такого визуального анализа будут рассмотрены ниже.

Алгоритм поиска аномалий в данных

Выборка, которая находилась в распоряжении авторов, как уже было отмечено ранее, зашумлена, что подтверждено визуальным анализом корреляционных полей. Сырая выборка содержит некоторые нечисловые признаки, которые могут быть использованы экспертом в предметной области для отбраковки некоторых образцов. Чтобы сократить количество образцов, которые должен рассмотреть эксперт, предложен алгоритм поиска аномальных объектов, подозрительных на выбросы. Алгоритм имеет два параметра: K – количество разбиений исходной выборки на тестовую и обучающую, T – допустимая величина отклонения модельного значения от реального, измеряемая в градусах.

Шаг 1. Разбиваем выборку K – раз на обучающую и тестовую в соотношении 3:1:

$$(X_{train}^{l_j}, X_{test}^{l-l_j})_{j=1}^K. \quad (3)$$

Шаг 2. Строим соответственно K регрессионных моделей по подвыборкам $X_{train}^{l_j}$ из (3).

Шаг 3. Для каждой модели $f_j, j = \overline{1, K}$ вычисляем значение $\hat{y}_i = f_j(x_i)$, $i = \overline{1, n}$ и определяем множество индексов образцов Ind_j , таких что

$$\hat{\varepsilon}_i = |y_i - \hat{y}_i| > T, i = \overline{1, n}.$$

Шаг 4. Вычисляем пересечение множеств индексов подозрительных на выбросы объектов каждой модели

$$Ind_{Outliers} = \bigcap_{j=1}^K Ind_j$$

В отличие от известных алгоритмов поиска выбросов, таких как IsolationForest или LOF [11], в предложенном алгоритме не нужно задавать в качестве параметра степень загрязненности выбросами. Отметим, что параметр K влияет на количество аномалий, с увеличением K их число уменьшается. Параметр T зависит от свойств данных и задается в соответствии с требованиями эксперта. В данной задаче $T = 8^\circ \text{C}$.

Данный алгоритм позволяет определить все выбросы, при условии, что вид восстанавливаемой функциональной зависимости выбран правильно. На рис. 3 изображены результаты работы предложенного алгоритма

для исследуемой выборки. Красным цветом выделены аномалии, подозрительные на выбросы.

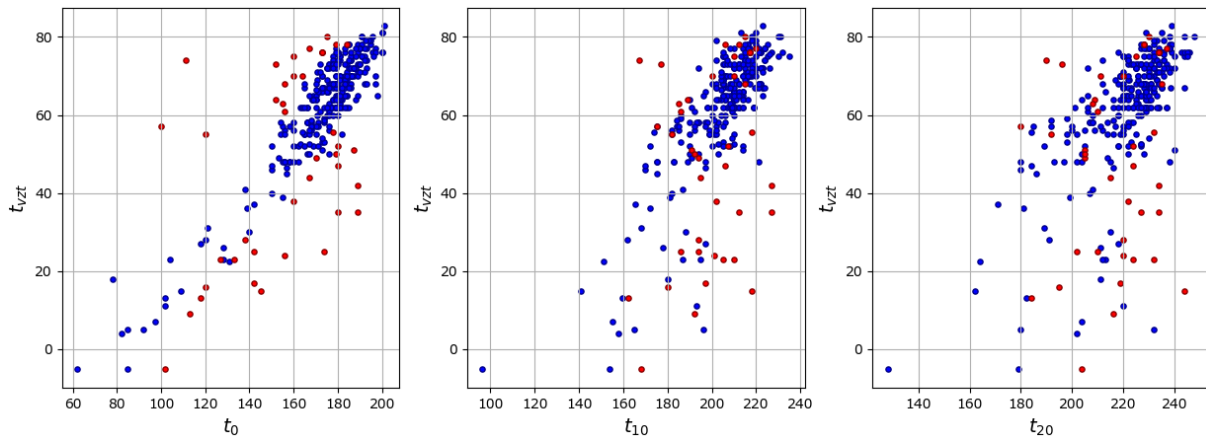


Рисунок 3 – Результат работы метода поиска аномалий

Далее такие аномальные объекты были направлены на исследование экспертам в данной предметной области, чтобы определить, нужно ли их удалять из рассмотрения.

В результате анализа аномальных образцов экспертом была подтверждена необходимость удаления этих объектов из выборки, так как они являлись фальсифицированными или нестандартными продуктами, то есть не являлись дизельным топливом.

Результаты экспериментов

Для решения задачи использовался язык программирования Python 3.5. Регрессионная модель строилась с использованием библиотеки scikit-learn v0.20.

Критерием качества для выбора наилучшей модели является среднеквадратичная ошибка $MSE = \sum_{i=1}^n (y_i - \hat{y}_i)^2$.

Процедура построения регрессионной модели состояла из трех этапов:

- построение модели на полной выборке, не очищенной от выбросов, количество случайных разбиений $K = 8$;
- поиск аномальных объектов, подозрительных на выбросы;
- построение модели на очищенной выборке, количество случайных разбиений 8.

В табл. 2 приведены результаты первого этапа. Модель с минимально ошибкой имеет вид:

$$f(x) = 0.66x_1 + 0.17x_2 - 0.08x_3 - 70.71. \quad (4)$$

Таблица 2 – Параметры регрессионной модели на полной выборке

Эксперимент ($K=8$)	MSE на обучающей выборке	MSE на тестовой выборке	w_0	w_1	w_2	w_3
1	-71.76	-83.11	-77.34	0.62	0.32	-0.16
2	-76.18	-69.64	-79.90	0.62	0.30	-0.12
3	-65.94	-101.56	-75.98	0.67	0.31	-0.19
4	-74.93	-74.04	-72.11	0.66	0.33	-0.22
5	-70.34	-85.79	-71.44	0.64	0.14	-0.02
6	-67.67	-96.71	-71.44	0.63	0.40	-0.27
7	-75.81	-69.26	-70.71	0.66	0.17	-0.08
8	-75.62	-72.30	-77.17	0.62	0.11	0.05
Среднее значение	-72.28	-81.55	-74.51	0.64	0.26	-0.13
Среднеквадратичное отклонение	3.73	11.65	03.27	3,01	0.10	0.10

Значения коэффициента детерминации и средней квадратичной ошибки на отложенной контрольной выборке X^k :

$$R^2(X^k) = 0.47, MSE(X^k) = 124,77. \quad (5)$$

Анализ средних значений и отклонений от среднего для MSE модели на обучении и на тесте по табл. 2 показал нестабильность исходной выборки, что связано с наличием выбросов в ней, а также с ее относительно малым объемом.

На рис. 4 приведены графики невязок построенной модели на обучающих и тестовых данных соответственно. Среднее значение

ошибки $E(\hat{\varepsilon}; X^k) = 6.48^\circ\text{C}$, среднее отклонение ошибки $D(\hat{\varepsilon}; X^k) = 9.27^\circ\text{C}$.

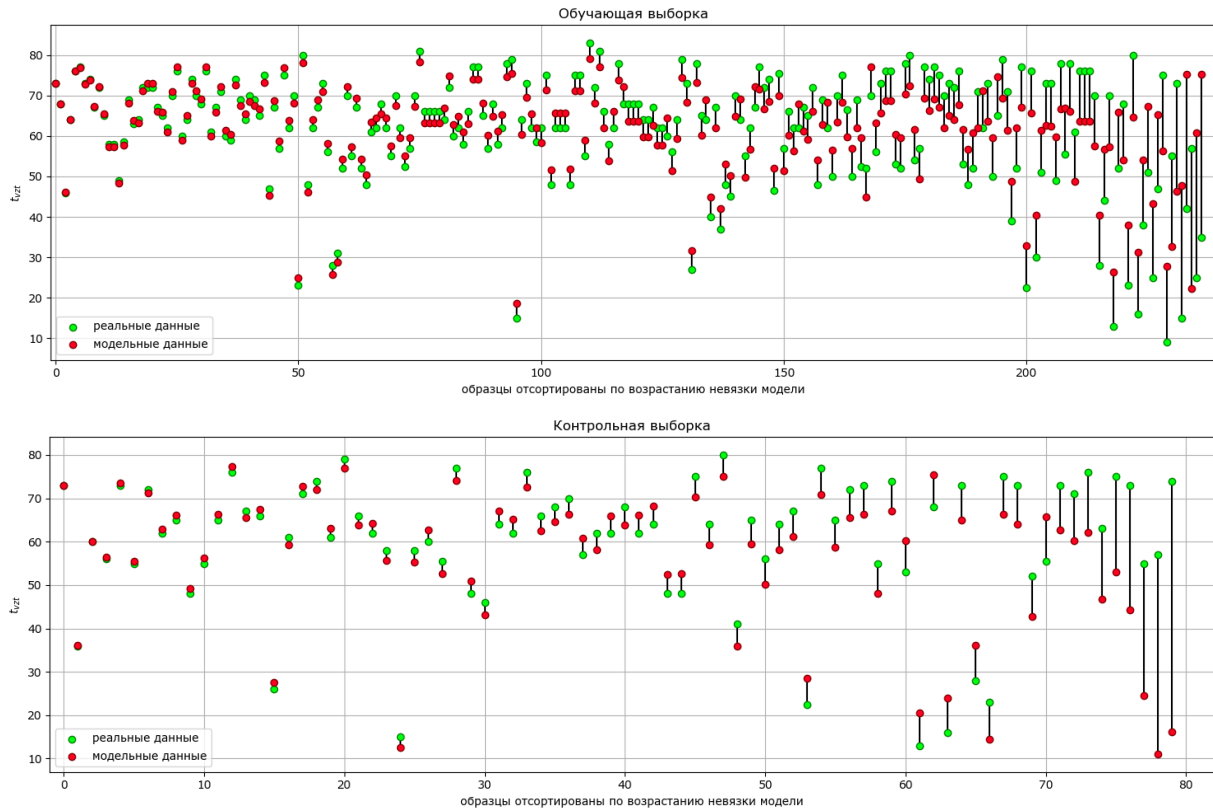


Рисунок 4 – Визуализация качества модели, построенной на полной выборке

Визуальный анализ графиков невязок модели, упорядоченных по возрастанию, позволяет оценить качество модели. Можно видеть, например, что модель ошибается на достаточно небольшом количестве образцов, которые скорее всего являются выбросами.

После выполнения алгоритма поиска аномальных объектов с параметрами $K = 1000$, $T = 8$ отобрано 39 подозрительных на выбросы

объектов, которые, после изучения экспертом, были отброшены (рис. 3).

В табл. 3 приведены результаты третьего этапа. Модель, построенная по очищенной от выбросов выборке с минимальной ошибкой имеет вид:

$$f(x) = 0.66x_1 + 0.19x_2 - 0.08x_3 - 74.29. \quad (6)$$

Таблица 3 – Параметры регрессионной модели на очищенной выборке

Эксперимент	MSE на обучающей выборке	MSE на тестовой выборке	w_0	w_1	w_2	w_3
1	-19.33	-22.41	-69.60	0.64	0.22	-0.11
2	-19.54	-23.14	-81.53	0.67	0.21	-0.07
3	-20.97	-17.08	-74.29	0.66	0.19	-0.08
4	-20.43	-19.35	-75.18	0.63	0.25	-0.10
5	-20.68	-18.22	-76.10	0.66	0.23	-0.11
6	-20.41	-18.88	-72.58	0.64	0.20	-0.08
7	-19.21	-22.54	-76.38	0.64	0.17	-0.03
8	-18.75	-23.92	-75.27	0.67	0.15	-0.05
Среднее значение	-19.92	-20.69	-75.11	0.65	0.20	-0.08
Среднеквадратичное отклонение	0.75	02.42	03.19	0.02	0.03	0.03

Значения коэффициента детерминации и средней квадратичной ошибки на отложенной контрольной выборке X^k :

$$R^2(X^k) = 0.90, MSE(X^k) = 21,25. \quad (7)$$

Анализ средних значений и отклонений от среднего для MSE модели на обучении и на тесте по табл. 3 показал стабилизацию модели. Значения коэффициентов функции регрессии практически не отличаются от средних коэффициентов w_i . На рис. 5 приведены графики невязок построенной модели на обучающих и тестовых данных соответственно.

Среднее значение ошибки $E(\hat{\varepsilon}; X^k) = 3.83^\circ\text{C}$,

среднее отклонение ошибки $D(\hat{\varepsilon}; X^k) = 2.68^\circ\text{C}$, что удовлетворяет требованиям действующих ГОСТов.

Сравним коэффициенты моделей, полученных по полной и по очищенной выборкам. Значения коэффициентов (4), (6) приведены в табл. 4. Как можно видеть, коэффициенты w_1 и w_3 совпадают, а w_0 отличается на 3.42 и w_2 на 0.2. Средние значения коэффициентов обеих моделей также очень близки. Это подтверждает тот факт, что построенная модель приближается некоторую неизвестную нам реальную зависимость, восстановление которого является основной задачей машинного обучения.

Таблица 4 – Коэффициенты моделей

	w_0	w_1	w_2	w_3
Коэффициенты лучшей модели, полная выборка	-70.71	0.66	0.17	-0.08
Средние знач. коэффициентов, полная выборка	-74.51	0.64	0.26	-0.13
Коэффициенты лучшей модели, очищенная выборка	-74.29	0.66	0.19	-0.08
Средние значения коэффициентов, очищенная выборка	-75.11	0.65	0.20	-0.08

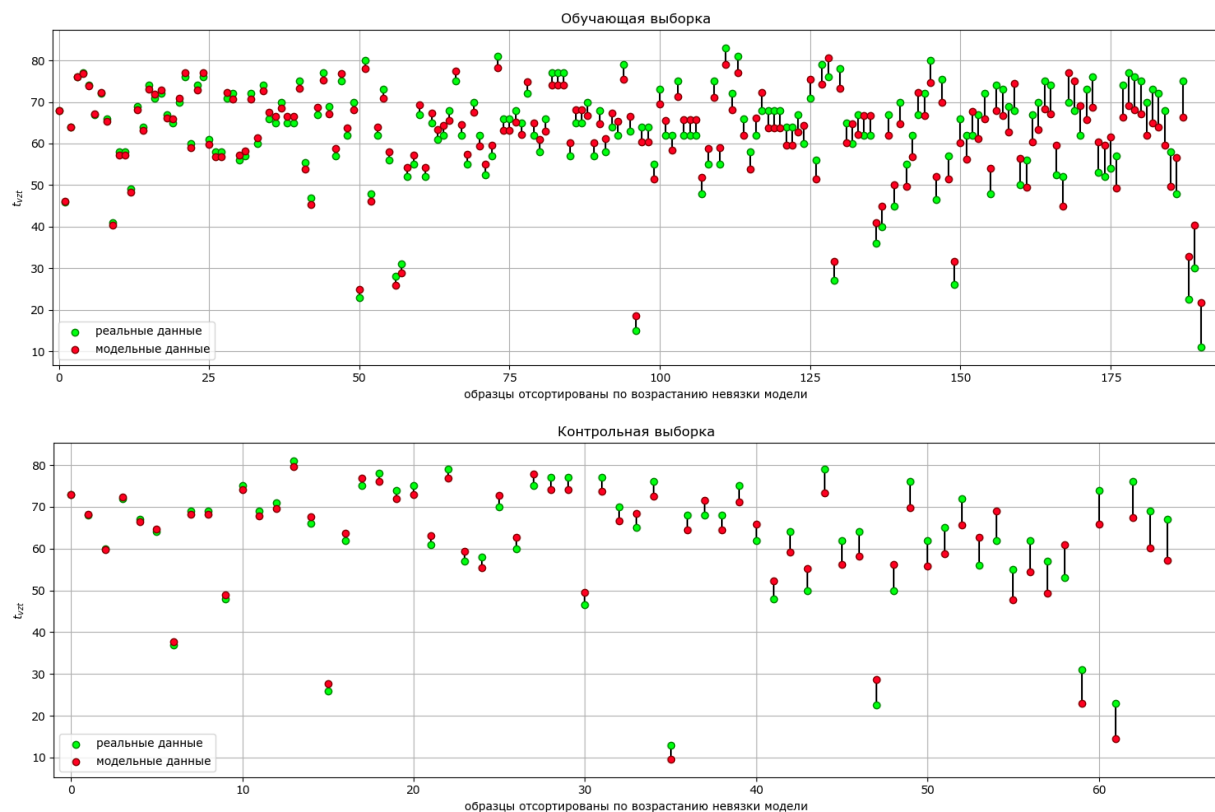


Рисунок 5 – Визуализация качества модели, построенной на очищенной выборке

Вероятностное представление результатов регрессионной модели

Из статистики известно, что коэффициенты модели многомерной линейной регрессии нормально распределены. Когда количество примеров обучающей выборки мало, оценки достоверности регрессионной модели становятся достаточно слабыми. В таких случаях

используют байесовский подход для представления коэффициентов модели как нормально распределенных случайных величин $\beta_0, \beta_1, \beta_2, \beta_3$.

Предлагается оценить коэффициенты w_i нормальным распределением. Так как метод кросс-валидации позволяет строить достаточно большое количество моделей, увеличив

количество случайных разбиений обучающей выборки на различные подвыборки до $K = 1000$, посчитаем оценки математических ожиданий и дисперсий данных величин. На рис. 6

представлены гистограммы коэффициентов w_i и восстановленные по статистически оцененным параметрам графики вероятностных распределений случайных величин $\beta_0, \beta_1, \beta_2, \beta_3$.

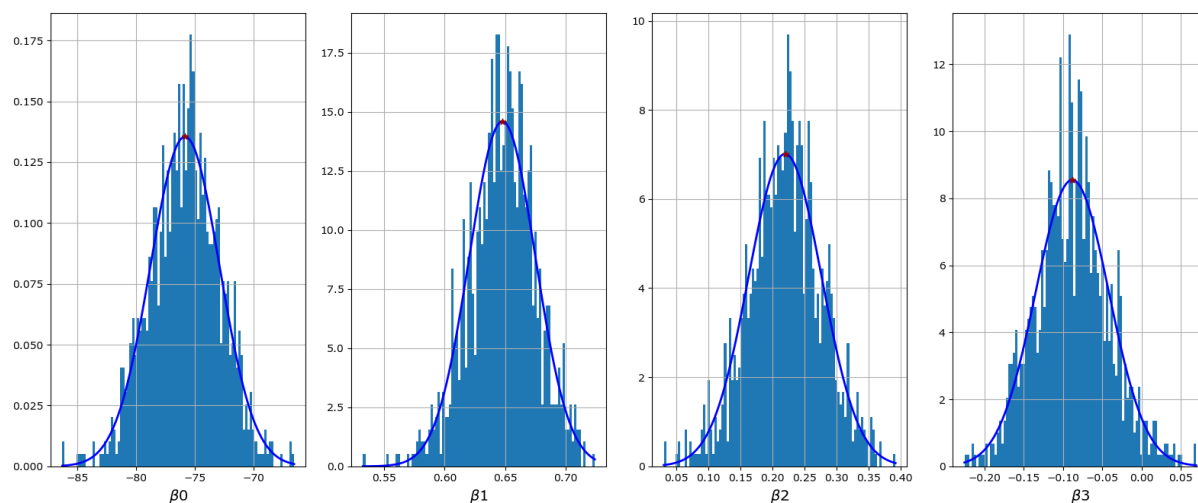


Рисунок 6 – Вероятностные оценки коэффициентов регрессионной модели

Выводы

В данной работе построена многомерная линейная регрессионная модель для вычисления ТВЗТ по температуре кипения 10 % и 20 % фракциям дизельного топлива.

Полученные результаты показали, что тяжелые фракции дизельного топлива не влияют на ТВЗТ, что опровергает результат, полученный в работе [4], где утверждается обратное.

Получить актуальные данные о топливе, которое используется в нашей стране, в настоящее время не представляется возможным. Однако имеющиеся в распоряжении данные за прошлые периоды позволили разработать общий подход к анализу таких данных и процессу построения регрессионной модели. Для усовершенствования модели потребуются новые качественные обучающие данные, которые могут быть получены при затрате определенных ресурсов.

Результат работы может быть использован при анализе похожих данных в области нефтехимии, например, для построения регрессионной модели взаимосвязанных свойств бензинов.

Предложенный метод поиска аномальных объектов показал высокую эффективность при работе с зашумленными данными, содержащими грубые ошибки, обусловленные, в том числе, человеческим фактором. Качество построенной модели полностью удовлетворяет требованиям существующих ГОСТов.

Так как выборка была сильно зашумлена, качество модели несколько хуже, чем в работах, проанализированных бразильскими авторами в [4]. В этой работе использованы чистые образцы,

полученные непосредственно на нефтеперерабатывающих заводах. В нашем случае даже привезенные в лабораторию «чистые» образцы были подвержены смешению с остатками в цистернах для транспортировки или в емкостях хранения. В целом контроль качества нефтепродуктов в нашей стране подвержен ослаблению, что вызвано нежеланием со стороны потенциальных заказчиков проводить мероприятия такого направления, но, с другой стороны, подтверждает актуальность исследований призванных повысить его эффективность, так как топлива высокого качества необходимы для стабильного развития экономики.

Литература

1. ГОСТ 32511-2013 Топливо дизельное ЕВРО. Технические условия [Текст]. Введ. 2015—01—01.— М.: Стандартинформ, 2014. - 20 с.
2. Пучков Н. Г. Дизельные топлива. — М.: Государственное научно-техническое издательство нефтяной и горно-топливной литературы. — 1953. — 194 с.
3. Arankalle A. Significance of flash point in diesel fuel specification. 3rd, International conference on synergy of fuel and automotive technology for a cleaner environment // SAE 2004 India Mobility Conference; India in International conference on synergy of fuel and automotive technology for a cleaner environment. — New Delhi: Allied Publishers. — 2004. — PP. 508-512.
4. Aleme H. G., Barberi P. J. S. Determination of flash point and cetane index in diesel using distillation curves and multivariate calibration // Fuel. — 2012. — V. 102. — PP. 129–134.

5. Александрова И. А. Перегонка и ректификация в нефтепереработке. – М. Химия, 1981 г. – 352 с., ил.
6. Ishida H., Iwarna A. Some Critical Discussions On Flash And Fire Points Of Liquid Fuels // Fire Safety Science 1, 1986: – P. 217-226. doi:10.3801/IAFSS.FSS.1-217
7. Laurent Catoire A Unique Equation to Estimate Flash Points of Selected Pure Liquids Application to the Correction of Probably Erroneous Flash Point Values // Journal of Physical and Chemical Reference Data, Vol. 33, №4, December, 2004. – P. 1083-1112.
8. ДСТУ 3868-99 «Топливо дизельное. Технические условия», 1993.
9. Айвазян С. А., Мхитарян В. С. Теория вероятностей и прикладная статистика. – М.: ЮНИТИ-ДАНА, 2001. – Т.1. – 656 с.
10. Айвазян С. А. Теория вероятностей и прикладная статистика. – Т.2 – М.: ЮНИТИ-ДАНА, 2001. – 432 с.
11. Breunig M. M., Kriegel H. P., Ng R. T., Sander J. LOF: identifying density-based local outliers // ACM sigmod record, 2000. – P. 1-12.

Максимова А. Ю., Иванова А. А., Лозинский Н. С. Регрессионная модель для прогнозирования температуры вспышки дизельного топлива в закрытом тигле. Рассматривается задача построения функциональной зависимости температуры вспышки дизельного топлива в закрытом тигле от его измеренных нормативных параметров с использованием подходов машинного обучения. Выполнен анализ сырых данных лаборатории контроля качества методами корреляционного анализа. Разработан метод поиска аномальных объектов, подозрительных на выбросы. Применена гребневая регрессионная модель, построены вероятностные оценки параметров линейной модели. Предложенный подход может быть применен при анализе похожих данных в области нефтехимии.

Ключевые слова: регрессионная модель, машинное обучение, температура вспышки в закрытом тигле.

Maksimova A., Ivanova A., Losinsky N. A regression model for predicting the flash point of diesel in a closed crucible. The problem of constructing a functional dependence of the flash point of diesel fuel in a closed crucible on its measured normative parameters using machine learning approaches is considered. The analysis of raw data of the laboratory of quality control by the methods of correlation analysis is performed. A method has been developed to search for anomalous objects suspicious of outliers. A ridge regression model is applied, and probabilistic estimates of the parameters of the linear model are constructed. The proposed approach can be applied in the analysis of similar data in the field of petrochemicals.

Keywords: regression model, machine learning, flash point of diesel in a closed crucible.

Статья поступила в редакцию 11.12.2019
Рекомендована к публикации профессором Миненко А. С.

Исследование устойчивости стеганографических методов для защиты информации при помощи цифровых водяных знаков

Т. В. Завадская, М. В. Крахмаль
Донецкий национальный технический университет, г. Донецк
meri_krakhmal@gmail.com

Аннотация

В данной статье рассматривается основная структура стеганографической системы, выполнен обзор и дана классификация методов внедрения скрытой информации в изображения. Приведена математическая модель внедрения цифровых водяных знаков с применением модификации путем устойчивости к различным внешним изменениям. Разработано программное обеспечение рассмотренного алгоритма для скрытия данных в растровых изображениях, показана общая структура выполнения работы. Произведено экспериментальное исследование приложения. Выполнен анализ полученных результатов и предложены пути дальнейших исследований.

Введение

В настоящее время одним из актуальных вопросов является обеспечение безопасности передачи данных по глобальной сети. Данные, которые передаются между пользователями, должны оставаться конфиденциальными. Кроме этого в Интернет должна гарантироваться защита данных от любого рода угроз и атак. Формирование и развитие новых информационных технологий является одной из ключевых характеристик информационных систем.

Основная задача стеганографии – скрыть секретное сообщение и передать его получателю таким образом, чтобы этот факт остался незамечен. Чаще всего в системах безопасности в качестве контейнеров для сообщений используют растровые изображения. При скрытии сообщения необходимо использовать такие методы, которые минимально влияют на контейнер – после внедрения сообщения в контейнер другой пользователь не обнаружит визуальных изменений [1].

Стеганографическая система включает в себя следующие основные компоненты: стеганоконтейнер, секретный ключ и сообщение, оригинальный файл. Существующие системы делятся на те, которым необходимо наличие первоначального изображения (из-за этого передача маркированного изображения становится неудобной и небезопасной), а также те, что требуют наличие секретного ключа для извлечения данных.

Системы первого типа вносят незначительные изменения в стеганоконтейнер, но появляется сложность передачи сообщения из-за необходимости наличия у получателя исходного изображения. Системы второго типа требуют наличия секретного ключа. Данные системы являются наиболее удобными для

передачи секретных данных. Анализ последних исследований и публикаций показал, что наибольшую популярность в компьютерной стеганографии получили методы, использующие в роли контейнера растровые изображения.

Цель и задачи работы

На сегодняшний день существуют надежные методы скрытия информации, минимально изменяющие внешний вид стеганоконтейнера, но имеющие недостатки или недоработки по определенным критериям (скрытие небольшого объема информации, значительные искажения после внедрения или невозможность модифицировать изображения-контейнеры).

Целью данной работы является улучшение эффективности защиты цифровых изображений, как объектов интеллектуальной собственности, от угроз хищения за счет внедрения в них цифровых водяных знаков (ЦВЗ).

Для достижения поставленной цели определены задачи работы: обзор и анализ существующих методов скрытия информации, разработка модели для защиты цифровых файлов, программная реализация и тестирование разработанной модели.

Обзор структуры стеганографической системы

Анализируя процесс развития компьютерной стеганографии можно сделать вывод, что в ближайшие годы интерес к развитию её методов будет увеличиваться всё больше и больше. Актуальность проблемы информационной безопасности постоянно растёт, что стимулирует поиск новых методов и алгоритмов защиты информации. Но с быстрым развитием информационных технологий

появляются различные возможности решения новых методов скрытия данных. Стеганографические методы, как и криптографические, занимают важное место среди методов защиты информации.

В любой стеганографической системе важную роль играет протокол – порядок действий, к которому обращается две или более сторон, с целью решения определенных задач. Обобщенная модель (стеганографическая система) скрытия информации в изображение и её извлечение изображена на рис. 1.

Передача файла от отправителя к получателю начинается с обработки секретной информации, которую необходимо скрыть. Это осуществляется с целью повышения защищенности и устойчивости стеганосистемы к различным искажениям. В большинстве случаев используется секретный ключ, который содержит в себе секретный алгоритм для определения порядка внесения сообщения в контейнер [2].

При скрытии данных размер контейнера должен в несколько раз превышать размер встраиваемых данных, так как это является важным требованием для улучшения маркировки. Также для повышения защищенности секретной информации сообщение можно зашифровать устойчивым криптографическим кодом (стего-ключ).

Скрытая информация заносится в соответствии с ключом в те биты, модификация которых не приведет к существенным искажениям контейнера. Эти биты образуют стеганографический канал (стегоканал) - канал передачи контейнера-результата [3]. Во время пребывания в стегоканале контейнер, содержащий скрытое сообщение, может подвергаться преднамеренным атакам или случайным помехам.

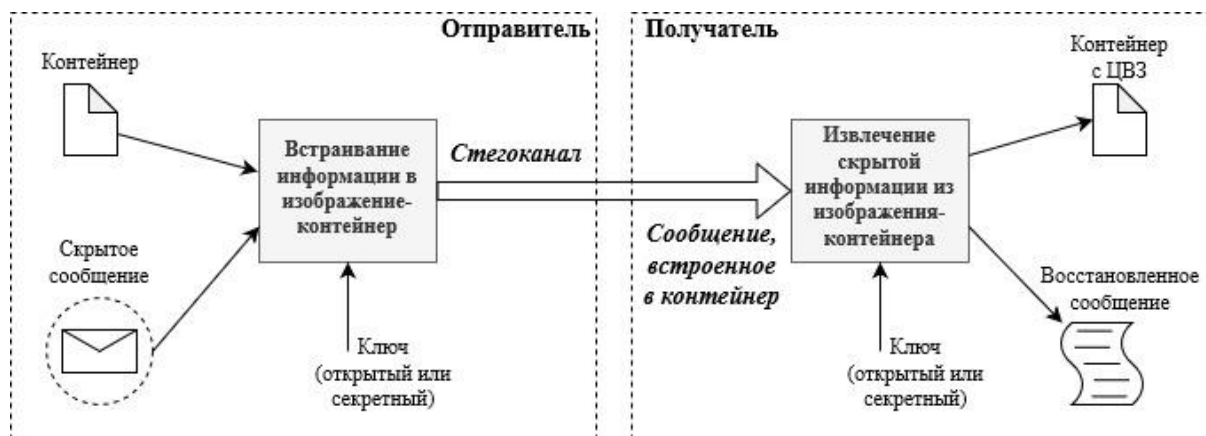


Рисунок 1 – Обобщенная структура стеганографической системы

Анализ существующих методов стеганографии

Одной из проблем, связанных с ЦВЗ, является многообразие требований, которые ставятся перед системой в зависимости от необходимых задач. Для защиты авторских прав необходимо, чтоб водяной знак был стойким к удалению или искажению, при этом объем данных может быть небольшим. Для скрытой связи важным параметром является количество передаваемой информации, а устойчивость к модификациям контейнера менее важна.

К популярным методам скрытия данных в растровых изображениях относятся: метод замены наименее значащего бита (Least Significant Bit – LSB), метод псевдослучайной перестановки, метод псевдослучайного интервала, метод замены палитры, метод блочного скрытия, метод Куттера-Джордана-Боссена (КДБ), метод Дармстедтера-Делейгла-Квисквотера-Макка (ДДКМ), метод Роджена.

В работе [4] приведено сравнение алгоритмов внедрения ЦВЗ в различные области изображения. Показана математическая модель алгоритмов LSB и Куттера-Джордана-Боссена. В [5] рассмотрен метод встраивания ЦВЗ на основе деления изображения-контейнера. Выполнено тестирование алгоритма в системе MATLAB.

Метод замены наименее значимого бита наиболее распространен среди методов замены в растровой области. Младший значимый бит изображения несет в себе меньше информации. Данный метод можно считать шумом и использовать для встраивания информации путем замены менее значимых битов пикселей изображения битами секретного сообщения. Популярность данного метода обусловлена его простотой и тем, что он позволяет скрывать в относительно небольших файлах достаточно большие объемы информации. Основной недостаток – чувствительность к искажениям

контейнера. Для улучшения данного метода часто дополнительно применяют помехоустойчивое кодирование [6].

Метод псевдослучайной перестановки сохраняет последовательности битов исходного сообщения. Он заключается в использовании генератора псевдослучайной последовательности индексов точек изображения $i_1, \dots, i_n$, где в соответствии k -бит сообщения встраивается в i_k -пиксель изображения. Так как у генератора псевдослучайных чисел ограниченный период, то есть возможность повторений индексов, сгенерированных генератором [7].

Метод псевдослучайного интервала заключается в случайном распределении битов текстового сообщения по контейнеру, в результате чего расстояние между двумя встроенными битами определяется псевдослучайно. Эта методика эффективна в случае, когда битовая длина текстового сообщения существенно меньше количества пикселей изображения. Расстояние между двумя соседними битами сообщения имеет случайных характер. Вместимость контейнера и устойчивость к атакам сильно зависят от функции распределения, определяющей расстояние между соседними пикселями. Недостатком является сохранение порядка битов сообщения в контейнере [8].

Метод изменения палитры заключается в том, что палитра цветов изображения сохраняется в виде списка пар индексов, который определяет соответствие между индексом и его вектором цвета. Каждому пикселю изображения ставится в соответствие определенный индекс в таблице. Так как порядок цветов в палитре не важен для восстановления изображения, конфиденциальная информация может быть скрыта путем перестановки цветов в палитре [9].

Метод блочного скрывания заключается в разбиении контейнера на блоки, которые не пересекаются. Данный метод имеет малую устойчивость к модификациям контейнера-изображения. К основным преимуществам можно отнести то, что изменение одного пикселя в блоке не приводит к изменениям статической характеристики контейнера и влияние на контейнер можно уменьшить путем уменьшения размера блока [10].

Метод Дармстедтера-Делейгла-Квисквотера-Макка позволяет достичь компромисса между устойчивостью стегосистемы к искажениям, между качеством встраивания и вычислительной сложностью алгоритма. Метод ДДКМ основан на элементарном восприятии и позволяет осуществлять встраивания относительно текущего содержания блоков контейнера [11].

Метод Рождена заключается в том, что встроенная информация представляет собой двумерную матрицу, состоящую из единиц и нулей. С помощью функции, анализирующей соседние пиксели, распознаются и вычисляются те пиксели, в которые может быть нанесена встроенная информация. Количество таких пикселей не превышает значения 0,01 от общего значения. Поэтому встраивание единиц будет выполнено не во все пиксели изображения. Для увеличения встроенной информации можно перед всеми преобразованиями выполнить небольшое искажение изображения-контейнера (поворот или добавление шума) [12].

Описание и применение метода Куттера-Джордана-Боссена

Куттер, Джордан и Боссен предложили алгоритм (метод «креста») встраивания в канал синего цвета изображения, имеющего RGB кодирование, так как к синему цвету зрительная система человека наименее чувствительна по сравнению с красным и зеленым [13].

На сегодняшний день проводятся исследования данного алгоритма. В работе [14] показан один из возможных способов модифицирования алгоритма путем использования трех составляющих цвета для повышения устойчивости к атакам. В статье [15] предлагается модифицировать этот алгоритм с помощью дополнительных правил для устранения проблем при извлечении данных. Также в статье [16] предложен метод уменьшения изменений в контейнере для повышения надежности алгоритма.

Благодаря своим характеристикам данный метод позволяет передавать сообщение в файлах формата JPEG и устойчив к разрушению младших бит контейнера, что невозможно для более известного метода LSB.

Встраивание информации в изображение-контейнер происходит следующим образом: один i -й бит m_i сообщения в один псевдослучайный пиксель контейнера $p = (x, y)$. Секретный ключ K_0 задает координаты точек, в которые будет внедряться информация.

При встраивании яркость красного и зеленого цветов не меняется, а яркость синего – меняется по формуле:

$$B_{x,y} = \begin{cases} B_{x,y} - v \cdot Y_{x,y}, & s_i = 0; \\ B_{x,y} + v \cdot Y_{x,y}, & s_i = 1; \end{cases} \quad (1)$$

где s_i – i -тый бит сообщения, которое необходимо нанести;

$Y_{x,y}$ – яркость пикселя;

v – коэффициент, задающий энергию встраиваемого биты данных (чем выше значение,

тем выше стойкость внедренной информации, но и тем выше её визуальная заметность).

Одной из проблем извлечения является то, что получатель не имеет оригинального изображения, поэтому точно узнать, как изменилась яркость синего цвета, нет возможности. В данном случае для извлечения информации из необходимого пикселя необходимо спрогнозировать значение яркости исходного цвета на основе яркости соседних элементов по формуле:

$$\overline{B_{x,y}} = \frac{\sum_{i=1}^{\sigma} (B_{x,y+i} + B_{x,y-i} + B_{x+i,y} + B_{x-i,y})}{4\sigma}, \quad (2)$$

где $\sigma = 1 \dots 3$ – размер области, по которой прогнозируется значение яркости.

«Пиксель в центре» – это пиксель, яркость синего цвета которого необходимо спрогнозировать на основании пикселей, которые обозначены светло-серым цветом (рис. 2).

	x-1	X	x+1
y-1			
y			
y+1			

Рисунок 2 – Прогнозирование значения яркости пикселя при $\sigma = 1$

При извлечении скрытого бита вычисляется разница δ между текущим ($B_{x,y}$) и спрогнозированным ($\overline{B_{x,y}}$) значениями интенсивности пикселя $p = (x, y)$ по формуле:

$$\delta = B_{x,y} - \overline{B_{x,y}}, \quad (3)$$

где $m_i = 0$, если $\delta < 0$ и $m_i = 1$, если $\delta > 0$.

Для оценки степени визуальных изменений изображения после внедрения скрытой информации, используется выражение отношения пикового уровня сигнала к шуму (PSNR) по формуле:

$$PSNR = 20 \log_{10} \frac{255xy}{\sum_{x,y} (c_{x,y} - s_{x,y})^2}, \quad (4)$$

где x – ширина изображения;

y – высота изображения;

$c_{x,y}$ – пиксель пустого контейнера с координатами (x, y) ;

$s_{x,y}$ – пиксель контейнера с координатами (x, y) .

Восстановленное изображение с ЦВЗ является приемлемым, если полученное значение $PSNR \geq 28-30$ дБ.

Преимуществом этого метода является высокая пропускная способность, устойчивость к несанкционированному доступу, к разрушению младшего бита контейнера и к атакам сжатия.

Недостатком является то, что извлечение скрытой информации имеет вероятностный характер. Один или несколько соседних пикселей при вычислении могут дать неправильное значение центрального пикселя, в результате чего при извлечении скрытой информации некоторые символы могут отсутствовать. Для уменьшения вероятности ошибки используется помехоустойчивое кодирование. Также возможен вариант многократного встраивания – в процессе маркировки каждый бит необходимо повторять несколько раз.

Цветность растровых изображений определяется с помощью палитры, состоящей из различных цветов.

Расширение палитры с использованием синих оттенков является эффективным способом для скрытия данных, но размер секретного сообщения должен быть достаточно коротким, так как невозможно скрыть большой объем информации в голубых оттенках. Именно по этой причине целесообразней использовать красные, зеленые оттенки вместе с синими, что значительно позволит увеличить размер скрытого сообщения. Добавление трех составляющих цвета позволит скрывать информацию большого размера (рис. 3).

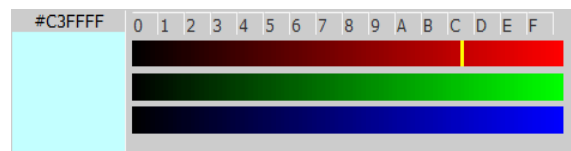


Рисунок 3 – Палитра для подбора RGB цвета

Для программной реализации данного метода были использованы три составляющих цвета (для увеличения объема встраиваемых данных) и многократное внедрение в каждый бит (для более точного извлечения скрытой информации).

Программная реализация внедрения скрытой информации методом КДБ

Разработанная программа осуществляет внедрение и извлечение цифровых водяных знаков, считанных из текстового файла, в изображения, которые могут быть представлены в любом из типов файлов (JPEG, PNG, BMP, TIFF). В результате выполнения маркировки изображение сохраняется в формате несжатого растрового изображения BMP (BitMap Pictures).

После заголовка файла идет цветовая палитра, которая содержит в себе набор данных о пикселях, которые идентифицируют положение каждого пикселя и его цвет. Данный формат поддерживает несколько различных глубин цвета – от черно-белого изображения (1 бит) до глубокого цвета (64 бит). Достоинством формата

является то, что изображение хранится без потерь качества.

Программа делится на два основных модуля: функцию шифрования, которая содержит подфункцию считывания данных из файла, и функцию дешифрования. Общая структура системы представлена на рис. 4.



Рисунок 4 – Обобщенная структура разработанной программы

При выборе функции шифрования работа программы осуществляется следующим образом.

После загрузки и открытия исходного изображения-контейнера необходимо выбрать текстовый файл, содержащий секретную информацию. Если размер контейнера меньше размера скрытой информации, необходимо или увеличить изображение, или уменьшить встраиваемые данные. При появлении данной ошибки будет показано соответствующее сообщение.

Далее выполняется встраивание ЦВЗ в изображение. Во время выполнения данной операции происходит считывание информации из текстового документа и её преобразование в оптимальный для встраивания вид: 1 бит информации преобразовывается в 1-м пикселе изображения с помощью изменения яркости синего канала. Функция определяет размер скрытого сообщения и изображения, в результате чего генерируется два массива (в первом массиве записаны координаты по вертикали, во втором – по горизонтали), которые содержат координаты пикселей для выполнения маркировки. Для удобной работы массивы объединяются в один трехмерный массив.

Также считывается числовое значение, которое обозначает, сколько раз будет встроен бит. Этот параметр влияет на зрительное восприятие – чем больше раз будет перезаписан ЦВЗ в изображение-контейнер, тем более отображаемым становится цифровой знак на результирующем изображении.

Трехмерный массив содержит координаты пикселей, в которые были записаны данные, и является ключом для последующего извлечения встроенной информации. При считывании скрытого сообщения используются сохраненные координаты пикселей по горизонтали и вертикали.

При выборе функции дешифрования изображения выполняются следующие действия.

Для извлечения ЦВЗ из изображения в функцию передается изображение-контейнер со встроенной информацией и массив с координатами пикселей, в которые осуществлялось встраивание.

На каждом этапе прогнозируется яркость синего канала и сравнивается с фактической яркостью путем сравнения соседних пикселей. Так как каждый бит перезаписывается для повышения устойчивости к атакам, в результате получается вектор, состоящий из 2-х цифр (два раза выполняется нанесение бита). Для получения исходного значения бита необходимо найти среднее арифметическое всех элементов данного вектора и округлить.

Извлеченное сообщение может незначительно отличаться от исходного. Данные ошибки появляются в результате прогнозирования яркости пикселей. Для их устранения необходимо увеличить число внедрений одного бита (повторное нанесение информации) и расширить диапазон соседних ячеек при прогнозировании до $\sigma = 3$, что даст более точный результат.

Рассматриваемый в данной работе метод Куттера-Джордана-Боссена был реализован в виде программы на языке программирования C#

в среде разработки Microsoft Visual Studio 2017. Главное окно программы со встроенным ЦВЗ показано на рис. 5.

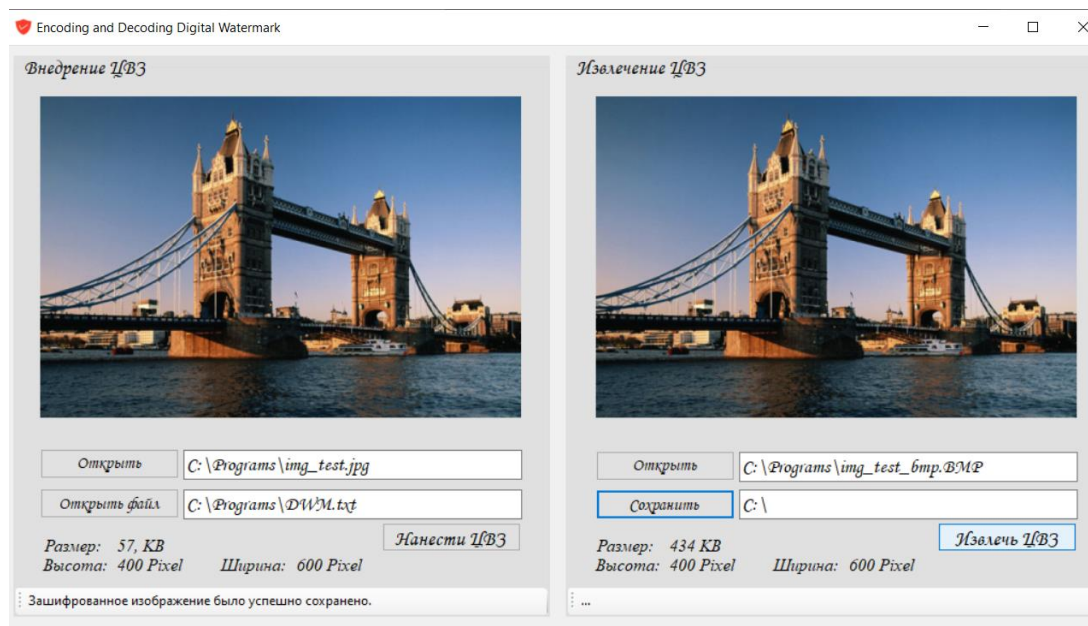


Рисунок 5 – Главное окно программы

Тестирование разработанной программы

Для демонстрации работы данного алгоритма в качестве изображения-контейнера было выбрано растровое изображение формата JPEG, размером 400×600 пикселей. В качестве ЦВЗ был использован текст «Цифровой водяной знак – ЦВЗ», сохраненный в текстовом файле DWM.txt.

При нанесении ЦВЗ на изображение, содержащее встроенную информацию, текущие данные перезаписывались. Если новое скрытое сообщение имеет небольшой размер, а

изображение-контейнер заполнено полностью, то ненужная информация в пикселях очищается и сохраняется только новое сообщение.

Было выполнено исследование алгоритма на устойчивость к таким атакам, как зашумление. В процессе работы контейнер был подвержен данной атаке, интенсивность которой изменялась. После применения сравнивался извлеченный ЦВЗ с исходным, который первоначально был встроен в изображение-контейнер. На рис. 6 показано два варианта зашумления изображения после встраивания в него ЦВЗ.

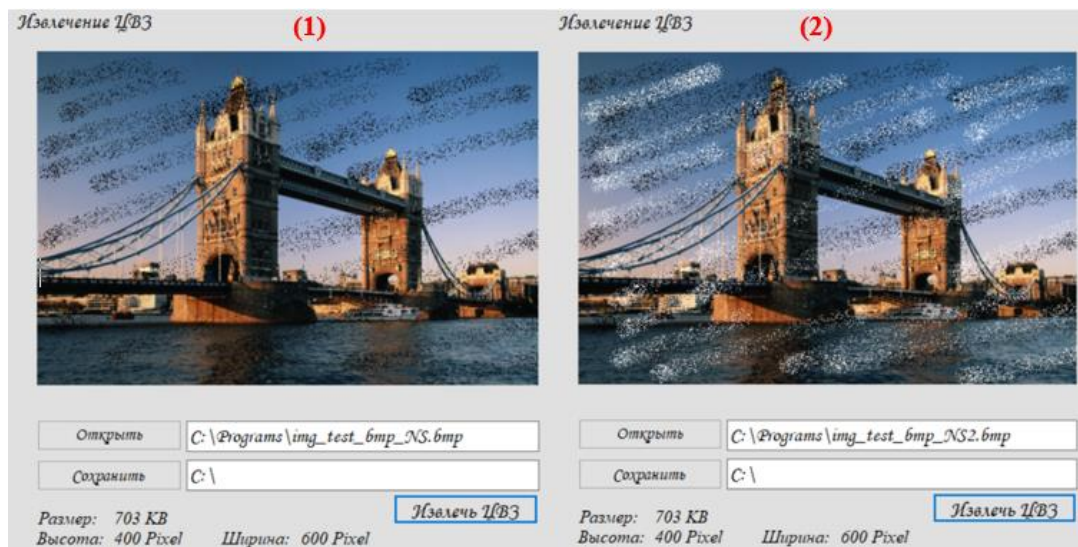


Рисунок 6 – Примеры изображения-контейнера с ЦВЗ после добавления шума

После добавления шума на изображение «1» было выполнено извлечение скрытого ЦВЗ. В результате извлечения ЦВЗ скрытая информация была получена полностью (рис. 7а). Далее было выполнено зашумление большей

части изображения-контейнера, в результате чего полученные данные отличались от исходного ЦВЗ (рис. 7б). Некоторые символы скрытого сообщения были удалены, так как большая часть пикселей синего цвета была изменена.

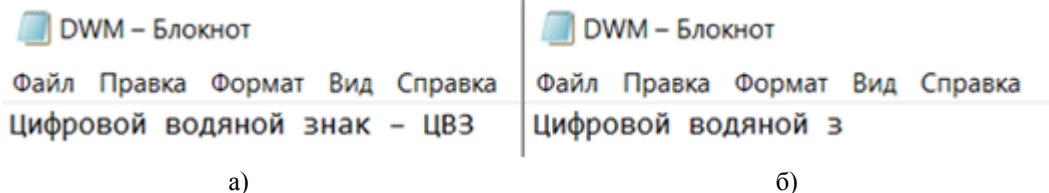


Рисунок 7 – Извлеченные ЦВЗ после добавления шума на изображение-контейнер:
а) допустимая степень зашумления (менее 30 %); б) неприемлемая степень зашумления (более 50 %).

На основе проведенных исследований, можно сделать вывод, что алгоритм КБД позволяет определить скрытую информацию после модификаций контейнера. После значительного искажения изображения-контейнера также можно извлечь секретное сообщение, но некоторые символы не будут отображаться.

Выводы

Целью данной работы являлось создание программного обеспечения для встраивания ЦВЗ на основании стеганографического метода Куттера-Джордана-Боссена. В статье были рассмотрены возможные методы внедрения информации в пространственную область.

Метод КДБ был реализован с двумя модификациями на языке программирования C#. Для исследования алгоритма были произведены эксперименты с добавлением атаки в виде шума, в результате чего выполнялось сравнение полученных результатов. Данный метод является «слепым», так как не требует для извлечения ЦВЗ исходного изображения.

Для улучшения полученных результатов необходимо реализовать устойчивость к таким атакам, как сжатие и масштабирование. Также возможно изменить область встраивания скрытой информации путем применения различных преобразований.

Дальнейшие исследования темы направлены на реализацию алгоритма устойчивого к различным изменениям и искажениям изображения-контейнера.

Литература

1. Коханович Г. Ф. Компьютерная стеганография: теория и практика / Г. Ф. Коханович, А. Ю. Пузыренко. – Киев: МК-Пресс, 2006. – 286 с.
2. Грибунин В. Г. Цифровая стеганография / В. Г. Грибунин, И. Н. Оков, И. В. Туринцев. – М.: СОЛОН-ПРЕСС, 2009. – 277 с.

3. Хорошко В. О. Основы компьютерной стеганографии: Учебное пособие для студентов и аспирантов / В. О. Хорошко, О. Д. Азаров, М. Э. Шелест. – Винница: ВДТУ, 2003. – 143 с.

4. Крахмаль М. В. Исследование методов цифровой стеганографии для маркировки изображений цифровыми водяными знаками / М. В. Крахмаль, Т. В. Завадская // Современные тенденции развития и перспективы внедрения инновационных технологий в машиностроении, образовании и экономике: VI Международная научно-практическая конференция, 13-14 мая 2019, г. Азов.

5. Крахмаль М. В. Исследование метода встраивания цифровых водяных знаков на основе деления изображения-контейнера / М. В. Крахмаль, Т. В. Завадская // 69-я Международная студенческая научно-техническая конференция, Астрахань, 15-19 апреля 2019 года [Электронный ресурс]: материалы / Астрахан. гос. техн. ун-т. — Астрахань: Изд-во АГТУ, 2019. – Режим доступа: <http://astu.org/Content/Page/5833> (дата обращения: 13.09.2019).

6. Fridrich J. Detecting LSB steganography in color and gray-scale images / J. Fridrich, M. Goljan, Rui Du // MultiMedia IEEE. – vol. 8, no. 4. – 2001. – P. 22-28.

7. Chen Yu Long How to Build Pseudorandom Function From Public Random Permutation / Yu Long Chen, E. Lambooi, B. Mennink // IACR Cryptology ePrint Archive. – 2019. – P. 266-293.

8. Bacher A. Generating Random Permutations by CoinTossing: Classical Algorithms, New Analysis, and Modern Implementation / A. Bacher, O. Bodini, H.-K. Hwang, T.-H Tsai // ACM Trans. Algorithms 13, 2, Article 24. – 2017. – 43 p.

9. Fridrich J. A New Steganographic Method for Palette-Based Images / J. Fridrich // Center for Intelligent Systems, SUNY Binghamton. – Binghamton. – NY 13902-6000. – [Электронный ресурс]. – Режим доступа: <https://pdfs.semantic->

scholar.org/576c/3a279223737d45291793cd1bfb019ce0e2a3.pdf (дата обращения: 13.10.2019).

10. Mahjabin T. A block based data hiding method in images using pixel value differencing and LSB substitution method / T. Mahjabin, S. M. Hossain, Md. S. Hague // 15th International Conference on Computer and Information Technology (ICCIT). – 22-24 Dec, 2012. – [Электронный ресурс]. – Режим доступа: <https://www.semanticscholar.org/paper/A-block-based-data-hiding-method-in-images-using-Mahjabin-Hossain/9a2bd47e498a3b699ee808fcfa045f9826280a8c#paper-header> (дата обращения: 13.10.2019).

11. Darmstaedter V. Low Cost Spatial Watermarking / V. Darmstaedter, J. F. Delaigle, J. J. Quisquater, B. Macq // Computers and Graphics, 1998. – Vol. 5. – P. 417-423.

12. Fkirin A. Steganography Literature Survey, Classification and Comparative Study / A Fkirin, G. Attiya, A. El-Sayed // Communications on Applied Electronics. – 5(10). – September 2016. – P. 13-22.

13. Kutter M. Digital Signature of color image using amplitude modulation / M. Kutter, F. Jordan, F. Bossen // Proc. of the SPIE Storage and Re-trieval for Image and Video Databases. – 1997. – Vol. 3022. P. 518-526.

14. Защелкин К. Усовершенствование метода скрытия данных Куттера-Джордана-Боссена / К. В. Защелкин // МНПК «Современные информационные и электронные технологии». – 2013. – С. 214-216.

15. Дизер А. Модификация метода Куттера-Джордана-Боссена скрытого хранения информации в изображениях формата JPEG / А. Е. Дизер, Е. С. Дизер, Т. М. Опарина // Математические структуры и моделирование. – 2016. – №3(39). – С. 177-183.

16. Фомин Д. Модификация метода скрытия информации Куттера-Джордана-Боссена / Д. В. Фомин // Вестник Амурского государственного университета. – Выпуск 65, Серия: естественные и экономические науки. – 2014. – С. 58-62.

Завадская Т. В., Крахмаль М. В. Исследование устойчивости стеганографических методов для защиты информации при помощи цифровых водяных знаков. В данной статье рассматривается основная структура стеганографической системы, выполнен обзор и дана классификация методов внедрения скрытой информации в изображения. Приведена математическая модель внедрения цифровых водяных знаков с применением модификации путем устойчивости к различным внешним изменениям. Разработано программное обеспечение рассмотренного алгоритма для скрытия данных в растровых изображениях, показана общая структура выполнения работы. Произведено экспериментальное исследование приложения. Выполнен анализ полученных результатов и предложены пути дальнейших исследований.

Ключевые слова: стеганография, информационная безопасность, встраивание информации, цифровые водяные знаки, скрытие данных.

Zavadskaya T., Krakhmal M. Investigation of the stability of steganographic methods for protecting information using digital watermarks. This article is considered the basic structure of a steganographic system, reviews and provides a classification of methods for implementation hidden information into images. A mathematical model of the implementation of digital watermarks using modifications by means of resistance to various external changes is presented. The software of the considered algorithm for hiding data in raster images is developed, the general structure of the work is shown. An experimental study of the application has been performed. The analysis of the results received has been completed and ways of further research is proposed.

Keywords: steganography, information security, embedding information, digital watermarks, data hiding.

Статья поступила в редакцию 29.10.2019
Рекомендована к публикации профессором Шепелевым В. Ю.

Синтез математической модели управления процессом функционирования железнодорожных переездов на основе новых способов формирования извещений

А. М. Трунаев, М. Н. Чепцов, С. А. Радковский
Донецкий институт железнодорожного транспорта, г. Донецк
davidoff-a@mail.ru

Аннотация

В статье на основе анализа способов формирования подачи извещения на переезд определено, что они не учитывают реальную скорость и местоположения поезда. Из-за чего увеличивается время простоя автотранспорта. Как следствие нервозность водителей и проезд переездных заградительных устройств, при их закрытом состоянии. Авторами разработана математическая модель управления процессом функционирования железнодорожных переездов, которая учитывает местоположение и скорость поезда. Проведено численное моделирование предложенной модели на основании реальных данных, определено, что простой автотранспорта сократится в среднем в 2-5 раз.

Введение

Пересечения железнодорожных путей и автомобильных дорог в одном уровне приводят к задержкам движения автотранспорта и характеризуются продолжительными непроизводительными простоями. В свою очередь это может спровоцировать водителей на нарушение правил проезда переезда и как следствие возникновение аварий, в том числе с особо тяжкими последствиями. В настоящее время в 98 % случаев аварий на переездах происходит по вине водителей [1]. Так как время включения заградительной сигнализации для автотранспорта зависит от максимально разрешенной скорости движения поездов по участку извещения, то переезды закрываются заблаговременно без учета реальной скорости следования подвижного состава. В большинстве случаев реальная скорость ниже максимально разрешенной, из-за чего увеличивает время простоя автотранспорта. Как следствие нервозность водителей и проезд переездных заградительных устройств, при их закрытом состоянии.

Согласно требованиям безопасности и методам контроля автоматики и телемеханики на железнодорожных переездах, формирование извещения на переезд определяется строго оговоренными условиями [2]. При этом, ни одно из условий не учитывает реальную скорость движения подвижной единицы по участку извещения к переезду.

По динамике известных значений технической скорости железнодорожного транспорта определено, что средняя скорость по России на сегодняшний день составляет 47 км/ч [3]. Разрешенная максимальная скорость, на большинстве участков составляет 90-120 км/ч.

Эта скорость учитывается при расчете длины участка извещения приближения поезда к переезду. Определяется ордината (точка извещения), при проследовании которой, подается сигнал на переезд о включении заградительных устройств [4]. При движении поезда с меньшей скоростью, увеличивается время нахождения поезда на участке извещения. Что приводит к увеличению времени простоя автотранспорта перед заградительными устройствами до тех пор пока поезд не покинет участок извещения.

Таким образом, существующие способы формирования извещения не являются универсальными и гибкими. А задача формирования извещения с учетом реальных скоростей и местоположения поезда, для систем автоматического управления процессом функционирования железнодорожных переездов является весьма актуальной.

Цель статьи

Разработать математическую модель управления процессом функционирования железнодорожных переездов на основе новых средств формирования извещений, учитывающую местоположение и скорость поезда.

Основная часть

При движении поезда возникают вынужденные колебания рельсовой линии [5], которые можно использовать как источник информации о приближающемся поезде. В работе [6] автором были разработаны алгоритмы и устройства системы оповещения о приближении поезда по виброакустическим колебаниям рельса. Проанализировав данные

исследования, в качестве контролируемого параметра местоположения подвижной единицы на участке предлагается использовать реакцию рельсовой линии, которая фиксируется с помощью акселерометра прикрепленного к боковой поверхности, либо подошве рельса. В работе [7] авторы промоделировали колебания железнодорожного рельса при воздействии на него подвижной вертикальной динамической нагрузки. Было определено, что существует амплитудно-частотная зависимость величины виброускорения рельсовой линии от расстояния между источником и местом съема колебаний рельсовой линии. В статье [8] был описан эксперимент отклика рельсовой линии на единичное импульсное воздействие, который подтвердил результаты моделирования.

Результаты выше изложенных исследований применим для формирования информации о приближающемся поезде к переезду по вынужденным колебаниям рельсовой линии. А именно по реакции рельсовой линии определим местоположение поезда на участке. По изменению расстояния определим его скорость. Что позволит определить время приближения поезда к переезду. Как только это время станет меньше типового времени закрытия заградительных устройств (рассчитывается индивидуально для каждого переезда), включаются заградительные устройства.

Разработаем структурно-функциональную модель управления переездом на основе нового метода формирования извещения (рис. 1).

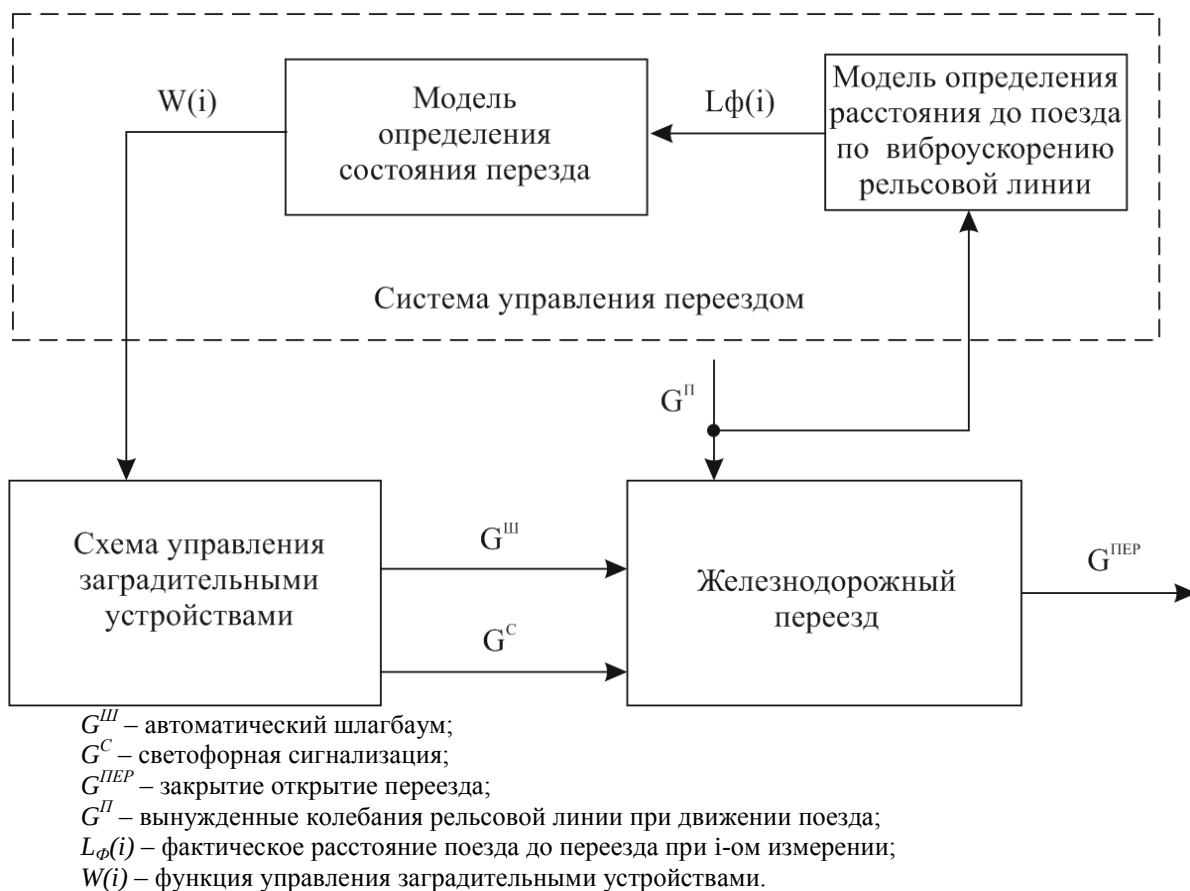


Рисунок 1 – Структурно-функциональная модель управления переездом

При движении поезда по участку извещения, возникают вынужденные колебания рельсовой линии G^{II} , которые обрабатываются моделью определения расстояния до поезда $L\phi(i)$ по виброускорению рельсовой линии. Каждый цикл измерений обрабатывается моделью, которая решает, в какой момент времени включить заградительные устройства $W(i)$.

Для структурно-функциональной модели разработаем алгоритм процесса функционирования

заградительных устройств переезда, который изображен на рис. 2.

Рассмотрим подробно алгоритм работы заградительных устройств, учитывающий время приближения поезда к переезду в реальном масштабе времени.

Исходное состояние заградительных устройств переезда, в исправном состоянии, открытое [9].

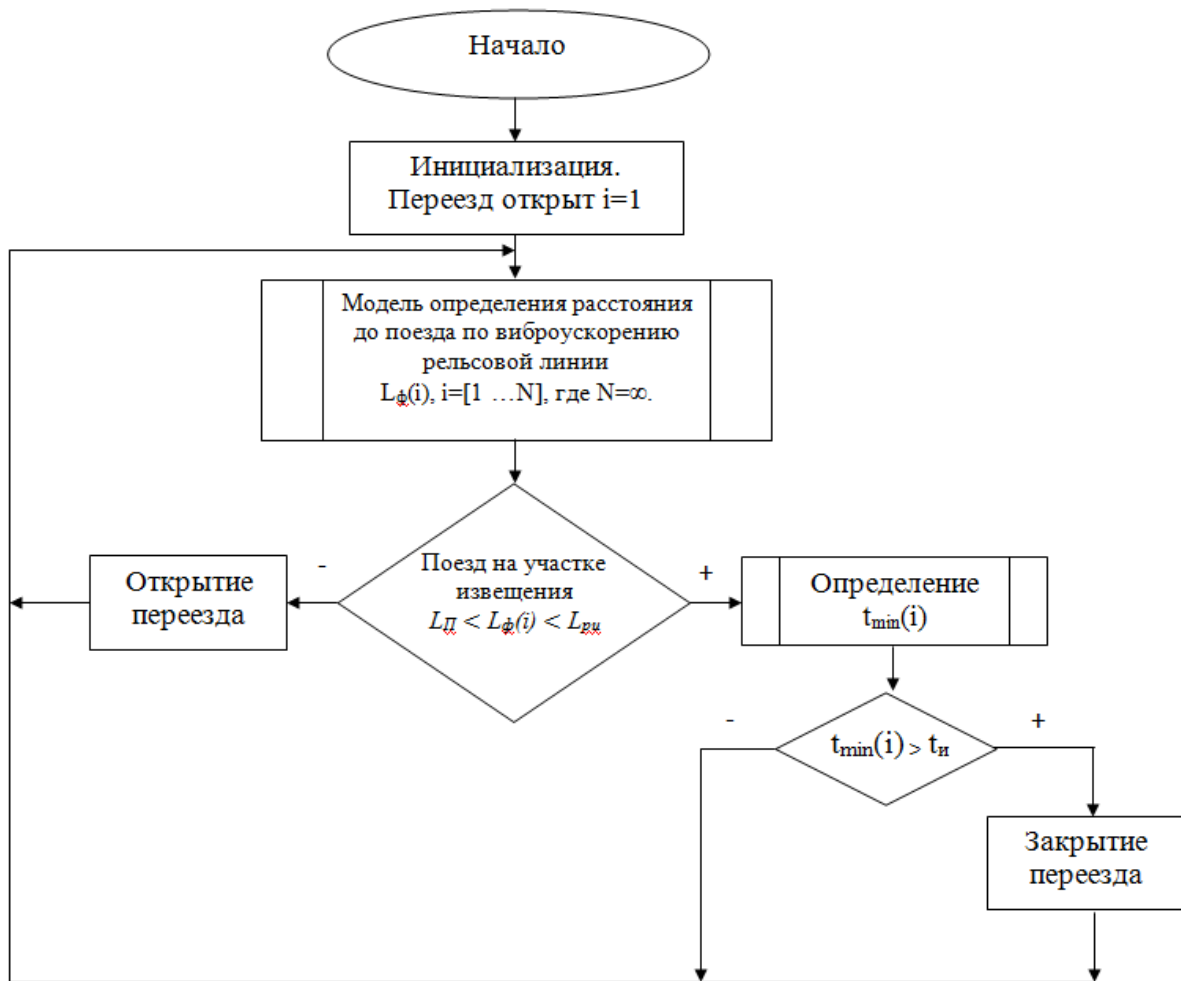


Рисунок 2 – Алгоритм управления заградительными устройствами переезда

Модель обработки виброускорения в режиме реального времени анализирует вынужденные колебания рельсовой линии. По амплитудно-частотным зависимостям определяет расстояние до поезда (L_ϕ – фактическое расстояние от ординаты переезда подвижного объекта). Сравнивает с $L_{пу}$ – расчетным расстоянием извещения (для каждого переезда индивидуально).

При условии $L_\phi < L_{пу}$, рассчитывается t_{min} – минимальное время движения поезда до переезда при максимально возможной скорости и ускорении.

Полученное время сравнивается с t_u – const, расчетное время закрытия заградительных устройств (для каждого переезда индивидуально)

Как только условие $t_{min} < t_u$ верно, переезд закрывается.

Открытие произойдет, как только поезд покинет участок извещения, т. е. условие $L_\phi < L_\phi < L_{пу}$ не выполняется.

Допустим, поезд движется по участку в соответствии с кривой скорости, изображенной на рис. 3.

В режиме реального времени по виброускорению рельсовой линии определяются

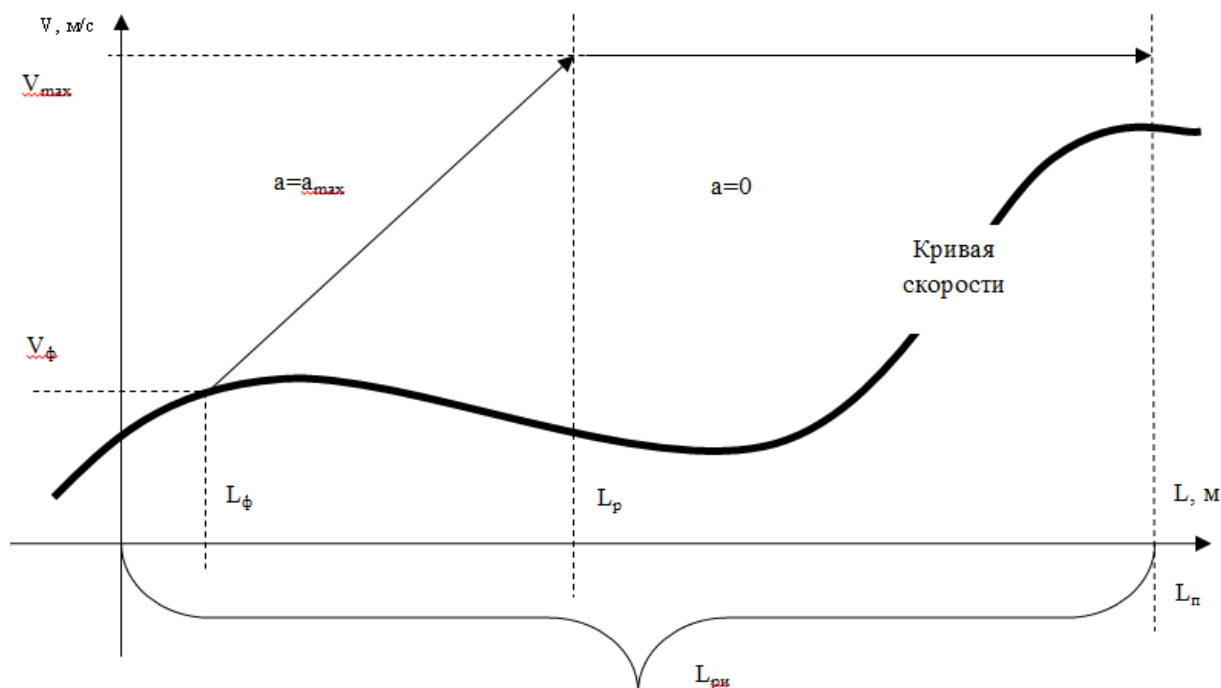
значения $V_\phi(i)$ и $L_\phi(i)$, для конкретного перегона и переезда значения: V_{max} , – максимальная скорость на участке, $L_{пу}$, – длина участка извещения, a_{max} – максимально ускорение подвижного единицы, известны и являются постоянными величинами. Как только, поезд въезжает на участок извещения $L_{пу} > L_\phi$, определяется t_{min} минимальное время приближения поезда к переезду, при максимально допустимых параметрах движения для текущей скорости и координаты.

Рассмотрим более детально этот процесс. Минимальное время приближения поезда до ординаты переезда при максимальных параметрах движения определяется:

$$t_{min}(i) = t_1(i) + t_2(i) \quad (1)$$

где $t_1(i)$ – время, которое проедет поезд от точки $L_\phi(i)$ до точки $L_p(i)$, при максимальном ускорении a_{max} , с текущей скорости до максимально разрешенной скорости по участку V_{max} ;

$t_2(i)$ время равномерного движения поезда с $L_p(i)$ до L_ϕ , при движении по участку с максимально разрешенной скоростью.



$L_{\phi}(i)$ – координата поезда;
 $V_{\phi}(i)$ – скорость поезда;
 $L_p(i)$ – определяется при известных значениях $V_{\phi}(i)$ и $L_{\phi}(i)$;
 V_{max} – максимально разрешенная скорость поезда по перегону (для каждого перегона индивидуальна);
 L_n – координата переезда;
 a – ускорение поезда;
 a_{max} – максимально возможное ускорение поезда;
 $L_{ри}$ – расчетное расстояние извещения (для каждого переезда индивидуально)

Рисунок 3 – График определения минимального времени приближения к переезду

Известно, что время, за которое поезд увеличит скорость с $V_{\phi}(i)$ до V_{max} при максимальном ускорении a_{max} определяется по формуле [10]:

$$t_1(i) = \frac{V_{max} - V_{\phi}(i)}{a_{max}}. \quad (2)$$

Получив значение времени t_1 , определим расстояние $L_{\phi}L_p(i)$, которое проедет поезд по формуле [9]:

$$L_{\phi}L_p(i) = V_{\phi}(i) + \frac{a_{max} \cdot t_1^2(i)}{2}. \quad (3)$$

Получив значение отрезка $L_{\phi}L_p(i)$ и зная фактическое расстояние до переезда $L_{\phi}L_n(i)$ определим $L_pL_n(i)$:

$$L_pL_n(i) = L_{\phi}L_n(i) - L_{\phi}L_p(i). \quad (4)$$

При максимально разрешенной скорости движения по участку $L_pL_n(i)$ рассчитывается время $t_2(i)$:

$$t_2(i) = \frac{L_pL_n(i)}{V_{max}}. \quad (5)$$

Подставив формулы (2-5) в (1) получим уравнение расчета минимального времени приближения текущего поезда к переезду для i -го измерения:

$$t_{min}(i) = \frac{V_{max} - V_{\phi}(i)}{a_{max}} + \frac{L_pL_n(i) - V_{\phi}(i) \cdot \left(\frac{V_{max} - V_{\phi}(i)}{a_{max}} \right) + \frac{a_{max} \cdot \left(\frac{V_{max} - V_{\phi}(i)}{a_{max}} \right)^2}{2}}{V_{max}} \quad (6)$$

Уравнение (6) является конечным в определении минимального времени движения к переезду. Полученное значение t_{min} необходимо сравнить с расчетным временем закрытия заградительных устройств t_u .

Как только это условие $t_{min} < t_u$ выполняется, включаются заградительные устройства (переезд зарывается).

На основании разработанной структурной схемы, алгоритма работы переезда и уравнения расчета минимального времени приближения текущего поезда к переезду, разработаем математическую модель функционирования заградительных устройств переезда, которая учитывает текущую координату и скорость движения поезда (7).

$$\left\{ \begin{aligned} t_{\min}(i) &= \frac{V_{\max} - V_{\phi}(i)}{a_{\max}} + \frac{L_p L_n(i) - V_{\phi}(i) * \left(\frac{V_{\max} - V_{\phi}(i)}{a_{\max}} \right) + \frac{a_{\max} * (V_{\max} - V_{\phi}(i))^2}{2}}{V_{\max}} \\ W(i) &= \begin{cases} 1, \text{ если } t_i > t_{\min}(i), L_{\phi}(i) < L_{\text{ри}}; \\ 0, \text{ если } L_{\Pi} > L_{\phi}(i) > L_{\text{ри}}; \end{cases} \\ V_{\phi}(i) &= \frac{L_{\phi}(i) - L_{\phi}(i-1)}{\Delta t} \end{aligned} \right. \quad (7)$$

где $L_{\phi}(i), L_p(i)$ – координата поезда;
 L_{Π} – координата переезда;
 $V_{\max}$ – максимально разрешенная скорость поезда по перегону;
 $a_{\max}$ – максимально возможное ускорение поезда;
 $V_{\phi}(i), V_p(i)$ – скорость поезда;
 $t_u = \text{const}$, время закрытия заградительных устройств (для каждого переезда индивидуально);
 $i = [1 \dots N]$, где $N = \infty$.

Для определения функционирования математической модели заградительными

устройствами по выражению (7), промоделируем её с реальными значениями, применяемыми на железных дорогах: $V_{\max} = 120 \text{ км/ч}$; $a_{\max} = 0.7 \text{ м/с}^2$; $t_u = 58 \text{ с}$.

Произведем расчет минимального времени движения поезда до переезда, для всего участка приближения от 0 до 2000 м с шагом 100 м и всех возможных скоростей от 0 до 120 км/ч с шагом 12 км/ч.

Получена кривая зависимости реальных скоростей и расстояний (рис. 4). Определены зависимости граничных значений скоростей и расстояний при $t_{\min} = t_u$.

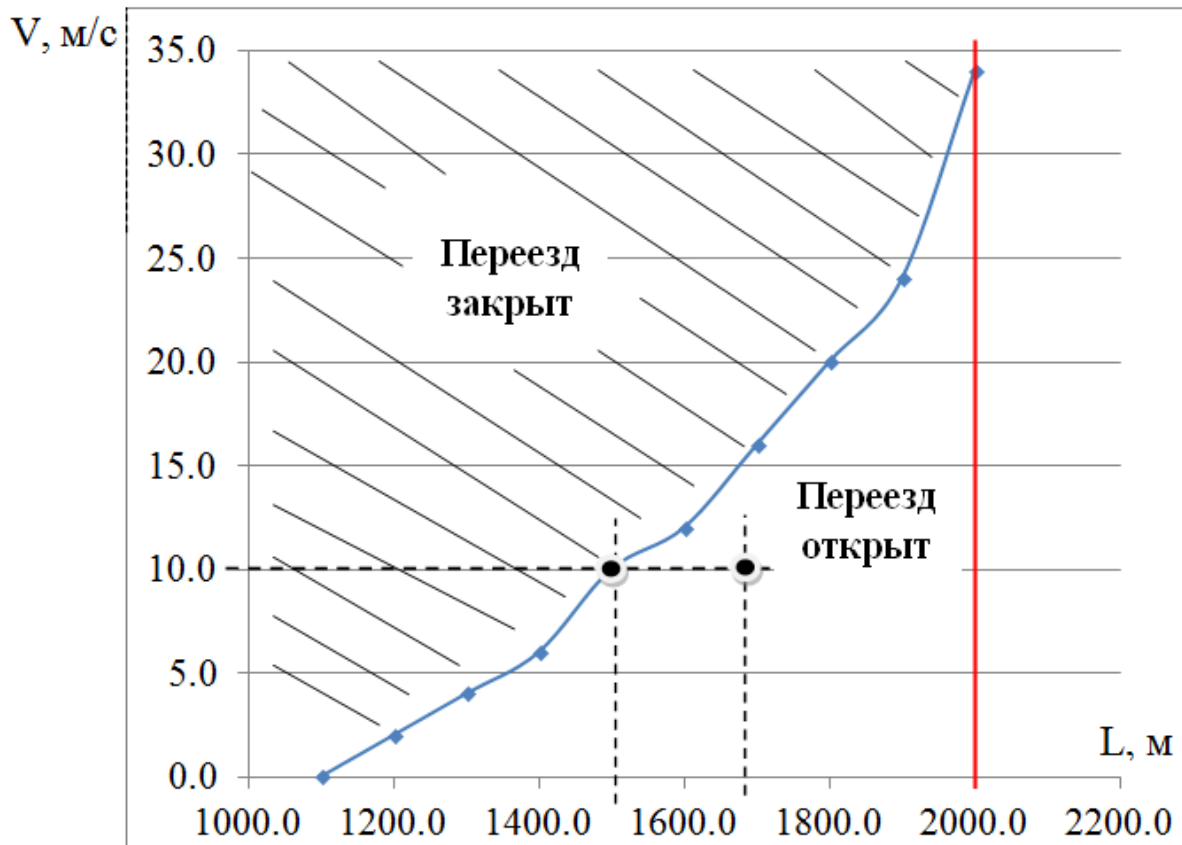


Рисунок 4 – График граничных значений скоростей от расстояний при $t_u = \text{const}$, для предлагаемой математической модели функционирования заградительными устройствами переезда.

Из графика на рис. 4 можно определить, например, что при приближении поезда к переезду со скоростью в 10 м/с, на расстоянии 1700 м, переездные устройства останутся открытыми. А при движении этого же поезда с равномерной скоростью и при пересечении точки 1500 м, включатся переездные

заградительные устройства. И останутся закрытыми до тех пор пока поезд не покинет участок извещения (требования безопасности работы заградительных устройств [2]).

Для сравнения предложенной модели функционирования заградительных устройств с существующими системами, рассчитаем

граничные значения скоростей от расстояний для существующих систем подачи извещений, при тех же реальных значениях. На рис. 4 изображена вертикальная прямая в точке 2000 м, которая является границей включения заградительных устройств. Т. е. при нахождении поезда в точке пути больше 2000 м переезд останется открытым.

Из выше изложенного можно сделать вывод, что модель предусматривает полную остановку поезда на некоторой части участка извещения, при которой переезд останется открытым. Что является недопустимым в существующих системах.

По динамике известных значений технической скорости железнодорожного транспорта определено, что средняя скорость по России на сегодняшний день составляет 47 км/ч. А разрешенная максимальная скорость, на большинстве участков составляет 120 км/ч. Эта скорость учитывается при расчете времени приближения поезда в существующих системах.

Учитывая динамику значений технической скорости железнодорожного транспорта простой автотранспорта сократится в среднем в 2-5 раз по сравнению с существующими системами извещения о приближении поезда.

Выводы

В работе разработана модель управления процессом функционирования железнодорожными переездами на основе новых методов формирования извещения. В режиме реального времени вычисляется время приближения поезда к ординате переезда с учетом текущего местоположения и скорости. Как только время закрытия превысит время приближения, включатся заградительные устройства. Что позволит минимизировать простой автотранспорта.

В результате численного моделирования времени приближения поезда к переезду на основании реальных данных, определено, что простой автотранспорта сократится в среднем в 2-5 раз.

Из этого следует, что предложенная модель управления процессом функционирования железнодорожными переездами является актуальной.

Литература

1. Федухин А. В. Информационный подход к повышению безопасности движения на железнодорожных переездах / А. В. Федухин,

Ар. А. Муха // Математические машины и системы, 2015. – № 4. – С. 145 – 151.

2. ГОСТ 33893-2016. Системы железнодорожной автоматики и телемеханики на железнодорожных переездах. Требования безопасности и методы контроля.

3. Хусаинов Ф. И. Показатели скорости как аналитические инструменты для оценки работы железных дорог / Ф. И. Хусаинов // Транспорт Российской Федерации, 2017. - № 4 (71). - С. 19–22.

4. Василенко М. Н. Автоматизация расчета параметров перегонной переездной сигнализации / М. Н. Василенко, С. И. Валиев, Т. А. Тележенко // Известия Петербургского университета путей сообщения. – СПб.: ФГБОУ ВПО «Петербургский государственный университет путей сообщения Императора Александра I», 2010. - №3

5. Трунаев А. М. Путьевой вибрационный датчик контроля наличия и определения параметров подвижных единиц / А. М. Трунаев, С. А. Радковский, М. Н. Чепцов, А. Б. Бойник // 36. наук. праць Донецького інституту залізничного транспорту. – Донецьк: ДонІЗТ, 2012. – Вип. 32. – С. 73–78.

6. Бибииков С. В. Алгоритмы и устройства системы оповещения о приближении поезда по виброакустическим колебаниям рельса: диссертация кандидата технических наук: 05.13.01 / С. В. Бибииков. - Санкт-Петербург, 2015. - 160 с.

7. Радковский С. А. Моделирование колебаний железнодорожного рельса при воздействии на него подвижной вертикальной динамической нагрузки / С. А. Радковский, А. М. Трунаев, В. Д. Пойманов // Сборник научных трудов Донецкого института железнодорожного транспорта. – Донецк: ДонИЖТ, 2016. – Вип. 43. – С. 4–9.

8. Радковский С. А. Исследование отклика рельсовой линии на единичное импульсное воздействие / С. А. Радковский, А. М. Трунаев // Сборник научных трудов. – Донецк: ДонИЖТ, 2017.- Вип. 47.– С.10 – 16.

9. Горелик А. В. Системы железнодорожной автоматики, телемеханики и связи: учебник: в 2 ч. / А. В. Горелик, Д. В. Шалягин, Ю. Г. Боровков, В. Е. Митрохин и др.; под ред. А. В. Горелика. – М.: ФГБОУ «Учебно-методический центр по образованию на железнодорожном транспорте», 2012.

10. Трофимова Т. И. Физика: учебник для студ. учреждений высш. проф. образования / Т. И. Трофимова. – 2-е изд., перераб. и доп. – М. : Издательский центр «Академия», 2013. – 352 с.

Трунаев А. М., Чепцов М. Н., Радковский С. А. Синтез математической модели управления процессом функционирования железнодорожных переездов на основе новых средств формирования извещений. В статье на основе анализа способов формирования подачи извещения на переезд определено, что они не учитывают реальную скорость и местоположения поезда. Из-за чего увеличивается время простоя автотранспорта. Как следствие нервозность водителей и проезд переездных заградительных устройств, при их закрытом состоянии. Авторами разработана математическая модель управления процессом функционирования железнодорожных переездов, которая учитывает местоположение и скорость поезда. Проведено численное моделирование предложенной модели на основании реальных данных, определено, что простой автотранспорта сократится в среднем в 2-5 раз.

Ключевые слова: железнодорожный переезд, способ подачи извещения, математическая модель, заградительные устройства, виброускорение, минимальное время приближения поезда к переезду.

Trunaev A., Cheptsov M., Radkovsky S. Synthesis of a mathematical model for controlling the process of functioning of level crossings based on new means of generating notifications. Based on an analysis of the methods for generating a notice of moving, the article determines that they do not take into account the real speed and location of the train. Because of what the idle time of vehicles increases. As a result, the nervousness of the drivers and the passage of traveling barrage devices when they are closed. The authors developed a mathematical model for controlling the process of functioning of level crossings, which takes into account the location and speed of the train. A numerical simulation of the proposed model was carried out on the basis of real data, it is clear that idle vehicles will be reduced by an average of 2-5 times.

Keywords: railway crossing, notification method, mathematical model, barrage devices, vibration acceleration, minimum time for a train to move.

Статья поступила в редакцию 12.12.2019
Рекомендована к публикации профессором Павлышом В. Н.

УДК 621.395

Применение генетического алгоритма для решения задач размещения базовых станций в сетях пятого поколения

К. А. Павловская

Донецкий национальный технический университет, г. Донецк

ks_pavlovskaya@rambler.ru

Аннотация

Поставлена задача территориального планирования сети пятого поколения с применением кластерного метода, определена целевая функция стоимости проектируемой сети. Предложено решение задачи оптимального размещения базовых станций в сетях пятого поколения с применением генетического алгоритма. Выделены особенности генетических алгоритмов, разработана структурная схема генетического алгоритма. Получено графическое представление и проведен анализ полученных результатов.

Введение

В связи с ростом потребностей абонентов в высокоскоростном интернете операторы мобильной связи вынуждены переходить на форматы, в которых мобильная связь отвечала бы необходимым критериям, как в вопросе качества услуг, так и в вопросе надежности связи. Кроме того, данные критерии затрагивают интересы не только потребителей, но и самого оператора. На данный момент форматом, который может удовлетворить все указанные требования, является новый формат 5G. При переходе на формат пятого поколения процесс успешной эксплуатации сетей непосредственно зависит от их планирования. В городах с хорошо развитой инфраструктурой операторам связи необходимо мгновенно реагировать на все изменения, рост передаваемой информации требует оптимизации структуры построения сети, поиска оптимального размещения базовых станций, учитывая требуемые критерии качества к проектируемой сети. Планирование размещения базовых станций является достаточно сложной задачей для операторов мобильной связи, но весьма актуальной с точки зрения затрат на построение и эксплуатацию сетей. Целью данной работы является разработка метода определения мест расположения базовых станций сотовых сетей 5G с оптимальными показателями качества QoS.

Постановка задачи

Задачей территориального планирования сети пятого поколения с применением кластерного метода [1] является расстановка на заданной территории минимального количества базовых станций с заданным числом абонентов и условий местности с учетом требуемого качества связи.

Исходя из принципа многокритериальности [2], предлагается модифицированная задача,

в которой учитывается не только потери распространения сигнала на основе канальной модели SUI, но и стоимостные характеристики модели. Решение задачи сводится к вычислению целевой функции стоимости (1) и ее дальнейшей минимизации [3].

$$\phi = \sum_k \sum_m \sum_n W_{kmn} X_{mn} Y_{kn} + \sum_m \sum_n c_n X_{mn} + \sum_k \sum_m \frac{1}{g_{km}} Y_{km} \rightarrow \min, \quad (1)$$

где m – число мест – кандидатов на расположение БС;

n – количество типов станций, $n = \overline{1, N}$;

k – число абонентов с заданными координатами, $k = \overline{1, K}$;

c_n – стоимость станции m -го типа, усл. ед.;

$X = (X_1, X_2, \dots, X_M)$ – вектор размещения базовых станций;

$Y = \|Y_{km}\|$ – матрица распределения клиентов по базовым станциям;

g_{km} – коэффициент учета потерь при распространении сигнала.

Дальнейшим шагом при решении поставленной задачи является разработка программного обеспечения для оптимального размещения базовых станций сетей пятого поколения.

Программная реализация решения задачи территориального планирования для сетей пятого поколения

На основе анализа различных способов нахождения оптимального размещения базовых станций принято решение: для поставленной задачи использовать стандартный генетический алгоритм и рекомендации к его применению. Генетические алгоритмы являются адаптивными методами поиска и используются для решения

задач оптимизации [4-7]. Естественный отбор и механизм генетического наследования являются основой генетического алгоритма, при этом используются биологическая терминология в упрощенном виде и основные понятия линейной алгебры [8].

Особенности генетических алгоритмов заключаются в следующем:

- исходные параметры задачи кодируются, и обрабатывается закодированная форма;

- процедура поиска оптимального решения производится из некоторой популяции, а не из одной точки;

- возможные варианты решения задачи используют вероятностные правила выбора.

Структурная схема генетического алгоритма для оптимального размещения базовых станций для сетей пятого поколения представлена на рис. 1.

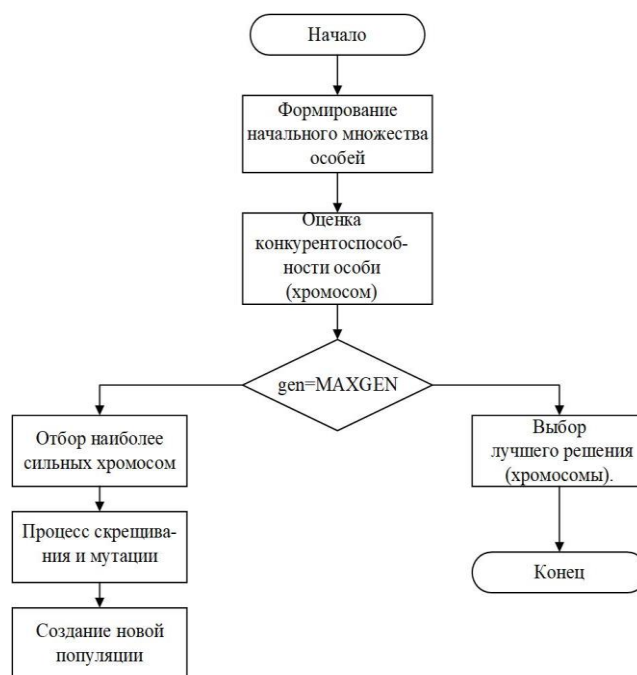


Рисунок 1 – Структурная схема генетического алгоритма для оптимального размещения базовых станций для сетей пятого поколения

В качестве программного комплекса был выбран пакет прикладных программ Matlab.

На первом этапе в разработке генетического алгоритма для оптимального размещения базовых станций для сетей пятого поколения происходит набор особей, которые называются популяцией, то есть определяются различные варианты расстановки базовых станций. При этом число вариантов расстановки может меняться в зависимости от требований к проектируемой сети. Особь представляет собой набор переменных, которые называют генами. Гены объединены в одну строку и формируют хромосому – решение задачи. Необходимо определить представленные хромосомы. В решении данной задачи используется бинарное представление в виде множества $\{0, 1\}$. Так же необходимо определить типы базовых станций (количество используемых переменных), в данной работе используется два типа базовых станций – микросота и пикосота.

На втором этапе происходит отбор наиболее сильных хромосом, то есть тех хромосом, которые будут принимать участие в

создании нового поколения. Суть отбора состоит в том, чтобы отобрать наиболее сильные особи и передать их гены следующему поколению особей. Определенное количество пар особей (родители) отбирается на основании их силы. Для отбора необходимо вычислить оценку родительских хромосом, самыми перспективными для дальнейшего участия в создании новых особей являются те хромосомы, у которых наименьший уровень силы, данный способ называется селекцией. Функция силы всегда принимает положительные значения. При создании начальной популяции определяется первый случайный вариант расстановки базовых станций, не учитывающий ограничения и требования к качеству услуг мобильной связи [9]. Из общего количества базовых станций двух типов производится выборка случайного числа БС1 и БС2.

Далее начинается главный цикл, где происходит оценка родительских хромосом, с учетом расстояния радиуса R1 и R2 от каждой базовой станции до каждого абонента. Если расстояние больше, чем радиус зоны покрытия

для принятой местности, то это расстояние объявляется переменной $\inf$ – бесконечно большой. Затем осуществляется поиск и подключение абонентов к базовой станции первого типа (микросота), при этом следует учитывать, что суммарный трафик всех абонентов не должен превышать производительность базовой станции. Далее происходит подключение оставшихся абонентов к базовой станции второго типа, в этом случае учитываются ограничения к поставленной задаче. Затем для данного варианта построения сети происходит оценка показателя качества QoS.

Необходимо отметить, что половина хромосом, участвующих в селекции, выбирается из числа наилучших особей, упорядоченных по степени ухудшения функции приспособленности. Другая половина родительских хромосом выбирается случайным образом из хромосом, не задействованных в селекции. Такой способ отбора предоставляет возможность создавать потомков как наилучшим, так и наихудшим особям, что увеличивает скорость сходимости на начальных итерациях алгоритма и позволяет в некоторых случаях выходить из локальных оптимумов. В результате процесса селекции создается новая родительская популяция.

Если бы отсутствовал этап отбора, то количество особей росло бы по экспоненциальному закону, вследствие чего требовалось бы большое количество оперативной памяти и времени на обработку каждого нового поколения популяции. Следовательно, после появления новых особей необходимо очистить популяцию от наименее удачных особей. Именно это происходит на этапе отбора.

После создания новой родительской популяции происходит процесс скрещивания и мутации, это приводит к возникновению новых потомков от предыдущей созданной популяции. Для скрещивания и мутации применяются генетические операторы, которые уже встроены в среду MATLAB: для скрещивания используется оператор `recombine individuals (crossover)`, а для мутации – `apply mutation` [10].

Основной оператор для создания новых хромосом в генетическом алгоритме – это `crossover` (скрещивание). Как и его аналог в природе, `crossover` приводит к созданию новых особей на основе существующих, в которых имеются некоторые части генетического материала обоих родителей.

Невозможно заранее предсказать, какие особи получатся в результате мутации, они могут быть как более удачные, так и менее.

Для задачи размещения базовых станций мутацию можно представить как изменение

подключения абонентов к базовым станциям. То есть, если до мутации клиент был подключен к БС 1-го типа, а после мутации переподключен к БС 2-го типа или к двум базовым станциям сразу, но при этом выполняются условия попадания в радиус действия. Затем происходит проверка условия завершения алгоритма, которое имеет следующий вид:

$$\text{gen} = \text{MAXGEN}, \quad (2)$$

где gen – номер текущего поколения;

MAXGEN – максимальное количество поколений (итераций) заданное в начале алгоритма (максимальное значение можно либо увеличивать, либо уменьшать).

Алгоритм может считаться завершенным, когда прекращается поиск минимального значения целевой функции, то есть, когда номер текущего поколения (итерации) достигнет максимального числа поколений (итераций), которое было задано в начале работы алгоритма.

Следует отметить, что генетические алгоритмы относят к эволюционным и эвристическим методам поиска решений, которые используют случайность как средство исследования, они являются неточными и часто называются «мягкими вычислениями».

Полученные результаты

Для тестирования и анализа разработанного алгоритма были решены задачи различных размерностей. Для обоснования необходимости учета в задаче оптимальной расстановки базовых станций для сетей пятого поколения коэффициента распространения сигнала, проведены исследования для двух случаев: с учетом коэффициента и без его учета.

Первый эксперимент заключается в анализе времени работы алгоритма при увеличении числа потенциальных абонентов, при этом количество потенциальных мест расстановки базовых станций остается постоянным. Результаты представлены на рис. 2.

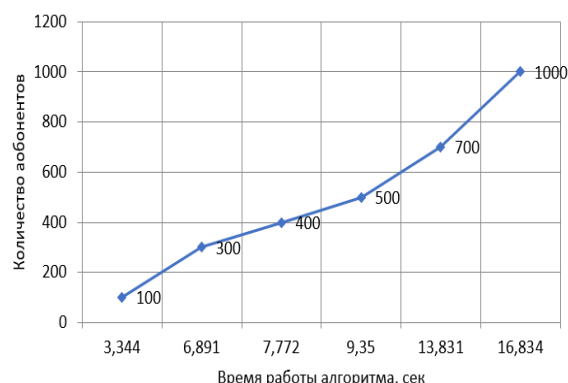


Рисунок 2 – Зависимость времени работы алгоритма от числа потенциальных абонентов

На основе анализа полученных результатов можно сделать вывод, что при увеличении количества абонентов увеличивается время работы алгоритма, причем в среднем на каждые сто абонентов время работы увеличивается на 1,5 с.

Следующий эксперимент заключается в определении зависимости времени работы алгоритма от увеличения количества мест расположения базовых станций, при этом число абонентов оставалось постоянной величиной ($P = 800$), а оценка стоимости построения сети производилась в зависимости от количества мест расстановки БС. Результаты представлены на рис. 3 и рис. 4.

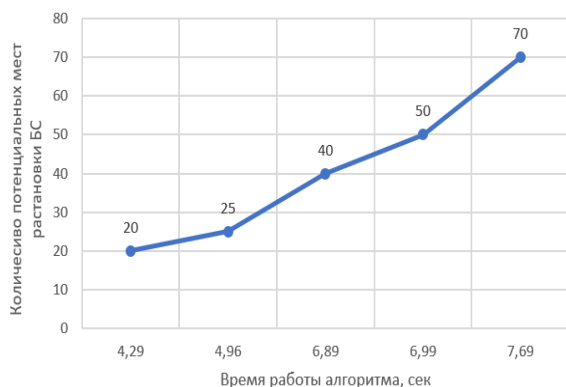


Рисунок 3 – Зависимость времени работы алгоритма от количества потенциальных мест расстановки базовых станций



Рисунок 4 – Зависимость стоимости проектируемой сети от количества потенциальных мест расстановки базовых станций

Анализируя полученные данные после двух экспериментов, можно сделать вывод, что рост числа мест кандидатов влияет на скорость решения задачи в меньшей степени, чем увеличение количества абонентов. Полученные

результаты свидетельствуют об экспоненциальной зависимости времени работы алгоритма от размерности исходных данных.

Оценивая стоимость проектируемой сети, можно сделать вывод, что при увеличении количества мест-кандидатов расстановки базовых станций уменьшается стоимость сети. Этому способствует сокращение неподключенных абонентов, а, следовательно, и уменьшение суммы штрафов за отсутствие подключения.

Анализ результатов проведенных экспериментов дает основание утверждать, что увеличение количества абонентов оказывает большее влияние на скорость решения задачи, чем рост числа мест-кандидатов. Полученные результаты также свидетельствуют об экспоненциальной зависимости времени работы алгоритма от размерности исходных данных.

Кроме того, проведенные эксперименты показали, что при увеличении количества мест-кандидатов расстановки базовых станций происходит уменьшение стоимости всей сети, сокращение неподключенных абонентов, снижение суммы штрафов за отсутствие подключения. Графическое представление полученных результатов приведено на рис. 5 и рис. 6. Из схемы расстановки базовых станций (рис. 5) видно, что распределение подключения 800 потенциальных абонентов к базовым станциям следующее:

- базовых станций первого типа – 17, подключенных абонентов – 725;
- базовых станций второго типа – 6, подключенных абонентов – 75.

При этом стоимость такой сети составляет 287800,00 у. е., что является минимальным значением из всей выборки. Подключения базовых станций второго типа происходит только в том случае, если исчерпан ресурс производительности макро базовых станций. На рис. 6 приведен пример расстановки базовых станций без учета коэффициента распространения сигнала. Анализ приведенных результатов позволяет сделать вывод, что расстановка происходит неравномерно, поскольку при этом не учитывается радиус действия базовых станций. Таким образом, 257 абонентов из 800 оказались не подключены, вследствие чего значение целевой функции (стоимость проектируемой сети) из-за штрафов за каждого такого абонента составила 12073663,00 у. е.

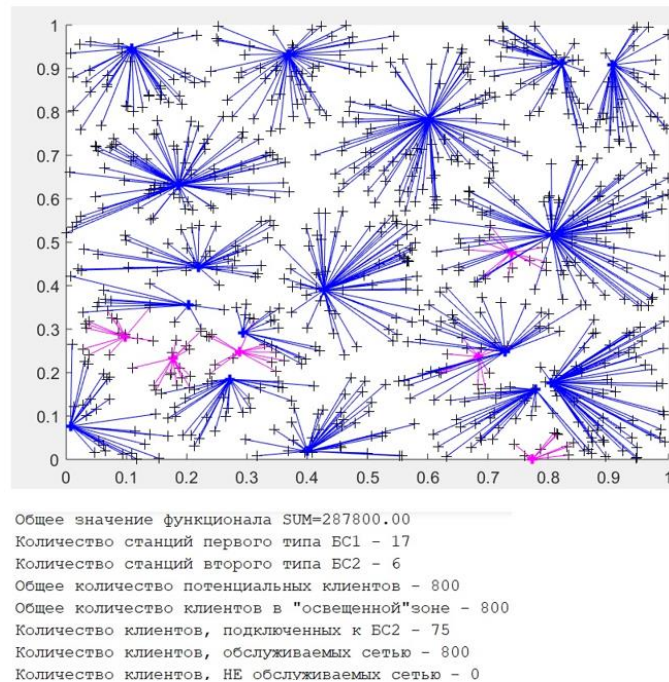


Рисунок 5 – Схема расстановки базовых станций для сети пятого поколения с учетом коэффициента распространения

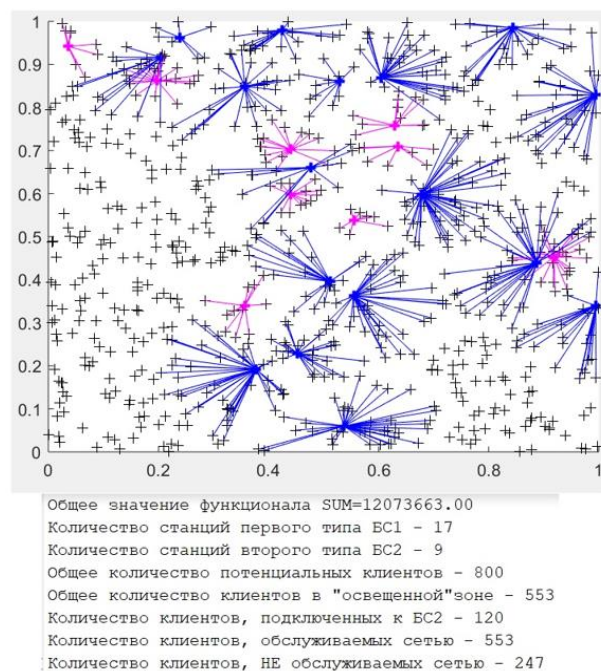


Рисунок 6 – Схема расстановки базовых станций для сети пятого поколения без учета коэффициента распространения и типа местности

Выводы

Тенденции роста сетевого трафика говорят о том, что возможностей сетей четвертого поколения уже недостаточно для удовлетворения потребительских услуг, страны, лидирующие в сфере систем мобильной связи, занимаются разработкой и внедрением сетей пятого поколения.

В работе поставлена задача территориального планирования сети пятого поколения с применением кластерного метода. Решение задачи предполагает использование генетического алгоритма, применение которого обеспечивает оптимальное соотношение между качеством полученного решения и временем, потраченным на его поиск.

Литература

1. Гимадинов Р. Ф. Кластеризация в мобильных сетях 5G. Случай частичной мобильности / Р. Ф. Гимадинов, А. С. Мутханна, А. Е. Кучерявый // Информационные технологии и телекоммуникации, 2015. - Т. 3. № 2. - С. 44-52
2. Чеботарёва Д. В., Безрук В. М. Многокритериальная оптимизация проектных решений при планировании сотовых сетей. - Харьков: СМИТ, 2013. - 148 с.
3. Павловская К. А. Задача оптимального размещения базовых станций в сетях пятого поколения / К. А. Павловская // Сборник научных трудов ДОНИЖТ. - Донецк, 2019. - №55. — С. 19-23.
4. Казаковцев Л. А. Модификация генетического алгоритма с жадной эвристикой для непрерывных задач размещения и классификации / Л. А. Казаковцев, А. А. Ступина, В. И. Орлов // Системы управления и информационные технологии, 2014. - №2. - С. 35-39.
5. Ермолаев С. Ю. Генетический подход к задаче оптимального размещения базовых станций в сетях IEEE 802.16-2004 / С. Ю. Ермолаев, В. Г. Карташевский // Труды 12-й Международной конференции «Цифровая обработка сигналов и ее применение» (DSPA-2010). - Москва, 2010. - С. 255-257.
6. Ермолаев С. Ю. Оптимальное размещение базовых станций / С. Ю. Ермолаев // Telecommunication Sciences, 2010. - Т. 1, №1. - С. 82-90.
7. Гладков Л. А. Генетические алгоритмы / Л. А. Гладков, В. В. Курейчик, В. М. Курейчик. - М.: 2-е изд. исправл. и доп. - ФИЗМАТЛИТ, 2010. - 320 с.
8. Казаковцев Л. А. Задача выбора оптимального размещения элементов беспроводной сети [электронное издание] / Л. А. Казаковцев, М. Н. Гудыма, А. А. Ступина, Ю. И. Кириллов // Современные проблемы науки и образования, 2013. - Вып.3.- URL <http://elibrary.ru/download/29613015.pdf>.
9. Павловская К. А. Многокритериальная оптимизация построения сетей пятого поколения на основе системного анализа / К. А. Павловская, В. Н. Лозинская, И. Н. Яремко // Вестник Академии гражданской защиты: научный журнал. - Донецк: ГОУВПО «Академия гражданской защиты» МЧС ДНР, 2019. - Вып. 4 (20). - С. 111-117.
10. Дьяконов В. П., В. В. Круглов MATLAB 6.5 SPI/7/7 SPI1/7 SP2 +Simulink5/6. Инструменты искусственного интеллекта и биоинформатики / Серия «Библиотека профессионала». — М: СЛОН-ПРЕСС, 2010. - С. 249-258.

Павловская К. А. Применение генетического алгоритма для решения задач размещения базовых станций в сетях пятого поколения. Поставлена задача территориального планирования сети пятого поколения с применением кластерного метода, определена целевая функция стоимости проектируемой сети. Предложено решение задачи оптимального размещения базовых станций в сетях пятого поколения с применением генетического алгоритма. Выделены особенности генетических алгоритмов, разработана структурная схема генетического алгоритма. Получено графическое представление и проведен анализ полученных результатов.

Ключевые слова: базовая станция, сеть, генетический алгоритм, оператор, задача, построение, сигнал

Pavlovskaya K. A genetic algorithm is used to solve the problems of placing base stations in fifth-generation networks. The task of territorial planning of the fifth generation network using the cluster method is posed, the objective cost function of the designed network is determined. It is proposed to solve the problem of optimal placement of base stations in fifth-generation networks, using a genetic algorithm. The features of genetic algorithms are highlighted, a structural diagram of the genetic algorithm is proposed. Visualization is presented; the analysis of the results is carried out.

Keywords: base station, network, genetic algorithm, operator, task, construction, signal.

Статья поступила в редакцию 16.12.2019
Рекомендована к публикации профессором Миненко А. С.

Повышение эффективности методов и средств защиты пользовательских мобильных приложений в операционной системе Android

В. Е. Лебедев, А. В. Чернышова

Донецкий национальный технический университет, г. Донецк

chernyshova.alla@rambler.ru, alent.1301@gmail.com

Аннотация

В данной работе был выполнен анализ методов и средств, позволяющих сохранять пользовательские данные в операционной системе Android. Проанализировав возможные атаки на стандартные хранилища, а также возможность использования root-доступа, были сделаны выводы о возможных атаках и предложено решение для защиты пользовательских данных с использованием шифрования. Спроектирован и разработан модуль, позволяющий получить возможность шифрования с использованием пароля пользователя.

Введение

Учитывая растущую популярность мобильных приложений, существуют серьезные опасения относительно их безопасности, например, бреши в системах защиты банковских приложений могут повлечь за собой финансовые потери для большого количества пользователей [1].

На данный момент операционная система Android является наиболее популярной операционной системой, а количество ее пользователей насчитывает больше двух миллиардов активных устройств. Данная операционная система используется на множестве устройств, таких как мобильные телефоны, планшеты, телевизоры, умные часы, ноутбуки, автомобили. Техническая грамотность пользователей постепенно растет, а вместе с ней понимание необходимости обеспечения конфиденциальности и безопасности личных данных.

Способы хранения пользовательских данных в ОС Android

Android – свободный и открытый проект. Большинство исходного кода распространяется под свободной лицензией Apache 2.0. Android основан на ядре Linux, однако, значительно отличается от большинства других Linux-систем.

Как и в других Linux-системах, ядро Linux обеспечивает такие низкоуровневые операции, как управление памятью, защиту данных, поддержку мультипроцессности и многопоточности.

Некоторые компоненты системы были заменены аналогами, которые лучше предназначены для использования в условиях ограниченной памяти, низкой скорости процессора и минимального потребления энергии – таким образом, Android больше похож

на встраиваемую (embedded) Linux-систему, чем на GNU/Linux.

Кроме того, ядро Linux в Android было модифицировано: добавлено несколько компонентов, в том числе ashmem (anonymous shared memory), Binder driver (часть фреймворка Binder), wakelocks (управление спящим режимом) и low memory killer.

В качестве libc (стандартной библиотеки языка C) в Android используется не GNU C library (glibc), а собственная минималистичная реализация под названием bionic, оптимизированная для встраиваемых (embedded) систем – она значительно быстрее, меньше и менее требовательна к памяти, чем glibc, которая обросла множеством слоёв совместимости.

Многие операционные системы делятся на собственно операционную систему и установленные приложения, ничего друг о друге не знающие и не умеющие взаимодействовать [2]. Система компонентов и intent'ов Android позволяет приложениям, по-прежнему абсолютно ничего друг о друге не зная, составлять для пользователя один интегрированный системный user experience (пользовательский опыт) – установленные приложения реализуют части одной большой системы (рис. 1), они составляют из себя систему.

Поставщики контента (Content providers) управляют доступом к структурированному набору данных. Они инкапсулируют данные и предоставляют механизмы обеспечения их безопасности. Поставщики контента представляют собой стандартный интерфейс для объединения данных в одном процессе с кодом, который выполняется в другом процессе.

В состав системы Android входят поставщики контента, которые управляют такими данными, как аудио, видео, изображения и личная контактная информация.

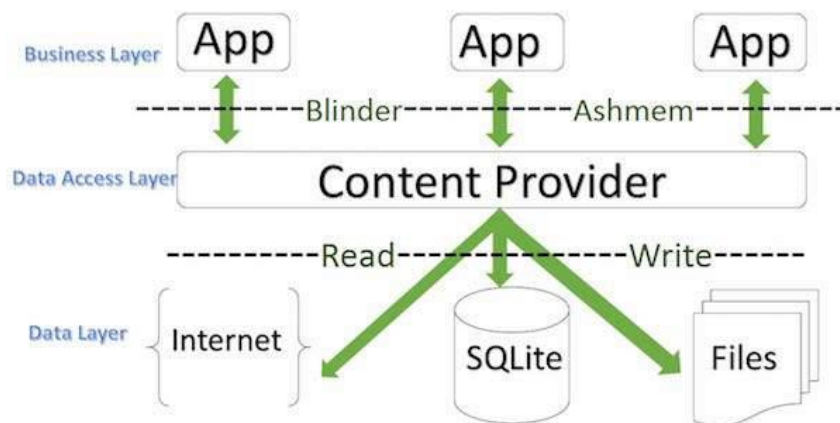


Рисунок 1 – Схема взаимодействия с использованием ContentProvider

Основная задача мобильного приложения, как и любой программы – получение, хранение и обработка пользовательской информации. Обеспечение защиты данных является одним из ключевых моментов разработки приложений для OS Android.

ОС Android предоставляет стандартные средства хранения пользовательских данных:

- Shared Preferences – инструмент для хранения небольших объемов информации, таких как настройки приложения;
- Account Manager – инструмент для безопасного хранения данных об аккаунтах пользователей (логин, email, пароль);
- хранение файлов в файловой системе;
- Sqlite – база данных для хранения данных приложения.

Shared Preferences позволяет сохранять данные в виде «ключ-значение». Используется для хранения небольшого, фиксированного объема информации [3]. Схема взаимодействия предусматривает чтение данных либо изменение значений с последующим сохранением. После сохранения изменений данные помещаются в файл в области данных приложения. Данные хранятся в формате XML (рис. 2).

```
<?xml version='1.0' encoding='utf-8' standalone='yes' ?>
<map>
  <string name="username">Nickname</string>
  <string name="password">pass1234</string>
</map>
```

Рисунок 2 – Формат хранения данных в SharedPreferences

Также позволяет сохранять данные примитивных типов данных: булевы значения, целые (int, long), вещественные числа, строки и массивы строк.

В случае использования SQLite в системе создается структурированный файл базы данных, к которому обращается приложение для сохранения и чтения данных.

По умолчанию, доступ к файлам, созданным во внутреннем хранилище, может

получить только приложение, создавшее эти файлы. Если необходимо передать данные другому приложению, следует использовать поставщики услуг.

Уязвимость AccountManager

Account Manager позволяет упростить хранение данных пользователя, таких как пароль, токен и другие строковые данные. Кроме того, позволяет автоматически обновлять токен при устаревании.

В случае использования Account Manager может возникнуть ситуация: на устройстве существует приложение А с accountType = «someType», создается аккаунт и сохраняется токен и пароль; далее на устройство устанавливается приложение В с таким же accountType = «someType» (рис. 3). Уязвимость заключается в том, что, если удалить приложение А, то все созданные аккаунты перейдут во власть приложения В с полным доступом. Обратное утверждение тоже верно, т.е. при удалении приложения В аккаунты передадутся приложению А. Данная уязвимость была исправлена лишь в Android 7.0 (API 24) [4].

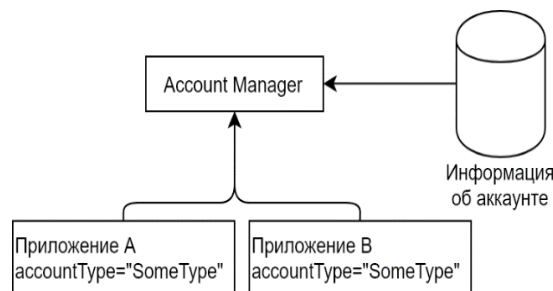


Рисунок 3 – Уязвимость AccountManager

Можно рассмотреть два варианта решения данной проблемы: продолжать пользоваться АМ, но хранить в нем все важные данные в зашифрованном виде либо отказаться от АМ и пользоваться хранилищем, которое можно шифровать.

Использование root-доступа

В Android приложения исполняются в собственной «песочнице», изолирующей его данные от других элементов системы. Таким образом файлы, сохраняемые в локальном хранилище, защищены от других приложений. Однако, эта особенность архитектуры ОС не в состоянии обеспечить должную защиту информации в случае использования root-доступа.

В Unix-подобных операционных системах существует понятие root-пользователя. Пользователь с UID (userid) равным 0, также

известный как суперпользователь, имеет доступ ко всем ресурсам системы, а также на него не распространяются никакие ограничения прав [5].

Практически любое устройство предоставляет возможность получения root-доступа, позволяющего управлять всеми данными приложения. Несмотря на ряд опасностей при использовании, большое количество пользователей решаются на модификацию системы (рис. 4). По статистике «Лаборатории Касперского» 6 из 10 стран, в которых чаще всего случаются атаки на мобильные устройства совпадают со странами, в которых чаще всего «рутуют» Android [6].

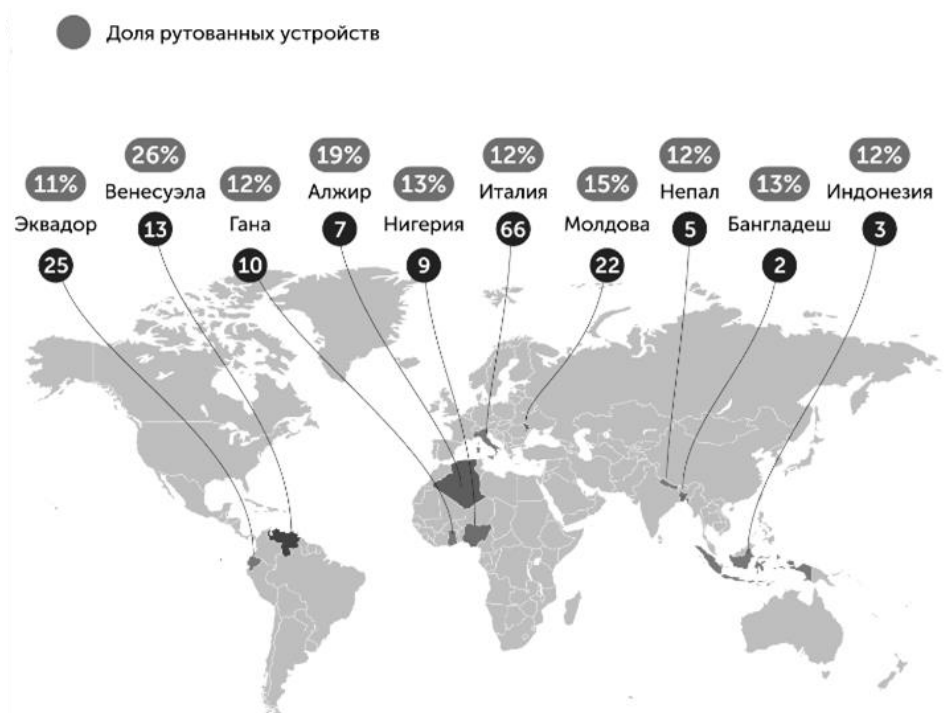


Рисунок 4 – Доля устройств с root-доступом

Разработка модуля защиты пользовательских данных

Исходя из того, что в случае использования root-доступа операционная система Android не может обеспечить защиту данных при использовании стандартных методов и средств, появляется задача улучшения способа защиты.

Root-доступ позволит получить доступ к любым файлам, размещенным в системе. Поэтому для защиты информации необходимо зашифровать, что значительно усложнит доступ к данным.

Следующей задачей будет сохранение ключей шифрования. Так как сохранение в локальный файл небезопасно, следует разработать другой способ хранения.

На данный момент в системе Android реализована возможность сохранить ключ

шифрования в хранилище, зашифрованном с использованием пароля пользователя. Данный способ предусматривает, что пользователь заранее использовал один из нескольких способов для блокировки доступа в меню настроек безопасности. На данный момент современные версии системы предоставляют следующие варианты защиты: 4-значный PIN-код; графический ключ (последовательность точек на поле 3×3); пароль; биометрические данные (отпечаток пальца или распознавание лица). Биометрические данные используются только в сочетании с одним из паролей, так как на данный момент датчики не могут гарантировать корректную работу во всех возможных ситуациях. Так, некоторые датчики можно обмануть, используя муляж пальца, либо он может быть поврежден.

Однако, часто пользователи не устанавливают пароль для использования

мобильного устройства, в таком случае разработчик должен сам обеспечить возможность безопасного хранения ключа шифрования.

Решением для данной проблемы может стать создание ключа шифрования из пароля пользователя. Таким образом, пропадает необходимость хранить ключ шифрования, а при необходимости прочитать данные из защищенного файла пользователь должен будет ввести пароль.

Для генерации ключа используется стандарт формирования ключа на основе пароля PBKDF2 (Password-Based Key Derivation

Function). Он использует псевдослучайную функцию для получения ключей, в качестве псевдослучайной может быть выбрана криптографическая хеш-функция, шифр, HMAC.

Данный стандарт использован в таких системах как: Wi-Fi Protected Access (WPA и WPA2) и Microsoft Windows Data Protection API (DPAPI). Применяется при шифровании в формате OpenDocument и AES в WinZip, а также для хранения паролей в LastPass [7].

Схема использования ключа шифрования с использованием пароля пользователя представлена на рис. 5.

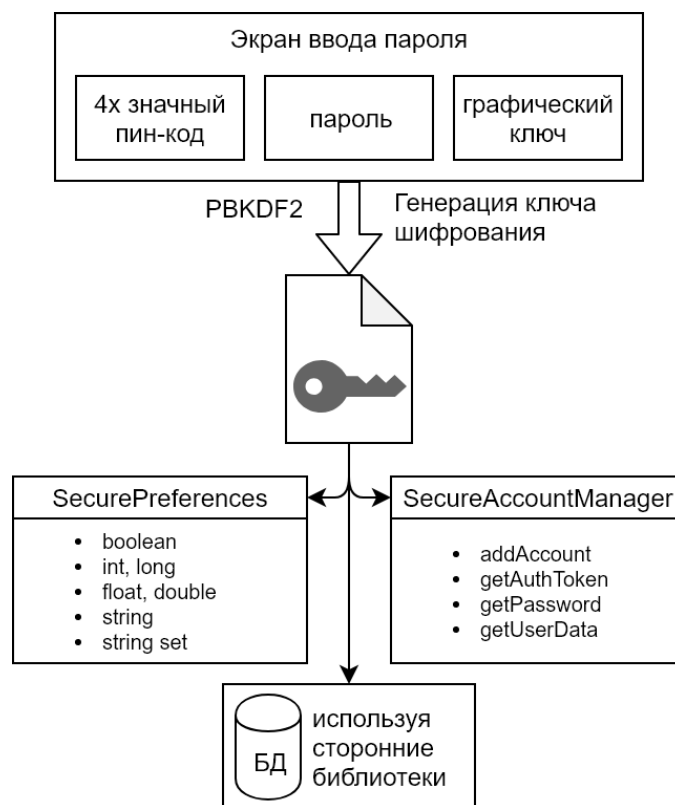


Рисунок 5 – Схема использования ключа шифрования исходя из пароля пользователя

Разработанная библиотека позволяет преобразовать примитивные типы данных (Boolean, Integer, Long, Float, Double, String), а также набор строк (Set<String>) в массив байт, используя специальные сериализаторы. Далее эти данные передаются в реализацию интерфейса ValueEncrytor, который выполняет шифрование массива байт. Также, в случае использования SharedPreferences необходимо сериализовать строковое название поля и зашифровать его.

Разработана собственная реализации AccountManager и SharedPreferences, которая соответствует интерфейсу встроенных средств, но перед сохранением в стандартное хранилище выполняет шифрование данных.

Общая схема сохранения информации в приложении будет выглядеть следующим образом (рис. 6).

Шифрование базы данных

В дополнение к вышеперечисленному, такая система генерации ключа шифрования из пароля пользователя позволит шифровать базу данных. Для Android уже существует ряд библиотек для работы с базой данных, которые позволяют выполнить шифрование данных.

Библиотека greenDAO позиционируется как инструмент для объектно-реляционного отображения (ORM) данных в приложениях на Android, являясь прослойкой, реализующей шаблон проектирования DAO (Data Access Object) (рис. 7).

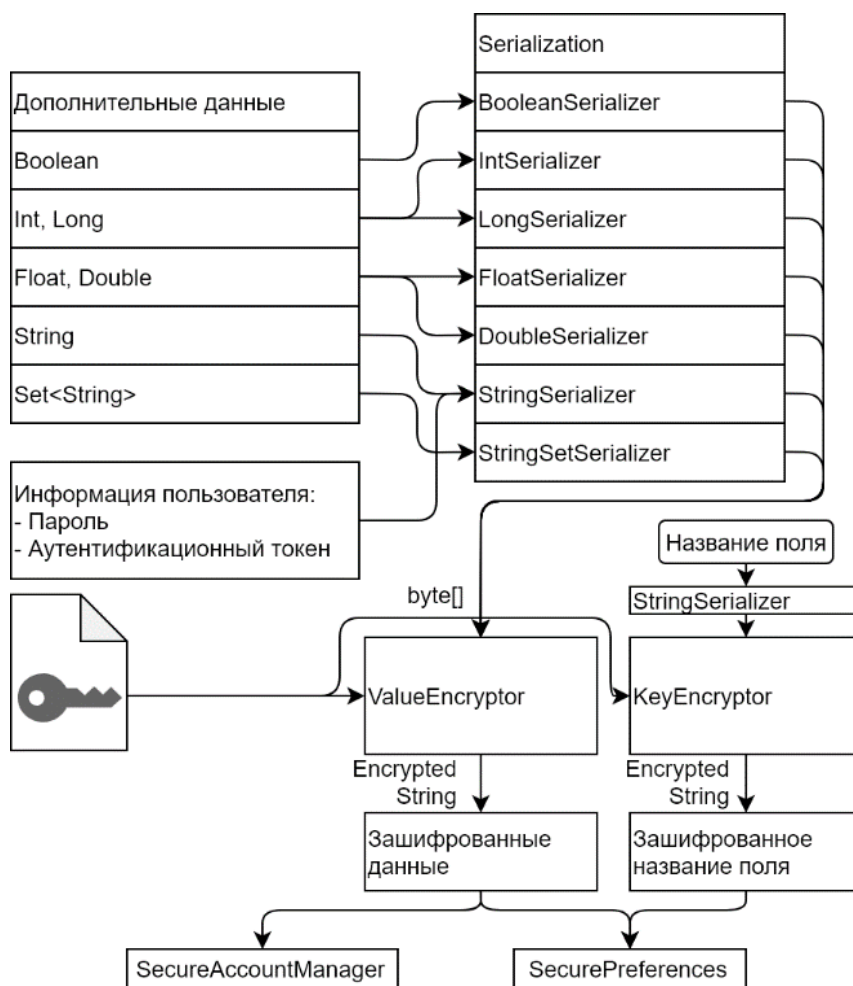


Рисунок 6 – Схема сохранения информации в SecureAccountManager и SecurePreferences

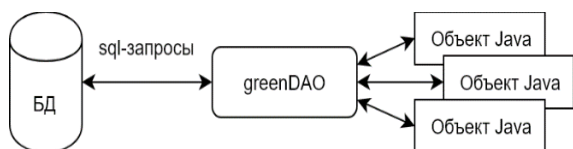


Рисунок 7 – Схема взаимодействия greenDAO с базой данных

Библиотека greenDAO позволяет использовать SQLCipher – библиотеку с открытым исходным кодом, созданную компанией Zetetic, для прозрачного 256-битного AES шифрования баз данных SQLite [8].

SQLCipher применяется во множестве коммерческих разработок и продуктов с открытым исходным кодом. Кроме того, SQLCipher – одна из самых популярных систем шифрования баз данных для мобильных, встраиваемых и настольных платформ.

SQLCipher имеет небольшой вес и отличную производительность, поэтому он идеально подходит для защиты встроенных баз приложений и любых для мобильных разработок [9].

Можно выделить следующие плюсы при использовании SQLCipher:

- высокая производительность и быстрое шифрование с 5-15 % временными затратами;
- база данных шифруется полностью;
- меры безопасности (режим CBC, вывод ключей);
- криптография с нулевой конфигурацией и уровнем приложения;
- алгоритмы, предоставленные экспертной криптографической библиотекой OpenSSL [10];
- SQLCipher поддерживает большинство современных языков программирования и платформ: C/C++, Obj-C, Java, Python и пр.

Выводы

Предложенная система для защиты пользовательских данных позволит защитить приложение от ряда атак, направленных на получение файлов приложения.

Основными пользователями являются разработчики мобильных приложений под операционную систему Android, которые хотят улучшить защиту в приложении.

В дальнейшем пользователи системы могут расширить возможности системы за счет разработки собственных реализаций интерфейсов KeyEncrptor и ValueEncrptor для реализации других алгоритмов шифрования. Кроме того, при необходимости, пользователи могут разработать собственные сериализаторы для других типов данных.

Библиотека реализует шифрование данных в качестве прослойки, при этом сохраняя стандартный интерфейс для взаимодействия, используя наследование стандартных классов Account Manager и Shared Preferences. Это позволит разработчикам легко использовать библиотеку без необходимости дополнительного изучения.

В дальнейшем планируется провести исследование эффективности предложенной системы.

Литература

1. Security tips // Android Developers [Электронный ресурс]. – Режим доступа: <https://developer.android.com/training/articles/security-tips>. – Загл. с экрана.
2. Как работает Android, часть 3 // Хабрахабр. [Электронный ресурс]. – Режим доступа: <https://habrahabr.ru/company/solarsecurity/blog/338494/>. – Загл. с экрана.
3. Сохранение данных // Developer Android. [Электронный ресурс]. – Режим доступа: <https://developer.android.com/training/basics/data-storage/index.html?hl=ru>. – Загл. с

экрана.

4. AccountManager //Android Data Residue Vulnerability. [Электронный ресурс]. – Режим доступа: <https://developer.android.com/training/basics/data-storage/index.html?hl=ru>. – Загл. с экрана.

5. Как работает Android, часть 4 // Хабрахабр. [Электронный ресурс]. – Режим доступа: <https://habr.com/ru/company/solarsecurity/blog/427431/>. – Загл. с экрана.

6. Root в Android: плюсы, минусы, подводные камни // Блог Лаборатории Касперского. [Электронный ресурс]. – Режим доступа: <https://www.kaspersky.ru/blog/android-root-faq/17785/>. – Загл. с экрана.

7. Итерации пароля (PBKDF2) //LastPass. [Электронный ресурс]. – Режим доступа: <https://helpdesk.lastpass.com/ru/account-settings/general/password-iterations-pbkdf2/>. – Загл. с экрана.

8. greenDAO 2.2 with Database Encryption //greenrobot. [Электронный ресурс]. – Режим доступа: <http://greenrobot.org/release/greendao-2-2-with-database-encryption/>. – Загл. с экрана.

9. Пакет: sqlcipher (3.4.1-1 и другие)//Debian. [Электронный ресурс]. – Режим доступа: <https://packages.debian.org/sid/sqlcipher>. – Загл. с экрана.

10. Standards//OpenSSL. [Электронный ресурс]. – Режим доступа: <https://www.openssl.org/docs/standards.html>. – Загл. с экрана.

Лебедев В. Е., Чернышова А. В. Повышение эффективности методов и средств защиты пользовательских мобильных приложений в операционной системе Android. В данной работе был выполнен анализ методов и средств, позволяющих сохранять пользовательские данные в операционной системе Android. Проанализировав возможные атаки на стандартные хранилища, а также возможность использования root-доступа, были сделаны выводы о возможных атаках и предложено решение для защиты пользовательских данных с использованием шифрования. Спроектирован и разработан модуль, позволяющий получить возможность шифрования с использованием пароля пользователя.

Ключевые слова: защита, шифрование, данные пользователя, Android, root-доступ, проектирование, БД.

Lebedev V., Chernyshova A. Improving the effectiveness of protection methods and means for custom mobile applications in the Android operating system. In this paper, the methods and tools that allow users to save user data in the Android operating system were analyzed. Also conclusions about possible attacks were made and a solution to protect data using encryption was proposed. A module has been designed and developed to enable encryption using a user password.

Keywords: security, encryption, user's data, Android, root-access, design, shared preferences, DB.

Статья поступила в редакцию 03.12.2019
Рекомендована к публикации профессором Зори С. А.

Системный анализ закономерностей мирового развития компьютерных систем

А. Я. Аноприенко, К. А. Сидоров, Н. С. Максименко, А. А. Койбаш
Донецкий национальный технический университет, г. Донецк
ska@donntu.org

Аннотация

Проведен анализ концептуальных, экспоненциальных, экономических закономерностей мирового развития компьютерных систем. В ходе анализа закономерностей выявлено, что темпы развития компьютерных систем остаются на высоком уровне. В области концептуальных закономерностей наблюдается тенденция к усложнению задач, решаемых информационными системами. В области экспоненциальных сохраняется стойкий рост производительности компьютерных систем. В сфере экономических наблюдается тенденция постепенного увеличения доли цифровой экономики в мире, что также требует увеличения количества и качества обслуживающих информационных систем. Выявлены предпосылки к постоянному усложнению компьютерных систем и связей между ними.

Введение

В настоящее время развитие компьютерных систем переживает очередную революцию. Происходит постоянный рост мощности и сложности компьютерных систем. Для определения точного вектора эволюции в данной сфере и своевременного реагирования на все изменения необходимо провести анализ закономерностей развития компьютерных систем. Результат анализа даст возможность экстраполировать выявленные закономерности на годы вперед, что позволит прогнозировать уровень развития компьютерных систем в будущем.

Цель статьи

Целью данной работы является анализ закономерностей мирового развития компьютерных систем. Для выявления основных тенденций и определения направления дальнейшего развития в статье рассмотрены концептуальные, экспоненциальные и экономические закономерности. Каждая из закономерностей показала устойчивые темпы роста в проанализированной области.

Концептуальные закономерности развития компьютерных технологий

Для того, чтобы отследить изменения концепций применения компьютерных систем, необходимо описать пять поколений развития вычислительной техники.

Первое поколение – компьютеры на электронных лампах (1946–1956 гг.). Основной целью создания компьютеров этого поколения было решение задач, связанных с созданием атомной бомбы. Машина занимала площадь 135 м³, весила 30 тонн и потребляла 150 кВт

электроэнергии. Основным рабочим элементом системы были 18 тысяч электронных ламп, а производительность составляла порядка тысячи операций в секунду.

Второе поколение – компьютеры на транзисторах (1956 – 1964 гг.). Габариты и энергопотребление данных систем стали значительно ниже, при этом вычислительная мощность выросла на порядок. Быстродействие машин составляло 1 миллион операций в секунду. В основе таких систем находилось уже 25 тысяч транзисторов. Целью компьютеров данного поколения по-прежнему являлось решение задач в военной отрасли.

Третье поколение – компьютеры на микросхемах с малой степенью интеграции (1964 – 1971 гг.). Сохраняя тенденцию к уменьшению размеров и энергопотребления, данная компьютерная система по габаритам не превосходила размеров холодильника. В основе уже лежали микросхемы IBM-360. Модели этого поколения имели вычислительную способность от 32 тысяч до 1,7 миллиона операций в секунду. В основном, это поколение компьютеров по-прежнему использовалось в военной промышленности. Однако некоторые модели компьютеров нашли применение в науке и образовании.

Четвертое поколение – компьютеры на микропроцессорах (1971 – настоящее время). Внедрение этого поколения во все сферы жизнедеятельности человека длилось достаточно долгое время. Началось оно с военной промышленности, науки, образования, а также коммерческих проектов. Позже благодаря компаниям Apple, Microsoft и IBM, появилось такое понятие как персональный компьютер. С этого момента, компьютеры начали расширять зону своего применения. На данный момент не

существует ни одной сферы, где не применялся бы компьютер.

Также эксперты выделяют пятое, так называемое перспективное поколение — это ЭВМ, использующие новые технологии и новую элементную базу. В пример можно привести квантовые компьютеры, ДНК компьютеры, суперсенсорные компьютеры [1].

В табл. 1 отражена последовательность компьютерных систем, и задач, для решения которых они применяются.

Таблица 1 – Развитие компьютерных систем

Года	Цели, задачи
1950-1960	ЭВМ, ускорение счета. Мейнфреймы и базы данных.
1960-1970	Настольные системы, персональные компьютеры. Управленческие ИС, ускорение создания отчетов.
1970-1980	Системы поддержки принятия решений. Автоматизация бизнеспроцессов.
1980-1990	Интернет, электронная коммерция, обеспечение процветания организации.
1990-2000	Мобильный широкополосный доступ в интернет.
2000-2010	Социальные сети, цифровая реклама, маркетинг.
2010-2020	Большие массивы данных, прогнозная аналитика, интернет вещей.
2020	Виртуальная реальность, искусственный интеллект, робототехника.

Экспоненциальные закономерности развития компьютерных систем

Одним из законов, описывающих темпы развития компьютерных систем, является закон Гордона Мура. Изначальная формулировка данного закона, в начале 60х годов, звучала следующим образом: «Количество элементов на кристаллах электронных микросхем будет и далее удваиваться каждый год».

Однако через 10 лет пришлось внести уточнения в данную формулировку. Она приобрела следующий вид: «Число транзисторов на кристалле удваивается каждые два года» [2].

Очевидно одно, суть самого закона не меняется, а именно — постоянный экспоненциальный рост количества транзисторов на микросхемах, а как следствие аналогичный рост производительности компьютерных систем.

Именно поэтому при попытке прогнозировать производительную мощность «компьютеров будущего», эксперты всё чаще говорят о том, что уже в ближайшее десятилетие произойдет революция в мире компьютерных систем. На рис. 1 отображены темпы развития микропроцессоров для компьютеров. На данной диаграмме указаны два пороговых значения «мышинный мозг» и «человеческий мозг».

Из диаграммы видно, что вычислительная мощность современных процессоров вплотную приблизилась к эквиваленту мышиного мозга. При продолжении диаграммы, с учетом применения закона Мура, уже к 2025 году должен произойти качественный скачок в развитии вычислительных мощностей. Мощность компьютера станет эквивалентной производительности человеческого мозга.

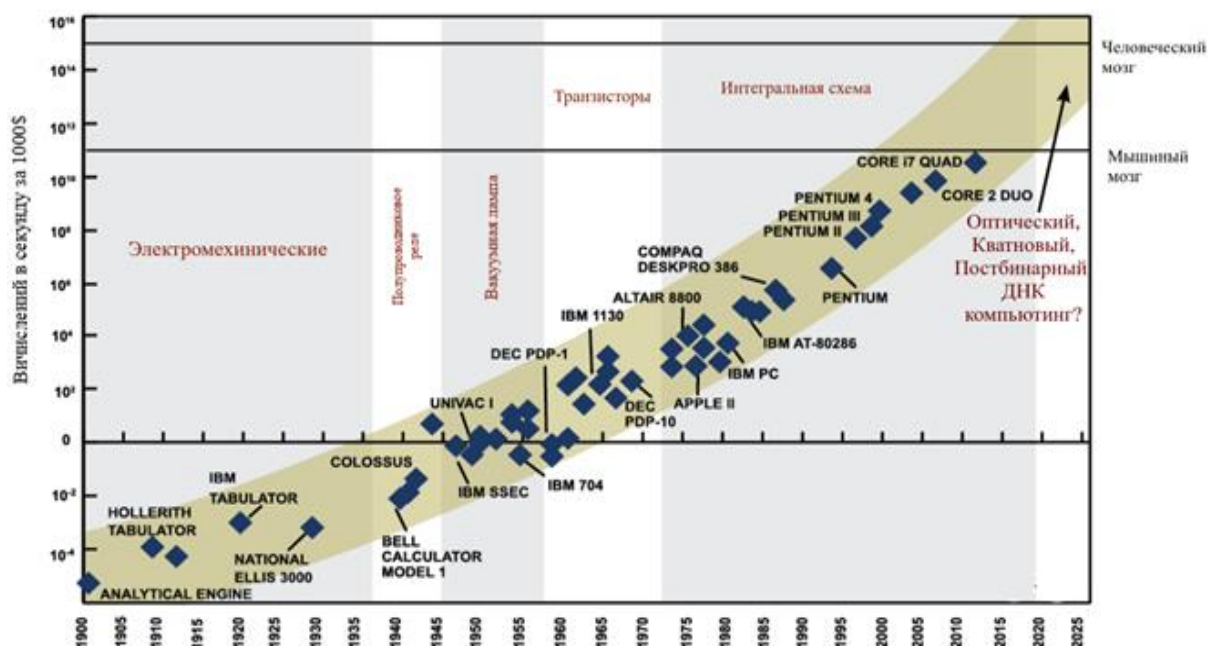


Рисунок 1 – Темп развития микропроцессоров [3]

На данном этапе многие ученые предлагают свои теории о развитии компьютеринга в целом. Существуют теории о ДНК компьютеринге, квантовых компьютерах, оптических компьютерингах. Помимо этих относительно давно распространившихся идей, так же существует и идея, разрабатывающаяся в

стенах Донецкого Национального Технического Университета – идея о постбинарном компьютеринге [4], который так же может занять свою нишу уже в ближайшем будущем.

Также стоит помнить о постоянном и экспоненциальном росте числа пользователей Интернет (табл. 2).

Таблица 2 – Мировая статистика использования интернета

Мировые регионы	Население млн.	% населения земли	Пользователи сети интернет млн. 31.12.2017	Прирост 2000-2018	Пользователи сети интернет %
Африка	1 288	16,9 %	453	9,941 %	10,9 %
Азия	4 207	55,1 %	2 023	1,670 %	48,7 %
Европа	827	10,8 %	704	570 %	17,0 %
Латинская Америка / Карибы	652	8,5 %	437	2,318 %	10,5 %
Ближний восток	254	3,3 %	164	4,893 %	3,9 %
Северная Америка	363	4,8 %	345	219 %	8,3 %
Океания / Австралия	41	0,6 %	28	273 %	0,7 %
Всего	7 634	100,0 %	4 156	1,052 %	100,0 %

По официальным данным сайта Internet World Stats [5], на 31 декабря 2017 года, почти 55 % населения земли тем или иным способом являются пользователями интернет.

Здесь специально ставится акцент, что для пользования интернетом не обязательно наличие

полноценного компьютера. Также не стоит забывать, что мобильный интернет уже преодолел психологическую отметку в 50 % от всемирного веб-трафика, что подтверждается данными сайта Statista [6] (рис. 2).

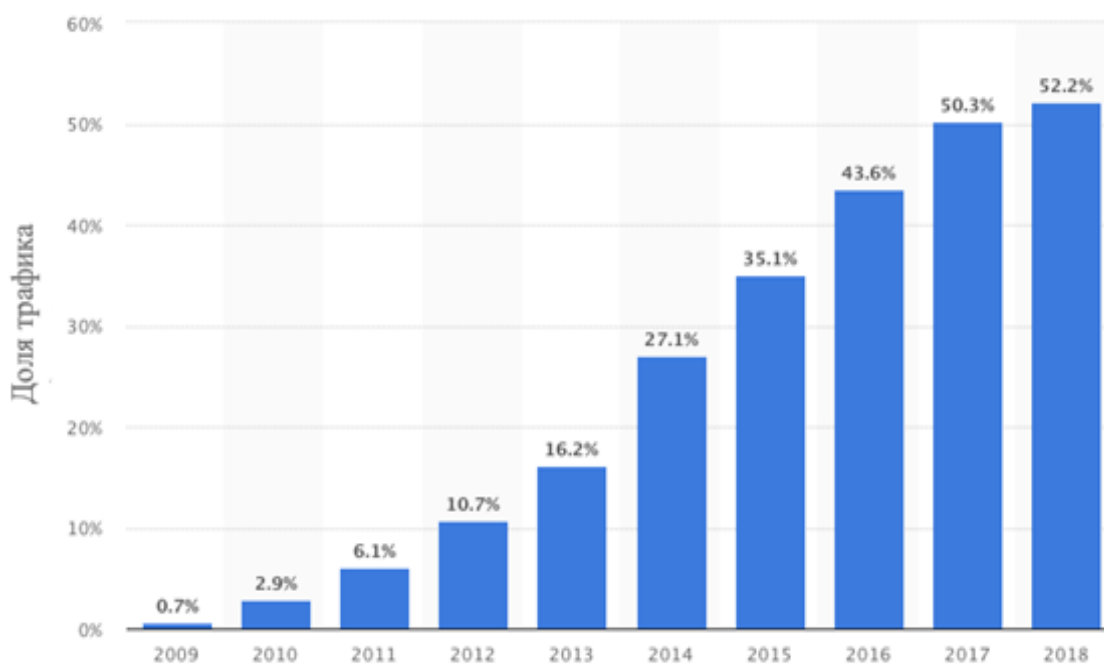


Рисунок 2 – Процент мобильного трафика [6]

Более подробную статистику предоставляет компания Яндекс, которая проводила исследования [7] в 2016 году на территории России.

На графике, который изображен на рис. 3, представлены показатели количества часов в неделю, проведенных в интернете с различных устройств.



Рисунок 3 – Время, проведенное в сети интернет, с различных устройств [7]

Однако нужно понимать, что все зарегистрированные пользователи – это далеко не весь перечень устройств, подключенных к сети Интернет.

Если обратить внимание на периодическую таблицу роста производительности различных классов компьютерных систем [8], то станет очевидным, что фиксировать трафик, создаваемый такими классами как суперкомпьютеры, ультрамобильные и

нанокомпьютеры, не возможно. Обусловлено это тем, что они используют протоколы передачи данных, которые сложно либо невозможно отследить. Однако количество нанокомпьютеров или встраиваемых компьютеров растёт с каждым днем, и все они имеют доступ к сети интернет. Этот выход предназначен для управления или оповещения, а не для веб-сёрфинга, что усложняет возможность подсчета количества таких устройств (рис. 4).

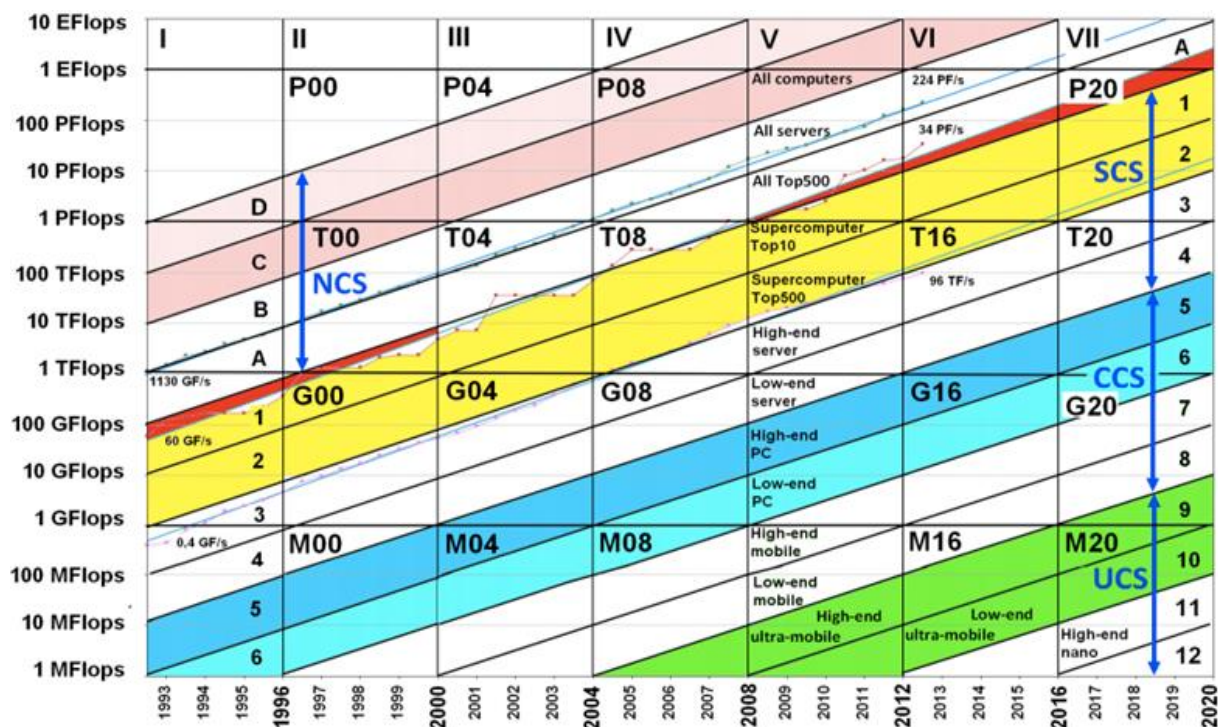


Рисунок 4 – Периодическая таблица роста производительности различных классов компьютерных систем [8]

Экономические закономерности

Учитывая столь высокие темпы развития компьютерных систем, и высокую степень интеграции во все сферы человеческой жизни, нельзя оставить без внимания сферу экономики. В последние десять лет в мировой экономике началась медленная, но уверенная революция. Очень плавный переход от классической экономики к экономике будущего «цифровой экономике».

Цифровая (электронная) экономика – экономическая деятельность, основанная на цифровых технологиях, связанная с электронным

бизнесом и электронной коммерцией, а также с производимыми и сбываемыми ими электронными товарами и услугами. Расчёты за услуги и товары электронной экономики производятся зачастую электронными деньгами.

Стоит также обратить внимание на новость от сайта Финмаркет за 2016 год: «Объем мировых платежей по кредитным и дебетовым банковским картам в 2016 году составит \$23,2 трлн и впервые в истории превысит общую сумму наличных покупок (\$22,6 трлн)», - пишет MarketWatch со ссылкой на исследование компании Euromonitor International» [9] (табл. 3).

Таблица 3 – Рост доли цифровой экономики в ВВП стран G20 с 2010 г. по 2016 г.

Страна	2010 г.	2016 г.
Великобритания	8,3%	12,4%
Южная Корея	7,3%	8,0%
Китай	5,5%	6,9%
Индия	4,1%	5,6%
Япония	4,7%	5,6%
США	4,7%	5,4%
Мексика	2,5%	4,2%
Германия	3,0%	4,0%
Саудовская Аравия	2,2%	3,8%
Австралия	3,4%	3,7%
Канада	3,0%	3,6%
Италия	2,1%	3,5%
Франция	2,9%	3,4%
Аргентина	2,0%	3,3%
Россия	1,9%	2,8%
ЮАР	1,9%	2,5%
Бразилия	2,2%	2,4%
Турция	1,6%	2,3%
Индонезия	1,3%	1,5%
Итого:	64,6%	84,9%

Из таблицы видно, что темп прироста суммарной доли цифровой экономики разных стран в период с 2010 по 2016 год составил 131,43 %. Рассматривая отдельно Россию, из графика (рис. 5) становится видно, что начиная с 2010 по 2017 гг. идет медленный, но уверенный рост доли цифровой экономики.

Составляя прогноз, можно увидеть, что к 2020 г. доля цифровой экономики в России должна увеличиться до 10%.

По мнению экспертов из DEC (digital economy center) темпы развития цифровой экономики будут расти по всему миру [10] (рис. 6).

К примеру, из инфографики видно, что уже в 2025 г. Китай, США и ЕС поровну поделят лидирующие позиции в мировой экономике.

Такое стремительное развитие использования компьютерных ресурсов в экономической сфере неизбежно привело к вопросам о защите и безопасности всей экономической информации, находящейся в мировой сети.

Логическим решением данного вопроса стали криптовалюты. Денежные средства, являющиеся ничем другим, как цифровыми записями во «всемирной записной книжке» – блокчейн, которая к тому же является прозрачной, открытой, общедоступной, но при этом устойчивой ко взломам и подменам информации. Удобства блокчейна заключаются в защищенности; анонимности; распределенном хранении информации; возможности использовать не только для обмена криптовалютами.

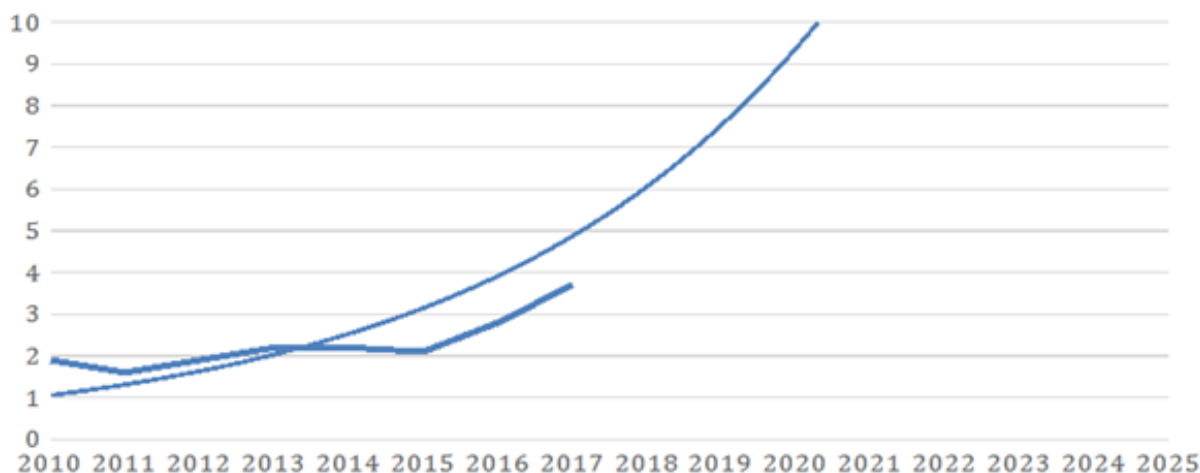


Рисунок 5 – Рост доля цифровой экономики России



Рисунок 6 – Доля цифровой экономики в мире [9]

На рис. 7 представлена схема работы блокчейн, на примере криптовалют.

Принимая во внимание высокую надежность и эффективность, некоторые эксперты утверждают, что блокчейн может стать некоторым стандартом для передачи ценностей, денег, облигаций, договоров и т. п.

Говоря о блокчейне нельзя не упомянуть о том, на чём он основан, а именно функция хеширования данных. На данный момент в основе наиболее популярных блокчейнов лежит алгоритм хеширования SHA-3 (Кессак). Данный алгоритм в течение 5 лет участвовал в конкурсе алгоритмов хеширования, в ходе которого более сотни алгоритмов прошли проверку на скорость выполнения, на коллизии и на детерминированность. По результатам конкурса алгоритм Кессак, удостоился звания «Лучший хэш-алгоритм». Принцип работы алгоритма состоит в том, что все входные данные

переводятся в битовый массив 5×5 . Далее последовательно выполняются перестановки и запутывания этих ячеек. После чего происходит преобразование результатов в хэш.

С учетом роста цифровой экономики, и появлением всё большего количества мобильных и встраиваемых устройств, подключенных к интернету, аналитиками и экспертами всё чаще используется такое понятие как «Интернет вещей», в котором для большинства действий отсутствует необходимость непосредственного участия человека.

Это понятие вполне обосновано, так как количество устройств Интернета вещей уже превысило таковые показатели для обычных устройств, используемых людьми для выхода в сеть [11]. На графике, изображенном на рис. 8, четко видна та тенденция, которая сохраняется на протяжении последних лет.

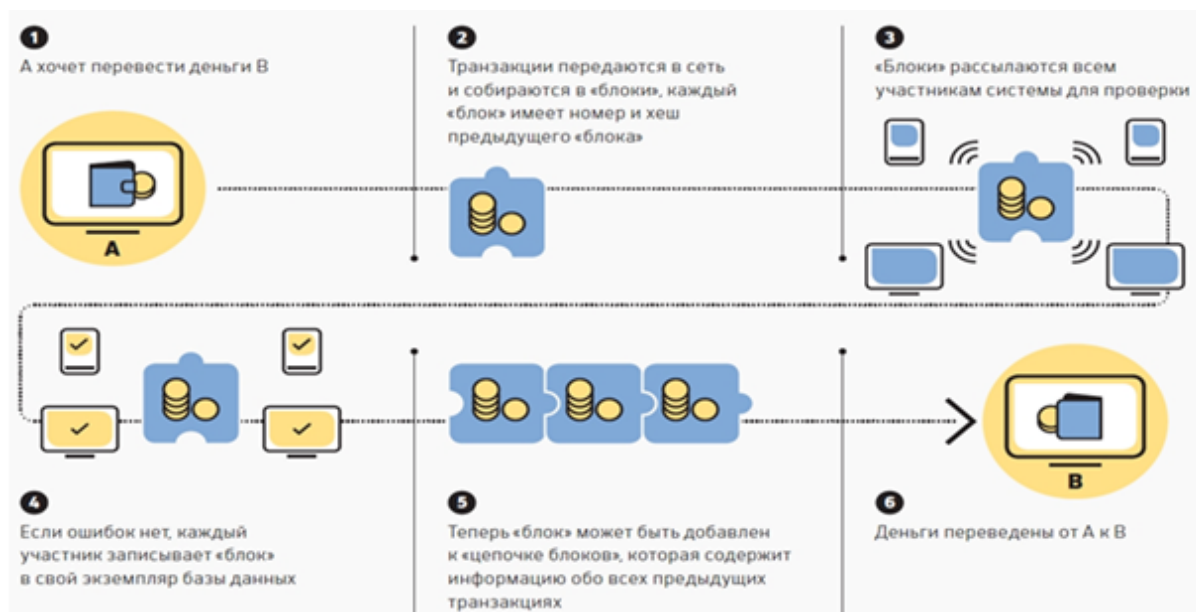


Рисунок 7 – Демонстрация работы блокчейна на примере криптовалют



Рисунок 8 – Число продаваемых во всём мире интернет устройств [12]

Линейно растут только продажи персональных компьютеров. А вот продажи других устройств, имеющих выход в интернет, растут уже экспоненциально. Безусловными лидерами в этой сфере являются носимые устройства, смарт-телевизоры, другая бытовая техника и промышленное оборудование.

Такие системы требуют поиска новых решений, подходов к созданию разумной среды в цифровой экономике [12].

Выводы

В заключении можно предположить, что темпы роста количества устройств, которые имеют выход в интернет, будут постоянно возрастать. Это неизменно приведет к усложнению информационных систем.

В дальнейшем будут появляться новые типы компьютерных систем, которые, в свою очередь, будут решать проблемы, возникающие на почве чрезмерной сложности самих систем и усложненной коммуникаций между ними.

Также не исключено, что подобные изменения в информационных системах потребуют качественного изменения персонала, занимающегося обслуживанием данных систем. Поэтому необходимо делать упор на подготовку квалифицированных кадров, которые через несколько лет смогут разрабатывать и управлять подобными системами.

Литература

1. Информационные технологии в управлении предприятием: Информация. – Санкт-Петербург: СПбГУ. – URL: <https://www.intuit.ru/studies/curriculums/16336/courses/1230/lecture/24061> (дата обращения: 03.11.2018). – Текст: электронный.
2. Аноприенко А. Я. Закономерности развития компьютерных систем // Сб. статей по материалам XVIII международной заочной научно-практической конференции «Научная дискуссия: инновации в современном мире». – М.: «Международный центр науки и образования», 2013. – №10 (18). – С. 19-29.
3. ExtremeTech: What is Moore's Law? By Graham Templeton. – URL: <https://www.extremetech.com/extreme/210872-extremetech-explains-what-is-moores-law> (дата обращения: 06.11.2018). – Текст: электронный.
4. Аноприенко, А. Я. Введение в постбинарный компьютинг. Арифметико-логические основы и программно-аппаратная реализация / А. Я. Аноприенко, С. В. Иваница. - Донецк: ГОУВПО «ДонНТУ», 2017. – 308 с.

5. Internet World Stats: Internet usage statistics. – URL: <https://www.internetworldstats.com/stats.htm> (дата обращения: 16.03.2019). – Текст: электронный.

6. The statistics portal: Percentage of all global web pages served to mobile phones from 2009 to 2018. – URL: <https://www.statista.com/statistics/241462/global-mobile-phone-website-traffic-share/> (дата обращения: 03.05.2019). – Текст: электронный.

7. Яндекс исследование: Развитие интернета в регионах России. – URL: https://yandex.ru/company/researches/2016/ya_internet_regions_2016 (дата обращения: 03.05.2019). – Текст: электронный.

8. Аноприенко А. Я. Периодическая система развития компьютерных систем и перспективы нанокompьютеризации // Инновационные перспективы Донбасса: Материалы международной научно-практической конференции. Донецк, 20-22 мая 2015 г. Том 5. Компьютерные науки и технологии. – Донецк: Донецкий национальный технический университет, 2015. С. 5-13.

9. В 2016 году объем платежей по картам в мире впервые опередит покупки за наличные. – URL: <https://www.interfax.ru/business/530116> (дата обращения: 16.03.2019). – Текст: электронный.

10. Digital Economy Center: Концепция развития цифровой экономики России. – URL: <http://dpfund.ru/consept> (дата обращения: 03.05.2019). – Текст: электронный.

11. Аноприенко А. Я. Закономерности и особенности развития компьютерных систем в контексте «революции криптехнологий» и перспектив постбинарного компьютинга // Материалы IX международной научно-технической конференции в рамках IV международного научного форума Донецкой Народной Республики. Донецк, 22-24 мая 2018 г. – Донецк: Донецкий национальный технический университет, 2018. С. 41-45.

12. Эксперт-Online: Цифроград. – 2015. – «Эксперт Урал» № 35. – URL: <http://expert.ru/ural/2015/35/tsifrograd/media/269105/> (дата обращения: 08.05.2019). – Текст: электронный.

Аноприенко А. Я., Сидоров К. А., Максименко Н. С., Койбаш А. А. Системный анализ закономерностей мирового развития компьютерных систем. Проведен анализ концептуальных, экспоненциальных, экономических закономерностей мирового развития компьютерных систем. В ходе анализа закономерностей выявлено, что темпы развития компьютерных систем остаются на высоком уровне. В области концептуальных закономерностей наблюдается тенденция к усложнению задач решаемых информационными системами. В области экспоненциальных сохраняется стойкий рост производительности компьютерных систем. В сфере экономических наблюдается тенденция постепенного увеличения доли цифровой экономики в мире, что так же требует увеличения количества и качества обслуживающих информационных систем. Выявлены предпосылки к постоянному усложнению компьютерных систем и связей между ними.

Ключевые слова: компьютерные системы, системный анализ, развитие компьютерных систем.

Anoprienko A., Sidorov K., Maksimenko N., Koibash A. System analysis of the laws of the world development of computer systems. The analysis of conceptual, exponential, economic laws of the world development of computer systems is carried out. During the analysis of patterns revealed that the pace of development of computer systems remain at a high level. In the exponential domain, the performance of computer systems continues to grow steadily. In the economic sphere, there is a tendency to gradually increase the share of the digital economy in the world, which also requires an increase in the number and quality of service information systems. Identified prerequisites to a continuous complication of the computer systems and the connections between them.

Keywords: computer systems, system analysis, development of computer systems.

Статья поступила в редакцию 21.11.2019
Рекомендована к публикации профессором Миненко А. С.

Основные принципы и подходы при разработке системы управления профессиональными знаниями сотрудников вуза

Н. К. Андриевская

Донецкий национальный технический университет, г. Донецк

nataandr@yandex.ua

Аннотация

В статье проведен анализ основных подходов и принципов при разработке систем управления знаниями на примере организации – высшего учебного заведения. Приведена и описана функциональная модель системы управления профессиональными знаниями сотрудников вуза в виде Use Case диаграммы, а также выделены основные пользователи информационной системы и разработана их иерархия. Разработана укрупненная метамодель системы управления знаниями сотрудников вуза, ядром которой является онтология. Предложена абстрактная архитектура программного комплекса, базирующегося на разработанной метамодели. В дальнейшем планируется продолжить исследования в заданном направлении.

Постановка задачи исследований

Современные потоки информации характеризуются непрерывным ростом количества электронных документов и их общедоступности, в том числе и в среде Интернет. Слабая структурированность информационных фондов осложняет управление информацией и работу пользователей с ней. Существующие подходы к работе с информацией становятся недостаточно эффективными. Определенным образом это относится к потокам информации, с которыми встречаются сотрудники вузов. Совершенствование существующих и разработка новых подходов к сбору, хранению, обработке и распространению информации является неотъемлемой частью процесса развития информационных технологий (ИТ) и информационных систем (ИС).

В результате своей профессиональной деятельности преподаватели кафедр вуза выполняют подготовку лекционных курсов, подбор материалов для самостоятельного изучения и формирования списка литературы, выполняют поиск информации для своей научной деятельности, а также ведут архивирование документации по результатам учебного процесса и методических разработок. Данные могут быть структурированными (мета описания, онтологии, метаданные библиотечных систем, БД, анкеты, тесты, семантически размеченные файлы и т. п.), слабоструктурированными (данные социальных сетей, шаблонные документы) и неструктурированными (отчеты, ВКР, рефераты, публикации электронных СМИ, электронные письма, сообщения в форумах, информация с сайтов конференций). Сотрудники работают с информационными ресурсами в виде

разнообразных документов, презентаций, наборов данных и БД различного типа, программных пакетов, изображений, аудио и видеофайлов, обеспечивающих качественную организацию учебного процесса.

В настоящее время кафедрами вуза уже накоплен большой объем знаний и информационных ресурсов по различным курсам, по результатам научной и методической работы. Однако эти данные слабо структурированы, плохо систематизированы, рассредоточены по различным ресурсам, библиотекам и архивам, что существенно ограничивает к ним доступ. Более того, по историческим, техническим и другим причинам, тематически связанные данные сохраняются в разных форматах под управлением различных систем хранения и обработки данных. Такое положение дел приводит к тому, что разнообразные коллекции, базы персоналий и публикаций, даже расположенные на одном физическом сервере, зачастую имеют различные логические входы и представляют собой разрозненные автономные информационных ресурсы. Отсутствие связанности информационных ресурсов и унифицированного доступа к ним приводят к неполноте рассмотрения и учета существующих данных и знаний.

Возникает необходимость создания современного интеллектуального инструмента, поддерживающего повседневную профессиональную деятельность преподавателя. Для решения этой задачи необходим переход на качественно новый уровень представления и обработки информации – семантический уровень, что позволит учитывать смысл (содержание) документов, извлекая из них важные для пользователя знания. Средства

представления и интерпретации информации в виде знаний могла бы обеспечивать информационная система, использующая как общие знания о мире, так и знания о предметной области (ПО), которую она обслуживает. Подобные системы относятся к классу систем управления знаниями (Knowledge Management System). Кроме возможностей таких систем по классификации знаний различного вида, они также могут извлекать и даже формировать новые знания. Подобные системы актуальны не только в учебной работе преподавателя, но и в методической работе, в научных исследованиях, а также при разработке различных ИТ проектов. Крайне важно, что подобные системы накапливают корпоративные знания, составляя некоторый интеллектуальный капитал, с которым могут работать другие сотрудники организации.

Для вузов актуальность сохранения интеллектуального потенциала кафедры очевидна в связи с частыми изменениями учебных планов, высокой степенью изменчивости курсов для современных инженерных и компьютерных специальностей из-за быстрых темпов роста прогресса, наличием «высокого порога вхождения» в некоторые дисциплины, а также в связи с текучестью кадров и иногда недостаточным количеством высококвалифицированных кадров. Кроме всего, из-за параллельной подготовки материалов по нескольким курсам, а также из-за узких временных рамок для подготовки, разработчикам дисциплин трудно «идти в ногу» с появлением новых технологий на ежедневной основе. Таким образом, система управления профессиональными знаниями (СУПЗ) должны облегчить процессы подготовки к учебному процессу, предоставляя разработчикам курсов и научным исследователям некоторые рекомендации и знания в их рабочей среде.

Как следует из выше сказанного, основной задачей СУПЗ становится извлечение структурированной информации из неструктурированной и разно-структурированной и ее интеграция, а также кодификация неявных знаний в структурированные БД знаний.

Внедрение такой системы управления знаниями сотрудников вуза призвано повысить эффективность профессиональной деятельности кафедр вуза, обеспечить единообразие извлечения и представления данных и знаний, организационно-административное сопровождение информационного взаимодействия, управление учебной и научно-исследовательской деятельностью.

Целью данной работы является описание функциональной структуры системы управления профессиональными знаниями сотрудников вуза, разработка обобщенной метамодели системы и

абстрактной архитектуры программного комплекса, реализующего основные функции и модели проектируемой системы.

Основные функции системы

Управление знаниями (Knowledge Management) – совокупность процессов и технологий, предназначенных для создания, поиска, распространения, обработки, хранения, извлечения, производства и предоставления для использования знаний. Система управления знаниями (СУЗ, или Knowledge Management Systems – KMS) является корпоративной информационной системой, предназначенной для хранения, генерирования и доставки пользователям полезной информации по вопросам профессиональной деятельности организации.

В последние годы активно ведутся исследования по разработке систем управления научными и архитектурными знаниями. В течение последних нескольких лет было предложено и применено несколько подходов, фреймворков и метамоделей для сбора и управления знаниями как в научных исследованиях, так и в промышленности [1]. Эта тенденция обусловлена главным образом тем, что знания все больше становятся организационным активом [2]. Использование этого актива позволяет повторно использовать знания, поддерживает принятие решений и позволяет избежать «испарения» знаний в организации.

Перспективность данного направления подтверждается результатами исследований таких зарубежных учёных, как D. Ameller and X. Franch [3], M. Bhat [4], P. Kruchten [5], R. Capilla [6], I. Lytra, H. Tran, U. Zdun [7] и др., а также российских ученых: А. Ф. Тузовского [8], В. З. Ямпольского [8], В. А. Лапшина [9], А. Г. Олейника [10], Т. А. Гавриловой [11] и др.

Таким образом, можно констатировать, что актуальность и важность проблематики управления знаниями осознается большинством специалистов, занимающихся корпоративным управлением и информационными технологиями для целей управления. Концептуальное единство распространяется и на существо процесса управления знаниями, важнейшими элементами (этапами) которого признается создание, хранение, поиск, передача (распространение) и использование знаний [8].

Для описания функциональности ИС и взаимодействия с пользователями и внешними информационными системами в нотации UML на практике разработчиками используются Use Case диаграммы [12].

Основные функции проектируемой системы представлены следующим образом (рис. 1):

- работа с онтологиями, мета описаниями и классификаторами;
- работа с семантическими сетями;
- экспертное наполнение хранилища знаний;
- приобретение знаний;
- поиск знаний в Интернете и по хранилищу;
- интеграция знаний и моделей;

- импорт как структурированной, так и не структурированной информации в хранилище;
- проверка данных на корректность;
- проверка документов на уникальность;
- извлечение знаний;
- формирование новых знаний;
- работа с пользователями системы (регистрация и аутентификация).

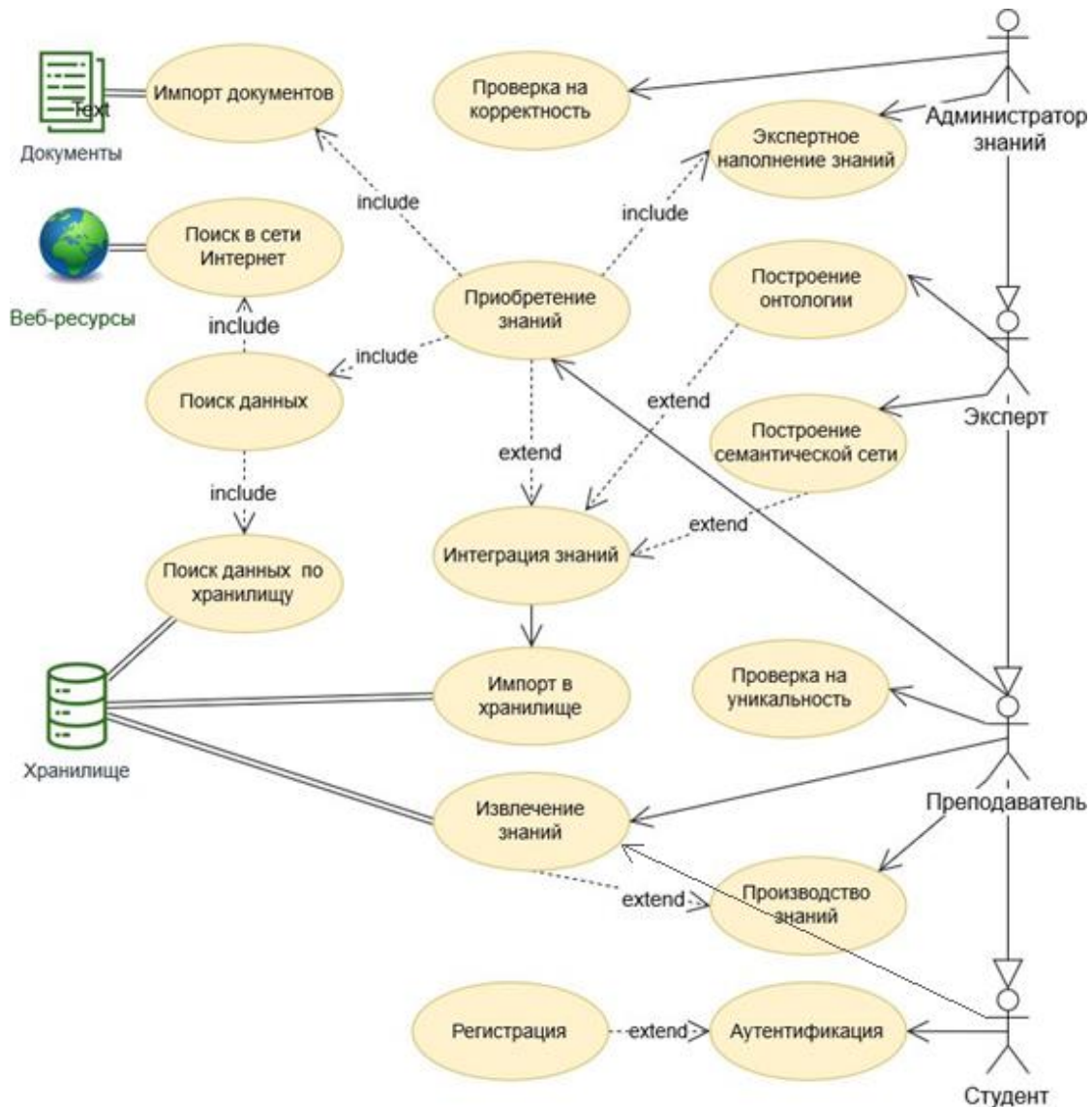


Рисунок 1 – Функциональная модель системы управления профессиональными знаниями сотрудников

С функциями системы работают несколько пользователей. Самым востребованным специалистом на этапе наполнения базы знаний является эксперт. Но на практике оказывается, что чем большим объёмом знаний обладает эксперт, тем сложнее ему структурировать данные и заполнять базу знаний. Помощь ему предоставляет инженер по управлению знаниями (администратор знаний). Но процедуры наполнения базы знаниями

становятся самыми эффективными тогда, когда в работу включается весь коллектив.

Таким образом, проектируемая подсистема содержит ряд функций, связанных с приобретением данных, с извлечением знаний из данных, с поиском знаний по запросам пользователей из хранилища знаний и выдачей результатов (должна предоставлять справочные и архивные материалы, контекстно-зависимую информацию), а также с производством (генерацией) новых знаний.

Мета модель системы управления знаниями сотрудников вуза

Следует отметить, что понятия данные, информация и знания отличаются друг от друга. Данные представляют собой фактическую информацию об объектах, процессах, информация – это результат преобразования и анализа данных, а знания – это структурированные зафиксированные данные плюс некоторые взаимосвязи между ними.

Для учета связанности знаний и обеспечения однородности представления данных в рамках единой тематики проектируемой системы предложена единая концептуальная схема – метамодель (рис. 2). Мета модель объединяет модели предметной и проблемной областей, на основе которой строятся внутренние хранилища знаний системы, а также выполняется поиск и извлечение знаний.



Рисунок 2 – Мета модель системы

Ядром, базовым компонентом метамодели системы является его онтология. Онтологический подход стал активно развиваться приблизительно с начала 90-х годов. Онтология является моделью знаний, которая может использоваться для описания семантики в различных информационных системах. Исследованию онтологий посвящены работы таких известных российских и зарубежных ученых, как: Грубера Т. [13], Гавриловой Т. А. [11], Гуарино Н. [14], Хорошевского Ф. В. [11], Тузовского А. Ф. [8], Ямпольского В. З. [8] и др. По назначению, согласно классификации в [15], онтологии подразделяют следующим образом: онтологии верхнего уровня (top-level ontology), онтологии предметных областей (domain ontology), онтологии задач (task

ontology), онтологии приложения (application ontology). Таким образом, онтология проектируемой системы будет иметь составную структуру, подразделяющуюся на онтологию верхнего уровня вуза и онтологию предметных областей проектируемой системы. В целом она служит для представления понятий, необходимых для описания научной и учебной деятельности, а также описания конкретных учебных дисциплин, в частности. Поскольку онтология системы может быть представлена в виде классов объектов и отношений между ними, то она же задает структуры для представления реальных объектов и событий, имеющих отношение к той же конкретной научной дисциплине или учебному курсу, и обеспечивает их взаимосвязь.

Информационное наполнение системы включает как знания общего характера (представлены в онтологии), так и конкретные знания о реальных объектах и событиях. Добавление новых знаний возможно вручную инженером по знаниям или экспертом, а также получены автоматически с помощью специальных моделей.

В качестве примера приведем основные онтологии разрабатываем системы: «Кафедра», «Сотрудник», «Студент», «Дисциплина», «Направление подготовки», «Образовательная программа», «План», «Научно-исследовательская работа», «Публикация» и др.

Процессы управления знаниями базируются на следующих основных принципах: приобретения, представления, хранения, извлечения знаний (рис. 2).

Разработанная укрупненная метамодель системы управления знаниями сотрудников вуза может быть представлена в шести уровнях.

Уровень выявления знаний – поиск данных из различных источников, таких, как документы и файлы различных форматов, веб-ресурсы, датасеты, БД и др. Информацию, представленную в структурированном виде (метаописания, онтологии, шаблонные факты, семантически размеченные файлы и т. п.) обрабатывать проще, но в таком виде представлено относительно мало информации. Гораздо больший объем информации содержится в неструктурированных данных.

Уровень приобретения знаний – получение знаний, извлечение неформализованных знаний из разнородных источников информации с помощью методов статистической обработки, семантического анализа, технологий TextMining и DataMining, а также экспертных моделей.

Уровень представления знаний – формализация знаний на основе создания онтологии, семантических и других моделей. Сначала выполняется идентификация знаний – первый этап определения системы междисциплинарных связей между элементами знаний используемых предметных областей. Так определяются те знания, которые можно рассматривать как уже приобретенные. Улучшение знаний происходит в процессе семантического поиска в неоднородных распределенных источниках знаний и энциклопедических справочных системах на основе онтологических, case-моделей и моделей успешных прецедентов поисковых запросов. Улучшение знаний можно рассматривать как подзадачу создания новых знаний и как задачу автоматического пополнения базы знаний. Осуществляется также классификация и систематизация знаний, без чего невозможно их

эффективное хранение знаний с целью обеспечения эффективного поиска.

Представление знаний в виде семантических сетей является одной из основных моделей представления знаний. Семантическая сеть – это ориентированная графовая структура, каждая вершина которой отображает некоторое понятие (объект, процесс, ситуацию), а ребра графа соответствуют положениям типа «это есть», «принадлежать», «быть причиной», «входить в», «состоять из», «быть как» и аналогичным между парами понятий. Достоинство использования семантических сетей – наглядность представления знаний. Недостаток – сложность вывода, поиска графа, соответствующего запросу.

Перед помещением явных знаний в репозиторий (хранилище) осуществляется их аннотирование – процесс создания метаописаний. Аннотирование может происходить как с участием человека, так и без него, с помощью специальных алгоритмов. Результатом аннотирования является набор метаописаний, который может помещаться в хранилище метаописаний.

В метаописаниях выделяют три типа:

- системные (служебные) метаданные;
- структурные метаданные;
- семантические метаописания.

Системные метаданные предназначены для функционирования информационных систем и систем управления знаниями. Они включают имена файлов и баз, даты их создания, тип и формат, размер файла и вид носителя и т. п. Структурные метаданные содержат, как правило, справочную информацию об объектах. Это могут быть наименование, статус, структурная принадлежность, профиль и т. п. То есть описания, используемые при идентификации и категоризации объектов в тех или иных целях. Семантические метаописания – особый вид описаний, включающий концептуальное (аннотированное) изложение содержания и смысла информации об объекте [8].

Уровень интеграции данных. На этом этапе выполняется занесение собранных структурированных материалов, онтологий и извлеченных знаний из данных в общее интегрированное хранилище, при необходимости выполняя трансформацию и слияние данных.

Уровень хранения знаний. Так как знания явные и неявные отличаются между собой, различаются и способы их хранения. Явные знания (текстовые документы различных форматов, электронные таблицы, базы данных, тестовые наборы данных, Web-страницы, чертежи, схемы, почтовые сообщения и т. п.) хранятся в специально создаваемых для этой цели хранилищах знаний. Для хранения используются все известные модели данных:

иерархическая, объектная, реляционная и файловая. Неявные знания хранятся, прежде всего, в нейронных структурах головного мозга сотрудников организаций. Перевод неявных знаний в явные происходит посредством процедуры экстернализации с помощью отдельного программного модуля [8].

Уровень извлечения и производства знаний. Проблема поиска решения в базе знаний с использованием моделей типа семантической сети сводится к задаче поиска фрагмента этой сети, соответствующего некоторой искомой подсети. Следует отметить, что в результате поиска мы ищем не сам документ, а сведения об информационном объекте, реальном или виртуальном, т. е. некоторые знания. Для извлечения и производства знаний используются следующие модели: нейронные сети, семантические сети, продукционные и формально-логические модели.

Для создания мультидисциплинарных знаний необходима системная интеграция уже разработанных онтологий различных предметных областей. Задача интеграции знаний связана с целым рядом других задач разработки: баз знаний; множеств объектных, онтологических, нечетких, семантических и аналитических моделей и др.

Архитектура программного комплекса СУПЗ

Абстрактная архитектура программного комплекса СУПЗ приведена на рис. 3 и включает в себя следующие модули:

- модуль поиска информации в сети Интернет;
- модуль поиска информации по хранилищу данных;
- модуль экспертного ввода данных;
- модуль импорта данных из семантически размеченных и шаблонизированных документов;
- модуль интеграции данных и модели;
- интеллектуальный модуль тематической классификации;
- модуль автоматического разбора текста;
- модуль проверки на уникальность;
- редактор онтологий;
- редактор классификаторов;
- редактор метаданных;
- модуль интеллектуального анализа данных и вывода знаний;
- модуль семантической сети;
- модуль продукции новых и комбинированных знаний;
- интерфейс пользователя.

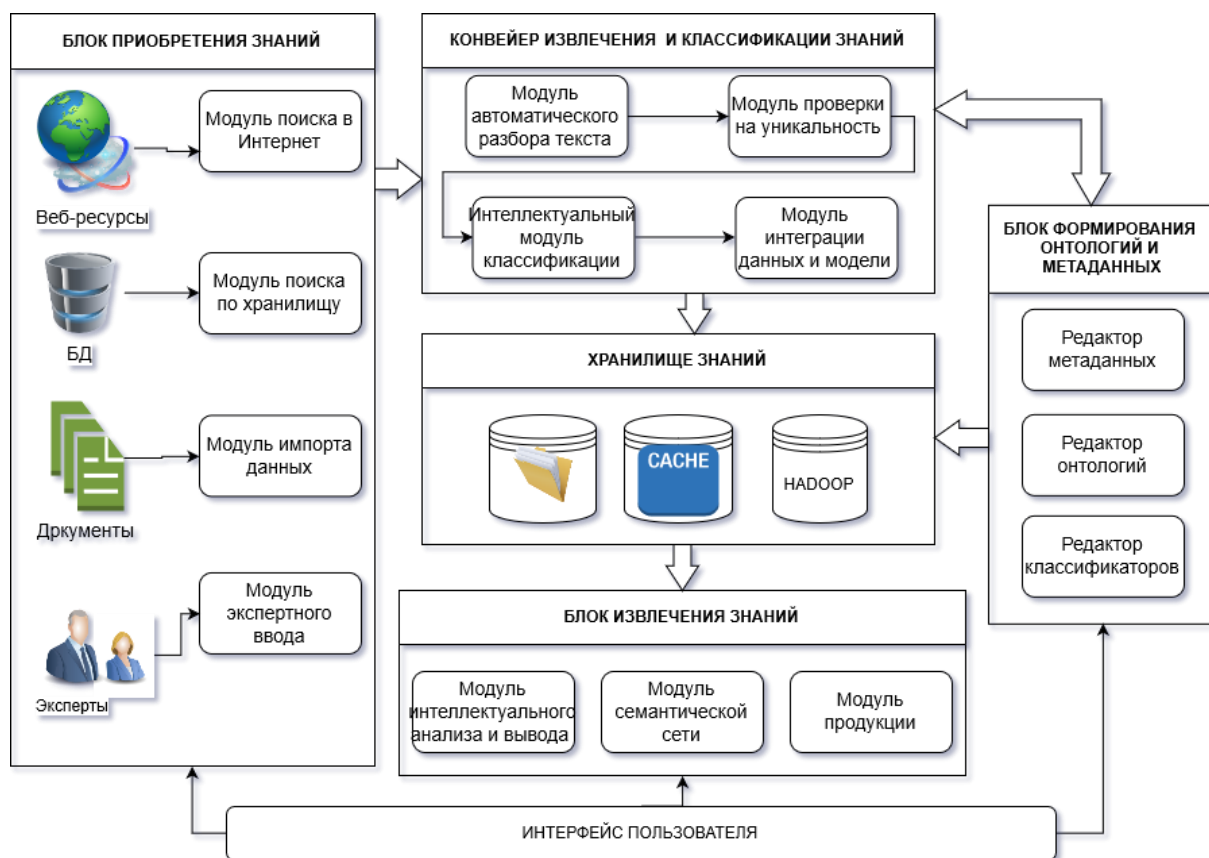


Рисунок 3 – Абстрактная архитектура программного комплекса СУПЗВ

Основными компонентами предлагаемой технологии создания системы являются: блоки приобретения данных, конвейер для обработки и классификации данных, а также блок выдачи и продукции знаний. Настройка и наполнение хранилища выполняется с помощью набора специализированных редакторов онтологий, классификаторов, метаданных. Основными функциональными возможностями редактора онтологий являются создание, модификация и удаление отдельных элементов онтологии, а также работа с иерархией классов.

Внутреннее хранилище знаний, а также навигация по нему, организуются на основе классов и отношений, определенных ранее в онтологии. Разработанные прикладные онтологии вуза будут использованы и при разработке других подсистем, что даст возможность различным системам интегрировать данные, обеспечивать обмен данными на основе онтологий, повторно использовать онтологию и ее совершенствовать.

Поиск в проектируемой системе также осуществляется в соответствии с содержанием онтологии. При поиске информации пользователю предоставляется возможность сформулировать запрос в терминах предметной области и из области метаданных. Навигация по хранилищу также осуществляется в соответствии с содержанием онтологии, позволяя переходить от понятий онтологии к ее экземплярам (информационным объектам), а затем осуществлять переход по онтологическим связям от конкретного экземпляра к спискам связанных с ним экземпляров.

Выводы и направления дальнейших исследований

В этой статье мы представили функциональную структуру и архитектуру основанной на метамодели системы управления профессиональными знаниями сотрудников вуза для информационной поддержки преподавательской и исследовательской работы. Предложен базирующийся на онтологии подход к построению такой системы, обеспечивающий представление и интерпретацию информации в виде знаний. В рамках этого подхода онтология не только обеспечивает представление знаний о предметной области проектируемой системы, но и служит для организации хранилища знаний, а также поддерживает поиск знаний и классификацию знаний. В дальнейшем онтология будет совершенствоваться, расширяться и легко настраиваться на другие области знаний в случае интеграции с другими подсистемами вуза.

Литература

1. Babar M. A. Software architecture knowledge management / M. A. Babar, T. Dingsøyr, P. Lago, H. van Vliet // Springer, 2009.
2. De Boer R. C. Experiences with semantic wikis for architectural knowledge management / R. C. De Boer and H. Van Vliet // WICSA. - IEEE, 2011. - PP. 32–41.
3. Ameller D. Ontology-based architectural knowledge representation: structural elements module / D. Ameller, X. Franch. // Advanced Inform. Syst. Eng. Workshops. - Springer, 2011. - PP. 296–301.
4. Bhat, M. Meta-model based framework for architectural knowledge management / M. Bhat, K. Shumaiev, A. Biesdorf, U. Hohenstein, M. Hassel, F. Matthes // Proceedings of the 10th ECSA Workshops. - ACM, 2016. - P. 12.
5. Kruchten, P. An ontology of architectural design decisions in software intensive systems // 2nd Groningen workshop on software variability. - Citeseer, 2004. - PP. 54–61.
6. Kruchten, P. The decision view's role in software architecture practice / P. Kruchten, R. Capilla, J. C. Duen˜as // IEEE software, 2009. - Vol. 26(2). - PP. 36–42.
7. Lytra, I. Supporting consistency between architectural design decisions and component models through reusable architectural knowledge transformations / I. Lytra, H. Tran, U. Zdun // ECSA.- Springer, 2013. - PP. 224–239.
8. Тузовский А. Ф. Системы управления знаниями (методы и технологии) / А. Ф. Тузовский, С. В. Чириков, В. З. Ямпольский // Под общ. ред. В. З. Ямпольского. - Томск: НТЛ, 2005. - 260 с.
9. Лапшин, В. А. Онтологии в компьютерных системах / В. А. Лапшин. - М.: Научный мир, 2010. - 222 с.
10. Олейник А. Г. Разработка онтологий интегрированного пространства знаний / А. Г. Олейник, П. А. Ломов // Онтология проектирования. - 2016. - Т.6. - № 4 (22). - С. 465–474.
11. Гаврилова Т. А. Инженерия знаний / Т. А. Гаврилова, Д. В. Кудрявцев, Д. И. Муромцев // Модели и методы. - СПб.: Лань, 2016. - 324 с.
12. Светличная В. А. Разработка функциональной структуры логистической системы формирования заказов для интернет-магазина / В. А. Светличная, Н. К. Андриевская, К. Ю. Чаленко // Информатика и кибернетика. - Д.: ДонНТУ, 2017. - № 3 (9). - С. 111–118.
13. Gruber T. Toward Principles for the Design of Ontologies Used for Knowledge Sharing // International Journal of Human-Computer Studies. November 1995. - Vol. 43. - Issues 5–6. PP. 907–928.

14. Guarino N. Formal Ontology in Information Systems / N. Guarino // Formal Ontology in Information Systems. Proceedings of FOIS'98, Trento, Italy, June 6–8. - Amsterdam: IOS Press, 1998. - PP. 3–15.

15. Загоруйко Ю. А. Инженерия знаний: учеб. пособие. / Ю. А. Загоруйко, Г. Б. Загоруйко. - Новосиб. гос. ун-т. – Новосибирск: РИЦ НГУ, 2016. – 93 с.

Андриевская Н. К. Основные принципы и подходы при разработке системы управления профессиональными знаниями сотрудников вуза. В статье проведен анализ основных подходов и принципов при разработке систем управления знаниями на примере организации – высшего учебного заведения. Приведена и описана функциональная модель системы управления профессиональными знаниями сотрудников вуза в виде Use Case диаграммы, а также выделены основные пользователи информационной системы и разработана их иерархия. Разработана укрупненная метамодель системы управления знаниями сотрудников вуза, ядром которой является онтология. Предложена абстрактная архитектура программного комплекса, базирующегося на разработанной метамодели. В дальнейшем планируется продолжить исследования в заданном направлении.

Ключевые слова: знания, управление, функциональная модель системы, метамодель, классификация, семантическая сеть.

Andrievskaya N. Basic principles and approaches for developing the management system of university staff professional knowledge. The article analyzes the main approaches and principles in the development of knowledge management systems on the example of such organization as a higher educational institution. The functional model of university staff professional knowledge management system in the form of a Use Case diagram is presented and described. The main users of the information system are identified and their hierarchy is developed. An enlarged metamodel of the university staff knowledge management system, the core of which is ontology, has been developed. An abstract architecture of the software package based on the developed metamodel is proposed. The future direction is to continue research of developed system.

Keywords: knowledge, management, functional model of the system, metamodel, classification, semantic network.

Статья поступила в редакцию 03.12.2019
Рекомендована к публикации профессором Миненко А. С.

Задача оптимизации энергопотребления гетерогенной сетью LTE в условиях крупного города

Дзюба А. В., Червинский В. В.

Донецкий национальный технический университет, г. Донецк
sevitch77@mail.ru, tscherwi@mail.ru

Аннотация

Предложен принцип снижения энергопотребления в гетерогенной сотовой сети четвертого поколения LTE, основанный на выборе оптимального сочетания включенных базовых станций различного типа при корректном определении метода выбора сот в каждый интервал времени. Проведен анализ особенностей развертывания гетерогенной сети LTE в условиях крупного города. Выделены основные факторы, влияющие на энергопотребление в такой сети. Сформулирован обобщенный критерий оптимизации по энергопотреблению гетерогенной сети LTE. Проанализированы особенности методов выбора сот в таких сетях.

Введение

Продолжающееся увеличение количества мобильных устройств приводит к значительному увеличению объема информации передаваемой между ними. Особенно ярко эта тенденция проявляется в крупных городах. В них регулярно возникают места массового скопления людей, в результате чего периодически резко возрастает нагрузка на базовые станции, обслуживающие абонентские устройства. Как правило, это происходит вблизи железнодорожных и автовокзалов, торговых центров, стадионов, мест проведения массовых мероприятий и в центральных районах городов. Решить задачу качественного обслуживания абонентов в таких условиях могут гетерогенные сети четвертого поколения стандарта LTE. Они представляют собой сложный объект, к которому необходимо применять методы системного анализа [1, 2].

Гетерогенные сети LTE состоят из сот различного размера. Макросоты обеспечивают глобальное покрытие. Пико- и микросоты используют для увеличения ёмкости в общественных местах скопления людей. Кроме того, существуют фемтосоты, покрывающие офисные и жилые помещения. Развёртывание малых сот является ключевой особенностью гетерогенных сетей и обеспечивает их гибкость [3-5].

Значительные расходы электроэнергии являются серьезной проблемой во многих системах. В последнее время вопросы потребления электроэнергии играют все более важную роль и в сотовых сетях. Большинство предложенных решений по повышению энергетической эффективности сотовых сетей основаны на принципе исключения той части системы, которая потребляет наибольшее количество электроэнергии. Например, на

временном отключении или уменьшении потребляемой мощности базовых станций [6].

Такой подход особенно актуален для гетерогенных сетей LTE, развёрнутых в местах периодического массового скопления людей. При этом основным фактором, влияющим на энергопотребление базовыми станциями, является текущая информационная нагрузка на сеть, которая может колебаться в широком диапазоне в зависимости от количества активных абонентов и затребованных ими инфокоммуникационных услуг [7].

Анализ публикаций по проблеме исследования

В последние пять лет проблема энергопотребления в сотовых сетях привлекает особое внимание. В большинстве публикаций отмечают, что подсистема базовых станций сотовой сети потребляет наибольшее количество электроэнергии. В проведенных исследованиях предлагается несколько путей энергосбережения. Например, М. Дерюк рекомендует отключать базовые станции [8], Э. Эмбрози – использовать соты меньшего размера [9], а С. Бойарди – частично отключать услуги [10]. Также К. Самданис, Д. Кутчер и М. Бруннер предлагают при наличии нескольких базовых станций в районах с высокой плотностью населения отключать одну из них в периоды низкой нагрузки [11].

Проведенные с использованием теории систем исследования рассматривают способы повышения энергоэффективности сети за счет отключения части оборудования. Такой подход не учитывает причины увеличения энергопотребления. Одной из них является повышение требований к предоставляемым абонентам услугам. В таких условиях использование стандартных методов

регулирования нагрузки на сеть может ограничить возможности снижения энергопотребления сетью.

Цели и задачи исследования

Целью данной статьи является постановка задачи оптимизации гетерогенной сети LTE для условий крупного города по критерию минимизации энергопотребления оборудованием базовых станций.

Для достижения данной цели необходимо решить следующие задачи:

- провести анализ особенностей развертывания гетерогенной сети LTE для условий крупного города на примере г. Донецк;
- используя системный анализ, выявить основные факторы, влияющие на энергопотребление в гетерогенной сети LTE;

– сформулировать обобщенный критерий оптимизации по энергопотреблению гетерогенной сети LTE

Гетерогенная сеть LTE в условиях г. Донецка

В таком крупном городе, как Донецк, гетерогенная сеть LTE может быть развернута в нескольких местах. Например, в районе площади Павших коммунаров, вблизи от которой находятся автовокзал «Южный», торгово-развлекательный центр «Золотое Кольцо», торговые центры «Юзовский Пассаж» и «Континент». Карта района предполагаемого развертывания гетерогенной сети приведена на рис. 1.

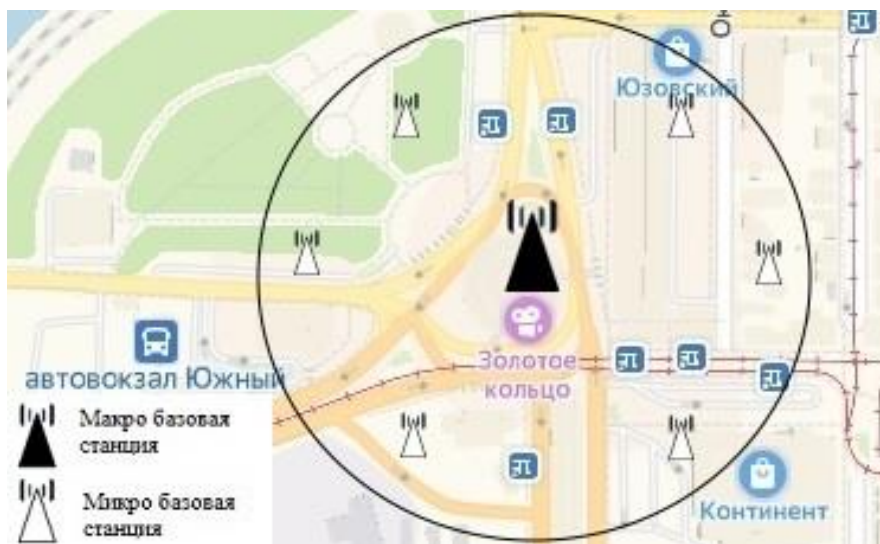


Рисунок 1 – Карта части города Донецка с предполагаемым местом развертывания гетерогенной сети

Рис. 1 демонстрирует одну макросоту с радиусом покрытия 700 м. Поскольку на данной территории в перспективе увеличения количества подключенных к сети LTE пользователей ожидается периодический наплыв большого числа активных абонентов, базовая станция макросоты не будет обеспечивать требуемый уровень качества их обслуживания. В связи с этим, на данной территории можно дополнительно развернуть несколько (на рис. 1 – шесть) микросот с радиусом покрытия менее 50 м.

В настоящее время в сетях LTE используются более 70 частотных полос,

лежащих в диапазоне от 450 до 5 900 МГц. Более низкая частота, как правило, увеличивает площадь покрытия и качество прохождения сигнала в условиях городской застройки, но уменьшает ёмкость сети и скорость передачи данных.

Операторы сетей LTE в Донецке используют преимущественно три частотных полосы: 3, 7 и 20. При этом существует тенденция предпочтительного использования полосы 3 для макро базовых станций и полосы 7 для микро базовых станций. Их сравнительная характеристика приведена в табл. 1.

Таблица 1 – Сравнительная характеристика частотных полос LTE

Полоса	Центральная частота, МГц	Дуплексный режим	Возможная ширина каналов, МГц
3	1800	FDD	1.4, 3, 5, 10, 15, 20
7	2600	FDD	1.4, 3, 5, 10
20	800	FDD	5, 10, 15, 20

Среди малых сот выделяют микро-, пико- и фемтосоты. Их сравнительная характеристика

с традиционными макросотами приведена в табл. 2.

Таблица 2 – Сравнительная характеристика малых сот

Тип соты	Мощность, Вт	Радиус действия, км	Количество пользователей	Расположение
Макро	10 — > 50	8 — 30	> 2000	Вне помещений
Микро	1 — 10	0,2 — 2	100 — 2000	Вне / внутри помещений
Пико	0,25 — 1	0,1 — 0,2	30 — 100	Вне / внутри помещений
Фемто	0,001 — 0,25	0,010 — 0,1	1 — 30	Внутри помещений

Энергопотребление в гетерогенной сети LTE

Для гетерогенной сети LTE, состоящей из сот различного размера общее потребление энергии является суммой потребления энергии всеми её составляющими, а также дополнительных расходов на обслуживание вспомогательного оборудования, вентиляции и охлаждения.

Потребление электроэнергии одной макро базовой станцией в каждый интервал времени составит:

$$E_{\text{macros}} = a \cdot (E_{\text{mac}} + W_{\text{adda}}), \quad (1)$$

где E_{macros} – общее потребление энергии системой макро базовой станции;

E_{mac} – потребление энергии макро базовой станцией;

W_{adda} – дополнительное потребление энергии вспомогательным оборудованием;

a – переменная, которая принимает значение 1, если станция включена, и значение 0, если станция выключена.

Для экономии потребления электроэнергии современные макро базовые станции имеют возможность пошагового диммирования, то есть снижения мощности радиоизлучения в зависимости от количества подключаемых пользователей. Таким образом формула (1) примет вид:

$$E_{\text{macros}} = a \cdot (E_{\text{mac}} + W_{\text{adda}}) \cdot K, \quad (2)$$

где K – коэффициент диммирования.

Тогда общее потребление энергии N_{macro} макро базовыми станциями в каждый интервал времени составит:

$$\sum_{i=1}^{N_{\text{macro}}} a \cdot (E_{\text{mac},i} + W_{\text{adda}}) \cdot K + N_{\text{macro}} \cdot W_{\text{per}}, \quad (3)$$

где W_{per} – потребление энергии постоянно работающими системами вентиляции и кондиционирования.

Потребление электроэнергии одной микро базовой станцией в каждый интервал времени составит:

$$E_{\text{micros}} = b \cdot (E_{\text{mic}} + W_{\text{addb}}), \quad (4)$$

где E_{micros} – общее потребление энергии системой микро базовой станцией;

E_{mic} – потребление энергии микро базовой станцией;

W_{addb} – дополнительное потребление энергии вспомогательным оборудованием;

b – переменная, которая принимает значение 1, если станция включена, и значение 0, если станция выключена.

Тогда общее потребление энергии N_{micro} микро базовыми станциями составит:

$$\sum_{i=1}^{N_{\text{micro}}} b \cdot (E_{\text{mic},i} + W_{\text{addb}}). \quad (5)$$

Общее потребление энергии пико и фемто базовыми станциями находим аналогично микро базовым станциям. При этом у фемто базовых станций отсутствует вспомогательное оборудование.

Таким образом, общее потребление электроэнергии гетерогенной сетью LTE в каждый интервал времени составит:

$$E_{\text{tot}} = \sum_{i=1}^{N_{\text{macro}}} a \cdot (E_{\text{mac},i} + W_{\text{adda}}) \cdot K + N_{\text{macro}} \cdot W_{\text{per}} + \sum_{i=1}^{N_{\text{micro}}} b \cdot (E_{\text{mic},i} + W_{\text{addb}}) + \sum_{i=1}^{N_{\text{pic}}} c \cdot (E_{\text{pic},i} + W_{\text{addc}}) + \sum_{i=1}^{N_{\text{fem}}} d \cdot E_{\text{fem},i}, \quad (6)$$

где E_{pic} – потребление энергии пико базовой станцией;

W_{addc} – дополнительное потребление энергии вспомогательным оборудованием;

c – переменная, которая принимает значение 1, если станция включена, и значение 0, если станция выключена;

N_{pic} – количество пико базовых станций;

E_{fem} – потребление энергии фемто базовой станцией;

d – переменная, которая принимает значение 1, если станция включена, и значение 0, если станция выключена;

N_{fem} – количество фемто базовых станций.

Поскольку макро базовые станции должны работать постоянно, то $a = 1$ всегда. Таким образом, влияя на суммарное потребление электроэнергии, можно, изменяя значение переменных, b , c и d с 0 на 1 при включении и с 1 на 0 при выключении микро, пико и фемто базовых станций соответственно,

а также, изменяя коэффициент диммирования макро базовых станций K .

Минимальное значение потребления электроэнергии будет достигнуто при условии, что, $b = 0$, $c = 0$, $d = 0$, а также коэффициент диммирования K примет минимальное значение. Но в этом случае будет невозможно обеспечивать заданное качество обслуживания абонентов QoS, в частности, соблюдать определенную скорость передачи данных в зависимости от выбранной услуги. Требования к скорости передачи данных для различных услуг представлены в табл. 3.

Таблица 3 – Требования к скорости передачи данных для различных услуг

Вид услуги	Скорость передачи данных, Мбит/с	Вероятность запроса услуги при нагрузке 0,2 Эрл
HTTP	2	0,5
SD-видео	3	0,25
HD-видео	5	0,15
FullHD-видео	8	0,05

Таким образом, для обеспечения заданного качества обслуживания абонентов необходимо в каждый интервал времени соблюсти следующие ограничения:

$$N_{abmac} \cdot a \geq N_{abmac}$$

где N_{abmac} – количество активных пользователей макро базовой станции;

$$N_{abmic} \cdot b \geq N_{abmic}$$

где N_{abmic} – количество активных пользователей микро базовой станции;

$$N_{abpic} \cdot c \geq N_{abpic}$$

где N_{abpic} – количество активных пользователей пико базовой станции;

$$N_{abfem} \cdot d \geq N_{abfem}$$

где N_{abfem} – количество активных пользователей фемто базовой станции;

$$BaseC \geq ExpC,$$

где $BaseC$ – требуемая скорость передачи данных;

$ExpC$ – достигаемая скорость передачи данных.

Достигаемая скорость передачи для каждого пользователя определяется по формуле:

$$ExpC = 0,75 \cdot ExpB \cdot \log\left(1 + \frac{SNR}{1,25}\right) \quad (7)$$

где $ExpB$ – полоса пропускания, выделяемая каждому пользователю, зависит от N_{abmac} ;

SNR – отношение сигнал / шум.

Отношение сигнал / шум влияет на пропускную способность радиосоединений в системах 4G. Это связано с тем, что такие функции, как модуляция более высокого порядка и функция пространственного мультиплексирования MIMO в LTE, требуют эффективного использования SNR .

Методы выбора сот в гетерогенных сетях LTE

Количество включенных станций каждого типа в гетерогенной сети LTE зависит от используемого метода выбора сот пользовательским оборудованием абонентских устройств. Причем эффективность метода может варьироваться в зависимости от прогнозируемого количества абонентов в разное время суток. В настоящее время выделяют четыре основных метода выбора сот в гетерогенных сотовых сетях четвертого поколения LTE:

- метод выбора сот по расстоянию;
- метод выбора сот по уровню сигнал / шум;
- метод выбора сот по полосе пропускания;
- метод выбора сот по энергосбережению.

Если гетерогенная сеть LTE использует метод выбора сот на основе расстояния, пользовательское оборудование подключается к конкретной малой базовой станции, когда оно находится в пределах заданного географического

расстояния от малой базовой станции независимо от расстояния до макро базовой станции и от соотношения сигнал / шум SNR получаемого сигнала от макро базовой станции и микро базовой станции [1]. Предполагается, что покрытие микро базовой станции определяется эффективным радиусом соты микро базовой станции.

В случае метод выбора сот по уровню сигнал / шум мобильные устройства подключаются к базовой станции, который имеет самый высокий уровень отношения сигнал / шум – SNR , независимо от расстояния между абонентским устройством и базовой станцией [2]. Этот метод дает максимально возможную достижимую скорость передачи данных.

Рассмотрим сеть LTE, состоящую из базовых станций в количестве N^{eNB} , поддерживающих пользователей в количестве N_{ab} . При выборе сот на основе SNR пользовательское оборудование UE_i , где $i \in \{1, \dots, N_{ab}\}$, подключается к базовой станции eNB_j , имеющей самый высокий уровень SNR , где $j \in \{1, \dots, N^{eNB}\}$. Таким образом, базовая станция eNB_j выбирается UE_i по условию:

$$eNB_j^i = \max_j \{SNR_j^i\}, \quad (8)$$

где SNR_j^i – уровень SNR , измеренный UE_i от базовой станции j .

В рассматриваемой гетерогенной сети базовые станции имеют различные уровни мощности передачи: макро базовые станции работают с высоким уровнем мощности, и микро базовые станции – с низким уровнем мощности.

Поскольку макро базовая станция передает сигнал на значительно более высоком уровне мощности, чем микро базовая станция, большинство абонентских устройств будут концентрироваться в макросоте.

Таким образом, абонентские устройства стремятся подключаться к макро базовой станции, оставляя микро базовые станции не полностью используемыми. Чтобы лучше использовать микро базовые станции, были предложены методы с добавлением некоторого значения смещения ΔSNR_j к SNR от микро базовой станции для увеличения логического уровня мощности сигнала микро базовой станции. Это приводит к тому, что UE_i выбирает базовую станцию с наибольшим SNR_j^i плюс смещение:

$$eNB_j^i = \max_j \{SNR_j^i + \Delta SNR_j\}. \quad (9)$$

Значение ΔSNR_j равно нулю для макро базовой станции и обычно выбирается в диапазоне 2-20 дБ для микро базовой станции

[3]. Для разных микро базовых станций смещение ΔSNR_j может быть различным. Добавление установленного значения смещения к показателю уровня принимаемого сигнала RSSI (received signal strength indicator) увеличивает диапазон географического охвата микро базовой станции, и, следовательно, к микро базовой станции подключается большее количество абонентских устройств, что делает сеть более сбалансированной. По мере увеличения количества абонентских устройств, подключающихся к микро базовой станции, увеличивается пропускная способность сети и общая скорость передачи данных. Однако такое улучшение зависит от выбранных значений смещения.

Если заданное значение смещения слишком мало, большинство абонентских устройств по-прежнему будут выбирать макро базовую станцию, что приведет к низкой эффективности его разгрузки. С другой стороны, когда заданное значение смещения слишком велико, абонентские устройства будут выбирать микро базовую станцию, даже если они имеют плохой SNR , и эффективность пострадает.

Однако, в результате абонентские устройства, которые подключаются к микро базовой станции, могут получать больше доступной пропускной способностью и, следовательно, более высокую достижимую скорость передачи данных.

Таким образом, при изменении логически заданного значения смещения, абонентские устройства будут подключаться к микро базовой станции даже при условии, что от макро базовой станции они получают более мощный сигнал. Таким образом, изменяя значения смещения, можно увеличить радиус микросот. В результате, трафик перенаправляется к микро базовой станциям, пропускная способность микросот используется больше, а спектральная эффективность возрастает. Следовательно, перевод трафика с мощных узлов (макро базовые станции) на маломощные узлы (микро базовые станции), обеспечивающих меньшее покрытие, может улучшить общее качество обслуживания в сети.

Однако, такое улучшение в значительной степени зависит от корректного выбора значения смещения. Если это значение слишком мало, большинство абонентских устройств останется подключенными к макро базовым станциям, что определит более низкую производительность сети для большинства абонентских устройств. То есть, если к одной и той же базовой станции подключается слишком много пользователей, для каждого из них становится доступной небольшая полоса пропускания. С другой стороны, когда значение смещения слишком велико, каналы связи тех абонентских устройств,

которые находятся далеко от микро базовых станций, будут испытывать большие помехи.

Однако, данный способ не может обеспечить сбалансированность гетерогенные сети LTE, когда доступная пропускная способность используется равномерно как для макро, так и для микро базовых станций.

Поэтому, чтобы поддерживать высокий уровень обслуживания, необходимо учитывать, как производительность сети с точки зрения достижимой скорости передачи данных для всех пользователей, так и возможность сокращения потребляемой энергии.

Таким образом, при выборе сот на основе SNR возникают проблемы перегрузки пользователей из макро в микро базовые станции.

Цель метода выбора сот по полосе пропускания состоит в том, чтобы переместить абонентские устройства из перегруженных базовых станций в те, которые менее нагружены, чтобы получить выгоду из более широкой доступной полосы пропускания и создать возможности для макро базовых станций не работать на полную мощность. Предположим, что все абонентские устройства получают одинаковую пропускную способность и одинаковый тип сервиса. Производительность сети и качество обслуживания измеряются путем сравнения достижимой пользователем скорости передачи данных со средней скоростью передачи данных при работе только с макро базовой станцией.

Абонентский терминал UE_i определяет самый высокий уровень SNR от всех базовых станций, сортирует их по убыванию и запрашивает подключение к базовой станции, которая обеспечивает самый высокий SNR (eNB_k , $k = 1$).

Базовая станция обрабатывает запрос и устанавливает соединение, если количество подключенных абонентских устройств N_k^{UE} не превышает некоторого порога (ПП). ПП устанавливается равным общему количеству абонентских устройств в зоне обслуживания, деленному на число базовых станций N_{UE} / N_{eNB} . Затем выбранная eNB_k обновляет количество подключенных пользователей N_k^{UE} . Если UE_i получает отказ, то попытается подключиться к базовой станции, которая обеспечивает следующий самый высокий уровень SNR. Если все варианты получили отказ, базовые станции (eNB_k , $k = 1, 2, \dots, N_{eNB}$) будут принимать решение в соответствии с количеством подключенных абонентских устройств и подключать абонентский терминал UE_i к той базовой станции, которая может обеспечить наиболее доступную полосу пропускания или

наименьшее количество подключенных пользователей $\min_k (N_k^{UE})$.

Таким образом, выбор сот можно осуществлять на основе полосы пропускания, учитывая количество пользователей для каждой базовой станции в дополнение к SNR. Учет полосы пропускания может решить проблему перегрузки пользователя в макро базовой станции в часы средней и высокой нагрузки и создать возможности для макро базовой станции для снижения энергопотребления. Однако в часы низкой нагрузки использование этого метода имеет недостаток. Выбор сот балансирует нагрузку между макро и микро базовыми станциями независимо от общего спроса. В некоторых случаях общий спрос может быть настолько низким, что не требуется, чтобы микро базовые станции были активными. Все пользователи могут подключаться к макро базовой станции и получать услуги на удовлетворительном уровне.

Целью метода выбора сот по принципу энергосбережения является создание возможности максимально снизить общее потребление энергии за счет максимально возможного выключения микро базовых станций и обеспечения достаточного уровня обслуживания абонентских устройств только при помощи макро базовых станций.

Рассмотрим сеть гетерогенную сеть LTE, которая содержит базовые станции eNB_j , где $j \in [1, N^{eNB}]$ и N^{eNB} – общее количество базовых станций в гетерогенной сети. Гетерогенная сеть LTE содержит одну центральную базовую станцию, которая является макро базовой станцией, обозначенной как eNB_j . Остальные базовые станции являются микро базовыми станциями и обозначаются eNB_j , где $j \in [2, N^{eNB}]$. Обозначим i -й абонентский терминал в момент времени t через UE_i , где $i \in [1, N_t^{UE}]$.

Учитывая модифицированную формулу мощности Шеннона:

$$C = 0,75 \cdot B \cdot \log\left(1 + \frac{SNR}{1,25}\right) \quad (10)$$

где B – полная полоса пропускания [1],

определяется ресурсный блок (RU – Resource Unit) как процент использования ресурса спектра как во временной, так и в частотной области:

$$RU = \frac{reqC}{0,75 \cdot \log\left(1 + \frac{SNR}{1,25}\right)}, \quad (11)$$

где $reqC$ – запрашиваемая скорость передачи данных по требованию пользователя.

Согласно (11), требуемый RU для UE^i это RU^i , когда запрашиваемая услуга требует достижимой скорости передачи данных $regC^i$ и соотношения сигнал / шум $regSNR^i$. При использовании метода выбора сот по энергосбережению, к макро базовой станции подключаются и предоставляются услуги всем пользователям, когда выполняются два условия.

Макро базовая станция eNB_1 должна предоставить услугу всем пользователям:

$$\sum_{i=1}^{N_{UE}} RU^i < 100 \% \quad (12)$$

SNR , которые все пользователи получают при минимальной мощности передачи макро базовой станции, должны соответствовать требованию:

$$SNR_1^i \geq reqSNR^i. \quad (13)$$

Если одно из условий не выполняется, макро базовая станция может переключиться на метод выбора соты по полосе пропускания и включить соседние микро базовые станции. Когда все абонентские устройства подключаются к макро базовой станции по методу выбора сот по энергосбережению, можно отключить все микро базовые станции, то есть для всех микро базовых станций $b = 0$, чтобы минимизировать общее потребление энергии в соответствии с (6). Кроме того, для снижения общего потребления энергии можно уменьшить $QoS (regC^i)$ для некоторых пользователей, чтобы оставить мощность передачи макро базовой станции на минимуме. Макро базовая станция должна увеличить свою мощность передачи для увеличения SNR_1^i , чтобы компенсировать ограничение, если число пользователей увеличивается. Кроме того, благодаря снижению QoS система может поддерживать общее количество $\sum_{i=1}^{N_{UE}} RU^i$ до 100 % и воздерживаться от включения микро базовых станций.

Выводы

В статье показано, что качественное обслуживание абонентов в местах периодического скопления людей могут предоставить гетерогенные сотовые сети четвертого поколения LTE. Одной из главных проблем развертывания таких сетей является высокое потребление электроэнергии. Решить ее можно, используя методы системного анализа, за счет выбора оптимального сочетания включенных микро, пико и фемто базовых станций и коэффициента диммирования макро базовой станции в каждый интервал времени. При этом необходимо избегать недопустимого

снижения качества предоставляемых абонентам услуг. Основой для снижения энергопотребления является корректное определение актуального метода выбора сот в гетерогенных сетях LTE в зависимости от количества активных абонентов, их распределения по территории макросоты и запрашиваемых ими услуг.

Литература

1. Степутин А. Н. Мобильная связь на пути к 6G. В 2 томах. Том 1 / А. Н. Степутин, А. Д. Николаев. – 2-е изд. – Москва-Вологда: Инфра-Инженерия, 2018. – 384 с.: ил.
2. Степутин А. Н. Мобильная связь на пути к 6G. В 2 томах. Том 2 / А. Н. Степутин, А. Д. Николаев. – 2-е изд. – Москва-Вологда: Инфра-Инженерия, 2018. – 420 с.: ил.
3. Соколов Н. Оценка параметров зоны обслуживания сети LTE. Часть 2 (Исправленная) / Н. Соколов [Электронный ресурс]. URL <http://omoled.ru/publications/view/737> (дата обращения: 04.12.2019).
4. Рыжков А. Е. Системы и сети радиодоступа 4G: LTE, WiMax. [Текст] / А. Е. Рыжков, М. А. Сиверс, В. О. Воробьев, А. С. Слышков, Р. В. Шуньков - Санкт-Петербург: Линк, 2012. – 226 с.
5. Тихвинский В. О. Сети мобильной связи LTE. Технологии и архитектура [Текст] / В. О. Тихвинский, С. В. Терентьев, А. Б. Юрчук – Москва: Эко-Трендз, 2010. – 284 с.
6. Бабков В. Ю. Сотовые системы мобильной радиосвязи. (2-е изд. перер. и доп.) [Текст] / В. Ю. Бабков - Санкт-Петербург: БХВ-Петербург, 2013. - 432 с.
7. Скрынников В. Г. Радиоподсистемы UMTS/LTE. Теория и практика [Текст] / В. Г. Скрынников – Москва: Издательство «Спорт и Культура – 2000», 2012. – 864 с.
8. Deruyck M. Modelling and Optimization of Power Consumption in Wireless Access Networks [Электронный ресурс] / M. Deruyck, E. Tanghe, J. Wout, L Martens. – URL: <https://biblio.ugent.be/publication/2040174/file/2040183> (дата обращения: 04.12.2018).
9. Ambrosy A. Dynamic Bandwidth Management for Energy Savings in Wireless Base Stations [Электронный ресурс] / A. Ambrosy, Wilhelm M., Wajda W., Blume O. // Globecom 2012, SessionSAC-GNCS8, Anaheim, Dec. 6th 2012. - URL: <https://pdfs.semanticscholar.org/698e/a8a8e4e3825c449a537be4fead1bf4703fc5.pdf> (дата обращения: 03.12.2019).
10. Boiardi S. Radio planning and management of energy-efficient wireless access networks [Электронный ресурс] / S. Boiardi. – URL: https://www.politesi.polimi.it/bitstream/10589/99707/1/2014_07_PhD_Boiardi.pdf (дата обращения: 04.12.2019).

11. Samdanis K. Self-organized energy efficient cellular networks [Электронный ресурс] / K. Samdanis, D. Kutscher, M. Brunner // Conference: Personal Indoor and Mobile Radio Communications (PIMRC). IEEE 21st International Symposium,

2010. – URL: https://www.researchgate.net/publication/224206537_Self-Organized_Energy_Efficient_Cellular_Networks (дата обращения: 04.12.2019).

Дзюба А. В., Червинский В. В. Задача оптимизации энергопотребления гетерогенной сетью LTE в условиях крупного города. Предложен принцип снижения энергопотребления в гетерогенной сотовой сети четвертого поколения LTE, основанный на выборе оптимального сочетания включенных базовых станций различного типа при корректном определении метода выбора сот в каждый интервал времени. Проведен анализ особенностей развертывания гетерогенной сети LTE в условиях крупного города. Выделены основные факторы, влияющие на энергопотребление в такой сети. Сформулирован обобщенный критерий оптимизации по энергопотреблению гетерогенной сети LTE. Проанализированы особенности методов выбора сот в таких сетях.

Ключевые слова: оптимизация, энергопотребление, критерий, гетерогенная сеть, базовая станция, сота.

Dziuba A., Chervynskiy V. The issue of optimizing energy consumption heterogeneous network LTE in a large city condition. The principle of reducing energy consumption in the heterogeneous cellular fourth generation network LTE is proposed, based on choosing an optimal combination of included different types base stations with the correct cell selection method in each time interval. The deployment features analysis of heterogeneous network LTE in a large city condition was done. The main factors, influencing energy consumption in such a network were allocated. The generalized criteria of optimizing for heterogeneous network LTE energy consumption was formulated. The cell selection methods features in such networks were analyzed.

Keywords: optimization, energy consumption, criteria, heterogeneous network, base station, cell.

Статья поступила в редакцию 14.12.2019
Рекомендована к публикации профессором Миненко А. С.

УДК 519.725, 625.7, 681.3

Альтернативная реализация циклических несистематических кодов на основе регистров конфигурации Галуа и Фибоначчи

О. Н. Дяченко

Донецкий национальный технический университет, г. Донецк

do@donntu.org

Аннотация

Предложена альтернативная реализация циклических несистематических кодов на основе регистров конфигурации Галуа и Фибоначчи. Выполнено моделирование на основе средств САПР Active-HDL различных вариантов реализации кодера, декодера и схемы восстановления информационных символов. На основе моделирования проиллюстрированы преимущества применения конфигурации Галуа и Фибоначчи.

Введение

Успехи современных информационных технологий неоспоримы. Трудно представить нашу жизнь без возможностей Интернета, без умных вещей, без цифрового телевидения, без множества услуг (телеработа, телемедицина, электронная торговля, дистанционное образование и т. д.) Сейчас человечество столкнулось с огромным количеством информации с одной стороны, и, с другой стороны, все возрастающей степенью интеграции микросхем. Все чаще используются термины цифровизация, искусственный интеллект, умные вещи (умный дом впервые описал Рэй Брэдбери более 60 лет назад). Об их популярности свидетельствуют не только технические проекты (в США уже занимаются их практической реализацией для умного дома, а Дубай по некоторым прогнозам станет одним из умных городов после EXPO 2020) и художественные произведения в литературе и кино, но даже в анекдотах.

Вместе с тем, существующие на сегодня интеллектуальные системы имеют очень узкие области применения. Например, программы, способные обыграть человека в шахматы, или могут отвечать на вопросы, но не могут поддержать разговор и т. д. Достаточно вспомнить ботов Алису или Софию и др. Им, конечно, далеко до героев фильма «Лучше чем люди» или до Голема XIV из «Кибериады» Станислава Лемма [1]. В своей вступительной лекции робот рассказывает о человеке с трех точек зрения – исторической, технологической и в замыкании на себя. И человек, несмотря на миллионы лет эволюции, оказывается с точки зрения робота, мягко говоря, несовершенным.

Тем не менее, началась гонка за то, чтобы стать мировым лидером в области искусственного интеллекта (ИИ). За последнее время свои национальные стратегии в области ИИ и продвижения ИИ-технологий разрабатывают и реализуют Австралия, Канада,

Китай, Дания, ЕС, Финляндия, Франция, Германия, Индия, Италия, Япония, Кения, Малайзия, Мексика, Новая Зеландия, Скандинаво-Балтийский регион, Польша, Россия, Сингапур, Южная Корея, Швеция, Тайвань, Тунис, ОАЭ, Великобритания, США [2]. 10 октября 2019 года президент России подписал указ «О развитии искусственного интеллекта в Российской Федерации» [3]. Этим указом утверждается Национальная стратегия развития ИИ до 2030 года. Указ подписан в целях обеспечения ускоренного развития искусственного интеллекта в Российской Федерации, проведения научных исследований в области искусственного интеллекта, повышения доступности информации и вычислительных ресурсов для пользователей, совершенствования системы подготовки кадров в этой области. В Москве 08-09.11.2019 прошла международная конференция по искусственному интеллекту AI Journey 2019.

Проблемам ИИ уже давно уделялось много внимания. Научные труды Алана Тьюринга – общепризнанный вклад в основания информатики (и, в частности, – теории искусственного интеллекта). Академик Виктор Михайлович Глушков (в 1964 был избран академиком АН СССР) стал легендой еще при жизни, за рубежом его даже называли Богом советской кибернетики. Институт кибернетики АН Украины пользовался огромнейшим авторитетом, в 60-70 годах его называли «Меккой» кибернетиков. Многие помнят художественный фильм «Кто за стеной?», снятый еще в семидесятые годы прошлого столетия. Многие ученые в области технологической сингулярности делают прогнозы о будущем человечества и мыслящих машин, но с оговоркой: «Что ждёт нас в ближайшем будущем, если нам хватит ума дожить до него?» [4]. Например, американский изобретатель и футуролог Рэй Курцвейл стал автором множества технологических прогнозов.

Свои первые предсказания он опубликовал в книге «Эпоха мыслящих машин», вышедшей в 1990 году. Рэй Курцвейл лауреат множества премий и автор книги «Как создать разум», последний прогноз до 2099 года он сделал в 2015 году. Речь идёт о системах дополненной и виртуальной реальности, носимой электронике, «умной» одежде, электронных ассистентах (таких, как Google, Siri и Cortana), автомобилях с автопилотом и десятках других изменений в привычной жизни.

В условиях зеттабайтной эпохи развиваются нанотехнологии. Приставки мини- и микро-, судя по всему, для цифровой техники, скоро станут на втором месте по сравнению с нано-. Трудно делать прогнозы. Но иногда они сбываются, чаще всего, пессимистичные. Уже многие предвещают мрачные времена для микропроцессоров и «смерть» закона Мура. Очень серьёзные проблемы связаны уже на этапе литографии. При освоении технологических норм от 5 нм и менее, размеры элементов микросхем становятся сравнимы с десятками и даже с единицами диаметров атомов кремния (радиус 111 пм). В настоящее время исследователи AttoLab изучают характеристики при фотонной бомбардировке во временных рамках пико- и даже не фемто-, а сразу аттосекунд. При этом появляются квантовые явления.

Однако есть и оптимистичные прогнозы. Технологический процесс, основанный на применении транзисторов Bizen (Bipolar Zener), намного лучше, чем CMOS (КМОП). Разработчики полагают, что этот техпроцесс перевернет отрасль. Кроме того, «гибель» закона Мура будет отсрочена лет на 10 назад.

В условиях сверхплотной степени интеграции СБИС и огромного количества данных при их обработке, передаче и хранении неизбежны ошибки, вызванные либо дефектами и помехами различного происхождения, либо людьми или роботами. Они могут привести или к выходу из строя, например, какого-нибудь смартфона, планшета, ноутбука, робота или непилотируемого космического аппарата, или к таким последствиям, как, например, авиа-, автокатастрофы или ядерная война.

Таким образом, теория и практика помехоустойчивого кодирования актуальны как для современных, так и для ИТ-систем обозримого будущего.

Для устранения возможных ошибок из-за естественных природных явлений, либо искусственных причин, или дефектов аппаратных информационных средств, для защиты от разрушений, возникающих под действием жесткого космического излучения, используются современные технологии помехоустойчивого кодирования при

проектировании микросхем памяти, методы и средства встроенного самотестирования цифровых систем [5-13].

Одним из способов повышения надежности и тестопригодности СБИС микропроцессоров, устройств на ПЛИС является применение встроенных средств контроля, реализующих методы компактного тестирования. Методы сканирования — это другой подход снижения трудоемкости тестирования цифровых устройств (рис. 1).

В большинстве случаев рассматриваются регистры сдвига с линейными обратными связями (РСЛОС) в конфигурации Галуа. Однако для реализации встроенного самотестирования более предпочтительны регистры конфигурации Фибоначчи.

Цель статьи

Целью статьи является разработка и исследование способов циклического кодирования и самотестирования цифровых схем на основе РСЛОС конфигурации Галуа и Фибоначчи.

Основная идея альтернативной аппаратной реализации циклических кодов заключается в минимизации сумматоров по модулю два, расположенных между разрядами регистра сдвига (внутренние сумматоры по модулю два).

Для этого, в декодерах следует исключить операцию умножения информационных символов на полином X^p (для укороченных кодов — на остаток от деления полинома X^{p+i} , где i — параметр укорачивания, на порождающий полином).

Кроме того, это можно сделать заменой РСЛОС конфигурации Галуа на РСЛОС конфигурации Фибоначчи.

Конфигурация Галуа и Фибоначчи

В настоящее время большинство циклических кодов, используемых для самотестирования электронных устройств, в частности, в ракетно-космических информационных системах, выполняют обнаружение двойных и исправление одиночных ошибок [6]. Это такие коды как циклические коды Хэмминга и не двоичные коды Рида-Соломона. Их актуальность связана с минимальными затратами на декодирование. Вероятность появления множественных ошибок по сравнению с одиночными отказами и сбоями гораздо меньше.

Кроме того, циклические коды Хэмминга — это фундамент, на котором строятся более мощные коды.

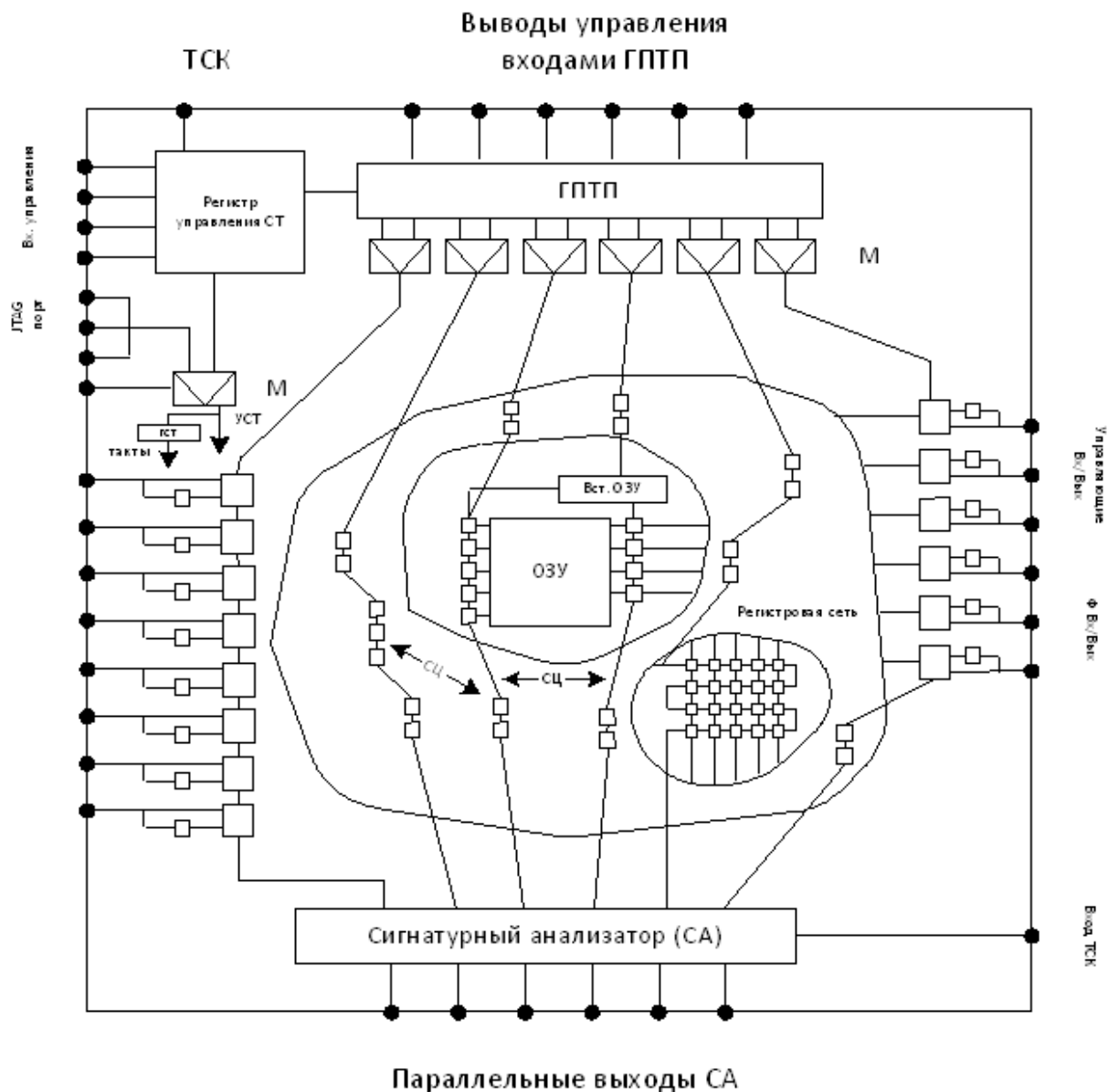


Рисунок 1 – Структура встроенной системы тестового диагностирования микропроцессора КМОП IBM S/390

Поэтому в данной работе рассматриваются методы построения и аппаратная реализация циклических кодов Хэмминга. Большинство результатов будут применимы для циклических кодов с другими обнаруживающими и исправляющими возможностями.

Различают систематические и несистематические коды. В качестве порождающих полиномов кодов используют неприводимые полиномы, для кодов максимальной длины – примитивные.

Кодер систематического кода – РСЛОС выполняющий функцию умножения на полином X^p и деления на порождающий полином, $p = n - k$ количество проверочных символов, n – длина кода, k – количество информационных символов.

Декодер представляет собой также РСЛОС выполняющий функцию умножения

информационных символов на полином X^p (для укороченных кодов – на остаток от деления полинома X^{p+i} , где i – параметр укорачивания, на порождающий полином) и деления на порождающий полином (генератор синдрома).

Кроме того, декодер содержит буферный регистр и схему «И», с помощью которой определяется момент появления на выходе буферного регистра искаженного символа и его исправление. Кодер несистематического кода – регистр, выполняющий функцию умножения информационных символов на порождающий полином.

Для восстановления информационных символов декодер дополнительно содержит РСЛОС, выполняющий функцию деления на порождающий полином.

Рассмотрим кодер и декодер, реализованные в виде регистров сдвига с внутренними сумматорами в цепях обратной связи с порождающим примитивным полиномом пятой степени $X^5 + X^2 + 1$.

Кодер несистематического (или систематического) кода представляет собой схему умножения информационных символов на порождающий полином $X^5 + X^2 + 1$ (рис. 2).

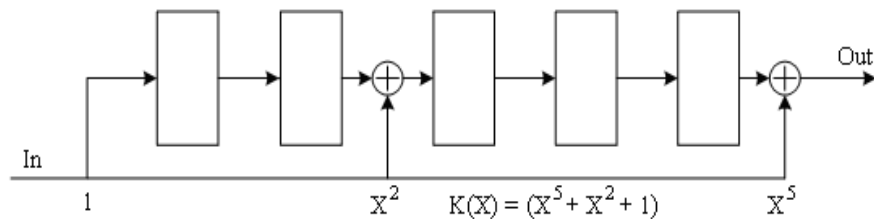


Рисунок 2 – Кодер несистематического кода

Схема декодера несистематического (31, 26) кода Хэмминга выполняет операцию умножения принятого кодового слова на полином

X^5 и деления на порождающий полином $X^5 + X^2 + 1$ (рис. 3).

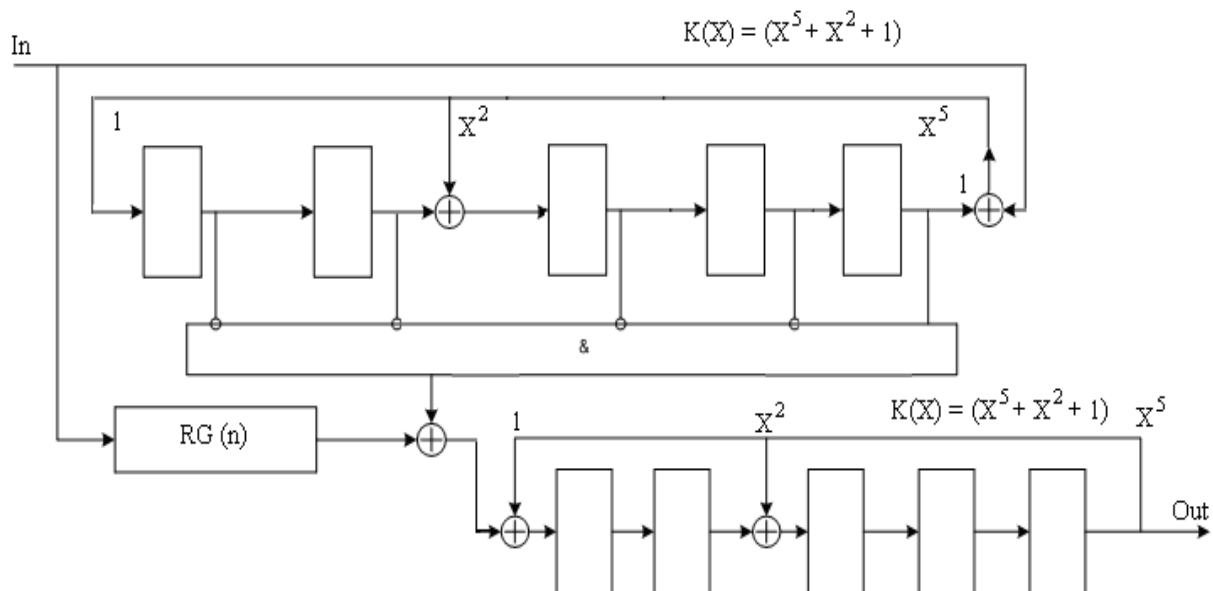


Рисунок 3 – Декодер несистематического кода

Далее исправленное кодовое слово снова делится на порождающий полином для восстановления информационных символов. В данной реализации каждый из трех регистров необходимо разрывать на две части для ненулевого коэффициента X^2 . Следует отметить, что для другого порождающего полинома, например, $X^5 + X^4 + X^3 + X + 1$ таких разрывов будет три для коэффициентов X^4, X^3, X .

На рис. 4 приведен пример реализации кодера, имитации ошибок и декодера (31, 26)-кода в САПР Active-HDL. Для моделирования используются специальные библиотеки, основной особенностью которых является применение элементов без задержек.

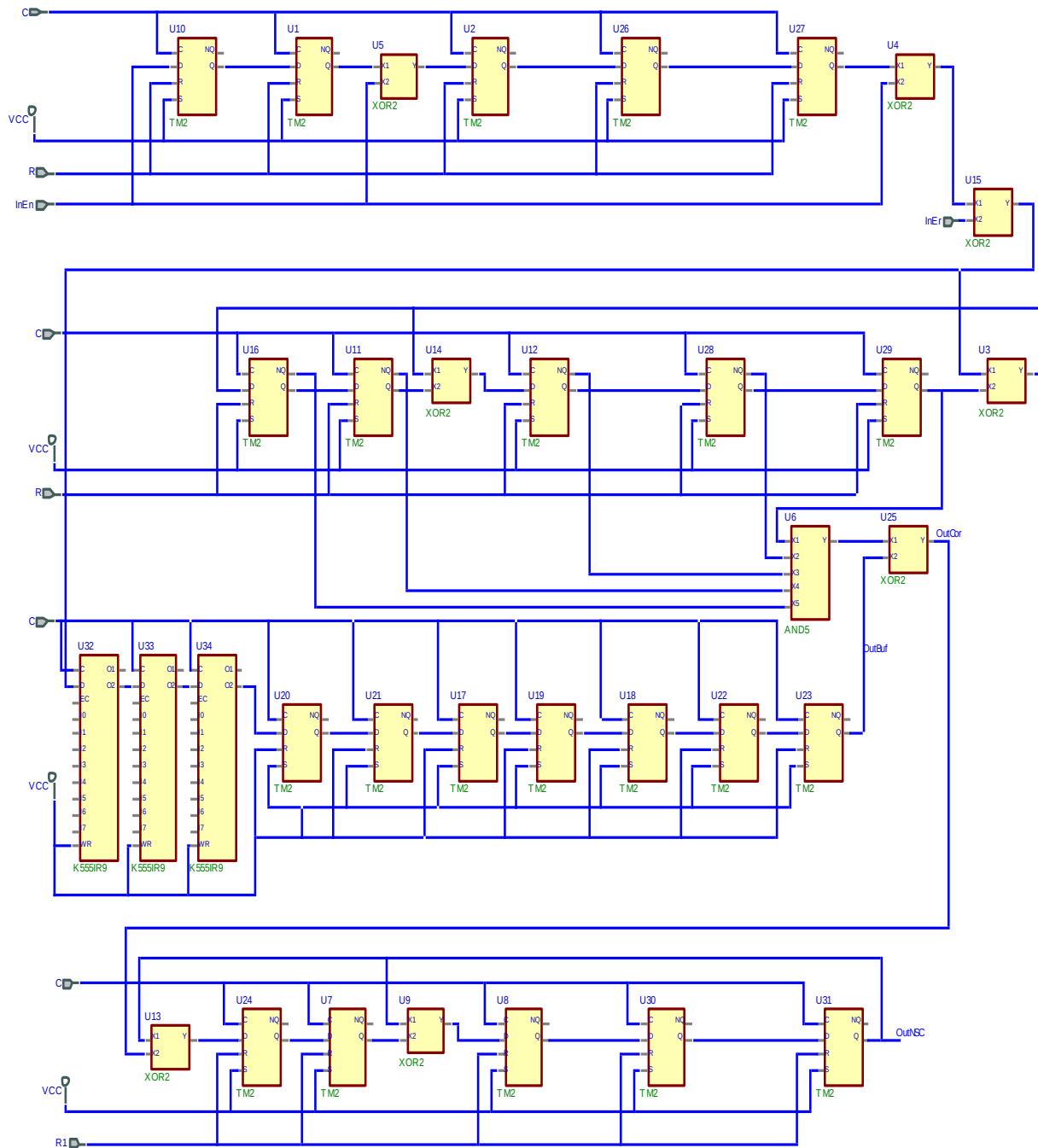


Рисунок 4 – Кодер и декодер несистематического (31, 26)-кода на основе РСЛОС конфигурации Галуа

Для иллюстрации моделирования (31, 26)-кода Хэмминга рассмотрим два варианта его реализации.

Для первого варианта выбран примитивный порождающий полином $X^5 + X^4 + X^3 + X + 1$, схема кодера и РСЛОС декодера – конфигурации Фибоначчи, а РСЛОС восстановления информационных символов конфигурации Галуа. Результаты моделирования показывают совместимость, например, таких вариантов для несистематических кодов:

Галуа – Галуа – Галуа,
Фибоначчи – Галуа – Галуа,
Фибоначчи – Фибоначчи – Галуа,
Фибоначчи – Фибоначчи – Фибоначчи.

На рис. 5 приведен кодер и декодер несистематического (31, 26)-кода на основе регистров конфигурации Фибоначчи – Фибоначчи – Фибоначчи, а рис. 6 демонстрирует организацию кодера и декодера несистематического (31, 26)-кода на основе регистров конфигурации Фибоначчи – Фибоначчи – Галуа. Результаты моделирования приведены на рис. 7-11.

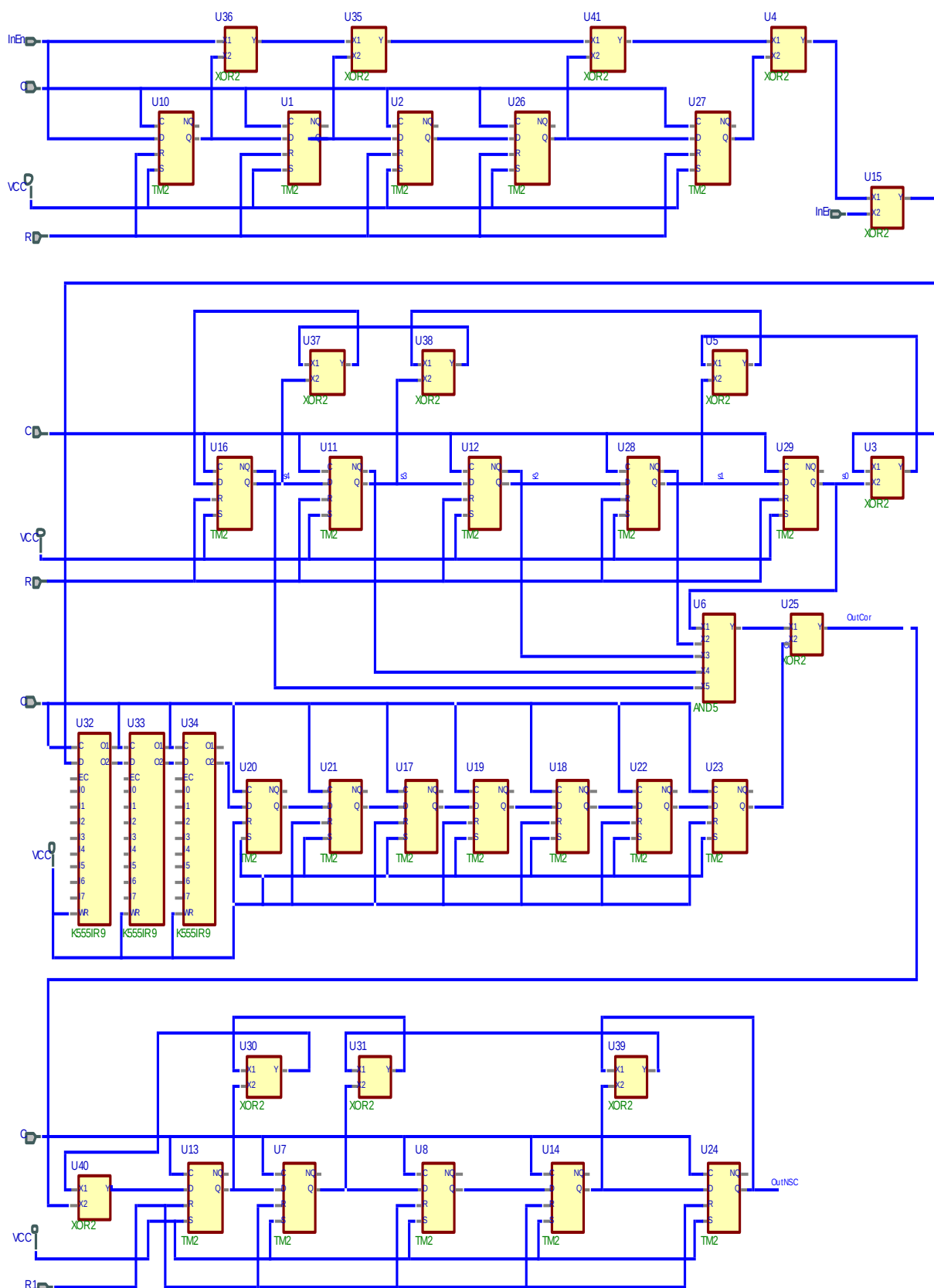
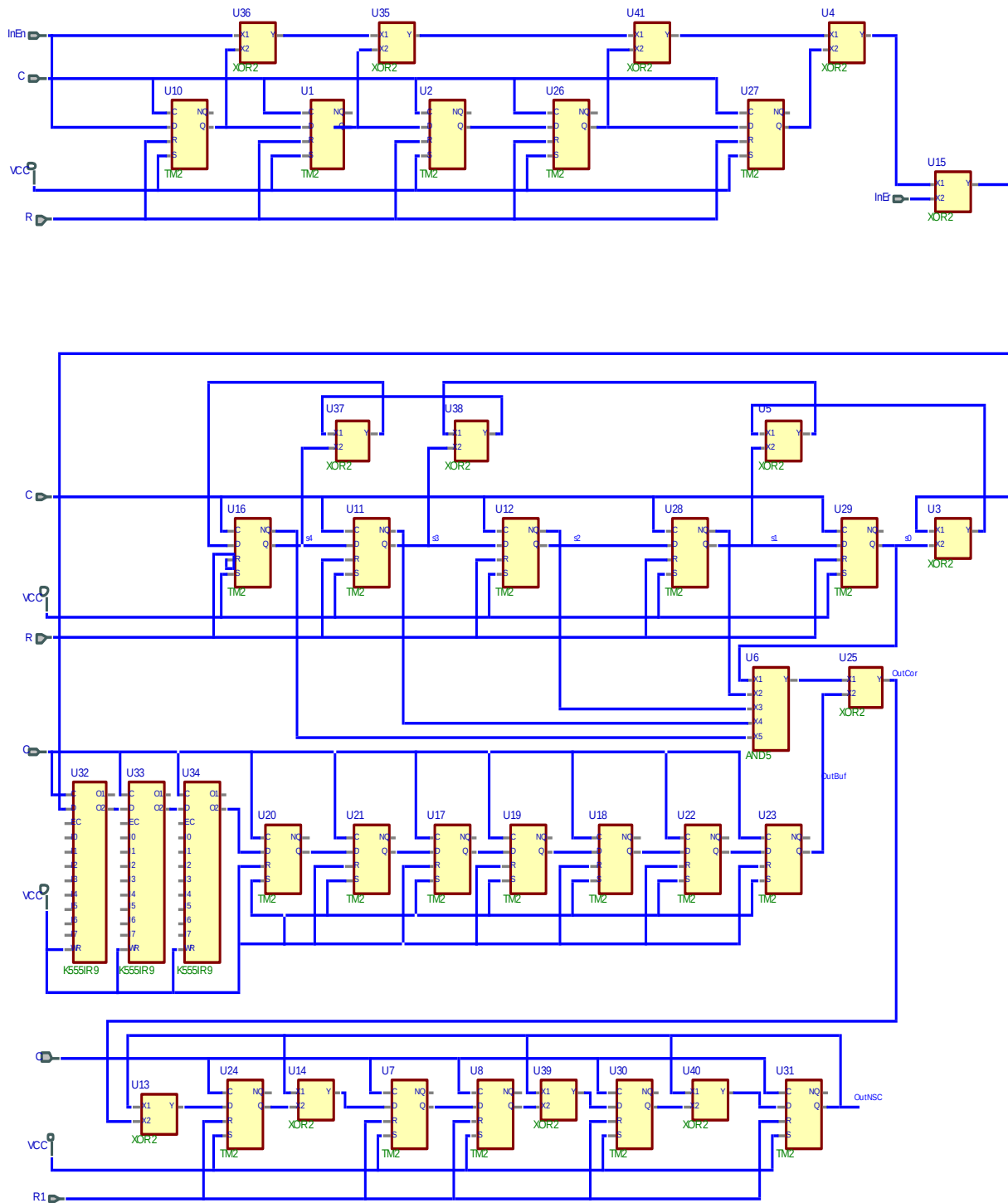


Рисунок 5 – Кодер и декодер несистематического (31, 26)-кода на основе регистров конфигурации Фибоначчи – Фибоначчи – Фибоначчи



Обозначение сигналов:

- C – (clock) синхросигналы;
- R – (reset) сигнал сброса (для всей схемы, кроме схемы восстановления информационных символов);
- R1 – (reset) сигнал сброса (для схемы восстановления информационных символов);
- InEn – вход кодера;
- InEr – вход имитации ошибок;
- OutBuf – выход буферного регистра;
- OutCor – выход буферного регистра после коррекции ошибки;
- OutNSC – выход декодера несистематического кода.

Рисунок 6 – Кодер и декодер несистематического (31, 26)-кода на основе регистров конфигурации Фибоначчи – Фибоначчи – Галуа

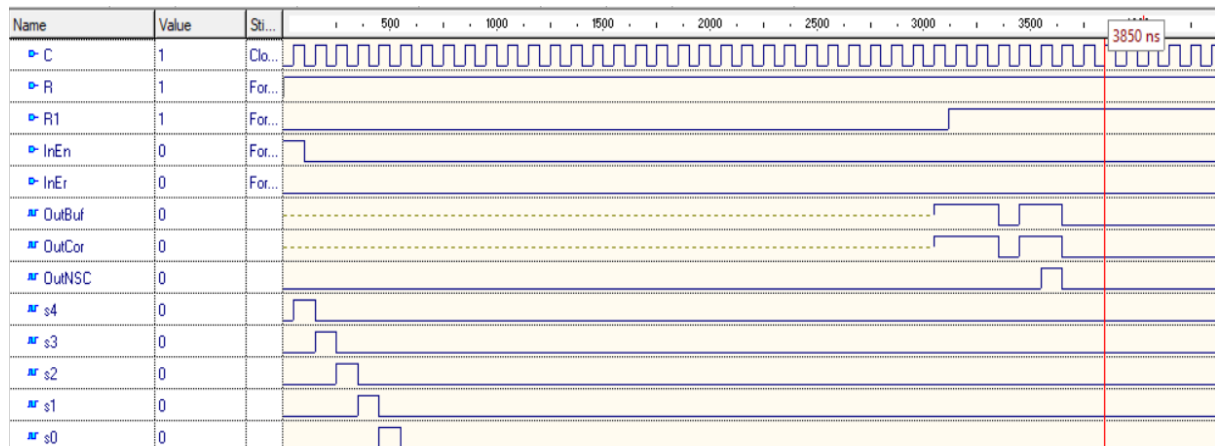


Рисунок 7 – Результат моделирования кодера и декодера несистематического (31, 26)-кода на основе регистров конфигурации Фибоначчи – Фибоначчи – Галуа и Фибоначчи – Фибоначчи – Фибоначчи без имитации ошибок

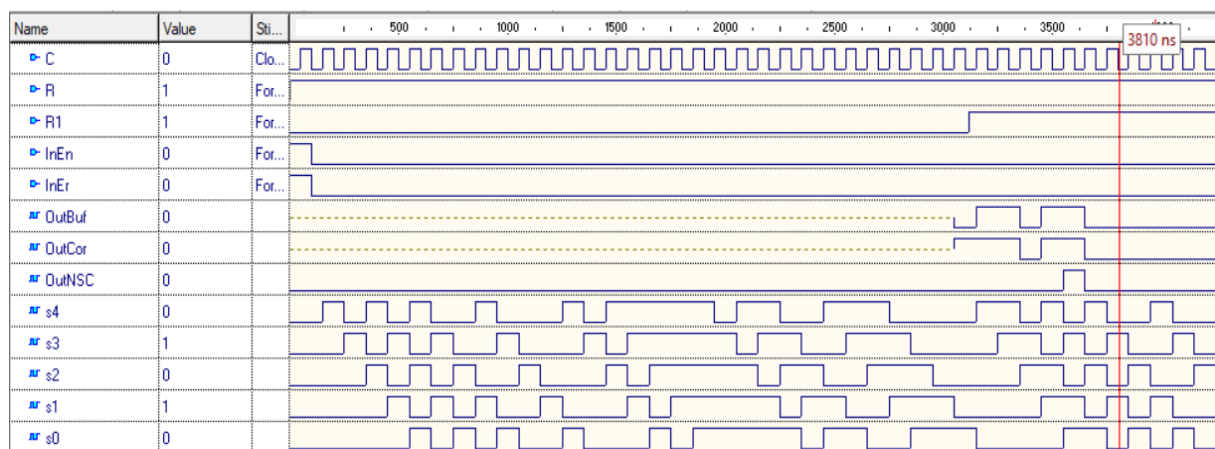


Рисунок 8 – Результат моделирования кодера и декодера несистематического (31, 26)-кода на основе регистров конфигурации Фибоначчи – Фибоначчи – Галуа и Фибоначчи – Фибоначчи – Фибоначчи с имитацией ошибки в первом символе

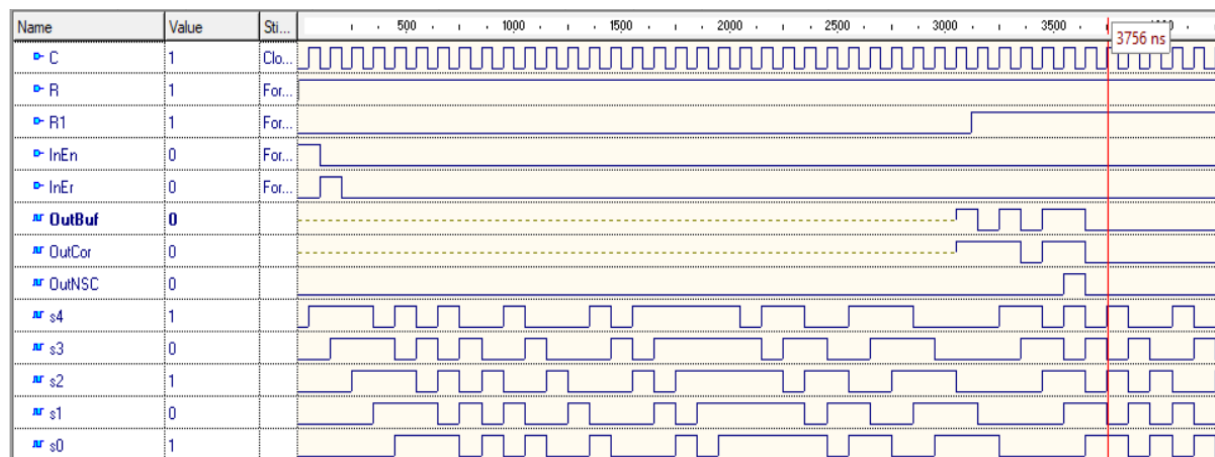


Рисунок 9 – Результат моделирования кодера и декодера несистематического (31, 26)-кода на основе регистров конфигурации Фибоначчи – Фибоначчи – Галуа и Фибоначчи – Фибоначчи – Фибоначчи с имитацией ошибки во втором символе

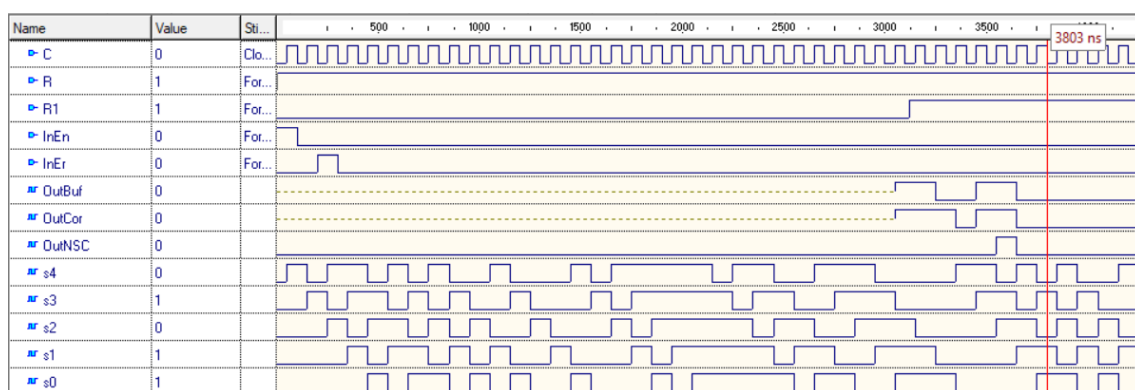


Рисунок 10 – Результат моделирования кодера и декодера несистематического (31, 26)-кода на основе регистров конфигурации Фибоначчи – Фибоначчи – Галуа и Фибоначчи – Фибоначчи – Фибоначчи с имитацией ошибки в третьем символе

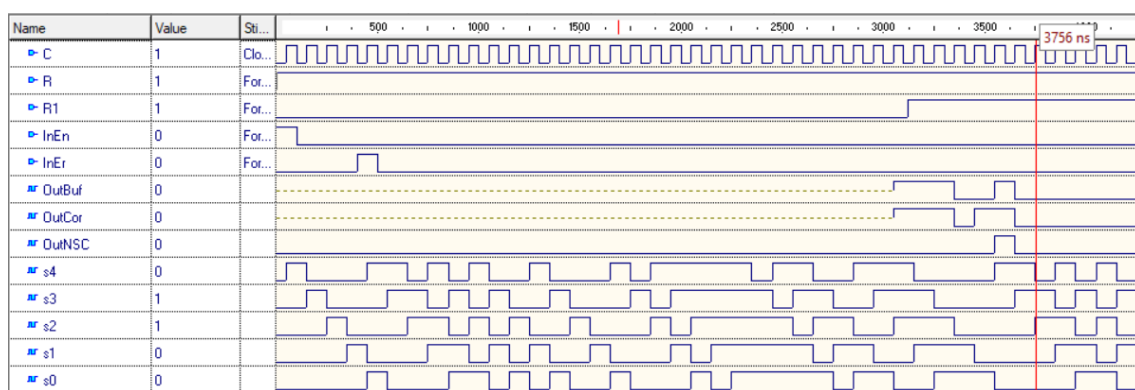


Рисунок 11 – Результат моделирования кодера и декодера несистематического (31, 26)-кода на основе регистров конфигурации Фибоначчи – Фибоначчи – Галуа и Фибоначчи – Фибоначчи – Фибоначчи с имитацией ошибки в пятом символе

Имитация ошибок InEr во всех вариантах – с помощью элемента суммы по модулю два. Используется свойство $A \text{ XOR } 0 = A$; $A \text{ XOR } 1 = \text{NOT } A$, т. е. инверсия символа. Так же выполняется и исправление ошибок.

В качестве входной информационной последовательности символов InEn используется нетривиальная: первая единица и все остальные нули. Таким образом, кодовое слово, которое должно быть на выходе буферного регистра OutBuf при моделировании без имитации ошибок: 111011 для порождающего полинома $X^5 + X^4 + X^3 + X + 1$ и все остальные нули, или 100101 для порождающего полинома $X^5 + X^2 + 1$ и все остальные нули. На выходе декодера OutNSC для обоих случаев – одна единица и все остальные нули.

Буферный регистр не обнуляется. На его выходе после неопределенного состояния появляется принятое кодовое слово. Этот прием используется для наглядности анализа результата моделирования, и не влияет на реальную конвейерную реализацию декодера.

Для того, чтобы исключить неопределенность в схеме восстановления информационных символов, через n тактов с помощью сигнала R1 ее обнуляют.

На выходе буферного регистра после коррекции ошибки OutCor через n тактов появляется исправленное кодовое слово.

Выводы

Регистры сдвига с линейными схемами суммы по модулю два применяются для решения различных задач. Они используются для умножения, деления в отдельных схемах и одновременного умножения и деления в одной схеме, для обнаружения и/или исправления ошибок в запоминающих устройствах на различных носителях, в качестве генераторов псевдослучайных последовательностей, для задач связи с исправлением ошибок (помехоустойчивое кодирование), для задач криптографии и многих других. При разработке этих схем используется математический аппарат Галуа и в большинстве случаев применяются регистры конфигурации Галуа.

Анализ регистров различной конфигурации позволяет сделать вывод, что с точки зрения математической интерпретации более предпочтительна конфигурация Галуа, для аппаратной реализации – конфигурация Фибоначчи.

В дальнейшем представляет интерес анализ различной конфигурации РСЛОС для систематических кодов, а также укороченных кодов.

Литература

1. Станислав Лем. «Вступительная лекция Голема. О человеке трояко» – Режим доступа: http://rulibs.com/ru_zar/nonf_publicism/lem/2/j53.html
2. Tim Dutton. An Overview of National AI Strategies. Jun 28, 2018 – Режим доступа: <https://medium.com/politics-ai/an-overview-of-national-ai-strategies-2a70ec6edfd>
3. Указ Президента РФ от 10.10.2019 N 490 «О развитии искусственного интеллекта в Российской Федерации» (вместе с «Национальной стратегией развития искусственного интеллекта на период до 2030 года») – Режим доступа: <http://publication.pravo.gov.ru/Document/View/0001201910110003?index=0&rangeSize=1>
4. Андрей Васильков. Прогноз развития технологий до 2099 года – Режим доступа: <https://www.computerra.ru/226917/predictions-of-raymond-kurzweil/>
5. Дяченко О. Н., Дяченко В. О. Альтернативный метод укорачивания циклических кодов // Электронные информационные системы. 2017. № 1 (12). С.94–100.
6. Ершов А. Н., Петров С. В., Пятошин Ю. П., Коханько Д. В., Зяблов В. В. и др. Улучшение радиационной стойкости памяти с помощью помехоустойчивых кодов // Ракетно-космическое приборостроение и информационные системы. 2014, том 1, выпуск 4. - С.42–49.
7. Дяченко О. Н. Анализ сигнатурной тестируемости комбинационных схем // Автоматика и вычислительная техника. – 1990. – № 5. С.85-89.
8. Dyachenko O. N. Analysis of signature testability of combinational circuits // Automatic Control and Computer Sciences. - 05/1991; 24(5): 77-81.
9. Дяченко О. Н., Дяченко В. О. Укорачивание циклических кодов на основе альтернативного деления полиномов // Современные тенденции развития и перспективы внедрения инновационных технологий в машиностроении, образовании и экономике: материалы III Международной научно-практической конференции (Азов, 25 мая 2016 г.). – Азов: Изд-во: ООО «АзовПечать», 2016. – С. 45-50
10. Дяченко О. Н., Дяченко В. О. Альтернативный метод укорачивания циклических кодов // Электронные информационные системы. 2017. № 1 (12). – С. 94–100.
11. Дяченко О. Н., Дяченко В. О. Аппаратная реализация кодов БЧХ и кодов Рида-Соломона // Современные тенденции развития и перспективы внедрения инновационных технологий в машиностроении, образовании и экономике: материалы IV Международной научно-практической конференции (Азов, 25 мая 2017 г.). – Ростов н/Д, ДГТУ, 2017. – С. 30-34.
12. Дяченко В. О., Дяченко О. Н. Комплексная оценка компактного тестирования цифровых схем на основе минимальных полиномов // Информатика и кибернетика. Донецк: ДонНТУ, 2018. № 1(11). – С. 36–41.
13. Дяченко В. О., Дяченко О. Н., Зинченко Ю. Е. Анализ эффективности самотестирования цифровых схем на основе неприводимых полиномов. // Информатика и кибернетика. Донецк: ДонНТУ 2018. № 3(13). – С. 5–13.

Дяченко О. Н. Альтернативная реализация циклических несистематических кодов на основе регистров конфигурации Галуа и Фибоначчи. Предложена альтернативная реализация циклических несистематических кодов на основе регистров конфигурации Галуа и Фибоначчи. Выполнено моделирование на основе средств САПР Active-HDL различных вариантов реализации кодера, декодера и схемы восстановления информационных символов. На основе моделирования проиллюстрированы преимущества применения конфигурации Галуа и Фибоначчи.

Ключевые слова: циклический несистематический код, конфигурация Галуа, конфигурация Фибоначчи, порождающий полином, примитивный полином.

Dyachenko O. Alternative implementation of cyclic nonsystematic codes based on Galois and Fibonacci configuration registers. Alternative implementation of cyclic nonsystematic codes based on Galois and Fibonacci configuration registers is proposed. Modeling of various variants of implementation of the encoder, decoder and information symbols recovery scheme was performed on the basis of Active-HDL CAD tools. Based on the simulation, the advantages of using the Galois and Fibonacci configurations are illustrated.

Keywords: nonsystematic cyclic code, generator polynomial, Galois configuration, Fibonacci configuration, generator polynomial, primitive polynomial.

Статья поступила в редакцию 20.11.2019
Рекомендована к публикации профессором Мальчевой Р. В.

**Требования к статьям,
направляемым в редакцию научного журнала
«Информатика и кибернетика»**

Редколлегией принимаются к рассмотрению статьи, в которых рассматриваются важные вопросы в области информатики и кибернетики. Научный журнал издаётся с 2015 года, периодичность издания – 4 раза в год.

В журнале предусмотрены следующие рубрики:

- информатика и вычислительная техника;
- компьютерные и информационные науки;
- инженерное образование.

В соответствии с номенклатурой специальностей научных работников МОН ДНР первые две рубрики соответствуют следующим укрупненным группам специальностей научных работников:

05.01 – «Инженерная геометрия и компьютерная графика»,

05.13 – «Информатика, вычислительная техника и управление».

С 01.02.2019 Научный журнал включён в Перечень рецензируемых научных изданий, в которых должны быть опубликованы основные научные результаты диссертаций на соискание учёной степени кандидата наук, на соискание учёной степени доктора наук (приказ МОН ДНР № 135) по группам специальностей 05.01.00 и 05.13.00.

Рубрика «Инженерное образование» предназначена опубликования сотрудниками научно-методических статей.

Журнал также включён в базу данных РИНЦ (Российский индекс научного цитирования) (лицензионный договор № 425-07/2016 от 14.07.2016).

Статьи, представляемые в данный сборник, должны отвечать следующим требованиям. **Содержание статьи** должно быть посвящено актуальным научным проблемам и включать следующие необходимые элементы:

- постановку проблемы в общем виде, её связь с важными научными и практическими задачами;
- анализ последних исследований и публикаций, в которых решается данная задача и на которые опирается автор, выделение нерешенных ранее частей общей проблемы, которым посвящается статья;
- формулировка цели статьи и постановка задач, решаемых в ней;
- изложение основного материала с полным обоснованием полученных научных результатов;
- выводы и перспективы последующих исследований в данном направлении.

Каждый элемент должен быть выделен соответствующим названием раздела, например, «введение», «постановка задачи», «цель и задачи работы», «цель статьи», «цель исследования», «цель разработки», «анализ ... », «сравнительная оценка ... », «разработка ... », «проектирование ... », «программная реализация», «тестирование ... », «полученные результаты», «выводы», «литература». Разделы «введение», «выводы», «литература» являются обязательными. Включать в названия разделов нумерацию не разрешается.

В основном тексте статьи формулируются и обосновываются полученные авторами утверждения и результаты. Выводы должны полностью соответствовать содержанию основного текста. Языки публикаций: русский, английский.

Объём статьи, формат страницы

Для оформления статьи следует использовать листы формата А4 (210x297 мм) с полями по 2,5 см со всех сторон. Нумерацию страниц выполнять не нужно.

Рекомендуемый объём статьи – 6-12 страниц. Рукописи меньшего объёма могут быть рекомендованы к публикации в качестве коротких сообщений.

Последняя страница текста статьи должна быть заполнена не менее чем на две трети, но содержать не менее трёх пустых строк в конце.

Форматирование текста

Подготовка статьи осуществляется в текстовом редакторе Microsoft Office Word.

Весь текст статьи оформляется шрифтом Times New Roman 10 пт с одинарным междустрочным интервалом, если ниже в требованиях не сказано иного. Абзацный интервал «перед» – 0 пт, «после» – 0 пт.

На первой строке с выравниванием по левому краю располагается УДК.

Заголовок (название) статьи оформляется шрифтом Times New Roman 14 пт, полужирное начертание, с выравниванием по центру (без абзацных отступов). Заголовок статьи следует печатать с прописной буквы без точки в конце, переносы слов не допускаются. Абзацный интервал «перед» – 12 пт, «после» – 12 пт.

После названия статьи следует информация об авторах, которая выравнивается по центру (без абзацных отступов). На одной строке указываются инициалы и фамилии всех авторов через запятую. Между двумя инициалами ставится пробел. С новой строки указывается название вуза (организации) и город (для каждого автора, если не совпадают). На следующей строке указываются адреса электронной почты (один адрес либо каждого автора – по желанию). Адрес электронной почты оформляется в виде гиперссылки.

К тексту аннотации применяется курсивное начертание, с выравниванием по ширине, отступы слева и справа по 1 см. Заголовок «Аннотация» выделяется полужирным начертанием. Объем аннотации – 450-550 символов (без пробелов). Абзацный интервал «перед» – 12 пт, «после» – 12 пт.

Основной текст статьи разбивается на две колонки шириной по 7,5 см (промежуток между столбцами – 0,99 см), выравнивается по ширине. Абзацный отступ первой строки – 1 см. Автоматический перенос слов не применяется.

Заголовки разделов выполняются шрифтом Arial 10 пт, полужирное курсивное начертание. Абзацный отступ отсутствует, интервал перед абзацем – 12 пт, после абзаца – 6 пт. Для заголовка «Введение» установить интервал «перед» – 0 пт, «после» – 6 пт.

Таблицы в тексте статьи

Название следует помещать над таблицей с абзацного отступа (1 см) в формате: слово «Таблица», пробел, номер таблицы, пробел, тире, пробел, название таблицы. Название таблицы записывают с прописной буквы без точки в конце строки и выравнивают по ширине. В ячейках таблицы устанавливается выравнивание текста по центру по вертикали. По горизонтали текст выравнивается по центру либо по левому краю. Границы ячеек таблицы должны быть только чёрного цвета, толщина линии – 1 пт. На все таблицы должны быть приведены ссылки в тексте статьи, при ссылке следует писать слово «табл.» с указанием её номера, например, «... данные приведены в табл. 5». Таблицы нумеруются в пределах статьи. Таблица располагается сразу после ссылки на неё, если это возможно (например, после окончания абзаца). Если же таблица не помещается на текущей странице, то она должна быть расположена в начале следующей страницы (или колонки). При необходимости допускается включение в статью таблицы, ширина которой превышает ширину колонки. В этом случае таблица и её название размещаются по центру страницы. Таблица не должна выступать за границы полей страницы. Таблица и её название отделяются от основного текста статьи одной пустой строкой до и после.

Рисунки в статье

Ссылки на иллюстрации по тексту статьи обязательны и оформляются в виде «... на рис. 2» и т. п. Рисунок и его подпись выравниваются по центру колонки (без абзацных отступов), положение рисунка – «в тексте». Размещается рисунок после его первого упоминания в тексте, если это возможно (например, после окончания абзаца). Если же иллюстрация не помещается на текущей странице, то она должна быть расположена в начале следующей страницы (или колонки). При необходимости допускается включение в статью рисунка, ширина которого превышает ширину колонки. В этом случае рисунок и его подпись выравниваются по центру страницы. Иллюстрация не должна выступать за границы полей страницы. Подпись рисунка оформляется в формате: слово «Рисунок», пробел, номер иллюстрации, пробел, тире, пробел, название рисунка. Название рисунка записывают с прописной буквы без точки в конце строки. Для подписи иллюстрации применяют курсивное

начертание. Иллюстрация и её подпись отделяются от основного текста статьи одной пустой строкой до и после. Не допускается выполнять рисунки с помощью встроенного графического редактора Microsoft Office Word. Если на иллюстрации имеется текст, размер шрифта должен быть не менее чем аналогичный текст, набранный шрифтом Times New Roman 10-го размера. Иллюстрация не должна содержать много незаполненного пространства.

Формулы

Формулы и уравнения рекомендуется набирать с использованием MathType (предпочтительно) или MS Equation. Формулы и математические символы не должны существенно отличаться по размеру от основного текста. Обязательной является нумерация формул, на которые имеется ссылка в тексте статьи. Ссылки в тексте на порядковые номера формул дают в скобках, например, «... согласно формуле (2)». Формулы размещаются по центру колонки, а их номера – по правому краю. Как для строки с формулой, так и для первой строки пояснений (при наличии), абзацный отступ убирается. Первая строка пояснения начинается со слова «где», после которого следует поставить табуляцию на 1 см, затем само пояснение в формате: символ, подлежащий объяснению, пробел, тире, пробел, поясняющий текст, запятая, обозначение единицы измерения физической величины. Пояснения перечисляются через точку с запятой, выравниваются по ширине. Вторая и последующие строки пояснений начинаются с абзацного отступа (1 см). Весь блок текста, связанный с формулой (только формула, несколько формул подряд или формула с пояснениями), отделяется от основного текста одной пустой строкой до и после. Переносить формулы на следующую строку допускается только на знаках выполняемых операций, причем знак в начале следующей строки повторяют. При переносе формулы на знаке умножения применяют знак « $\times$ ». Формулы и математические уравнения могут быть записаны в тексте документа, если их высота не превышает высоту строки. При этом следует учитывать, что знаки математических операций отделяются от чисел или символов пробелами с обеих сторон. Например, «Если учесть, что $y < 0$ и $2x + y = 1$, то из формулы (3) можно выразить x ...». К символам, которые приведены в формуле, при дальнейшем их употреблении (в том числе в пояснениях к формуле) должно применяться курсивное начертание. При этом к любым числам (верхние и нижние индексы, содержащие цифры и т.п.), а также к математическим знакам курсивное начертание не применяется. Не допускается вставлять формулы, выполненные в виде рисунков.

Перечисления: оформление списков

Основной текст статьи может содержать перечисления, оформленные в виде маркированного списка. В качестве маркера элемента списка разрешается использовать только короткое тире «—». Каждый элемент перечисления записывается с новой строки с абзацного отступа, равного 1 см. После символа короткого тире текст располагается с отступом в 1,5 см от левой границы строки, выравнивается по ширине, при переносе на новые строки располагается без отступов. Нумерованные и многоуровневые списки включать в статью не разрешается.

Литература

В тексте статьи обязательны ссылки на все литературные источники, номер источника указывается в квадратных скобках. Ссылки на неопубликованные работы не допускаются. Рекомендуемое количество источников, на которые ссылается автор, не менее 10. Перечень источников приводится в порядке их упоминания в статье. Библиографическое описание каждого литературного источника оформляется в соответствии с ГОСТ Р 7.0.100–2018. Перечень литературных источников оформляется в виде нумерованного списка. В качестве маркеров элементов списка используют порядковые арабские цифры с точкой. Каждый источник представляет собой отдельный элемент перечисления, записывается с новой строки с абзацного отступа, равного 1 см. После порядкового номера с точкой текст располагается с отступом в 1,5 см от левой границы строки, выравнивается по ширине, при переносе на новые строки располагается без отступов.

В конце статьи обязательно приводятся аннотации на русском и английском языках, каждая заканчивается перечнем 5-6 ключевых слов.

К тексту аннотации применяется курсивное начертание, с выравниванием по ширине, отступы слева и справа по 1 см. Слово «Аннотация» опускается. Текст аннотации начинается с ФИО авторов и названия статьи, выделяемых полужирным начертанием. Аннотация на русском языке совпадает с аннотацией, приведенной в начале статьи. В тексте аннотации на английском языке после фамилии автора указывается только первая буква имени с точкой. Абзацный интервал «перед» – 12 пт, «после» – 12 пт. Ключевые слова оформляются с новой строки аналогично тексту аннотации. Заголовок «Ключевые слова:» (англ. «Keywords:») выделяется полужирным начертанием. Ключевые слова перечисляются через запятую.

Порядок представления статьи и сопроводительные документы

В редакцию необходимо представить:

- файл с текстом статьи;
- файл, содержащий фамилию, имя и отчество авторов полностью; ученую степень, ученое звание; место работы с полным указанием должности, подразделения и наименования организации, города (страны); номера телефонов и e-mail для связи;
- экспертное заключение о возможности публикации статьи, подписанное руководителем и заверенное печатью организации, в которой работает автор статьи;
- выписка из заседания кафедры или письмо организации с просьбой об опубликовании и указанием, что изложенные в статье результаты ранее не публиковались.

Статьи и сопроводительные документы следует высылать на электронный адрес infcyb.donntu@yandex.ru.

К сведению авторов

Если статья оформлена с нарушением указанных выше требований и правил, редакция после предварительного рассмотрения может отклонить статью.

На рецензирование статьи направляются членам редакционной коллегии журнала. Все статьи публикуются при наличии положительной рецензии.

В статью могут быть внесены изменения редакционного характера без согласования с автором. Ответственность за содержание статьи и качество перевода аннотаций несут авторы.

Публикация статей в научном журнале «Информатика и кибернетика» осуществляется на некоммерческой основе.

Все номера Научного журнала размещаются на сайте <http://infcyb.donntu.org/>.

CONTENT

Informatics and computer engineering

A regression model for predicting the flash point of diesel in a closed crucible <i>Maksimova A., Ivanova A., Losinsky N.</i>	5
Investigation of the stability of steganographic methods for protecting information using digital watermarks <i>Zavadskaya T., Krakhmal M.</i>	14
Synthesis of a mathematical model for controlling the process of functioning of level crossings based on new means of generating notifications <i>Trunaev A., Cheptsov M., Radkovsky S.</i>	22
A genetic algorithm is used to solve the problems of placing base stations in fifth-generation networks <i>Pavlovskaya K.</i>	29
Improving the effectiveness of protection methods and means for custom mobile applications in the Android operating system <i>Lebedev V., Chernyshova A.</i>	35
System analysis of the laws of the world development of computer systems <i>Anoprienko A., Sidorov K., Maksimenko N., Koibash A.</i>	41
Basic principles and approaches for developing the management system of university staff professional knowledge <i>Andrievskaya N.</i>	49
The issue of optimizing energy consumption heterogeneous network LTE in a large city condition <i>Dziuba A., Chervynskiy V.</i>	57
Alternative implementation of cyclic nonsystematic codes based on Galois and Fibonacci configuration registers <i>Dyachenko O.</i>	65
<u>Requirements to articles which are sent to the editors office of the scientific journal "Informatics and Cybernetics"</u>	75

Электронное периодическое издание

Научный журнал

ИНФОРМАТИКА И КИБЕРНЕТИКА

(на русском, английском языках)

№ 4 (18) - 2019

Ответственный за выпуск Р. В. Мальчева

Технический редактор Р. В. Мальчева

Компьютерная верстка А. И. Воронова

Подписано к выпуску 18.12.2019. Усл. печ. лист. 9,2. Уч.-изд. лист. 5,3.
Адрес редакции: ДНР, 83001, г. Донецк, ул. Артема, 58, ГОУ ВПО «ДонНТУ»,
4-й учебный корпус, к. 36., ул. Кобозева, 17.
Тел.: +38 (062) 301-07-35, +38 (071) 334-89-11
E-mail: infcyb.donntu@yandex.ru, URL: <http://infcyb.donntu.org>