ISSN 2225-5435

Вестник УрФО.

БЕЗОПАСНОСТЬ В ИНФОРМАЦИОННОЙ СФЕРЕ

№ 4(42) / 2021

УЧРЕДИТЕЛИ

ФГАОУ ВО «ЮЖНО-УРАЛЬСКИЙ ГОСУДАРСТВЕННЫЙ УНИВЕРСИТЕТ»

ООО «ЮЖНО-УРАЛЬСКИЙ ЮРИДИЧЕСКИЙ ВЕСТНИК»

ПРЕДСЕДАТЕЛЬ РЕДАКЦИОННОГО СОВЕТА

ЧУВАРДИН О. П.,
руководитель Управления Федеральной службы по техническому и экспортному контролю России по Уральскому федеральному округу

ГЛАВНЫЙ РЕДАКТОР

СОКОЛОВ А. Н.,
к. т. н., доцент, зав. кафедрой «Защита информации», Южно-Уральский государственный университет (национальный исследовательский университет) (г. Челябинск)

ВЫПУСКАЮЩИЙ РЕДАКТОР

СОГРИН Е. К.

ВЁРСТКА

ШРЕЙБЕР А. Е.

КОРРЕКТОР

ФЁДОРОВ В. С.

Подписной индекс 73852 в каталоге «Почта России»

Журнал зарегистрирован Федеральной службой по надзору в сфере связи, информационных технологий и массовых коммуникаций.

Свидетельство ПИ № ФС77-65765 от 20.05.2016

Издатель: **ООО «Южно-Уральский юридический вестник»**

Адрес редакции и издателя: Россия, 454080, г. Челябинск, пр. Ленина, д. 76.
Тел./факс (351) 267-97-01.

Электронная версия журнала в Интернете:
www.info-secur.ru,
e-mail: urvest@mail.ru

16+

Journal of the Ural Federal District. Information security

ISSN 2225-5435

№ 4(42) / 2021

Subscription index 73852

in the «Russian Post» catalog

The journal is registered by the Federal service in the field of communication, information technology and mass communications.

Certificate
PI No. ФС77-65765 dd. 05/20/2016

Publisher: OOO « South Ural Legal Newsletter»

Editorial and publisher address: Russia, 454080, Chelyabinsk, Lenin Avenue, 76
Phone / fax (351) 267-97-01.

Electronic version of the magazine in the Internet:

www.info-secur.ru,
e-mail: urvest@mail.ru

16+

В НОМЕРЕ

ИССЛЕДОВАНИЕ И ПРОЕКТИРОВАНИЕ ТЕХНИЧЕСКИХ СРЕДСТВ

ИНФОРМАЦИОННЫЕ ТЕХНОЛОГИИ И КОМПЬЮТЕРНАЯ БЕЗОПАСНОСТЬ

МЕТОДЫ АНАЛИЗА ДАННЫХ

ОРГАНИЗАЦИОННО-ТЕХНИЧЕСКАЯ И ПРАВОВАЯ ЗАЩИТА ИНФОРМАЦИИ

АКТУАЛЬНЫЕ ПРОБЛЕМЫ КИБЕРБЕЗОПАСНОСТИ

IN THIS ISSUE

RESEARCH AND DESIGN OF TECHNICAL FACILITIES

INFORMATION TECHNOLOGY AND COMPUTER SECURITY

METHODS OF DATA ANALYSIS

ORGANIZATIONAL, TECHNICAL AND LEGAL PROTECTION OF INFORMATION

TOPICAL PROBLEMS OF CYBERSECURITY

УДК 004.056.5

Вестник УрФО № 4(42) / 2021, с. 5–14

Савельев И.А., Антипенко А.О.

DOI: 10.14529/secur210401

ФАЗОВЫЕ ХАРАКТЕРИСТИКИ ГОЛОСОВЫХ ВОКАЛИЗМОВ КАК СЛЕДУЮЩИЙ ВОЗМОЖНЫЙ ЭТАП РАЗВИТИЯ СИСТЕМ ЗАЩИТЫ РЕЧЕВОЙ СВЯЗИ

В статье отмечается, что экспоненциальное увеличение мощности вычислительных систем, в том числе домашнего использования, а также стремительное развитие алгоритмов машинного обучения, особенно в части анализа и синтеза речи, уже привели к появлению новых относительно доступных видов угроз обеспечению конфиденциальности и целостности системам речевой связи. Вместе с тем использование удобных речевых технологий в некоторых специфических задачах, например, биометрической аутентификации говорящего, сильно ограничено из-за возникающих проблем безопасности. В статье подробно рассмотрены характеристики речевых сигналов и приведены их основные модели (Гильбертовская модель, а также модель синусоидального описания речевого сигнала МакАуэля и Куатьери). Автор предлагает обратить внимание на фазовые характеристики голосовых вокализмов как на один из возможных путей купирования новых угроз системам речевой связи. Внимание при написании статьи уделено их практическому использованию – описаны основные области применения, а также представлено программное обеспечение, позволяющее анализировать и использовать фазовые характеристики во многих задачах защиты речевой информации. В выводах автор подчёркивает необходимость продолжения исследования возможной области применения фазовых характеристик, в том числе для надёжной аутентификации диктора.

Ключевые слова: *голосовые вокализмы, защита речевой информации, разборчивость, речевой сигнал, фазовые характеристики, синусоидальная модель.*

PHASE CHARACTERISTICS OF VOICE VOCALISMS AS THE NEXT POSSIBLE STAGE IN THE DEVELOPMENT OF SPEECH COMMUNICATION PROTECTION SYSTEMS

The article notes that the exponential increase in the power of computing systems, including home use, as well as the rapid development of machine learning algorithms, especially in terms of speech analysis and synthesis, have already led to the emergence of new relatively affordable types of threats to confidentiality and integrity of speech communication systems. At the same time, the use of convenient speech technologies in some specific tasks, for example, biometric speaker authentication, is severely limited due to emerging security problems. The article discusses in detail the characteristic of speech signals and presents their main models (Hilbert's model and the model of the sinusoidal description of the speech signal by McAulay and Quatieri). The author proposes to pay attention to the phase characteristics of voice vocalisms as one of the possible ways to stop new threats to speech communication systems. Attention when writing this article is intended for their practical use - the possibility of using software that allows you to analyze and use phase characteristics in many problems of protecting speech information. In the conclusions, the author emphasizes the need to continue the study of a possible field of application of phase characteristics, including for reliable speaker authentication.

Keywords: *voice vocalisms, speech information protection, intelligibility, speech signal, phase characteristics, sinusoidal model.*

Введение

В последние несколько лет в мире наблюдается взрывной рост количества сервисов, использующих в своём составе речевые технологии – это могут быть как распространённые виртуальные ассистенты, способные качественно понимать человеческую речь и синтезировать голосовые ответы на простые вопросы, так и относительно новые системы, способные в интерактивном режиме общаться с человеком (заказать столик в ресторане, записать в парикмахерскую) [1]. Такие технологии легки в освоении и удобны для пользователя. В большинстве своём рост качества анализа и синтеза речи связан с развитием области нейронных сетей, а также экспоненциальным увеличением мощности обрабатывающих серверов.

Вместе с тем речевые технологии могут применяться злоумышленниками для нарушения конфиденциальности и целостности систем речевой связи [2]. При этом использование фазовых характеристик голосовых вокализмов позволит не только купировать эти угрозы, но и добиться повышения качества синтеза речи, а также возможности использовать речевые технологии в новых областях, например, для голосовой аутентификации диктора.

Речь и основные показатели речевых сигналов

Одним из основополагающих аспектов развития человечества стало использование устной речи. Она позволяет передавать прошлый и настоящий общечеловеческий опыт, то есть, фактически, обмениваться знаниями и

достижениями между людьми. Пусть речь – форма общения, сложившаяся под воздействием совокупности исторических правил построения языковых конструкций [3]. Главным компонентом устной речи являются звуки. Благодаря их эффективному распространению в среде передачи, например в воздухе, мы органами слуха (либо иными приёмниками) можем улавливать некоторую информацию.

Из многих характеристик речевого сигнала чаще всего выделяют следующие основные:

1) Временные характеристики;
2) Частотные характеристики;
3) Амплитудные характеристики;
4) Энергетические характеристики.

Под временными характеристиками речи зачастую понимают её темп – скорость произношения различных элементов речи, таких как звуки, слоги, слова [3]. Темп речи считается одним из важных компонентов интонации (существенные части разговора стараются произносить медленнее, чем второстепенные моменты), он сильно зависит от эмоциональной окраски разговора, а также особенностей речевого аппарата оратора. Обычно его измеряют либо количеством произнесённых элементов речи в момент времени, либо средней продолжительностью произнесения отдельного элемента речи. Следует отдельно отметить, что при быстрой речи часто снижается её разборчивость.

Частотная характеристика (диапазон) речевого сигнала среднестатистического человека лежит в области от 120 до 400 герц для женщин и от 80 до 150 герц для мужчин. Вместе с тем основную информацию несут лишь 10% от этих интервалов – данный диапазон называют диапазоном разговорного голоса [4]. Анализ частотного диапазона индивидуума позволяет косвенно оценить как его психическое, так и эмоциональное состояние.

Амплитудные характеристики речи отражаются на громкости звука, то есть, фактически, на субъективном восприятии его силы. Международной организацией по стандартизации (ISO) введена единица абсолютной шкалы громкости – сон, который равен громкости чистого непрерывного синусоидального тона частотой в 1 кГц и создающего звуковое давление равное 2 мПа. На практике чаще в качестве величины звукового давления используют не паскали, а децибелы (дБ):

$$P = 2 * 10^{-5}\ \text{Па}$$

$$P_{\text{дБ}} = 20\lg \frac{P_{\text{Па}}}{P_0}\ \text{дБ}$$

Исходя из вышеприведённых формул под децибелами понимают отношение величины

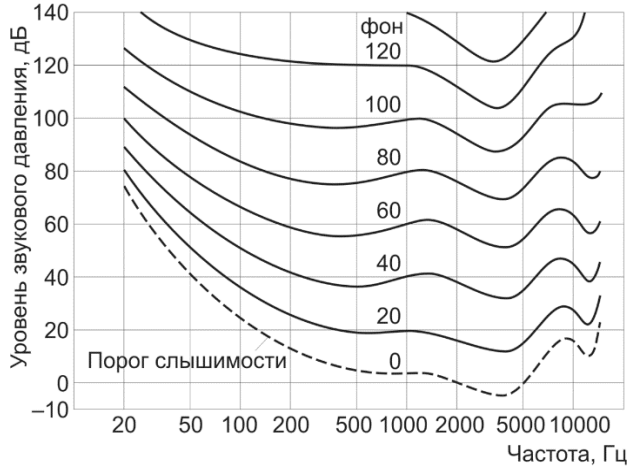

Рис. 1. Зависимость уровня громкости от звукового давления и частоты

звукового давления P к некоторому пороговому значению. Существует также и относительная величина громкости звука, выражаемая в фонах. Согласно российскому стандарту ГОСТ Р ИСО 226–2009 «Акустика. Стандартные кривые равной громкости» (аналог международного стандарта ISO 226) вводятся изофоны – кривые равной громкости, значения которых схематично представлены на рисунке 1.

Представленный график отражает зависимость уровня звукового давления от частоты при определённом уровне громкости, позволяя, тем самым, легко сопоставить уровень создаваемого звукового давления и уровень громкости тона определённой частоты. Например, нам дана звуковая волна с частотой 500 Гц и уровнем звукового давления в 120 дБ, тогда по графику легко определить, что громкость звука равна 120 фон. Под

нулём фон понимают порог слышимости человеческого уха, на графике его изображают пунктирной линией [5].

В группу энергетических характеристик речевого сигнала обычно включают плотность энергии и интенсивность звука. Сумму энергии звуковых колебаний, находящейся в единице объёма, называют плотностью энергии и измеряют в Дж/м^3. Количество энергии, проходящей за единицу времени через единицу площади, перпендикулярной к направлению распространения, называют интенсивностью (силой) звука – её измеряют в Вт/м^2. Интенсивность речевых сигналов зависит от амплитуды колебания голосовых связок человека и их напряжённости. Интенсивность звука снижается в случае уменьшения амплитуды колебаний. При этом различают несколько уровней интенсивности – от низкого до высокого. Интенсивность может быть как постоянна на отрезке времени, так и динамически меняться (плавно или резко).

Вместе с тем к основным характеристикам речи зачастую не относят фазовые характеристики, о которых и пойдёт речь далее – до недавнего времени исследователи практически не уделяли им внимания, за исключением нескольких работ [6]. Это связано со сложностью их вычисления и неясностью перспектив практического использования. Однако некоторые косвенные данные свидетельствуют, что включение именно этих характеристик в соответствующие математические модели речи позволит существенным образом повысить точность этих моделей, расширить их область применения, а также решит некоторые другие задачи защиты речевой информации.

Модели описания речевых сигналов

В целях выделения характеристик речевого сигнала и последующего приведения его к подходящей форме для анализа-синтеза разработано большое количество различных математических моделей, таким образом каждая из моделей удобна для решения своей узкой технической задачи.

Например, для целей искусственного синтеза голоса широкое распространение получили нейролингвистические модели, основанные на описании и моделировании участков речи с применением преобразования Фурье [7]. Однако необходимо отметить их существенный недостаток (как и подавляющего большинства других математических моделей речевого сигнала) – они либо вовсе не учитывают фазовые характеристики, либо моделируют их искусственно.

Для исключения вышеназванного недостатка авторы предлагают использовать преобразования Гильберта. В результате вокализованные участки речи можно представить следующим выражением [8]:

$$S(t) = \sum_{k=1}^{\infty} a_k(t) \cos k2\pi ft$$

где k – номер гармоники, $a_k(t)$ – амплитуда k-ой гармоники основного тона, af – частота основного тона.

Для невокализованных звуков необходимо использовать другую формулу:

$$S(t) = \int_0^{\infty} a_\omega(t) \cos \omega t \, dt$$

в которой – спектральная амплитудная плотность на частоте .

Кроме того, часто используют иное описание речевого сигнала [8]:

$$S(t) = a(t)\cos\varphi(t)$$

в котором a(t) и $\varphi(t)$ – мгновенная амплитуда и мгновенная фаза соответственно, определяемые преобразованиями В.И. Коржикова, Гильберта или В.И. Тихонова.

С помощью преобразования Гильберта речь можно представить в аналитическом виде и таким образом найти значения параметров (опорных точек), отвечающих за разборчивость, однако при таком подходе фаза не является определённой по времени и частоте, что сильно усложняет вычисление начальной фазы. Фактически традиционные модели описания речевых сигналов не подходят для решения задач вычисления и дальнейшего анализа фазовых характеристик–для построения эффективного процесса требуется нахождение и использование нового алгоритма.

Наиболее исчерпывающее объяснение синусоидальных моделей описания речевых сигналов, которые также включают низкие частоты, были представлены в исследованиях МакАуэля и Куатьери. Синусоидальная модель, предложенная МакАуэлем и Куатьери, представляет речь как линейную комбинацию синусоид с изменяющимися во времени амплитудами, фазами и частотами [9]:

$$S_{\widetilde{SR}}(n) = \sum_{K=1}^{L} A_k \cos(\Omega_k n + \varphi_k)$$

где $\widetilde{SR}$ обозначает синусоидальное представление. Заметим, что число синусоид L изменяется во времени. Возможность уменьшить скорость передачи данных с использованием этой модели связана с тем, что голо-

совая речь как правило высокопериодична и, следовательно, она может быть представлена ограниченным набором синусоид. Основным предположением является то, что параметры синусоидальной модели медленно изменяются во времени по отношению к длительности импульса голосовая тракта.

МакАуэль и Куатьери показали, что высококачественное восстановление речевого сигнала может быть достигнуто путём использования синусоид с амплитудами, частотами и фазами, соответствующими пикам коротко-временного преобразования Фурье. Ширина окна Хемминга, равная 2.5 средним высотам, подходит и гарантирует, что синусоидальные волны хорошо определены. Более того голос, начиная со средней высоты, используется только для определения длины анализируемого окна.

Вышепредставленная синусоидальная модель описана в своей наиболее общей форме. Основные вклады в работе МакАуэля и Куатьери лежат в анализе минимального параметра синусоидальной модели, а также в разработке алгоритма отслеживания синусоидальных параметров от окна к окну. Прежде всего, поскольку число синусоид меняется с высотой, было установлено понятие «жизни» и «смерти» синусоидальных компонентов для обеспечения соответствия динамических параметров. В добавок к этому разработаны новые алгоритмы интерполяции фазы и амплитуды для соответствия этим параметрам от одного окна к другому. Эксперименты с моделью показали, что при использовании адаптированного окна Хемминга шириной в 2.5 от средней высоты, и 1024 точки быстрого преобразования Фурье, которое обновляется каждые 10 мс, уже 80 синусоид могут быть использованы для синтеза голоса. Модель хорошо работает с речью в присутствии фонового шума.

Для низкочастотных приложений частоты синусоидальных волн могут быть ограничены, в итоге они получаются целыми, кратными основной частоте, то есть [9]:

$$S_{\widetilde{HR}}(n) = \sum_{k=1}^{L(\Omega_0)} \cos(\varphi_k + k\Omega_0 n) A_k$$

где $L(\Omega_0)$ – количество гармоник интересующей речевой полосы частот (обычно 4 кГц), Ω_0 – частота основного тона, а $\widetilde{HR}$ означает гармоническое представление сигнала. Гармоническое представление обеспечивает оптимальное множество частот только для идеально звучащих сегментов. Основное предположение в голосовой речи - шаг периода постоянен в течение всего периода анализа окна.

Простой пример реконструкции сегмента голосовой речи при помощи линейной комбинации гармонических синусоид показан на рисунке 2. Голосовой сегмент, сформированный с помощью 32 мс прямоугольных окон и амплитуд фаз синусоидальных волн, оценивался по пикам сегмента дискретного преобразования Фурье (рисунок 3).

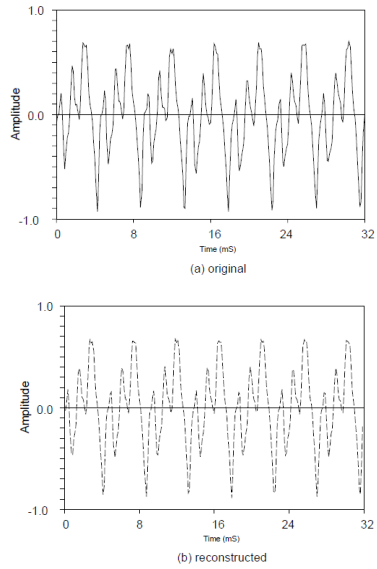

Рис. 2. Восстановление сигнала по набору гармоник

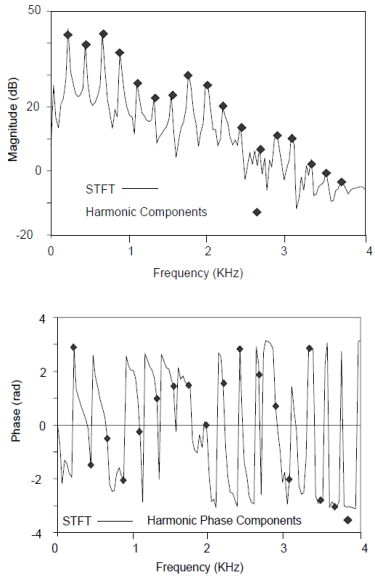

Рис. 3. Амплитуды и фазы участка речевого сигнала, выделенные с помощью дискретного преобразования Фурье

Алгоритм вычисления векторов приведённых начальных фаз на участках речевых вокализмов

Одной из главных задач фоноскопических исследований речи является поиск таких параметров и признаков исследуемого фонообъекта, которые были бы инвариантны к реально встречающимся искажениям речевых сигналов. Наиболее актуально такая задача возникает при решении проблемы верификации личности говорящего [6].

Предложен следующий способ учёта вариативности акустического описания речевого сигнала. На вокализованных участках протяжённостью Δt в точках анализа с шагом r, речевой сигнал $S_r(t)$ может быть представлен суммой M его составляющих гармоник по формулам, представленным ранее. Для данного временного участка речи Δt и точки анализа r вводятся понятия вектора начальных фаз φ_r и вектора приведённых начальных фаз $\overline{\varphi_r}$:

$$\varphi_r = \begin{bmatrix} \varphi_{M-1} \\ \varphi_{M-2} \\ \vdots \\ \varphi_1 \\ \varphi_0 \end{bmatrix}$$

$$\overline{\varphi_r} = \begin{bmatrix} \varphi_{M-1} - \varphi_0 \\ \varphi_{M-2} - \varphi_0 \\ \vdots \\ \varphi_1 - \varphi_0 \\ 0 \end{bmatrix}$$

В качестве опорной начальной фазы в $\overline{\varphi_r}$ можно взять фазу любой гармонической составляющей речевого сигнала. Такое приведение фаз всех имеющихся на анализируемом участке гармоник к одной опорной необходимо для снятия неопределённости, связанной с выбором точки начала отсчёта при выполнении процедур анализа.

Выдвигаются следующие три гипотезы. Первая - вектор приведённых начальных фаз можно считать инвариантным на всём протяжении вокализованного участка и значения приведённых фаз, кроме опорной, отличны от нуля. Вторая - вектор приведённых начальных фаз одинаков для одинаковых звуков и слов, произнесённых одним лицом. Третья - вектор приведённых начальных фаз одинаков для всех звукосочетаний данного лица.

Предлагается следующий метод нахождения вектора приведённых начальных фаз. Устранив из изначальной формулы $S_r(t)$ амплитуду гармоник, получаем следующее описание данного вокализованного участка речи протяжённостью Δt:

$$\overline{S_r(t)} = \sum_{k=0}^{M-1} \cos(\omega_k t + \varphi_k)$$

На рисунке 4 в полосе частот 4 кГц на временном интервале примерно равным 0.5 с при шаге анализа по времени 1.25 мс представлены сонограммы исходного сигнала $S_r(t)$ и сигнала $\overline{S_r(t)}$ с устранённой амплитудой гармоник.

Один из моментов времени t_0 на отрезке анализируемого вокализованного участка речи принимается за начало отсчёта. Далее строится система уравнений размерностью равной числу гармоник M, находящихся в интервале анализа:

$$\begin{cases} \overline{S_r(t_0)} = \sum_{k=0}^{M-1} \cos(\omega_k t_0 + \varphi_k) \\ \overline{S_r(t_0 + r)} = \sum_{k=0}^{M-1} \cos(\omega_k (t_0 + r) + \varphi_k) \\ \vdots \\ \overline{S_r(t_0 + (M-1)\cdot r)} = \sum_{k=0}^{M-1} \cos(\omega_k (t_0 + (M-1)\cdot r) + \varphi_k) \\ \overline{S_r(t_0 + M\cdot r)} = \sum_{k=0}^{M-1} \cos(\omega_k (t_0 + M\cdot r) + \varphi_k) \end{cases}$$

Результатом решения этой системы являются вектора $\{\cos\varphi_r\}$ и $\{\cos\overline{\varphi_r}\}$. Процедура решения системы повторяется на всём протяжении вокализованного участка Δt для более точного определения вектора косинусов приведённых начальных фаз $\{\cos\overline{\varphi_r}\}$.

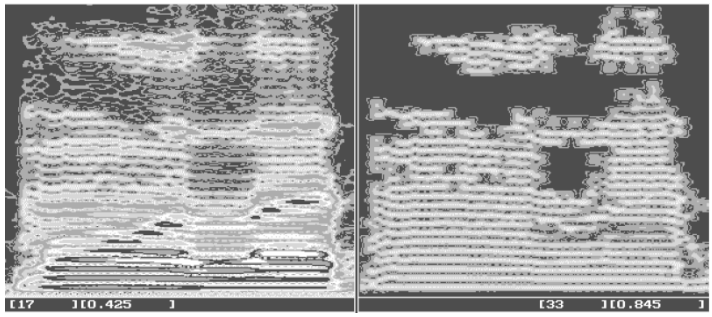

Рис. 4. Сонограмма исходного участка речи (левая половина) и сонограмма этого же участка с устранённым влиянием амплитудного спектра (правая половина)

Как пример выражение $\{cos\overline{\varphi_r}\}$ может использоваться в качестве эталонного описания голоса в системах верификации - процесс будет состоять из сравнения заранее вычисленного эталонного вектора косинусов приведённых начальных фаз с вектором, вычисляемым в реальном времени из анализируемой речи.

Практическое применение фазовых характеристик голосовых вокализмов

Требования к комплексному подходу при проектировании и построении систем защиты информации, в том числе речевой, продиктован, прежде всего, широким спектром возможных угроз, а также постоянным усложнением как самих информационных систем, так и совершаемых злоумышленниками на них атак [6].

Угрозы речевой информации делятся на угрозы её доступности, целостности и конфиденциальности, некоторые методы защиты от них представлены на рисунке 5. В последние несколько лет в виду как активного развития алгоритмов машинного обучения, так и стремительного увеличения мощности вычислительных машин, особое внимание злоумышленников привлекает именно нарушение целостности речевой информации. Например, в 2020 году преступники с помощью технологии машинного обучения смоделировали голос одного из состоятельных клиентов банка ОАЭ, а затем осуществили звонок от его лица управляющему, который, узнав голос своего доверителя, без дополнительных проверок перевёл около 400 тысяч долларов США на сторонний банковский счёт [2]. Вместе с тем это далеко не единичная история – аналогичный случай произошёл в 2019 году в Великобритании, и тогда мошенникам удалось похитить около 240 тысяч долларов США [10].

Рис. 5. Различные методы защиты речевой информации от угроз целостности, конфиденциальности и доступности

При этом фазовые характеристики голосовых вокализмов могут быть применены для купирования многих угроз речевой информации, в том числе и приведённых выше, путём реализации их в соответствующих методах защиты. В частности, на основе эталонного вектора косинусов приведённых начальных фаз возможно построить технологию речевой подписи.

В целях проведения дальнейших исследований разрабатывается программный комплекс, позволяющий автоматизировано проводить анализ голосовых вокализмов, вычислять их фазовые характеристики (в том числе эталонный вектор начальных фаз), на основе полученных данных строить фазограммы с наложением спектрограмм, а также совершать различные звуковые преобразования (шумоо-

чистку). Программное обеспечение разрабатывается на языке программирования C++ с применением фреймворка Qt 5. Такой подход позволяет с одной стороны эффективно распоряжаться ресурсами автоматизированного рабочего места, а с другой стороны разрабатывать приложение, работающее на всех популярных платформах и поддерживающее построение интерфейса под различные разрешения и соотношения экранов. Снимок начального экрана программы в среде MicrosoftWindows представлен на рисунке 6.

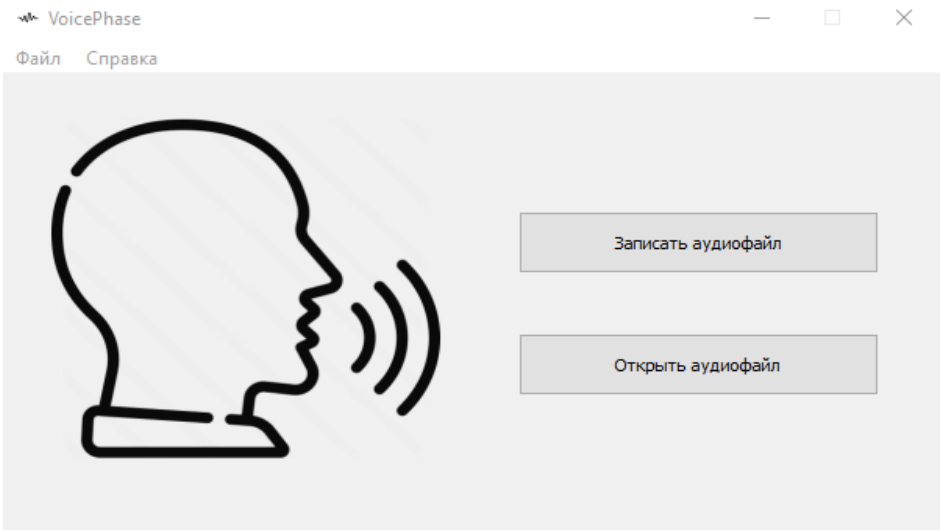

Рис. 6. Начальное окно программы в среде ОС MicrosoftWindows

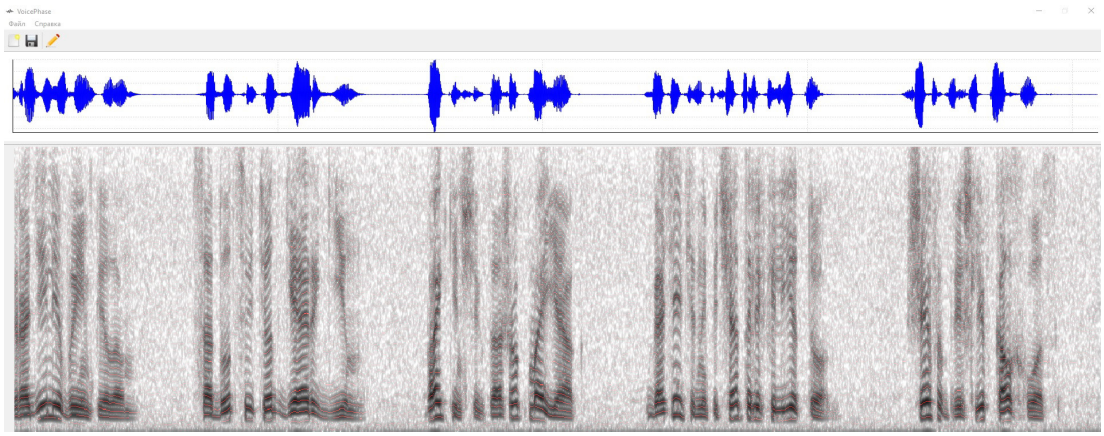

Рис. 7. Основное рабочее окно программы с анализом аудиофайла

На рисунке 7 приведён скриншот основного рабочего окна программы, в нижней его половине построена фазограмма аудиосигнала с наложением спектрограммы. Этот рисунок является многомерной визуализацией звука (в данном случае речи), поскольку кроме частоты и времени в нём в градациях серого выражены ещё мощность и фаза. Подобный формат представления кажется удобным, поскольку в него можно встроить биометрическую информацию о говорящем, например, рукописную подпись или изображение отпечатка пальца.

Авторы считают, что использование фазовых характеристик голосовых вокализмов может дать новый толчок к развитию различных сервисов, использующих технологии синтеза и анализа речи – повысить надёжность и удобство использования речевой подписи, уточнить модели систем распознавания и синтеза голоса, усовершенствовать защиту выделенных для конфиденциальных переговоров помещений от прослушивания.

Заключение

В статье рассмотрены характеристики речевых сигналов и приведены их основные модели (Гильбертовская модель, а также модель синусоидального описания речевого сигнала МакАуэля и Куатьери). Описан алго-

ритм вычисления векторов приведённых начальных фаз на участках речевых вокализмов, а также приведены примеры реконструкции сегмента голосовой речи при помощи линейной комбинации гармонических синусоид. Представлено специальное программное обеспечение, предназначенное для работы с фазовыми характеристиками голосовых вокализмов.

Необходимо отметить, что фазовые характеристики голосовых вокализмов могут найти широкое применение во многих прикладных областях – начиная с построения удобных и безопасных систем голосовой аутентификации говорящего, заканчивая системами автоматического анализа, а также синтеза голоса.

Литература

1. Chauhan S. EvolutionofSpeechRecognitionTechnology// Электронный ресурс https://readwrite.com/2019/10/22/evolution-of-speech-technology (дата обращения: 05.11.2021 года).

2. Quach K. Bank manager tricked into handing $35m to scammers using fake 'deep voice' tech // Электронный ресурс https://www.theregister.com/2021/ 10/16/ai_in_brief/ (дата обращения: 05.11.2021 года).

3. Кодзасов С.В., Кривнова О.Ф. Общая фонетика. Москва: РГГУ, 2001. 592 с.

4. Дидковский В.С. Акустическая экспертиза каналов речевой коммуникации. Монография. Киев: Имекс-ЛТД, 2008. 420 с.

5. Царегородцев А.В. Техническая защита информации. Москва: Финансовый университет при Правительстве РФ, 2014. 276 с.

6. Дворянкин С. В., Уленгов С. В., Устинов Р. А., Дворянкин Н. С., Антипенко А. О. Системное моделирование речеподобных сигналов и его применение в сфере безопасности, связи и управления // Безопасность информационных технологий. 2019. Т. 26, №4. С. 101-119.

7. Голубинский А.Н. Математические модели речевых сигналов для верификации и идентификации личности по голосу. Воронеж: Воронежский государственный университет, 2010. 363 с.

8. Коржик В. И. Расширенное преобразование Гильберта и его применения в теории сигналов // Проблемы передачи информации. 1969. Т. 5, №4. С. 3-18.

9. McAulay R. J., Quatieri T. F. Speech analysis/Synthesis based on a sinusoidal representation // Article in IEEE Transactions on Acoustics Speech and Signal Processing. 1986. ASSP-34(4). pp. 744-754.

10. Stupp C. Fraudsters Used AI to Mimic CEO's Voice in Unusual Cybercrime Case // Электронный ресурс https://www.wsj.com/articles/fraudsters-use-ai-to-mimic-ceos-voice-in-unusual-cybercrime-case-11567157402 (дата обращения: 05.11.2021).

References

1. Chauhan S. Evolution of Speech Recognition Technology // Electronic resource https://readwrite.com/2019/10/22/evolution-of-speech-technology (accessed: 05.11.2021).

2. Quach K. Bank manager tricked into handing $35m to scammers using fake 'deep voice' tech // Electronic resource https://www.theregister.com/2021/ 10/16/ai_in_brief/ (accessed: 05.11.2021).

3. Kodzasov S. V., Krivnova O. F. Obshchaya fonetika. Moskva: RGGU, 2001. 592 s

4. Didkovskij V.S. Akusticheskaya ekspertiza kanalov rechevoj kommunikacii. Monografiya. Kiev: Imeks-LTD, 2008. 420 s.

5. Caregorodcev A. V. Tekhnicheskaya zashchita informacii. Moskva: Finansovyj universitet pri Pravitel'stve RF, 2014. 276 s.

6. Dvoryankin S. V., Ulengov S. V., Ustinov R. A., Dvoryankin N. S., Antipenko A. O. Sistemnoe modelirovanie rechepodobnyh signalov i ego primenenie v sfere bezopasnosti, svyazi i upravleniya // Bezopasnost' informacionnyh tekhnologij. 2019. T. 26, №4. S. 101-119.

7. Golubinskij A. N. Matematicheskie modeli rechevyh signalov dlya verifikacii i identifikacii lichnosti po golosu. Voronezh: Voronezhskij gosudarstvennyj universitet, 2010. 363 s.

8. Korzhik V. I. Rasshirennoe preobrazovanie Gil'berta i ego primeneniya v teorii signalov // Problemy peredachi informacii. 1969. T. 5, №4. C. 3-18.

9. McAulay R. J., Quatieri T. F. Speech analysis/Synthesis based on a sinusoidal representation. Article in IEEE Transactions on Acoustics Speech and Signal Processing, 1986, ASSP-34(4), pp. 744-754.

10. Stupp C. Fraudsters Used AI to Mimic CEO's Voice in Unusual Cybercrime Case // Electronic resource

https://www.wsj.com/articles/fraudsters-use-ai-to-mimic-ceos-voice-in-unusual-cybercrime-case-11567157402 (accessed: 05.11.2021).

САВЕЛЬЕВ Иван Андреевич, кандидат технических наук, доцент Департамента информационной безопасности, Финансовый университет при Правительстве Российской Федерации. 125993, г. Москва, Ленинградский проспект, 49. E-mail: IASavelyev@fa.ru

АНТИПЕНКО Антон Олегович, аспирант Департамента информационной безопасности, Финансовый университет при Правительстве Российской Федерации. 125993, г. Москва, Ленинградский проспект, 49. E-mail: An-go-55@yandex.ru

SAVELYEV Ivan Andreevich, candidate of technical sciences, Associate Professor of the Department of Information Security, Financial University under the Government of the Russian Federation. 125993, Moscow, Leningradsky Prospekt, 49. E-mail: IASavelyev@fa.ru

ANTIPENKO Anton Olegovych, postgraduate student of the Department of Information Security, Financial University under the Government of the Russian Federation. 125993, Moscow, Leningradsky Prospekt, 49. E-mail: An-go-55@yandex.ru

УДК 004.056 Вестник УрФО № 4(42) / 2021, с. 15–23

Акбулякова Л.М., Шабуров А.С.

DOI: 10.14529/secur210402

О СОВЕРШЕНСТВОВАНИИ АРХИТЕКТУРЫ ИНФОРМАЦИОННОЙ СИСТЕМЫ ПЕРСОНАЛЬНЫХ ДАННЫХ ПРИ ВЗАИМОДЕЙСТВИИ ОПЕРАТОРА С СЕГМЕНТОМ «ЕГИСЗ»

В статье проанализирована структура и основные задачи, решаемые на основе внедрения Единой государственной информационной системой в сфере здравоохранения (ЕГИСЗ). Сформулирована проблема безопасного взаимодействия автоматизированных рабочих мест работников с ЕГИСЗ. Обозначены уязвимости действующей модели информационного взаимодействия. Предложена универсальная математическая модель поиска оптимальных структур информационных систем. Разработан алгоритм выбора варианта схемы подключения к ЕГИЗС при передаче персональных данных (ПДн). Сформированы основные схемы подключения, с помощью которых разработана усовершенствованная схема подключения ЕГИСЗ, позволяющая реализовать требования по защите информации.

Ключевые слова: *государственная информационная система, персональные данные, защита информации, морфологический подход.*

ON IMPROVING THE ARCHITECTURE OF THE PERSONAL DATA INFORMATION SYSTEM WHEN THE OPERATOR INTERACTS WITH THE EGISZ SEGMENT

The article analyzes the structure and main tasks solved on the basis of the introduction of a Unified state information system in the field of healthcare (EGISZ). The problem of safe interaction of automated workplaces of employees with EGISZ is formulated. The vulnerabilities of the current model of information interaction are identified. A universal mathematical model of the search for optimal structures of information systems is proposed. The article draws attention to the developed algorithm for selecting a variant of the connection scheme to the EGISZ when transferring personal data (PDt). The basic connection schemes have been formed, with the help of which an improved EGISZ connection scheme has been developed, which allows implementing information security requirements.

Keywords: *state information system, personal data, information protection, morphological approach.*

Активное развитие цифровых отношений в современном обществе обуславливают увеличение интенсивности информационного обмена, совершенствование механизмов обработки информации, качества предоставляемых сервисов и услуг. В то же время, процесс «цифровизации» способствует проявлению новых уязвимостей и угроз, связанных с внедрением передовых технологий, что требует выхода на новый качественный уровень решения проблемы обеспечения информационной безопасности.

Создание единых центров обработки данных, объединяющих многоуровневые и многофункциональные информационные системы, предполагает разработку и внедрение новых решений, сервисов и служб поддержки, основанных на современных технологиях и оптимальных алгоритмах. В тоже время, способы проникновения вредоносной информации в подобные информационные системы также стремительно модифицируются, что делает процесс защиты информации более трудоемким, требующим поиска и применения как новых средстви методов защиты, так и оптимизации архитектуры информационных систем.

Одним из примеров подобного класса сложных систем является Единая государственная информационная система в сфере здравоохранения (ЕГИСЗ), включающая множество подсистем, информационных реестров и ресурсов, в том числе конфиденциальной информации, относящейся к специальной категории персональных данных граждан Российской Федерации [1].

Основными задачами ЕГИСЗна сегодняшний день являются:

1) информационное обеспечение государственного регулирования в сфере здравоохранения;

2) информационная поддержка деятельности медицинских организаций, включая поддержку осуществления медицинской деятельности;

3) информационное взаимодействие по-

ставщиков информации в единую систему и пользователей информации, содержащейся в единой системе;

4) информирование населения по вопросам ведения здорового образа жизни, профилактики заболеваний, получения медицинской помощи, передачи сведений о выданных рецептах на лекарственные препараты из медицинских информационных систем медицинских организаций в информационные системы фармацевтических организаций;

5) обеспечение доступа граждан к услугам в сфере здравоохранения в электронной форме, а также взаимодействия информационных систем, информационных систем государственных внебюджетных фондов.

Ключевым принципом организации информационного взаимодействия в Едином цифровом контуре здравоохранения является обеспечение возможности обмена данными между информационными системами о случаях оказания медицинской помощи в электронном виде в объеме, необходимом и достаточном для обеспечения преемственности, и непрерывности процессов оказания медицинской помощи в отношении каждого отдельно взятого пациента [2].

Внедрение ЕГИСЗ является важной частью программы модернизации системы здравоохранения, главной задачей которой является создание целостногои доступного информационного пространства для всех участников информационного обмена в условиях обеспечения информационной безопасности.

В настоящее время ЕГИСЗ представляет собой информационную систему, обеспечивающую взаимодействие различных подсистем (рис. 1):

1) федерального сегмента и региональных сегментов ЕГИСЗ;

2) системы межведомственного электронного взаимодействия (СМЭВ);

3) web-порталов Министерства здравоохранения РФ и Правительства РФ - Единого портала государственных и муниципальных услуг (ЕПГУ);

4) системы удостоверяющих центров Министерства здравоохранения РФ;

5) защищенной сети передачи данных.

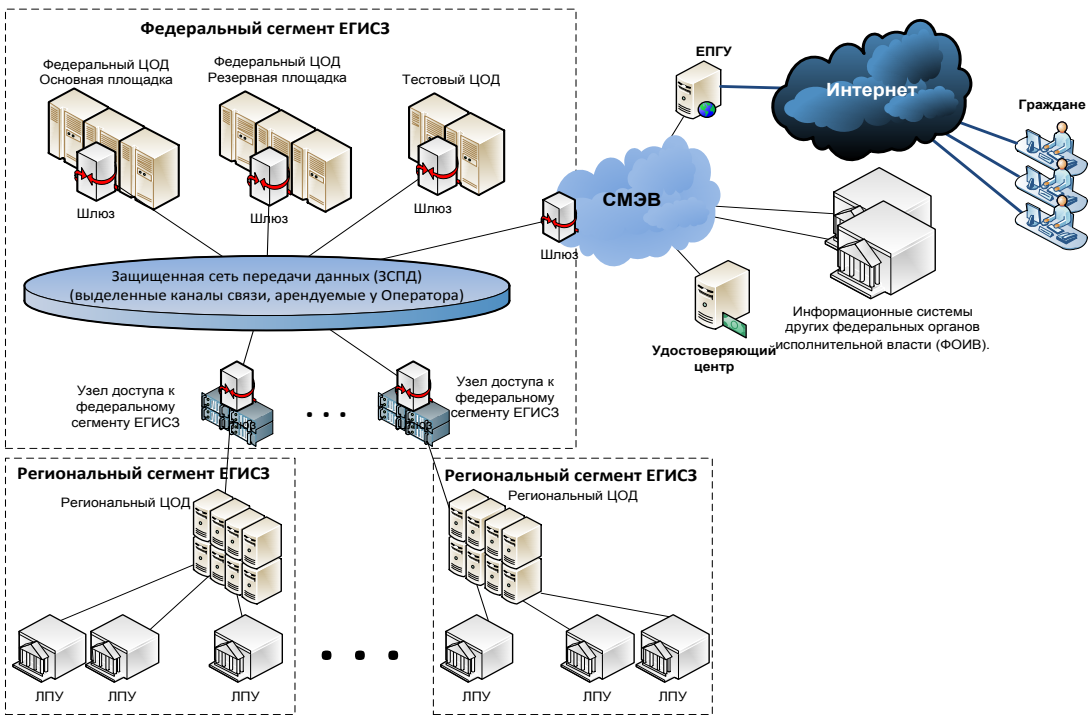

Рис. 1. Схема взаимосвязи компонентов ЕГИСЗ

Для решения задач обеспечения информационной безопасности АРМ медицинских организаций, подключенных к ЕГИСЗ необходима реализация оптимальной схемы подключения, которая смогла бы в полном объеме реализовать требования защищенной передачи персональных данных граждан [3]. При этом, на уровне федерального сегмента ЕГИСЗ предполагается обработка ПДн, примерно 145 млн. пользователей, являющихся

гражданами Российской Федерации. В приведенной схеме отражена логическая взаимосвязь основных компонентов федерального сегмента ЕГИСЗ с иными компонентами ЕГИСЗ и информационными системами.

Пользователями данной системы являются как граждане, так и сотрудники, осуществляющие взаимодействие через единый портал государственных и муниципальных услуг (ЕПГУ) и сети связи общего пользования. Подобное взаимодействие предполагает, что АРМ пользователей требуют защищенного соединения. Для этого данное взаимодействие осуществляется через СМЭВ, на выделенных для этой цели технических средствах Министерства здравоохранения и Правительства РФ, не имеющих прямого подключения к Федеральному центру обработки данных (ФЦОД).

В целях обеспечения достоверности передаваемых данных происходит взаимодействие с системой удостоверяющих центров Министерства здравоохранения при обеспечении участников информационного взаимодействия квалифицированными сертификатами электронной подписи, а также реализации функции по их проверке. Система удостоверяющих центров обеспечивает юридическую значимость передаваемых данных между участниками информационного обмена. Кроме того, для защиты информации на объекте информатизации реализованы организационные и технические меры, в соответствии с требованиями приказов ФСТЭК России. Перечень защитных мер адаптирован применительно к структурно-функциональным характеристикам выбранной информационной системы и особенностям её функционирования.

В то же время, при реализации базовых требований по защите информации на конкретных объектах не в полной мере учитываются особенности инфраструктуры и алгоритмов взаимодействия в подсистемах. Так, например, при удаленном подключении к АРМ работника, обрабатывающего конфиденциальную информацию, допускается подключения к общей сети внутри учреждения, где подключены и АРМ, которые не предназначенных для обработки персональных данных. При этом, допускается подключение к другим информационным системам, доступ которых осуществляется без использования программного комплекса средств защиты информации. Наличие подобных уязвимостей требует поиска новых, более совершенных, с точки зрения информационной безопасности, системных решений, а также оптимизации процедур взаимодействия элементов ЕГИСЗ.

В целях оптимизации информационных систем, при их разделении на подсистемы, целесообразно использовать морфологический подход, который широко применяется в проектировании сложных систем [3]. В этом случае предполагается, что любой вариант системы имеет определенную структуру, то есть состоит из конечного числа элементов (подсистем), и распределение, или перераспределение системных функций среди них могут быть выполнены с помощью конечного числа методов.

Процесс формирования множества допустимых вариантов системы можно представить посредством функциональной декомпозиции, в виде набора элементов в следующем виде:

$$\left\{ f_i, i = \overline{1, L}, \bigcup_{i=1}^{N} f_i = f \right\}. \quad (1)$$

Это предполагает разбиение конечного набора элементов системы S на N морфологических классов m(l), l=$\overline{1, L}$ таких, что $m(l) \cap m(l') = \emptyset$ при l ≠ l/.

Введем понятие морфологического пространства F ⊆ 2^{ε}, все элементы которого являются морфологическими вариантами системы $f = (f_1, f_2, ..., f_L)$. Каждый морфологический вариант f представляет собой определенный набор экземпляров класса $f(l) \in m(l)$. В данном случае для любого $f \in F$ и любого l=$\overline{1, L}$ множество $f \in F$ содержит один элемент.

Если предположить, что существует множество способов реализации каждой подсистемы $f_{lk}, k = \overline{1, K}, l = \overline{1, L}$, тогда общее количество возможных морфологических вариантов системы можно определить как:

$$Q = \prod_{l=1}^{L} K_l. \quad (2)$$

При формировании множества допустимых вариантов системы необходимо учитывать ограничения, накладываемые как на структуру, так и параметры, и техническую реализацию всех элементов системы. Кроме того, на систему в целом, а также допустимые варианты соединений элементов и ограничения на значения показателей качества системы. При учете всех этих показателей могут возникать противоречия в требованиях. С одной стороны, желательно, представить все

возможные варианты системы во всей их полноте, чтобы не пропустить потенциально лучшие варианты. С другой стороны, существуют ограничения, предусмотренные значением допустимых расходов (времени и средств) на проектирование системы. После определения множества возможных вариантов системы в терминах конкретной структуры, вычисляются значения показателей качества, и выделяются множество Парето-оптимальных вариантов, которое может сокращаться до единственного, наиболее предпочтительного, варианта [4]. Применение подобных оптимизационных подходов основывается на определении множества допустимых вариантов структуры системы, а также вариантов информационного взаимодействия ее элементов.

Анализ моделей информационного взаимодействия типовых элементов информационной системы, в первую очередь, предполагает изучение порядка функционирования и разработку моделей основных информационных связей, возникающих вследствие обработки информации. Для ЕГИС – функционирование при организации оказания различных медицинских услуг.

В первую очередь необходимо проанализировать варианты подключения медицинских организаций (МО), АРМ которых взаимодействуют непосредственно с медицинской информационной системой - ЕГИСЗ, при необходимости выполнении требований по безопасности информации [5]. При этом, выбор схемы подключения зависит от количества необходимых АРМ, потребностей пакетной обработки ПДн, а также наличии в организации защищенной сети. Варианты информационного взаимодействия медицинской информационной системы с АРМ работника характеризуется следующими признаками:

1. Масштаб медицинской организации.

2. Количество обрабатываемых персональных данных.

3. Применение программно-аппаратного комплекса.

Для оптимального выбора схемы подключения АРМ МО к ЕГИСЗ разработан алгоритм выбора варианта схемы подключения (рис.2).

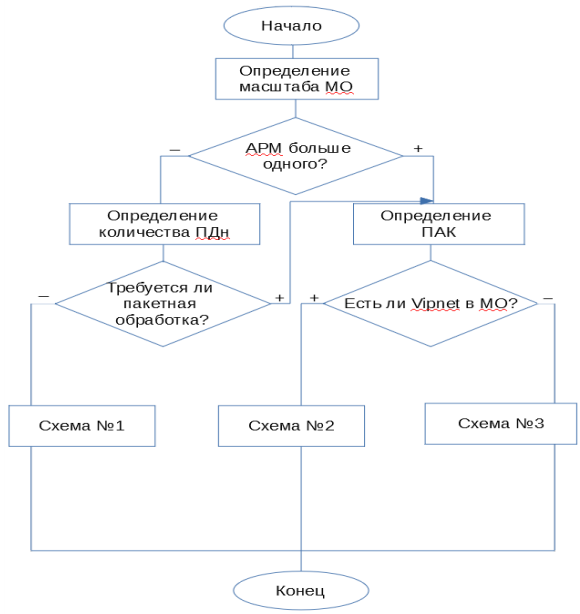

Рис. 2. Варианты информационного взаимодействия медицинской информационной системой с АРМ работника

Анализ различных вариантов информационного взаимодействия медицинской информационной системой с АРМ работника позволил сформировать несколько основных моделей подключения.

Схема подключения №1 ЕГИСЗ к АРМ работника для оказания медицинских услуг (рис.3) предполагает защиту информации конфиденциального характера, на основе применения сертифицированных шифровальных средств на базе продуктов семейства ViPNet. Данная схема предполагает под-

ключение с использованием изолированного автоматизированного рабочего места с локальным доступом в сеть Интернет и без доступа к локальной сети организации.

Данная схема предполагает следующие ограничения. ПДн обрабатываются только на выделенном АРМ, в режиме ручной обработки информации. В связи с данными ограничениями, подобная схема рекомендуется к использованию учреждениям с незначительным объемом обработки ПДн.

Схема подключения №2 (рис.4) является наиболее распространённой и рассчитана на подключение одного АРМ. Подключение и передача данных осуществляется с помощью программного комплекса ViPNet Client. Подключение по данной схеме обеспечивает большую защищенность данных при автоматизированной обработке ПДн, что позволяет произвести обмен данными с федеральными информационными системами. Подобная схема рассчитана на крупные медицинские учреждения и значительные потоки информационного взаимодействия.

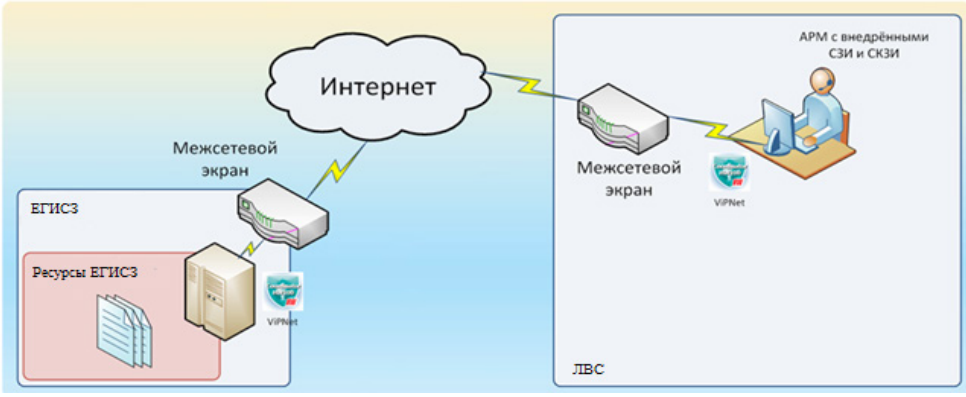

Рис. 3. Схема подключения №1

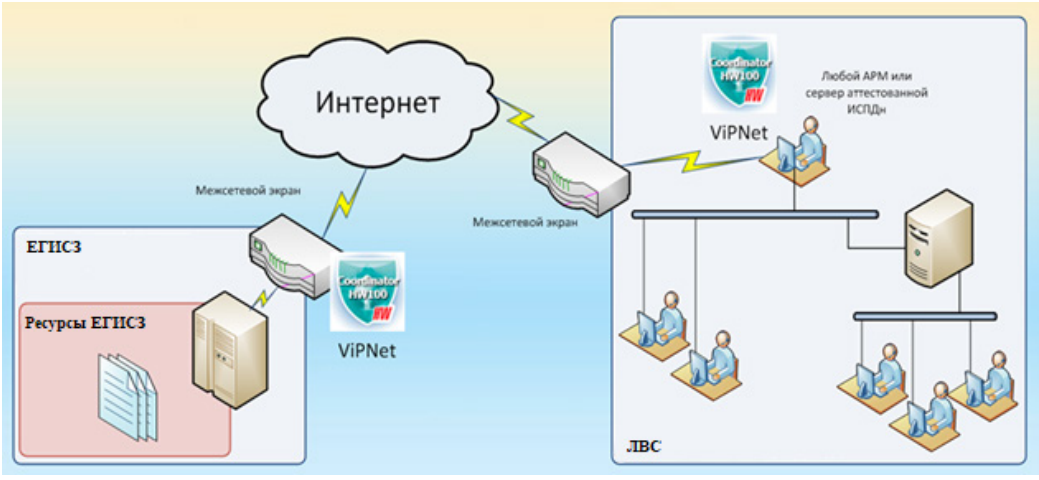

Рис. 4. Схема подключения №2

Схема подключения №3 (рис.5) предполагает возможность выделение отдельного сегмента, предназначенного для передачи данных, который может состоять из одного рабочего места и с которого может осуществляться загрузка данных. Подключение и передача данных осуществляется с помощью программно-аппаратного комплекса ViPNet Coordinator и рассчитана на количество от двух АРМ, поддерживает как ручной ввод информации, так и пакетную выгрузку персональных данных.

Таким образом, проанализированные схемы подключения позволяют систематизировать подходы к обработке информации и проанализировать основные методы защиты информации на типовых объектах информатизации учреждений здравоохранения.

В результате анализа характеристик можно выбрать оптимальную схему подключения, что обеспечивает необходимые процессы обработки ПДн, но не в полном объеме может гарантировать безопасность передачи

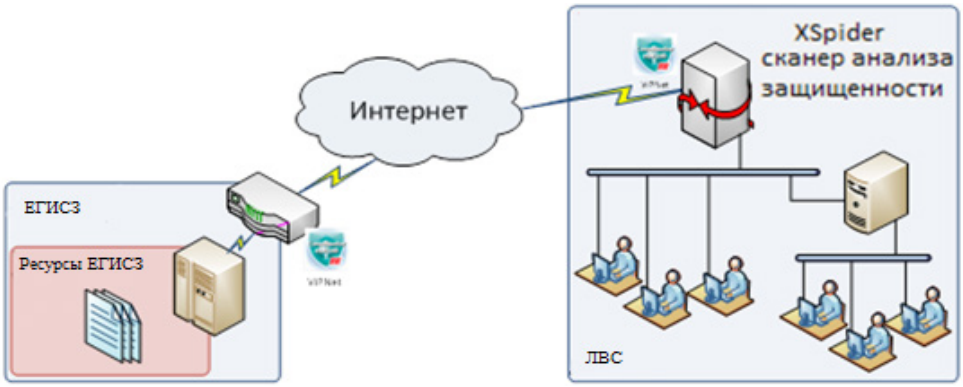

Рис. 5. Схема подключения №3

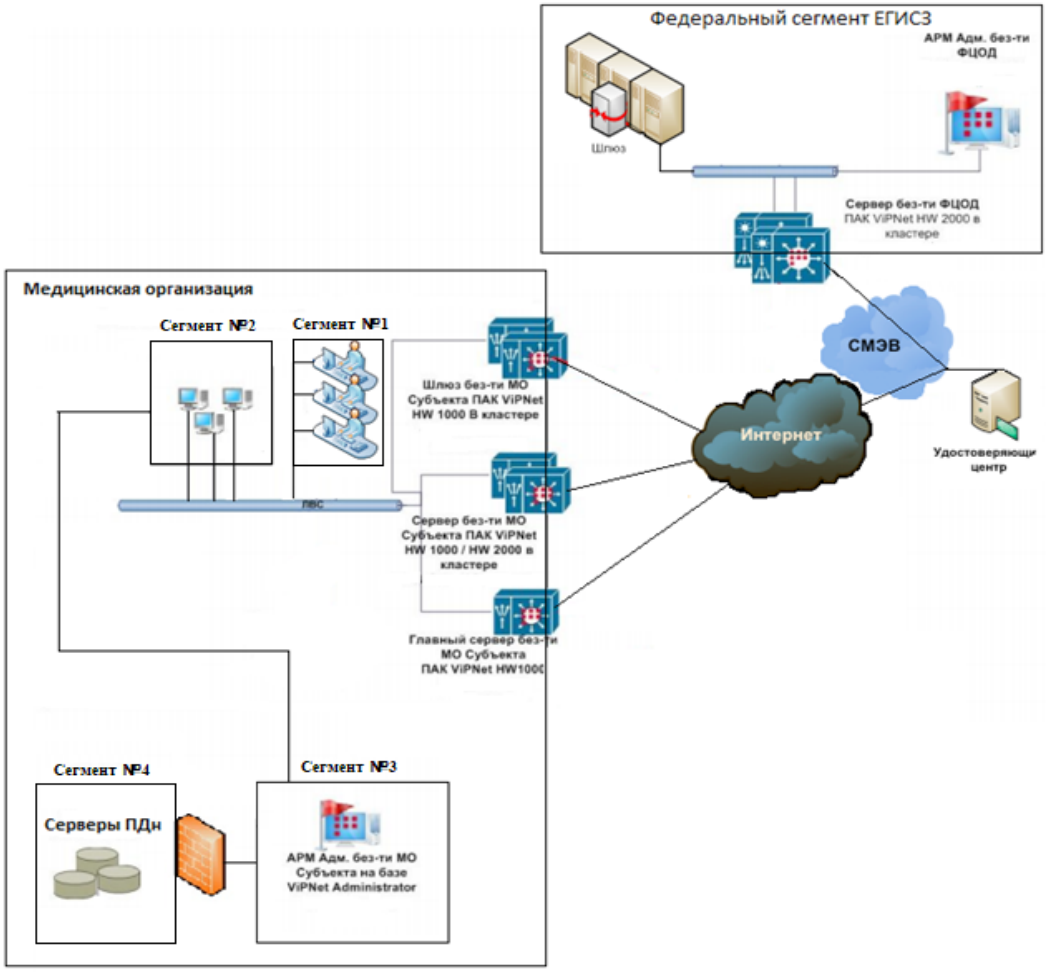

Рис. 6. Сегментированная структура сети

персональных данных из самих АРМ МО в сегмент ЕГИСЗ [6]. То есть, сохраняется недостаток архитектуры информационной системы, состоящий в сохранении возможности подключения АРМ МО к общей сети, в том числе и не предназначенных для обработки ПДн через сегмент ЕГИСЗ, подключенных к иным информационным системам, доступ которых осуществляется без требуемых средств защиты информации (VipNet) [7].

Для устранения данной уязвимости предлагается оптимизировать архитектуру информационной системы путем разделение ее на необходимое в конкретном случае количество сегментов (рис.6).

Обновленная архитектура отличается от

базовых схем подключения тем, что подсистема, предназначенная для обработки защищаемой информации (ПДн) разделена на 4 сегмента.

В первый сегмент включаются АРМ работников, не предназначенные для обработки защищаемой информации, или подключенные к другим информационным системам [8]. Выделение данного сегмента позволит устранить проблему, связанную с конфликтом доступа с одного физического устройства со своими установленными средствами защиты информации в разные информационные системы, требующих использования разных средств защиты информации для полноценного функционирования.

Во второй сегмент включаются АРМ сотрудников, на который поступает запрос от ЕГИСЗ и осуществляет перенаправление его на третий сегмент. Выделение данного сегмента позволяет разграничить права доступа сотрудников организации, а именно: определить тех, кто осуществляет работу с ЕГИСЗ, а кто нет. Это позволяет контролировать процесс распространения ПДн в организации. При этом, введенные ограничения не допускают обработку ПДн теми сотрудниками, которые не допущенных до данного процесса обработки.

Третий сегмент представляет собой АРМ администратора безопасности МО и получает перенаправленный запрос от второго сегмента. Кроме того, из данного сегмента может осуществляться доступ к серверным станциям (базам данных ПДн), в отличие от других АРМ. Данный сегмент также предполагает возможность определения прав доступа к базам данных.

В четвертый сегмент включаются серверные станции для хранения баз данных (ПДн). При этом, из четвертого сегмента не может осуществляться прямое взаимодействие с первым и вторым сегментом, или наоборот. Подобная конфигурация позволяет устранить противоречие в разграничении прав доступа для сотрудников, непосредственно работающих с ЕГИСЗ, а также позволяет осуществлять контроль процесса доступа и работы с ПДн.

Таким образом, проведенный анализ функционирования информационной системы, а также системы защиты ИСПДн позволил выявить уязвимости взаимодействия с ЕГИСЗ при передаче данных. Усовершенствованная архитектура взаимодействия в сети медицинской организации позволяет решить проблему одновременного доступа к разным информационным системам, при работе с разными категориями ПДн. При этом совершенствование архитектуры подобной системы позволяет устранить уязвимость без значительных ресурсных затрат, с сохранением требуемой технологии обработки информации.

Литература

1. Федеральный закон от 27 июля 2006 г. № 152-ФЗ «О персональных данных». -URL: https://www.consultant.ru/document/cons_doc_LAW_61801/ (дата обращения 15.11.2021).

2. Постановление Правительства РФ от 05.05.2018 № 555 «О единой государственной информационной системе в сфере здравоохранения» //Собрание законодательства РФ. – 25.05.2018. - № 28. - Ст. 4241.

3. Золотарев А. В. Методы оптимизации распределительных процессов / А. В. Золотарев: ЛитРес, 2014. С. 22-34.

4. Подиновский В.В. Парето-оптимальные решения многокритериальных задач / В.В. Подиновский. – М.: Наука, 1982. – 256 с.

5. Князюк Н.Ф. Методические подходы к внедрению международного стандарта iso/iec 27001:2005 при построении системы управления информационной безопасностью медицинской организации/ Н.Ф.Князюк, И.С. Кицул. – Москва: Юрайт, 2018. – 102 с. (Менеджер здравоохранения). ISBN 978-5-534-02989-5.

6. Методические рекомендации медицинским организациям по организации криптографической защиты каналов при взаимодействии в рамках единой государственной информационной системы в сфере здравоохранения: официальный сайт. – Москва. - 2018. - URL: https://portal.egisz.rosminzdrav.ru (дата обращения 15.11.2021).

7. Решения по комплексному обеспечению информационной безопасности в ЕГИСЗ: официальный сайт. – Москва. - 2019. - URL: https://portal.egisz.rosminzdrav.ru(дата обращения 15.11.2021).

8. Техническое задание «Оказание услуги комплексного сервиса в целях обеспечения сервисной поддержки функционирования медицинской организации в рамках регионального сегмента единой

государственной информационной системы в сфере здравоохранения»:официальный сайт. – Москва. - 2020. - URL: https://vkr.pspu.ru/uploads (дата обращения 15.11.2021).

References

1. Federal'nyj zakonot 27 iyulya 2006 g. № 152-FZ «O personal'nyh dannyh». - URL: https://www.consultant.ru/document/cons_doc_LAW_61801/ (data obrashcheniya 15.11.2021).

2. Postanovlenie Pravitel'stva RF ot 05.05.2018 № 555 "O edinoj gosudarstvennoj informacionnoj sisteme v sfere zdravoohranenija" // Sobranie zakonodatel'stva RF. – 25.05.2018. - № 28. - St. 4241.

3. Zolotarev, A. V. Metody optimizacii raspredelitel'nyh processov / A. V. Zolotarev: LitRes, 2014. S. 22-34.

4. Podinovskij, V.V. Pareto-optimal'nye reshenija mnogokriterial'nyh zadach / V.V. Podinovskij. – M.: Nauka, 1982. – 256 c.

5. Knjazjuk, N.F. Metodicheskiepodhody k vnedreniju mezhdunarodnogo standarta iso/iec 27001:2005 pripostroenii sistemy upravlenija informacionnoj bezopasnost'ju medicinskoj organizacii/ N.F. Knjazjuk, I.S. Kicul. – Moskva: Jurajt, 2018. – 102 s. (Menedzherzdravoohranenija). ISBN 978-5-534-02989-5.

6. Metodicheskie rekomendacii medicinskim organizacijam po organizacii kriptograficheskoj zashhity kanalov pri vzaimodejstvii v ramkah edinoj gosudarstvennoj informacionnoj sistemy v sfere zdravoohranenija: oficial'nyj sajt. – Moskva. - 2018. - URL: https://portal.egisz.rosminzdrav.ru (data obrashhenija 15.11.2021).

7. Reshenija po kompleksnomu obespecheniju informacionnoj bezopasnosti v EGISZ: oficial'nyj sajt. – Moskva. - 2019. - URL: https://portal.egisz.rosminzdrav.ru (data obrashhenija 15.11.2021).

8. Tehnicheskoe zadanie «Okazanie uslugi kompleksnogo servisa v celjah obespechenija servisnoj podderzhki funkcionirovanija medicinskoj organizacii v ramkah regional'nogo segmenta edinoj gosudarstvennoj informacionnoj sistemy v sfere zdravoohranenija»: oficial'nyj sajt. – Moskva. - 2020. - URL: https://vkr.pspu.ru/uploads (dataobrashhenija 15.11.2021).

ШАБУРОВ Андрей Сергеевич, кандидат технических наук, доцент, доцент кафедры автоматики и телемеханики, Пермский национальный исследовательский политехнический университет. 614990, Пермский край, г. Пермь, Комсомольский проспект, д. 29. E-mail: shans@at.pstu.ru.

АКБУЛЯКОВА Лилия Маратовна, студент Пермского национального исследовательского политехнического университета. 614990, Пермский край, г. Пермь, Комсомольский проспект, д. 29. E-mail: akbulyakowa@mail.ru.

SHABUROV Andrey Sergeevich, Candidate of Technical Sciences, Associate Professor of the Department of Automation and Telemechanics. Perm National Research Polytechnic University. 614990, Perm, Komsomolsky Ave, 29. E-mail: shans@at.pstu.ru

AKBULYAKOVA Liliya Maratovna, student, Department of Automation and Telemechanics. Perm National Research Polytechnic University. 614990, Perm, 29, Komsomolskypr. E-mail: akbulyakowa@mail.ru.

УДК 004.942:519.23/.25 Вестник УрФО № 4(42) / 2021, с. 24–38

Поршнев С.В., Рябко Н.Ю.

DOI: 10.14529/secur210403

ОБ ОСОБЕННОСТЯХ АППРОКСИМАЦИЙ РОЗЕНБЛАТТА-ПАРЗЕНА ЭМПИРИЧЕСКИХ ФУНКЦИЙ РАСПРЕДЕЛЕНИЙ И ПЛОТНОСТЕЙ РАСПРЕДЕЛЕНИЙ СЛУЧАЙНЫХ ВЫБОРОК

В статье обсуждаются результаты сравнительного анализа эмпирических функций распределения (ФР) и плотностей распределения (ПР) случайных выборок и их аппроксимаций Розенблатта-Парзена, зависящих от случайной выборки и параметра размытости ядерной функции h.

Для оценки качества аппроксимаций Розенблатта-Парзена были использованы накопленные среднеквадратические погрешности разностей между эмпирическими ФР и ПР и соответствующими аппроксимациями Розенблатта-Парзена.

Продемонстрировано, что при следовании существующим рекомендациям использовать для вычисления аппроксимации Розенблатта-Парзена значение параметра размытости h, равного абсциссе локального максимума информационного функционала, не удается обеспечить минимального возможного значения выбранного критерия качества аппроксимации эмпирических ФР и ПР.*

Предложено использовать в качестве h абсциссу минимума зависимости накопленной среднеквадратической погрешности разностей между эмпирической ПР и соответствующей аппроксимацией Розенблатта-Парзена от параметра размытости.*

Ключевые слова: *случайная выборка, функция распределения, плотность распределения, эмпирическая функция распределения, эмпирическая плотность распределения, аппроксимация Розенблатта-Парзена, критерий типа Колмогорова-Смирнова.*

THE FEATURES OF THE ROSENBLATT-PARSEN APPROXIMATIONS OF EMPIRICAL DISTRIBUTION FUNCTIONS AND DISTRIBUTION DENSITIES OF RANDOM SAMPLES

The article discusses the results of a comparative analysis of empirical Cumulative Distribution Functions (CDF) and Probability Density Function (PDF) of random samples and their Rosenblatt-Parsen approximations that depending on the random sample and the fuzzy parameter of the kernel function h.

For estimation the quality of the Rosenblatt-Parsen approximations were used the accumulated root-mean-square errors (ARMSE) of the differences between the CDF and PDF and the corresponding Rosenblatt-Parsen approximations.

It has been demonstrated that following the existing recommendations to use the value of the fuzzy parameter h for the Rosenblatt-Parsen approximation calculating equal to the abscissa of the local maximum of the information functional, it is not possible to provide the minimum possible value of the selected criterion of the quality of the empirical CDF and PDF approximation.*

It is proposed to use as h the minimum abscissa of the ARMSE of the difference between the empirical PDF and the corresponding Rosenblatt-Parsen approximation dependence on fuzzy parameter.*

Keywords: *Random Sample, Distribution Function, distribution density, Empirical Distribution Function, Empirical distribution density, Rosenblatt-Parsen approximation, Kolmogorov-Smirnov test*

Введение

Задача аппроксимации функций распределений (ФР) и плотностей распределений (ПР) случайных выборок теми или иными аналитическими моделями возникает в различных отраслях наук, например, в радиотехнике при анализе случайных сигналов [1,2], при проектировании телекоммуникационных систем аудиообмена, систем подвижной связи и систем телекоммуникации [3,4], при оптимизации параметров процедуры обратимого сжатия цифровых данных [5], при создании математических моделей систем управления [6], при моделировании случайных процессов [7], при анализе количественных характеристики Интернет-трафика [8,9], при оценке прочностной надежности магистральных газо- и нефтепроводов [10].

Напомним, что в наиболее общем виде данная задача имеет следующую формулировку. Имеется одномерная выборка независимых упорядоченных в порядке возрастания случайных величин $\{x_1, x_2, ..., x_N\}$, извлеченных из некоторого распределения, в общем случае, с неограниченной областью рассеяния с неизвестными непрерывными ФР $F(x)$ и плотностью распределения ПР $f(x)$ такими что $\int_{-\infty}^{x} f(\xi)d\xi = F(x),$

где $F(x) = \Pr[\xi \le x],$

здесь $\Pr\left[\xi \le x\right]$ – вероятность того, что случайная величина $\xi \le x$. Требуется на основе выборки $\{x_1, x_2, ..., x_N\}$ оценить вид ФР $F(x)$ и ПР $f(x)$

Методы решения рассматриваемой задачи являются объектом исследования параметрической и непараметрической статистики. Напомним, что в параметрической статистике на основе априорной информации о свойствах случайного процесса, из которого извлечена данная выборка, выбирается один из известных законов распределения (число которых на сегодняшний день превышает 100) $F^{\langle\text{theory}\rangle}(x, a, b, ...)$, *a, b*, ... – параметры распределения. Далее по анализируемой случайной выборке $\{x_1, x_2, ..., x_N\}$ находят оценки параметров $\tilde{a}, \tilde{b}$, ... выбранного закона распределения $F^{\langle\text{theory}\rangle}(x, a, b, ...)$, например, с помощью метода максимального правдоподобия. Далее проверяют статистическую гипотезу о соответствии найденной ФР $F^{\langle\text{theory}\rangle}(x, \tilde{a}, \tilde{b}, ...)$ случайной выборки $\{x_1, x_2, ..., x_N\}$ априори выбранному теоретическому закону распределения $F^{\langle\text{theory}\rangle}(x, a, b, ...)$, используя один из многочисленных известных статистических критериев (см. [11, Глава 3]), например, критерий типа Колмогорова–Смирнова.

В непараметрической статистике вычисляют те или иные аппроксимации эмпирической ФР распределения вероятностей $F_N^{\langle\text{emp}\rangle}(x)$:

$$F_N^{\langle\text{emp}\rangle}(x) = \frac{1}{N}\sum_{i=1}^{N}\Theta(x - x_i),$$

где $\Theta(s)$ – функция Хэвисайда,

$$\Theta(s) = \begin{cases} 1, s \ge 0, \\ 0, s < 0, \end{cases}$$

случайной выборки $\{x_1, x_2, ..., x_N\}$, такой, что

$$\Pr\left[\lim_{N\to\infty}\sup\left|F_N^{\langle\text{emp}\rangle}(x) - F^{\langle\text{theory}\rangle}(x)\right| = 0\right] = 1,$$

которая, в этой связи при $N \to \infty$ является оптимальной эмпирической оценкой теоретической функции распределения $F^{\langle\text{theory}\rangle}(x)$. Существование данной функции гарантируется центральной теоремой математической статистики (теорема Гливенко) [12]. Также в прикладной статистике рассматривают эмпирическую ПР случайной выборки $\{x_1, x_2, ..., x_N\}$

$$f_N^{\langle\text{emp}\rangle}(k\Delta x) = \frac{n_k}{N\Delta x},$$

где n_k – число членов случайной выборки, попавших в k-ый интервал $\left[x_1 + (k-1)\Delta x, x_1 + k\Delta x\right]$, $k = \overline{1, N_g}$, N_g – число интервалов гистограммы, $\Delta x = (x_N - x_1)/N_g$, представляющую собой кусочно-постоянную аппроксимацию непрерывной функции $F^{\langle\text{theory}\rangle}(x)$.

Для аппроксимации функции распределения вероятностей $F_N^{\langle\text{emp}\rangle}(x)$ на сегодняшний день разработано достаточно большое число различных методов, основанных на использовании некоторых линейных комбинаций (ядерных) функций, выбор которых не требует априорной информации о функции $F^{\langle\text{theory}\rangle}(x)$ (метод Розенблатта-Парзена [13,14]); ортогональных многочленов [15], в том числе, полиномов Чебышева-Эрмита [16]; многочленов Берштейна (см. [17] и приведенные там ссылки). При этом использование любого из методов сопряженно с решением проблем, связанных с выбором критериев для оценки качества аппроксимации функции распределения $F_N^{\langle\text{emp}\rangle}(x)$ и выбора таких параметров используемого метода, что обеспечивается оптимальное качество аппроксимации. Таковым параметром в методе Розенблатта-Парзена, одном из наиболее активно используемых сегодня непараметрических методов аппроксимации ФР и ПР случайных выборок, является, так называемый, параметр размытости *h*.

В статье проведен критический анализ известного подхода к оценке оптимального значения данного параметра.

Метод Розенблатта-Парзена

Напомним, следуя [13,14], что в соответствие с методом Розенблатта-Парзена аппроксимация ФР случайной выборки $\{x_1, x_2, ..., x_N\}$ $F_N^{\langle\text{approx}\rangle}\left(x, \{x_1, x_2, \ldots, x_N\}, h\right)$ вычисляется локально в любой точке *x* как сумма

$$F_N^{\langle\text{approx}\rangle}\left(x, \{x_1, x_2, \ldots, x_N\}, h\right) = \frac{1}{N}\sum_{i=1}^{N} K\left(\frac{x - x_i}{h}\right), \quad (1)$$

где

K(t) – так называемая ядерная функция, удовлетворяющая следующим условиям:

а) K(t) – монотонная неубывающая функция, область значений которой принадлежит интервалу [0,1];

б) K(t) = 1-K(t) –ядерная функция, симметричная относительно 0;

h – размытости параметр (называемый также полосой пропускания ядра), что $h \to 0$ при $N \to \infty$.

Соответственно, аппроксимация ПР $f^{\langle\text{approx}\rangle}\left(x, \{x_1, x_2, \ldots, x_N\}\right)$ вычисляется по формуле

$$f^{\langle\text{approx}\rangle}\left(x, \{x_1, x_2, \ldots, x_N\}, h\right) = \frac{1}{Nh}\sum_{i=1}^{N} k\left(\frac{y - x_i}{h}\right), \quad (2)$$

где

$$k(x) = \frac{d}{dx}K(x). \quad (3)$$

Таблица 1

Ядерные функции, используемые в методе Розенблатта-Парзена

№	Ядро	Функция
1	Нормальное	$k(t)=\frac{1}{\sqrt{2\pi}}e^{-\frac{t^2}{2}}$
2	Лапласа	$k(t)=\frac{1}{2}e^{-\lvert t\rvert}$
3	Фишера	$k(t)=\frac{1}{2\pi}\left(\frac{\sin\left(\frac{t}{2}\right)}{\frac{t}{2}}\right),\ \left\lvert\frac{t}{2}\right\rvert\le\pi$
4	Коши	$k(t)=\frac{1}{\pi}\left(\frac{1}{1+t^2}\right)$
5	Логистическое	$k(t)=\frac{e^{-t}}{\left(1+e^{-t}\right)^2}$
6	Епанечникова	$k(t)=\frac{3\cdot\left(1-\frac{t^2}{5}\right)}{4\sqrt{5}},\quad \lvert t\rvert\le\sqrt{5}$
7	Равномерное	$k(t)=\frac{1}{2},\quad \lvert t\rvert\le 1$
8	Треугольное	$k(t)=1-\lvert t\rvert,\quad \lvert t\rvert\le 1$
9	Квадратичное	$k(t)=\frac{3\cdot\left(1-t^2\right)}{4},\quad \lvert t\rvert\le 1$

Функции k(x), связанные с функцией K(x) соотношением (3), называемые ядреными функциями представлены в таблице 1.

При практическом использовании метода Розенблатта-Парзена приходится решать задачи:

1) выбора оптимальной ядерной функции k(x);

2) нахождение для выбранной ядерной функции оценки оптимального значения параметра размытости h*, обеспечивающего на выбранному критерию минимальное отклонение между аппроксимациями Розенблатта-Парзену и соответствующими эмпирическими ФР и ПР.

Отметим, что универсальных методов выбора ядерных функций и нахождения соответствующего оптимального значения размытости h*, определяющего вид ФР (1) и ПР (2) (см., например, [10] и рисунки 2,3, обсуждаемые далее), несмотря на непрерывно продолжающийся поиск данных методов (см. например, [18–24]), в настоящее время не создано. Основная проблема здесь связана с тем, что в теоретические формулы, предложенные для расчета оптимального значения параметра размытости h*, входит неизвестная теоретическая ПР $f^{\langle\text{theory}\rangle}(x)$ анализируемой выборки $\{x_1, x_2, ..., x_N\}$. В качестве примера

приведем формулу, полученную как результат минимизации среднего накопленного квадрат ошибки (англ. – Mean Integrated Squared Error, MISE)):

$$h^* = \arg\min_h \left[\int \left(f_N^{\langle \text{approx} \rangle}(\xi, h) - f^{\langle \text{theory} \rangle}(\xi) \right)^2 d\xi \right],$$

называемую в этой связи асимптотической MISE оценкой оптимального значения параметра размытости:

$$h^* = \frac{R(k)^{1/5}}{m_2(k)^{2/5} R\left(\left(f^{\langle \text{theory} \rangle} \right)'' \right)^{1/5} N^{1/5}}, \quad (4)$$

где $R(k) = \int k(\xi)^2 d\xi,$

$$m_2(k) = \int \xi^2 k(\xi) d\xi.$$

Из (4) видно, что асимптотическая MISE-оценка оптимального значения параметра размытости h* зависит от неизвестной плотности распределения $f^{\langle \text{theory} \rangle}(x)$. Отметим, что получить законченное аналитическое выражение для вычисления h*:

$$h^* = \left(\frac{4\hat{\sigma}}{3N} \right)^{1/5} = 1.06 \hat{\sigma} N^{-1/5},$$

где $\hat{\sigma}$ – среднеквадратическое отклонение анализируемой выборки $\{x_1, x_2, ..., x_N\}$, в даже в предположении о виде теоретической ПР $f^{\langle \text{theory} \rangle}(x)$ удается в единственном случае, когда удается функции $f^{\langle \text{theory} \rangle}(x)$ k(x) гауссовы. В остальных случаях найти аналогичные аналитические выражения, минимизируя средний накопленный квадрат ошибки разности теоретической ПР $f^{\langle \text{emp} \rangle}(x)$ и ее аппроксимации Розенблатта-Парзена $f_N^{\langle \text{approx} \rangle}(x)$

$$h^* = \arg\min_h \left[\int \left(f_N^{\langle \text{approx} \rangle}(\xi) - f_N^{\langle \text{emp} \rangle}(\xi) \right)^2 d\xi \right], \quad (5)$$

оказывается невозможным.

В этой связи на практике применяется следующий метод оценки эффективного значения параметра размытости h* аппроксимации Розенблатта-Парзена ФР и ПР случайной выборки $\{x_1, x_2, ..., x_N\}$ [25,10]. Для каждой из ядерных функций $k_m(t)$, $m = \overline{1,9}$, представленных в таблице 1, находятся значения h^*_m, обеспечивающие максимальные значения информационных функционалов

$$J_m\left(k_m(h_m)\right) = \int \ln\left(k_m(h_m, \xi)\right) k_m(h_m, \xi) d\xi.$$

Далее сравнивают максимальные значения информационных функционалов $\max_{h_m} \left(J(h_m) \right) = J_m\left(k_m(h_m^*)\right)$, и выбирают для вычисления аппроксимации Розенблатта-Парзена ту m_0-ую ядерную функцию, у которой значение информационного функционала $J_{m_0}\left(k_{m_0}\left(h^*_{m_0}\right)\right)$ оказывается наибольшим. Для дискретных данных информационный функционал принимает вид:

$$J\left(k_m(h_m)\right) = \frac{1}{N} \sum_{i=1}^{N} \ln\left[\frac{1}{(N-1) \cdot h_m} \sum_{j \neq i}^{N-1} k_m\left(\frac{x_i - x_j}{h_m} \right) \right], \quad (6)$$

соответственно, задачи нахождения абсциссы его максимального значения и выбора оптимальной ядерной функции записываются в виде:

$$h_m^* = \arg\max_{h_m} \left[J\left(k_m(h_m)\right) \right] =$$

$$\arg\max \left\{ \frac{1}{N} \sum_{i=1}^{N} \ln\left[\frac{1}{(N-1) \cdot h_m} \sum_{j \neq i}^{N-1} k_m\left(\frac{x_i - x_j}{h_m} \right) \right] \right\}, \quad (7)$$

$$\left\{ h^*_{m_0}, k_{m_0}, \right\} = \arg\max_{h_m^*, k_m(x)} \left(J_m\left(k_m\left(h_m^*\right)\right) \right). \quad (8)$$

Из (7) видно, что нахождение оптимального значения параметра размытости h^*_m для каждой из ядерных базисных функций $k_m(t)$ сводится к решению сложного нелинейного уравнения

$$\frac{\partial J\left(k_m(h_m)\right)}{\partial h_m} = \frac{\partial}{\partial h_m} \sum_{i=1}^{N} \ln\left[\frac{1}{(N-1) \cdot h_m} \sum_{j \neq i}^{N-1} k_m\left(\frac{x_i - x_j}{h_m} \right) \right] = 0, \quad (9)$$

найти которое оказывается возможным только численно с помощью какого-либо известного итерационного метода. Так как сходимость итерационных методов решения нелинейных уравнений к искомому решению зависит от выбора начального приближения, который, априори, не очевиден, на практике ищут не решение (9), но вычисляют на выбранном интервале $\left[h_m^{\langle \min \rangle}, h_m^{\langle \max \rangle} \right]$ значения информационного функционала (6) и далее находят значение аргумента h^*_m, соответствующее максимальному значению информационного функционала $J\left(k_m(h_m)\right)$ [10]. Однако, при использовании обсуждаемого подхода возникает целый ряд проблем, в том числе:

1) выбора интервала $\left[h_m^{\langle \min \rangle}, h_m^{\langle \max \rangle} \right]$, состоящая в том, что в условиях отсутствиях правил обоснованного выбора границ интервала поиска максимального значения информационного функционала (6), искомый локальный минимум может отсутствовать на выбранном интервале;

2) наличия у вычисленной зависимости $J\left(k_m(h_m)\right)$ разрывов первого и второго (пример подобной зависимости представлен на рисунке 1).

Из рисунка 1 видно, видно, что что при $h \in \left[10^{-4}, 10^{-1} \right]$ функция в точках отрезка $\left[1.1643 \cdot 10^{-3}; 1.1645 \cdot 10^{-3} \right]$ имеет разрывы первого рода, в точках отрезка $\left[2.6196 \cdot 10^{-3}; 2.6230 \cdot 10^{-3} \right]$ – разрывы второго рода. Для значений параметра размытости h, не принадлежащих данным отрезкам, функ-

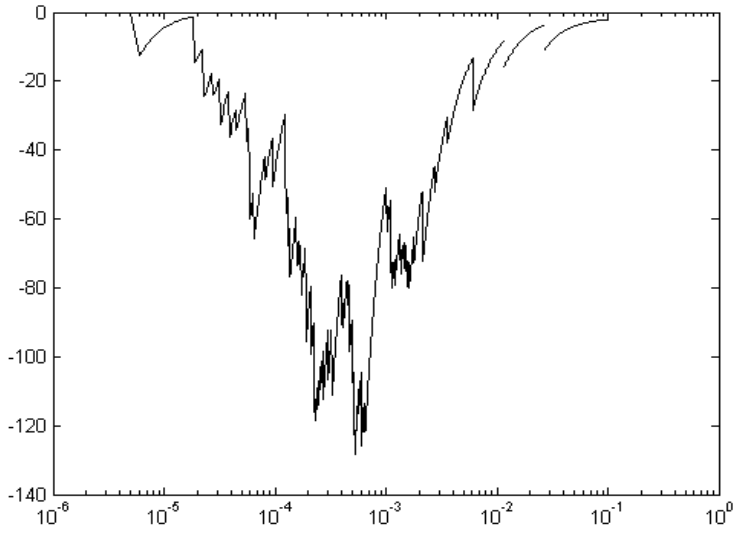

Рис. 1. График зависимости $\ln\left[J\left(k_1\left(h_1\right)\right)\right] = f\left[\ln\left(h_1\right)\right]$, вычисленной для случайной последовательности $x_i, i = \overline{1,100}$, сгенерированной в соответствие с нормальным законом распределения *N*(1,4), в пакете MATLAB

ция $\ln\left[J\left(k_1\left(h_1\right)\right)\right] = f\left[\ln\left(h_1\right)\right]$ оказывается непрерывной. Обсуждаемая особенность функции $J\left(k_1\left(h_1\right)\right)$, которая у ядерных функций с ограниченной область определения $\left(m = 3, 6, 7-9\right)$ отсутствует, обусловлена конечной точностью машинной арифметики в (см., [7, Глава 3]). Также в [7]:

– обоснован выбор интервалов поиска оптимальных значений параметров размытости h^*_m и предложен алгоритм его поиска для ядерных функций с неограниченной областью определения (1, 2, 4, 5);

– получены аналитические выражения для вычисления соответствующих оптимальных значений параметров размытостей для ядерных функций с ограниченной областью определения (m = 3, 6, 7, - 9):

$$h^*_3 = \frac{\max\left(x_i - x_j\right)}{2\pi},\ h^*_6 = \frac{\max\left(x_i - x_j\right)}{\sqrt{5}},$$
$$h^*_{7,8,9} \geq \max\left(x_i - x_j\right);$$

– предложена модифицированная формула для вычисления значений информационного функционала $J(k(h_m))$ целочисленных случайных последовательностей:

$$J\left(k\left(h_m\right)\right) = \frac{1}{N}\sum_{i=1}^{N}\ln\left[\frac{1}{(N-1)\cdot h_m}\sum_{x_i \neq x_j}^{Ns-1} k_m\left(\frac{x_i - x_j}{h_m}\right)\right].$$

Отметим, что рассмотренный выше алгоритм реализации метода Розенблатта-Парзена имеет очевидный недостаток, связанный с тем, что здесь, де-факто, отсутствуют количественные критерии позволяющие количественно оценить качество аппроксимации эмпирической ФР и ПР. Отмеченную проблему иллюстрирует рисунок 2, на котором представлены аппроксиммации Розенблатта-Парзена с нормальной ядерной функцией случайной выборки $\{x_1, x_2, ..., x_{50}\}$, сгенерированной в соответствие с логнормальным законом распределения

$$f_{LN}\left(x,\mu,\sigma\right) = \frac{1}{x\sigma\sqrt{2\pi}}e^{-\left(\ln(x)-\mu\right)^2/2\sigma^2},$$

у которого $\mu = 0.3,\ \sigma = 0.03$.

Из рисунка 1 видно, что форма аппроксимаций ПР и ФР случайной выборки анализируемой случайной выборки $\{x_1, x_2, ..., x_{50}\}$ определяется значением параметра размытости:

– график функция $F_{50}^{\langle\text{approx}\rangle}\left(x,\left\{x_1,x_2,\ldots,x_{50}\right\},0.0029\right)$ при выбранном масштабе координатных осей визуально неотличима от соответствующего графика эмпирической ФР, одновременно, функция $f_{50}^{\langle\text{approx}\rangle}\left(x,\left\{x_1,x_2,\ldots,x_{50}\right\},0.0029\right)$, имеющая 17 локальных максимумов, оказывается существенно отличной от гистограммы анализируемой случайной выборки $\{x_1, x_2, ..., x_{50}\}$,

– по мере увеличения h функция $F_{50}^{\langle\text{approx}\rangle}\left(x,\left\{x_1,x_2,\ldots,x_{50}\right\},h\right)$ стремится к ФР случайной величины с равномерным законом распределения на интервале $[x_1, x_{50}]$ при этом, одновременно, уменьшается число локальных максимумов функции $f_{50}^{\langle\text{approx}\rangle}\left(x,\left\{x_1,x_2,\ldots,x_{50}\right\},h\right)$, которое оказывается равным 1 при h=0.0349 и 0 при h=0.4028.

Следовательно, при выборе оптимального значения параметра размытости h* необходимо обеспечивать компромисс между

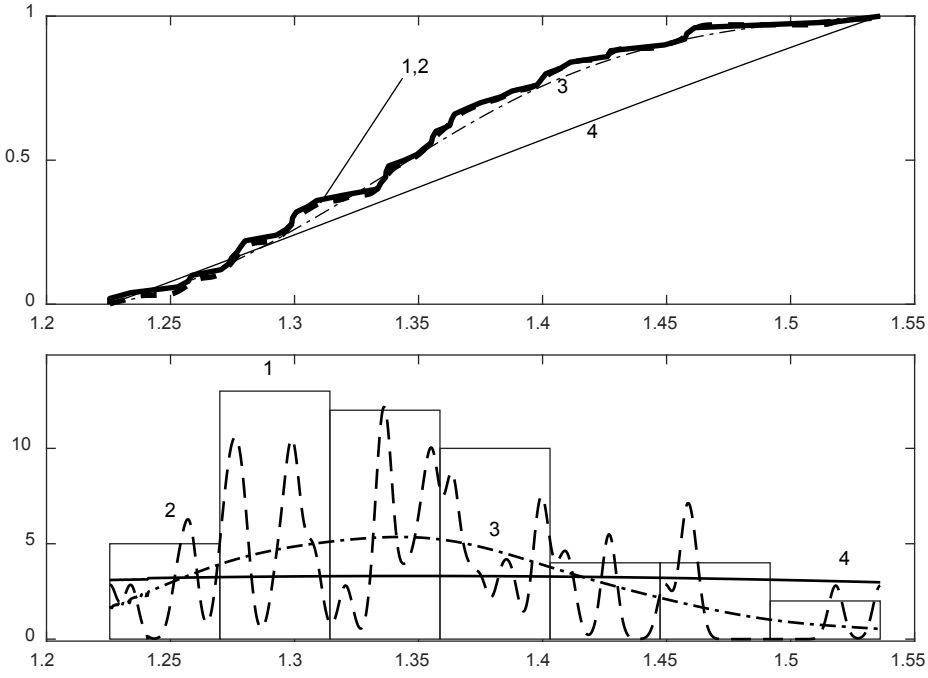

Рис. 2. Случайная выборка {x_1, x_2, ..., x_{50}}, сгенерированная в соответствие с логнормальным законом распределения $f_{LN}(x, 0.3, 0.03)$: сверху 1 – эмпирическая ПР, 2 – график функции $F_{50}^{\langle \text{approx} \rangle}\left(x, \{x_1, x_2, \ldots, x_{50}\}, 0.0029\right)$, 3 – график функции $F_{50}^{\langle \text{approx} \rangle}\left(x, \{x_1, x_2, \ldots, x_{50}\}, 0.0349\right)$, 3 – график функции $F_{50}^{\langle \text{approx} \rangle}\left(x, \{x_1, x_2, \ldots, x_{50}\}, 0.0349\right)$, 4 – график функции $F_{50}^{\langle \text{approx} \rangle}\left(x, \{x_1, x_2, \ldots, x_{50}\}, 0.4028\right)$; 1 – гистограмма случайной выборки {x_1, x_2, ..., x_{50}}; 2 – график функции $f_{50}^{\langle \text{approx} \rangle}\left(x, \{x_1, x_2, \ldots, x_{50}\}, 0.0029\right)$, 3 – график функции $f_{50}^{\langle \text{approx} \rangle}\left(x, \{x_1, x_2, \ldots, x_{50}\}, 0.0349\right)$, 4 – график функции $f_{50}^{\langle \text{approx} \rangle}\left(x, \{x_1, x_2, \ldots, x_{50}\}, 0.4028\right)$

требованием «похожести» функции $F_{50}^{\langle \text{approx} \rangle}\left(x, \{x_1, x_2, \ldots, x_{50}\}, h\right)$ и эмпирической ФР, а также «похожести» функции $f_{50}^{\langle \text{approx} \rangle}\left(x, \{x_1, x_2, \ldots, x_{50}\}, h\right)$ и эмпирической ФР, в то время как устоявшаяся точка зрения рекомендует находить h* как решение задачи (7), в постановке которой отсутствуют и эмпирическая ПР и эмпирическая ФР.

В этой связи авторы провели самостоятельные исследования и получили количественные качества аппроксимаций Розенблатта-Празена ПР и ФР случайных выборок, обсуждаемые далее.

Методика анализа качества аппроксимаций Розенблатта-Парзена ПР и ФР случайных выборок

В качестве критериев качества аппроксимации Розенблатта-Парзена были выбраны следующие критерии, традиционно используемые в задачах аппроксимации функций, заданных таблично, и функций распределений независимых случайных выборок с законами распределения $F_1(x)$ и $F_2(x)$:

– накопленная среднеквадратическая ошибка разности между функциями $F_1(x)$ и $F_2(x)$, вычисляемая по формуле

$$M_1 = \left[\frac{1}{x_N - x_1} \int_{x_1}^{x_N} \left(F_1(\xi) - F_2(\xi) \right)^2 d\xi \right]^{1/2},$$

которая для дискретных выборок записывается в виде:

$$M_1 = \left[\frac{1}{N} \sum_{i=1}^{N} \left(F_1(x_i) - F_2(x_i) \right)^2 \right]^{1/2}; \quad (10)$$

– критерий типа Колмогорова-Смирнова:

$$M_2 = \max_{\overline{1,N}, \overline{1,N_h}} \left(\left| F_1(x_i) - F_2(x_i) \right| \right), \quad (11)$$

где $i = \overline{1,N}, j = \overline{1,N_h}$, используемый для проверки статистических гипотез вида:

$$H_0 : F_1(x) = F_2(x),$$
$$H_1 : F_1(x) \neq F_2(x).$$

Исследование особенностей качества аппроксимаций ПР и ФП случайных выборок было проведено в соответствие с методикой, реализующейся выполнением следующей последовательности действий.

1. Выбор закона распределения случайной выборки $F^{\langle \text{theory} \rangle}(x, \boldsymbol{\alpha})$, $\boldsymbol{\alpha}$ – вектор параметров.

2. Выбор объема случайной выборки N.

3. Генерация случайной выборки $\{\tilde{x}_1, \tilde{x}_2, \ldots, \tilde{x}_N\}$ в соответствие в выбранным законом распределения $F^{\langle \text{theory} \rangle}(x, \boldsymbol{\alpha})$.

4. Вычисление в соответствие с методом максимального правдоподобия $\tilde{\alpha}$ – оценок параметров закона распределения $F^{\langle \text{estimation} \rangle}\left(x, \tilde{\alpha}\right)$ случайной выборки $\left\{\tilde{x}_1, \tilde{x}_2, \ldots, \tilde{x}_N\right\}$.

5. Упорядочивание случайной выборки $\left\{\tilde{x}_1, \tilde{x}_2, \ldots, \tilde{x}_N\right\}$ в порядке возрастания значений $\tilde{x}_i$, $i = \overline{1, N}$ (далее – случайная выборка $\left\{\tilde{x}_1, \tilde{x}_2, \ldots, \tilde{x}_N\right\}$).

6. Выбор диапазона изменения значений параметра размытости h, $h \in \left[h_{\min}, h_{\max}\right]$.

7. Выбор N_h.– числа узлов сетки по параметру размытости h.

8. Вычисление координат узлов сетки по параметру размытости h:

$$h_j = h_{\min} + \frac{h_{\max} - h_{\min}}{N_h - 1}\left(j - 1\right), \ j = \overline{1, N_h}.$$

9. Выбор N_x – числа узлов сетки по переменной x, используемой для вычисления значений аппроксимаций Розенблатта-Парзена ПР случайной выборки $\{x_1, x_2, ..., x_N\}$.

10. Вычисление координат узлов сетки по переменной x:

$$y_k = x_1 + \frac{x_N - x_1}{N_x - 1}\left(k - 1\right), \ k = \overline{1, N_x}. \quad (12)$$

11. Выбор ядерной функции ядерной функции k(t).

12. Вычисление аппроксимаций Розенблатта-Парзена с использованием выбранной ядерной функции ПР случайной выборки $\{x_1, x_2, ..., x_N\}$ в каждом узле сетки по параметру h $f^{\langle \text{approx} \rangle}\left(y_k, h_j\right)$, $j = \overline{1, N_h}$, y_k – координаты узлов сетки (12).

13. Вычисление численным интегрированием аппроксимаций Розенблатта-Парзена ФР случайной выборки $\{x_1, x_2, \ldots, x_N\}$ в каждом узле сетки по параметру h $F^{\langle \text{approx} \rangle}\left(y_k, h_j\right)$, $j = \overline{1, N_h}$:

$$F^{\langle \text{approx} \rangle}\left(y_k, h_j\right) = \int_{x_1}^{y_k} f^{\langle \text{appox} \rangle}\left(\xi, h_j\right) d\xi.$$

14. Нормировка аппроксимаций Розенблатта-Парзена ФР случайной выборки $\{x_1, x_2, ..., x_N\}$ $F^{\langle \text{approx} \rangle}\left(y_k, h_j\right)$, вычисленных в каждом узле сетки по параметру h, $k = \overline{1, N_x}$, $j = \overline{1, N_h}$:

$$F^{\langle \text{norm approx} \rangle}\left(y_k, h_j\right) = F^{\langle \text{approx} \rangle}\left(y_k, h_j\right) \Big/ F^{\langle \text{approx} \rangle}\left(x_N, h_j\right).$$

15. Вычисление на основе использования зависимости $F^{\langle \text{norm approx} \rangle}\left(y_k, h_j\right)$, $k = \overline{1, N_x}$, $j = \overline{1, N_h}$ с помощью линейной интерполяции значений аппроксимации Розенблатта-Парзена ФР случайной выборки $\{x_1, x_2, ..., x_N\}$ в точках $x_1, x_2, ..., x_N$ – $F^{\langle \text{norm approx} \rangle}\left(x_i, h_j\right)$, $i = \overline{1, N}$, $j = \overline{1, N_h}$.

16. Вычисление эмпирической ФР случайной выборки $\{x_1, x_2, ..., x_N\}$ $F^{\langle \text{emp} \rangle}\left(x_i, h_j\right)$, $i = \overline{1, N}, j = \overline{1, N_h}$.

17. Вычисление зависимости $M_1(h_i)$ в соответствие с (10) для следующих пар зависимостей $F_1(x_i)$, $F_2(x_i)$, представленных в таблице 2.

Таблица 2

Пары ФР и ПР, использованных для вычисления зависимостей критерия оценки качества аппроксимации Розенблатта-Парзена M_1 от параметра размытости h

$F_1\left(x\right)$	$F^{\langle \text{theory} \rangle}\left(x_i, \alpha\right)$	$F^{\langle \text{estimation} \rangle}\left(x_i, \tilde{\alpha}\right)$	$f^{\langle \text{theory} \rangle}\left(x_i, \alpha\right)$	$f^{\langle \text{estimation} \rangle}\left(x_i, \tilde{\alpha}\right)$
$F_2\left(x\right)$	$F^{\langle \text{norm approx} \rangle}\left(x_i, h_j\right)$	$F^{\langle \text{norm approx} \rangle}\left(x_i, h_j\right)$	$f^{\langle \text{norm approx} \rangle}\left(x_i, h_j\right)$	$f^{\langle \text{norm approx} \rangle}\left(x_i, h_j\right)$

Таблица 3

Пары зависимостей, использованных для вычисления зависимостей критерия оценки качества аппроксимации Розенблатта-Парзена M_1 от параметра размытости h

$F_1\left(x\right)$	$F^{\langle \text{theory} \rangle}\left(x_i, \alpha\right)$	$F^{\langle \text{estimation} \rangle}\left(x_i, \tilde{\alpha}\right)$
$F_2\left(x\right)$	$F^{\langle \text{norm approx} \rangle}\left(x_i, h_j\right)$	$F^{\langle \text{norm approx} \rangle}\left(x_i, h_j\right)$

18. Вычисление зависимостей $M_2\left(h_j\right)$ в соответствие с (11) на основе результатов проверки статистических гипотез для каждого из узлов выбранной сетки по параметру размытости h:

$$H_0 : F_1\left(x\right) = F_2\left(x\right), H_1 : F_1\left(x\right) \neq F_2\left(x\right),$$

если верна H_0 $H(h_j) = 0$ иначе $H(h_j) = 1$ для следующих пар зависимостей $F_1(x_i)$, $F_2(x_i)$, представленных в таблице 3.

19. Вычисление для выбранной ядерной функции в соответствие (6) зависимости информационного функционала J(h) от параметра размытости h.

20. Вычисление для выбранного ядра оптимального в смысле (7) значения параметра размытости h*.

21. Анализ зависимостей $M_1(h_j)$, $M_2(h_j)$ $f^{\langle \text{approx} \rangle}\left(x_k, h_j\right)$, $F^{\langle \text{norm approx} \rangle}\left(x_k, h_j\right)$, $H(h_j)$, $i = \overline{1, N}$, $j = \overline{1, N_h}$.

Анализ качества аппроксимаций Розенблатта-Парзена ПР и ФР случайных выборок, вычисляемых на основе анализа зависимости

$$J = J(h)$$

Рассмотрим типичные зависимости $M_1(h_j)$, $M_2(h_j)$ для случайной выборки $\{x_1, x_2, ..., x_{50}\}$, сгенерированной в соответствие с нормальным законом распределения N(1,3), вычисленные для аппроксимаций Розенблатта-Парзена эмпирических ПР и ФР с нормальной ядерной функцией (см. рисунок 3).

Зависимости

$$J = J\left(h\right),$$

$$M_1^{\langle 1 \rangle} = M_1\left(F^{\langle \text{theory} \rangle}\left(x_i, \boldsymbol{\alpha}\right), F^{\langle \text{approx} \rangle}\left(x_i, h_i\right), h_i\right),$$

$$M_1^{\langle 2 \rangle} = M_1\left(F^{\langle \text{estimation} \rangle}\left(x_i, \tilde{\boldsymbol{\alpha}}\right), F^{\langle \text{approx} \rangle}\left(x_i, h_i\right), h_i\right),$$

$$M_1^{\langle 3 \rangle} = M_1\left(f^{\langle \text{theory} \rangle}\left(x_i, \boldsymbol{\alpha}\right), f^{\langle \text{approx} \rangle}\left(x_i, h_i\right), h_i\right),$$

$$M_1^{\langle 4 \rangle} = M_1\left(f^{\langle \text{estimation} \rangle}\left(x_i, \tilde{\boldsymbol{\alpha}}\right), f^{\langle \text{approx} \rangle}\left(x_i, h_i\right), h_i\right),$$

представлены на рисунке 4.

Из рисунка 4 видно, что на рассматриваемом интервале зависимость J=J(h) имеет локальный максимум в точке h*=2.34; зависимости $M_1^{\langle 1 \rangle}(h_i)$, $M_1^{\langle 2 \rangle}(h_i)$ достигают локального минимума в точках $h_{\min}^{\langle 1 \rangle} = 3.01$, $h_{\min}^{\langle 2 \rangle} = 3.65$, соответственно; зависимости $M_1^{\langle 3 \rangle}(h_i)$, $M_1^{\langle 4 \rangle}(h_i)$ – в точках $h_{\min}^{\langle 3 \rangle} = 3.09$, $h_{\min}^{\langle 4 \rangle} = 2.67$. Значения зависимостей $J = J(h)\, M_1^{\langle 1 \rangle}(h_i), M_1^{\langle 2 \rangle}(h_i), M_1^{\langle 3 \rangle}(h_i)$, $M_1^{\langle 4 \rangle}(h_i)$ для указанных выше значений аргумента h^*, $h_{\min}^{\langle 1 \rangle}$, $h_{\min}^{\langle 2 \rangle}$, $h_{\min}^{\langle 3 \rangle}$, $h_{\min}^{\langle 4 \rangle}$ представлены в таблице 3.

Из таблицы 3 видно, что $M_1^{\langle 1 \rangle}\left(h^*\right) > M_1^{\langle 1 \rangle}\left(h_{\min}^{\langle 1 \rangle}\right)$, $M_1^{\langle 2 \rangle}\left(h^*\right) > M_1^{\langle 2 \rangle}\left(h_{\min}^{\langle 2 \rangle}\right)$,

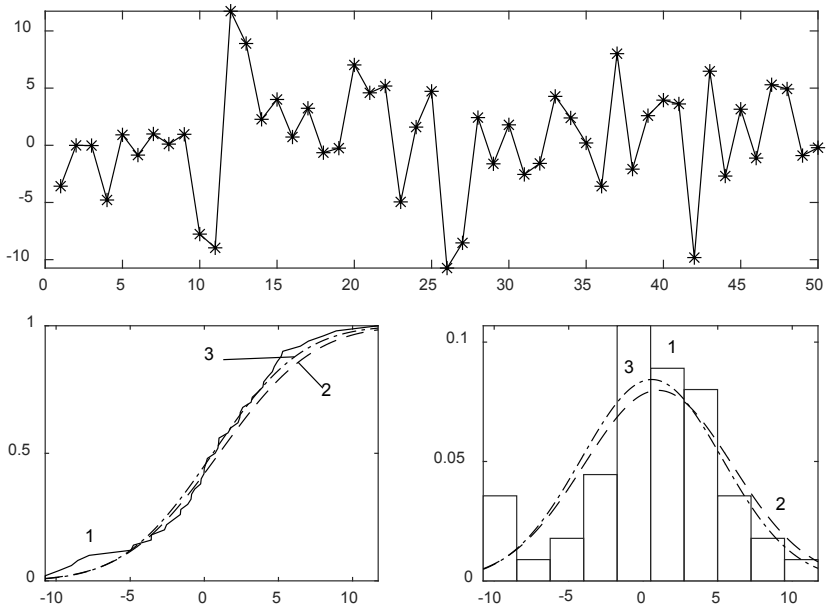

Рис. 3. Случайная выборка $\{x_1, x_2, ..., x_{50}\}$, сгенерированная в соответствие с нормальным законом распределения N(1,4), и ее статистические характеристики (на рисунке внизу слева: 1 – $F^{\langle \text{emp} \rangle}(x)$, 2 – $F^{\langle \text{theory} \rangle}(x, \boldsymbol{\alpha})$, 2 – $F^{\langle \text{estimation} \rangle}(x, \tilde{\boldsymbol{\alpha}})$; на рисунке внизу справа: 1– гистограмма случайной выборки $\{x_1, x_2, ..., x_{50}\}$, 2 – $f^{\langle \text{theory} \rangle}(x, \boldsymbol{\alpha})$, 3 – $f^{\langle \text{estimation} \rangle}(x, \tilde{\boldsymbol{\alpha}})$

$M_1^{\langle 3 \rangle}\left(h^*\right) > M_1^{\langle 3 \rangle}\left(h_{\min}^{\langle 3 \rangle}\right)$, $M_1^{\langle 4 \rangle}\left(h^*\right) > M_1^{\langle 4 \rangle}\left(h_{\min}^{\langle 4 \rangle}\right)$.

При этом минимальное значение вероятности принятия гипотезы H_0 для значений параметра размытости $h_{\min}^{\langle 1 \rangle}$, $h_{\min}^{\langle 2 \rangle}$, $h_{\min}^{\langle 3 \rangle}$, $h_{\min}^{\langle 4 \rangle}$, представленных в таблице 3, оказалось равным 0.91.

Аналогично результаты были получены и случайных выборок, генерируемых в соответствие с другими известными законами распределения, что иллюстрируют, например, результаты аппроксимации случайной выборки объемом N = 50, сгенерированной в соответствие с логнормальным законом распределения

$$f_{LN}(x,\mu,\sigma)=\frac{1}{x\sigma\sqrt{2\pi}}e^{-(\ln(x)-\mu)^2/2\sigma^2},$$

в котором $\mu=0.3,\ \sigma=0.03$, представленные на рисунках 5, 6 и таблице 4.

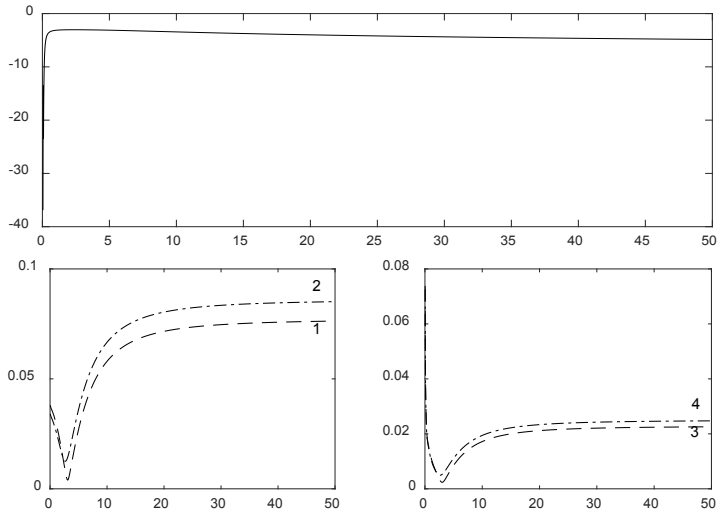

Рис. 4. Случайная выборка, сгенерированная в соответствие с нормальным законом N(1,4). Графики зависимостей: J = J(h) (сверху); $M_1^{\langle 1\rangle}(h_i)$, $M_1^{\langle 2\rangle}(h_i)$ (внизу слева); $M_1^{\langle 3\rangle}(h_i)$, $M_1^{\langle 4\rangle}(h_i)$ (внизу справа)

Таблица 3

Значения зависимостей $M_1^{\langle 1\rangle}(h_i), M_1^{\langle 2\rangle}(h_i), M_1^{\langle 3\rangle}(h_i), M_1^{\langle 4\rangle}(h_i)$ в выбранных точках (случайная выборка, сгенерированная в соответствие с нормальным законом N (1,4) распределения

hi	$M_1^{\langle 1\rangle}(h_i)$	$M_1^{\langle 2\rangle}(h_i)$	$M_1^{\langle 3\rangle}(h_i)$	$M_1^{\langle 4\rangle}(h_i)$
2.34	0.0129	0.0137	0.0042	0.0052
3.01	0.0040	–	–	–
3.65	–	0.0125	–	–
3.09	–	–	0.0024	–
2.67	–	–	–	0.0050

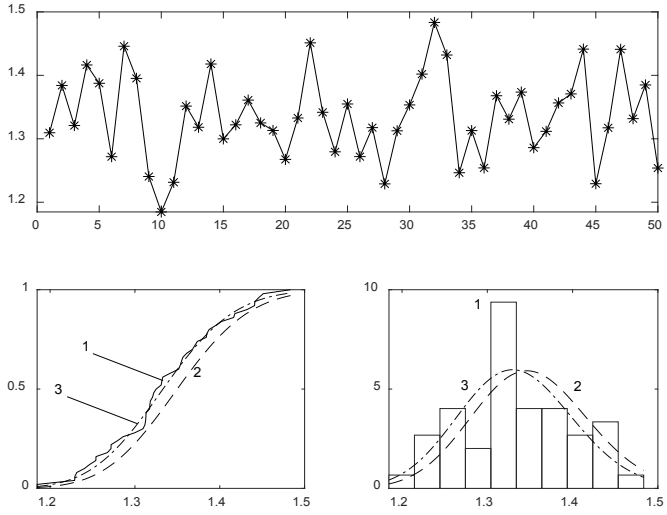

Рис. 5. Случайная выборка $\{x_1, x_2, ..., x_{50}\}$, сгенерированная в соответствие с логнормальным законом распределения $f_{LN}(x,0.3,0.03)$, и ее статистические характеристики (на рисунке внизу слева: 1 – $F^{\langle \text{emp}\rangle}(x)$, 2 – $F^{\langle \text{theory}\rangle}(x,\boldsymbol{\alpha})$, 2 – $F^{\langle \text{estimation}\rangle}(x,\tilde{\boldsymbol{\alpha}})$; на рисунке внизу справа: 1– гистограмма случайной выборки $\{x_1, x_2, ..., x_{50}\}$, 2 – $f^{\langle \text{theory}\rangle}(x,\boldsymbol{\alpha})$, 3 – $f^{\langle \text{estimation}\rangle}(x,\tilde{\boldsymbol{\alpha}})$

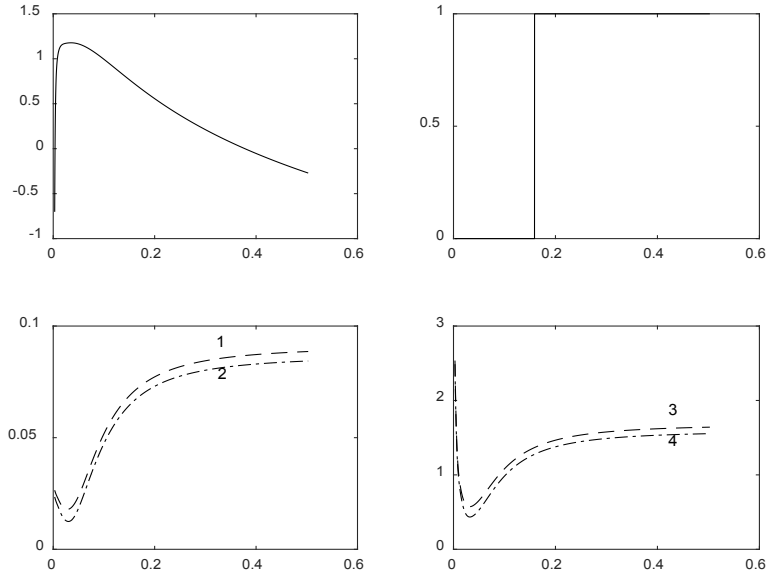

Рис. 6. Случайная выборка {x1, x2, ..., x50}, сгенерированная в соответствие с логнормальным законом распределения $f_{LN}(x, 0.3, 0.03)$. Графики зависимостей: J = J(h) (сверху слева), M2(hi) (сверху справа); $M_1^{\langle 1\rangle}(h_i)$, $M_1^{\langle 2\rangle}(h_i)$ (внизу слева); $M_1^{\langle 3\rangle}(h_i)$, $M_1^{\langle 4\rangle}(h_i)$ (внизу справа)

Таблица 4

Значения зависимостей $M_1^{\langle 1\rangle}(h_i), M_1^{\langle 2\rangle}(h_i), M_1^{\langle 3\rangle}(h_i), M_1^{\langle 4\rangle}(h_i)$ в выбранных точках (случайная выборка {x_1, x_2, ..., x_{50}}, сгенерированная в соответствие с логнормальным законом распределения $f_{LN}(x, 0.3, 0.03)$)

hi	$M_1^{\langle 1\rangle}(h_i)$	$M_1^{\langle 2\rangle}(h_i)$	$M_1^{\langle 3\rangle}(h_i)$	$M_1^{\langle 4\rangle}(h_i)$
0.0361	0.0299	0.0154	0.7102	0.4194
3.01	0.0293	–	–	–
3.65	–	0.0095	–	–
3.09	–	–	0.7065	–
2.67	–	–	–	0.3588

Из рисунка 6 видно, что значения параметров размытости h*, $h_{\min}^{\langle 1\rangle}$, $h_{\min}^{\langle 2\rangle}$, $h_{\min}^{\langle 3\rangle}$, $h_{\min}^{\langle 4\rangle}$ таковы, что они попадают в диапазон значений параметра размытости [0, 0.1589], в котором принимается гипотеза H_0.

Из таблицы 4 видно, что, как и в предыдущем случае, $M_1^{\langle 1\rangle}(h^*) > M_1^{\langle 1\rangle}(h_{\min}^{\langle 1\rangle})$, $M_1^{\langle 2\rangle}(h^*) > M_1^{\langle 2\rangle}(h_{\min}^{\langle 2\rangle})$, $M_1^{\langle 3\rangle}(h^*) > M_1^{\langle 3\rangle}(h_{\min}^{\langle 3\rangle})$, $M_1^{\langle 4\rangle}(h^*) > M_1^{\langle 4\rangle}(h_{\min}^{\langle 4\rangle})$. При этом минимальное значение вероятности принятия гипотезы H_0 для значений параметра размытости $h_{\min}^{\langle 1\rangle}$, $h_{\min}^{\langle 2\rangle}$, $h_{\min}^{\langle 3\rangle}$, $h_{\min}^{\langle 4\rangle}$, представленных в таблице 3, оказалось более 0.78.

Таким образом, обсуждаемые выше результаты исследований опровергают устоявшуюся точку зрения о том, что максимально возможная точность аппроксимации Розенблатта-Парзена ПР и ФР случайной выборки {x_1, x_2, ..., x_N}, оцениваемая накопленным среднеквадратическим отклонением разностей $F^{\langle \text{theory}\rangle}(x_i, \boldsymbol{\alpha}) - F^{\langle \text{norm approx}\rangle}(x_i, h_j)$, $F^{\langle \text{estimation}\rangle}(x_i, \tilde{\boldsymbol{\alpha}}) - F^{\langle \text{norm approx}\rangle}(x_i, h_j)$, $f^{\langle \text{theory}\rangle}(x_i, \boldsymbol{\alpha}) - f^{\langle \text{norm approx}\rangle}(x_i, h_j)$, $f^{\langle \text{estimation}\rangle}(x_i, \tilde{\boldsymbol{\alpha}}) - f^{\langle \text{norm approx}\rangle}(x_i, h_j)$, достигается при использовании оптимального значения параметра размытости h*, являющегося решением задачи (7).

В этой связи возникает необходимость использования альтернативных методов оценивания оптимального значения параметра размытости h*, основанных на сравнении доступных исследователю эмпирических ФР и ПР с их аппроксимациями Розенблатта-Парзена, например, на основе анализа накопленной среднеквадратической ошибки $M_1^{\langle 5\rangle}(h) = M_1\left(f^{\langle \text{approx}\rangle}(x_j, h), f_N^{\langle \text{emp}\rangle}(x_j)\right)$,

где $f_N^{\langle \text{emp} \rangle}(x_j)$ - кусочно-постоянная аппроксимация эмпирической ПР случайной выборки $\{x_1, x_2, ..., x_N\}$,

$$f_N^{\langle \text{emp} \rangle}(y) = \frac{F_N^{\langle \text{emp} \rangle}(y+h) - F_N^{\langle \text{emp} \rangle}(y)}{h} = \frac{1}{Nh}\sum_{i=1}^{N}\left[\theta(y+h-x_i) - \theta(y-x_i)\right] = \frac{\nu_y}{Nh}, \quad (13)$$

здесь ν_y – количество значений случайной выборки $\{x_1, x_2, ..., x_N\}$, попавших в интервал [y, y+h], обсуждаются в следующем разделе.

Их выбор обусловлен тем, что исследователю статистических характеристик случайных выборок оказываются доступными исключительно эмпирические ФР и ПР анализируемых выборок.

Анализ качества аппроксимаций Розенблатта-Парзена ПР и ФР случайных выборок на основе анализа зависимости $M_1^{\langle 5 \rangle}(h)$

Рассмотрим типичную зависимость $M_1^{\langle 5 \rangle}(h)$ случайной выборки $\{x_1, x_2, ..., x_{50}\}$, сгенерированной в соответствие с логнормальным законом распределения $f_{LN}(x, \mu, \sigma)$ $\mu = 0.3$, $\sigma = 0.03$, представленную на рисунке 6. Здесь число интервалов кусочно-постоянной аппроксимации эмпирической ПР (гистограммы) N_g выбиралось в соответствии с правилом правило Стерджеса:

$$N_g = \left[1 + \log_2 N\right],$$

где [ ] – целая часть выражения.

Из рисунка 6 видно, что анализируемая

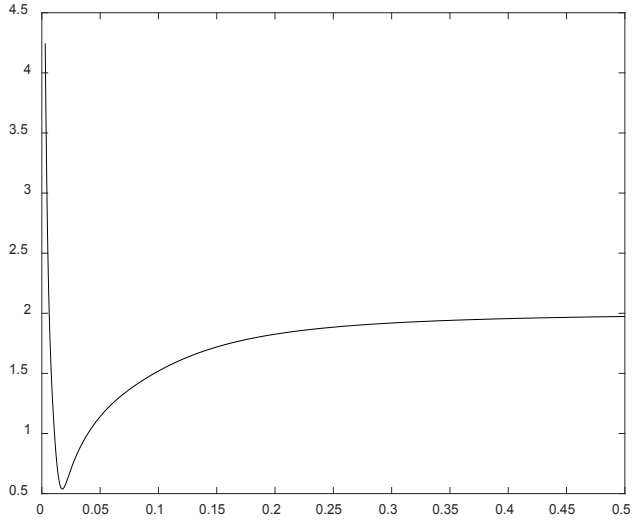

Рис. 6. Случайная выборка $\{x_1, x_2, ..., x_{50}\}$, сгенерированная в соответствие с логнормальным законом распределения $f_{LN}(x, 0.3, 0.03)$. График зависимости $M_1^{\langle 5 \rangle} = M_1^{\langle 5 \rangle}(h)$

зависимость достигает $M_1^{\langle 5 \rangle}(h)$ своего минимального 0.5385 в точке $h_{\min}^{\langle 5 \rangle} = 0.0176$. Оценка оптимального значения параметра размытости для данной выборки, вычисленная как решение задачи (7), оказалась равной $h^* = 0.0156$, значения анализируемых зависимостей в выбранных точках $M_1^{\langle 5 \rangle}(h^*) = 0.9208$, $M_1^{\langle 5 \rangle}(h_{\min}^{\langle 5 \rangle}) = 0.5385$. Так как $M_1^{\langle 5 \rangle}(h^*) > M_1^{\langle 5 \rangle}(h_{\min}^{\langle 5 \rangle})$, качество аппроксимации Розенблатта-Парзена эмпирической ПР для значения $h = h_{\min}^{\langle 5 \rangle}$ выше качества обсуждаемых аппроксимаций при h = h*. Соответствующие графики функций $F^{\langle \text{emp} \rangle}(x)$, $F^{\langle \text{norm approx} \rangle}(x, h_{\min}^{\langle 5 \rangle})$, $F^{\langle \text{norm approx} \rangle}(x, h^*)$, $f^{\langle \text{emp} \rangle}(x)$, $f^{\langle \text{approx} \rangle}(x, h_{\min}^{\langle 5 \rangle})$, $f^{\langle \text{approx} \rangle}(x, h^*)$ представлены на рисунке 7.

Аналогичные свойства имеют аппроксимации Розенблатта-Парзена эмпирических ФР и ПР с иными другими ядерными функциями случайных выборок, сгенерированных в соответствие с законами распределений, отличными от рассмотренными в статье.

Заключение

Результаты проведенных исследований особенностей аппроксимации Розенблатта-Парзена с помощью нормальной ядерной функции k(t) ПР и ФР упорядоченных случайных выборок $\{x_1, x_2, ..., x_n\}$, сгенерированных в соответствие с нормальным и логнормальным распределениями опровергают устоявшуюся точку зрения, рекомендующую использовать в качестве оптимального значения параметра размытости h* решение задачи (7).

Оценку данного параметра следует находить как значение параметра размытости $h_{\min}^{\langle 5 \rangle}$, которое обеспечивает достижение минимального значения накопленной случайной ошибки разности между эмпирической

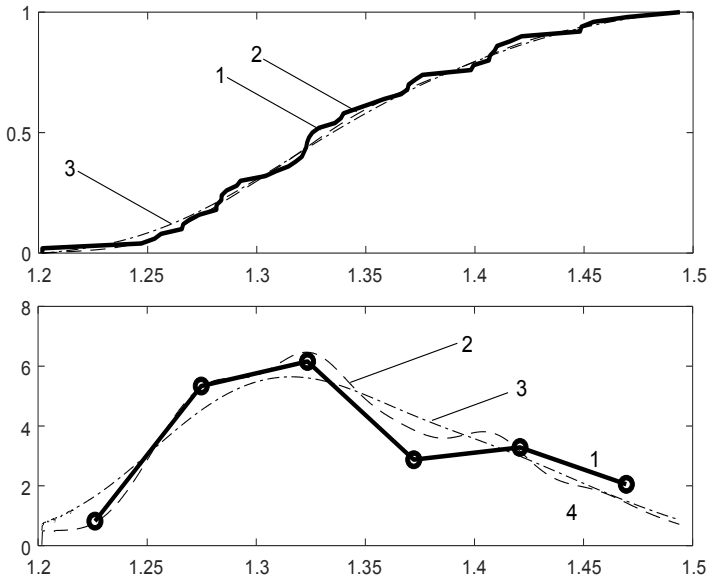

Рис. 7. Графики функций: сверху – 1 – $F^{\langle \text{emp} \rangle}(x)$, 2 – $F^{\langle \text{norm approx} \rangle}\left(x, h_{\min}^{\langle 5 \rangle}\right)$, 3 – $F^{\langle \text{norm approx} \rangle}\left(x, h^{*}\right)$; снизу – $f^{\langle \text{emp} \rangle}(x)$, 2 – $f^{\langle \text{approx} \rangle}\left(x, h_{\min}^{\langle 5 \rangle}\right)$, 3 – $f^{\langle \text{approx} \rangle}\left(x, h^{*}\right)$

ПР $f^{\langle \text{emp} \rangle}(x)$ и аппроксимацией Розенблатта-Парзена ПР $f^{\langle \text{approx} \rangle}(x, h)$ анализируемой случайной выборки $\{x_1, x_2, ..., x_N\}$. При этом следует проводить проверку с помощью критерия типа Колмогорова-Смирнова статистической гипотезы о том, что $H_0 : F^{\langle \text{emp} \rangle}(x) = F^{\langle \text{norm approx} \rangle}\left(x, h_{\min}^{\langle 5 \rangle}\right)$. При отклонении гипотезы H_0 в качестве оптимального значение параметра использовать левую границу интервала значений параметра размытости h – $h^{\langle 6 \rangle}$, на котором гипотеза H_0 оказывается верной.

Литература

1. Левин Б. Р. Вероятностные модели и методы в системах связи и управления/Б. Р. Левин, В. Шварц// – М.: Радио и связь, 1985. – 312 с.

2. Тихонов В. И. Статистическая радиотехника/В.И. Тихонов// –М.: Радио и связь, 1982. – 624 с.

3. Кропотов Ю. А Методы оценивания моделей плотности вероятностей акустических сигналов в телекоммуникациях аудиообмена/ Ю.А. Кропотов// Системы управления, связи и безопасности, 2017. № 1. С. 26–39.

4. Кропотов Ю.А. Моделирование и методы исследования акустических сигналов, шумов и помех в системах телекоммуникаций/ Ю.А. Кропотов, В.А. Ермолаев// – Берлин: Директ-Медиа, 2016. – 251 с.

5. Сушко Д.В. Оптимальная аппроксимация частотных вероятностей/Д.В. Сушко//Информационные процессы, 2018. Том 18. № 1. С. 40–54.

6. Новоселов А.А. Параметризация моделей управляемых систем // Вестник государственного аэрокосмического университета им. академика М.Ф. Решетнева, 2010. № 5 (31). С. 52–56.

7. Поршнев С.В., Копосов А.С. Случайные величины с ограниченной областью рассеяния: математическое и алгоритмическое обеспечение для оценивания плотностей вероятностей и функций распределений/С.В. Поршнев, А.С. Копосов// –М.: Горячая линия-Телеком, 2018. –184 с.

8. Киреева Н.В., Чупахина Л.Р Сравнение возможностей использования различных методов аппроксимации для анализа трафика с самоподобным распределением/Н.В. Киреев, Л.Н. Чупахина// Международный журнал прикладных и фундаментальных исследований, 2016. № 12. С. 1287–1289.

9. Поршнев С.В., Божалкин Д.А. Математическое и алгоритмическое обеспечение для анализа характеристик информационных потоков в магистральных интернет-каналах/С.В. Поршнев, Д.А. Божалкин// –М.: Горячая линия-Телеком, 2021. –214 с.

10. Сызранцев В.Н. Расчет прочностной надежности изделий на основе методов непараметрической статистики/В.Н. Сызранцев, Я.Н. Невелев, С.Л. Голофаст// –Новосибирск: Наука, 2008. –128 с.

11. Кобзарь А.И. Прикладная математическая статистика. Для инженеров и научных работников/ А.И. Кобзарь// –М.: ФИЗМАТЛИТ, 2006. –816 с.

12. Rosenblatt M. Remarks on Some Nonparametric Estimates of a Density Function // The Annals of Mathematical Statistics., 1956. – T. 27. Vol. 3. –P. 832–837.

13. Боровков А.А. Математическая статистика. – М.: Наука, 1984, 472 с.

14. Parzen E. On the estimation of probability density function and the mode // Ann. Math. Stat., 1962. – Vol. 33. – P. 1065–1076.

15. Суетин П.К. Классические ортогональные многочлены. – М.: Наука, 1976. – 328 с.

16. Димаки А.В., Светлаков А.А. Аппроксимация плотностей распределений случайных величин с применением ортогональных полиномов Чебышева–Эрмита/А.В. Димаки, А.А. Светлаков// Известия Томского политехнического университета, 2006. Т. № 8. С. 6–11.

17. Голик Ф. В. Аппроксимация эмпирических распределений вероятностей полиномами Бернштейна/ Ф.В. Голик// Журнал радиоэлектроники [электронный журнал]. 2018. № 7. Режим доступа: http://jre.cplire.ru/jre/jul18/5/text.pdf (дата обращения 16 октября 2021 г.) DOI 10.30898/1684-1719.2018.7.5.

18. Лемешко Б.Ю. О нахождении параметра размытости непараметрических оценок функции плотности/ Б.Ю. Лемешко, С.Н. Постовалов, А.В. Французов// Труды V международной конференции «Актуальные проблемы электронного приборостроения», АПЭП-2000. Новосибирск 26–29 сентября, 2000. В 7 тт. –Новосибирск: Новосибирский государственный технический университет, 2000. –Т. 6. –С. 17–20.

19. Botev. Z.I., Grotowski. J.F., Kroese. D.P. Kernel density estimation via diffusion, Annal Statistic, October 2010, Vol. 38 (5), pp. 2916–2957, DOI: https://doi.org/10.1214/10-AOS799.

20. Silverman B.W. Density Estimation for Statistics and Data Analysis. –Chapman and Hall/CRC, 1986. –175 p.

21. Leiva-Murillo J.M., Artés-Rodríguez A. Algorithms for maximum-likelihood bandwidth selection in kernel density estimators. Pattern Recognition Letters, Volume 33, Issue 13, 2012, pp. 1717–1724.

22. Xu Xiaoyuan, Yan Zheng, Xu Shaolun Estimating wind speed probability distribution by diffusion-based kernel density method, Electric Power Systems Research, Volume 121, April 2015, pp. 28–37.

23. Ouarda T.B.M.J., Charron C., Shin J.-Y., Marpu P.R., Al-Mandoos A.H., Al-Tamimi, Ghedira H., Al Hosary T.N, Probability distributions of wind speed in the UAE, Energy Conversion and Management, Volume 93, 2015, pp. 414-434.

24. Jones M.C., Marron J.S., Sheather S.J. A brief survey of bandwidth selection for density estimation, Journal of the American Statistical Association, Volume91, № 433, March 1996, pp.401–4071.

25. Симахин В.А. Робастные непараметрические оценки: адаптивные оценки взвешенного максимального правдоподобия в условиях статистической априорной неопределенности/ В.А. Симахин // –Saarbrucken, Germany: LAP LAMBERT Academic Publishing GmbH & Co. KG, 2011. –292 с.

References

1. Levin B. R. Veroyatnostnyye modeli i metody v sistemakh svyazi i upravleniya/B. R. Levin, V. Shvarts// – M.: Radio i svyaz', 1985. – 312 s.

2. Tikhonov V. I. Statisticheskaya radiotekhnika/V.I. Tikhonov// –M.: Radio i svyaz', 1982. – 624 s.

3. Kropotov YU. A Metody otsenivaniya modeley plotnosti veroyatnostey akusticheskikh signalov v telekommunikatsiyakh audioobmena/ YU.A. Kropotov// Sistemy upravleniya, svyazi i bezopasnosti, 2017. № 1. S. 26–39.

4. Kropotov YU.A. Modelirovaniye i metody issledovaniya akusticheskikh signalov, shumov i pomekh v sistemakh telekommunikatsiy/ YU.A. Kropotov, V.A. Yermolayev// – Berlin: Direkt-Media, 2016. – 251 c.

5. Sushko D.V. Optimal'naya approksimatsiya chastotnykh veroyatnostey/D.V. Sushko// Informatsionnyye protsessy, 2018. Tom 18. № 1. S. 40–54.

6. Novoselov A.A. Parametrizatsiya modeley upravlyayemykh sistem // Vestnik gosudarstvennogo aerokosmicheskogo universiteta im. akademika M.F. Reshetneva, 2010. № 5 (31). S. 52–56.

7. Porshnev S.V., Koposov A.S. Sluchaynyye velichiny s ogranichennoy oblast'yu rasseyaniya: matematicheskoye i algoritmicheskoye obespecheniye dlya otsenivaniya plotnostey veroyatnostey i funktsiy raspredeleniy/S.V. Porshnev, A.S. Koposov// –M.: Goryachaya liniya-Telekom, 2018. –184 s.

8. Kireyeva N.V., Chupakhina L.R Sravneniye vozmozhnostey ispol'zovaniya razlichnykh metodov approksimatsii dlya analiza trafika s samopodobnym raspredeleniyem/N.V. Kireyev, L.N. Chupakhina// Mezhdunarodnyy zhurnal prikladnykh i fundamental'nykh issledovaniy, 2016. № 12. S. 1287–1289.

9. Porshnev S.V., Bozhalkin D.A. Matematicheskoye i algoritmicheskoye obespecheniye dlya analiza kharakteristik informatsionnykh potokov v magistral'nykh internet-kanalakh/S.V. Porshnev, D.A. Bozhalkin// –M.: Goryachaya liniya-Telekom, 2021. –214 s.

10. Syzrantsev V.N. Raschet prochnostnoy nadezhnosti izdeliy na osnove metodov neparametricheskoy statistiki/V.N. Syzrantsev, YA.N. Nevelev, S.L. Golofast// –Novosibirsk: Nauka, 2008. –128 s.

11. Kobzar' A.I. Prikladnaya matematicheskaya statistika. Dlya inzhenerov i nauchnykh rabotnikov/A.I. Kobzar'// –M.: FIZMATLIT, 2006. –816 s.

12. Rosenblatt M. Remarks on Some Nonparametric Estimates of a Density Function // The Annals of Mathematical Statistics., 1956. – T. 27. Vol. 3. –P. 832–837.

13. Borovkov A.A. Matematicheskaya statistika. – M.: Nauka, 1984, 472 s.

14. Parzen E. On the estimation of probability density function and the mode // Ann. Math. Stat., 1962. – Vol. 33. – P. 1065–1076.

15. Suyetin P.K. Klassicheskiye ortogonal'nyye mnogochleny. – M.: Nauka, 1976. – 328 s.

16. Dimaki A.V., Svetlakov A.A. Approksimatsiya plotnostey raspredeleniy sluchaynykh velichin s primeneniyem ortogonal'nykh polinomov Chebysheva–Ermita/A.V. Dimaki, A.A. Svetlakov// Izvestiya Tomskogo politekhnicheskogo universiteta, 2006. T. № 8. S. 6–11.

17. Golik F. V. Approksimatsiya empiricheskikh raspredeleniy veroyatnostey polinomami Bernshteyna/ F.V. Golik// Zhurnal radioelektroniki [elektronnyy zhurnal]. 2018. № 7. Rezhim dostupa: http://jre.cplire.ru/jre/jul18/5/text.pdf (data obrashcheniya 16 oktyabrya 2021 g.) DOI 10.30898/1684-1719.2018.7.5.

18. Lemeshko B.YU. O nakhozhdenii parametra razmytosti neparametricheskikh otsenok funktsii plotnosti/ B.YU. Lemeshko, S.N. Postovalov, A.V. Frantsuzov// Trudy V mezhdunarodnoy konferentsii «Aktual'nyye problemy elektronnogo priborostroyeniya», APEP-2000. Novosibirsk 26–29 sentyabrya, 2000. V 7 tt. –Novosibirsk: Novosibirskiy gosudarstvennyy tekhnicheskiy universitet, 2000. –T. 6. –S. 17–20.

19. Botev. Z.I., Grotowski. J.F., Kroese. D.P. Kernel density estimation via diffusion, Annal Statistic, October 2010, Vol. 38 (5), pp. 2916–2957, DOI: https://doi.org/10.1214/10-AOS799.

20. Silverman B.W. Density Estimation for Statistics and Data Analysis. –Chapman and Hall/CRC, 1986. –175 p.

21. Leiva-Murillo J.M., Artés-Rodríguez A. Algorithms for maximum-likelihood bandwidth selection in kernel density estimators. Pattern Recognition Letters, Volume 33, Issue 13, 2012, pp. 1717–1724.

22. Xu Xiaoyuan, Yan Zheng, Xu Shaolun Estimating wind speed probability distribution by diffusion-based kernel density method, Electric Power Systems Research, Volume 121, April 2015, pp. 28–37.

23. Ouarda T.B.M.J., Charron C., Shin J.-Y., Marpu P.R., Al-Mandoos A.H., Al-Tamimi, Ghedira H., Al Hosary T.N, Probability distributions of wind speed in the UAE, Energy Conversion and Management, Volume 93, 2015, pp. 414-434.

24. Jones M.C., Marron J.S., Sheather S.J. A brief survey of bandwidth selection for density estimation, Journal of the American Statistical Association, Volume91, № 433, March 1996, pp.401–4071.

25. Simakhin V.A. Robastnyye neparametricheskiye otsenki: adaptivnyye otsenki vzveshennogo maksimal'nogo pravdopodobiya v usloviyakh statisticheskoy apriornoy neopredelennosti/ V.A. Simakhin // –Saarbrucken, Germany: LAP LAMBERT Academic Publishing GmbH & Co. KG, 2011. –292 s.

ПОРШНЕВ Сергей Владимирович, доктор технических наук, профессор, директор Учебно-научного центра «Информационная безопасность» федерального государственного автономного образовательного учреждения высшего образования «Уральский федеральный университет им. первого Президента России Б.Н. Ельцина». 620002, г. Екатеринбург, ул. Мира, 32. E-mail: s.v.porshnev@urfu.ru

РЯБКО Николай Юрьевич, аспирант федерального государственного автономного образовательного учреждения высшего образования «Уральский федеральный университет им. первого Президента России Б.Н. Ельцина». 620002, г. Екатеринбург, ул. Мира, 32. E-mail: N.Yu.Ryabko@urfu.ru

PORSHNEV Sergey Vladimirovich, Doctor of Technical Sciences, Professor, Director of the Educational and Scientific Center «Information Security» of the Federal State Autonomous Educational Institution of Higher Education «Ural Federal University named after the first President of Russia B.N. Yeltsin». 620002, Yekaterinburg, st. Mira, 32. E-mail: s.v.porshnev@urfu.ru

RYABKO Nikolay Yurievich, post-graduate student of the Federal State Autonomous Educational Institution of Higher Education "Ural Federal University named after the first President of Russia B.N. Yeltsin". 620002, Yekaterinburg, st. Mira, 32. E-mail: N.Yu.Ryabko@urfu.ru

УДК 004.896 Вестник УрФО № 4(42) / 2021, с. 39–45

Асяев Г.Д., Соколов А.Н.

DOI: 10.14529/secur210404

МОДЕЛИ ПРЕДИКТИВНОЙ ЗАЩИТЫ ИНФОРМАЦИИ АВТОМАТИЗИРОВАННОЙ СИСТЕМЫ УПРАВЛЕНИЯ ВОДОСНАБЖЕНИЕМ НА ОСНОВЕ ВРЕМЕННЫХ РЯДОВ С ИСПОЛЬЗОВАНИЕМ ТЕХНОЛОГИЙ МАШИННОГО ОБУЧЕНИЯ[1]

Применительно к задаче прогнозирования кибератак рассмотрены модели, основанные на методах предиктивного обслуживания, а также сформированы гипотезы о границах применимости метода предиктивной защиты информации. Для проверки гипотез проанализирован набор данных «Water_pump» (сайт Kaggle), состоящий из количественных и качественных характеристик автоматизированной системы управлением водоснабжением, записанных в течение шести месяцев. Проведена предобработка, очистка и группировка экспериментальных данных для обучения нейронной сети, а также проведен анализ основных свойств данных и поиск в них общих закономерностей, распределений и аномалий. Метод предиктивной защиты информации реализован с применением технологий машинного обучения. Для каждой модели выполнена настройка гиперпараметров и проведена оценка по метрикам качества «precision», «recall» и «accuracy». Полученные результаты позволяют сделать вывод о применимости реализованного метода предиктивной защиты информации на практике для анализа данных автоматизированных систем управления технологическими процессами.

Ключевые слова: *автоматизированная система управления технологическим процессом (АСУ ТП), временной ряд, задача прогнозирования, информационная безопасность, кибератака, машинное обучение, предиктивная защита информации.*

[1] Исследование выполнено при финансовой поддержке Минобрнауки России (грант ИБ МТУСИ) в рамках научного проекта № 40469-29/2021-К.

PREDICTIVE INFORMATION PROTECTION MODELS OF AUTOMATED WATER MANAGEMENT SYSTEM BASED ON TIME SERIES USING MACHINE LEARNING TECHNOLOGIES

Models based on predictive maintenance methods are considered in relation to the problem of predicting cyberattacks, and also formed hypotheses about the limits of applicability of the predictive information protection method. To test the hypotheses, the "Water_pump" dataset (Kaggle website) was analyzed, consisting of quantitative and qualitative characteristics of an automated water supply management system recorded over six months. Preprocessing, cleaning and grouping of experimental data for training a neural network, and also analyzed the main properties of the data and searched for general patterns, distributions and anomalies in them. The preprocessing, cleaning and grouping of experimental data for training the neural network was carried out, as well as the analysis of the main properties of the data and the search for general patterns, distributions and anomalies in them. The method of predictive information protection is implemented using machine learning technologies. For each model, the hyperparameters were adjusted and the quality metrics "precision", "recall" and "accuracy" were assessed. The results obtained allow us to draw a conclusion about the applicability of the implemented predictive information protection method in practice for analyzing data from Industrial Control Systems.

Keywords: *Industrial Control Systems (IDS), time series, forecasting problem, Information Security, cyberattack, machine learning, predictive information protection.*

Проблематике обнаружения и предотвращения кибератак на объекты автоматизированных систем управления технологическими процессами (АСУ ТП) посвящено большое количество публикаций. Однако описанные подходы, как правило, позволяют выявить аномальное поведение системы [1], когда злоумышленник уже проник в систему, либо совершает попытки несанкционированного доступа к ней. Однако более интересной и полезной с точки зрения практического применения задачей является задача обнаруживать кибератаку до того, как она началась, и прогнозировать время, через которое система даст сбой при её реализации. Подобный подход имеет много общего с новыми технологиями предиктивного обслуживания, основной целью которых является обеспечение надежности критичных для деятельности предприятия производственных и технологических процессов [2]. В отличии от традиционных подходов, направленных на поддержание каждой единицы оборудования в безупречном состоянии, технологии предиктивного обслуживания не требуют неоправданно высоких затрат. Под предиктивной защитой информации при этом понимается деятельность, позволяющая по косвенным признакам (параметрам) системы определить возможность наступления кибератаки, спрогнозировать время, через которое она наступит, а также выбрать адекватные превентив-

ные меры защиты. В рамках проведенного исследования разработана модель предиктивной защиты объекта автоматизированной системы управления водоснабжением, направленной на предотвращение нарушения одного из свойств информационной безопасности объекта – доступности информации. Под доступностью при этом понимают свойство информации быть готовой к использованию по запросу авторизованного субъекта, имеющего на это право [6].

Для исследования поставленной задачи был взят датасет «Water_pump» с сайта Kaggle. Данные представляют собой количественные значения, принимаемые с 51 датчика (сенсора). Размер датасета 220320x55. Столбец «timestamp» представляет собой временной интервал, показывающий, что измерения проводились каждую минуту (рис. 1). Анализ показал, что датасет имеет достаточно большое количество признаков с пропусками данных (более 50%). Вследствие появления зашумленных значений при обучении нейронной сети исключены столбцы, у которых процент пропущенных значений составлял более 60. Чтобы отразить максимально реалистичную ситуацию по работе с датчиками, выпадающие значения восполнены медианными значениями последних 20 минут. Кроме того, вследствие разной размерности датчиков/сенсоров (например, sensor_00 изменяет свои значения от 1 до 5, в то время как sensor_01 – от 47 до 48) данные были нормированы с использованием функции StadartScaler.

timestamp	sensor_00	sensor_01	sensor_02	sensor_03	sensor_04	sensor_05	sensor_06	sensor_07	...	sensor_43	sensor_44	sensor_45	s
2018-04-01 00:00:00	2.465394	47.09201	53.2118	46.310760	634.3750	76.45975	13.41146	16.13136	...	41.92708	39.641200	65.68287	
2018-04-01 00:01:00	2.465394	47.09201	53.2118	46.310760	634.3750	76.45975	13.41146	16.13136	...	41.92708	39.641200	65.68287	
2018-04-01 00:02:00	2.444734	47.35243	53.2118	46.397570	638.8889	73.54598	13.32465	16.03733	...	41.66666	39.351852	65.39352	
2018-04-01 00:03:00	2.460474	47.09201	53.1684	46.397568	628.1250	76.98898	13.31742	16.24711	...	40.88541	39.062500	64.81481	
2018-04-01 00:04:00	2.445718	47.13541	53.2118	46.397568	636.4583	76.58897	13.35359	16.21094	...	41.40625	38.773150	65.10416	
2018-04-01 00:05:00	2.453588	47.09201	53.1684	46.397568	637.6157	78.18568	13.41146	16.16753	...	42.70833	38.773150	63.65741	
2018-04-01 00:06:00	2.455556	47.04861	53.1684	46.397568	633.3333	75.81614	13.43316	16.13136	...	43.22916	38.194440	61.92130	
2018-04-01 00:07:00	2.449653	47.13541	53.1684	46.397568	630.6713	75.77331	13.25231	16.12413	...	42.96875	38.194443	59.60648	
2018-04-01	2.463426	47.09201	53.1684	46.397568	631.9444	74.58916	13.28848	16.13136	...	42.18750	38.194440	57.87037	

Рис. 1. Представление данных в датафрейме

Наличие в датасете информации с большого количества сенсоров приводит к существенному проценту зашумленности данных, что может сказаться на качестве распознавания будущей модели. Было решено проанализировать временное распределение каждого датчика и выявить взаимосвязь в их работе. На рис. 2 видно, что sensor_01, sensor_02, sensor_03 имеют похожее распределение. Аналогичный вывод можно сделать и для датчиков sensor _4 – sensor _12.

Для уменьшения размерности данных датчики были сгруппированы по степени схожести их временного распределения по 7 блокам (табл. 1).

Категориальным признаком в датасете является состояние датчика в текущий момент времени:

• «NORMAL» (нормальная работа система): 205836 значений;

• «RECOVERING» (восстановление работы датчика): 14477 значений;

• «BROKEN» (датчик не работает): 7 значений.

Как показано на гистограмме распределения категориальной переменной (рис. 3) выборка достаточно несбалансированная, и преобладают записи с нормальным поведением системы. Однако приоритетом формируемой модели является умение предсказывать имен-

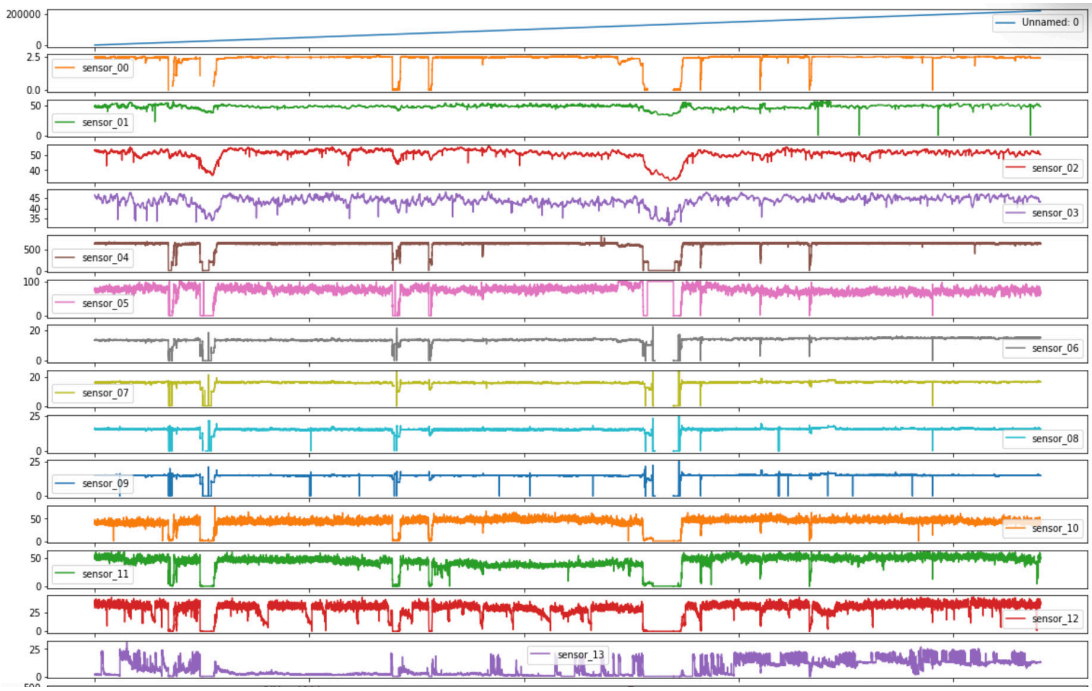

Рис. 2. Временное распределение различных сенсоров

Таблица 1

Группировка сенсоров по блокам

№ блока	Номер сенсора
I блок	sensor_01, sensor_02, sensor_03
II блок	sensor_04, sensor_05, sensor_06, sensor_07, sensor_08, sensor_09
III блок	sensor_10, sensor_11, sensor_12
IV блок	sensor_14, sensor_16, sensor_17, sensor_18
V блок	sensor_19, sensor_20, sensor_21, sensor_22, sensor_23, sensor_24
VI блок	sensor_25, sensor_26, sensor_28, sensor_29, sensor_30, sensor_31, sensor_32, sensor_33
VII блок	sensor_34, sensor_35

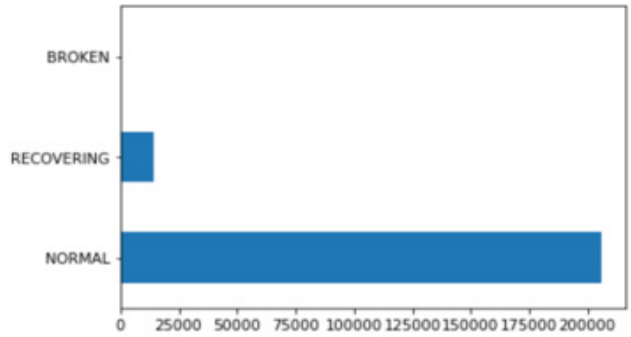

Рис. 3. Гистограмма распределения категориального признака

но выход датчика из строя (нарушение доступности информации). Данные были разбиты на обучающую и тестовую выборку. В обучающей выборке помимо нормальных состояний системы, состояний восстановления были использованы 2 записи состояний поломки датчика.

Так как модели машинного обучения не умеют работать с категориальными признаками (за исключением модели Catboost от компании Яндекс), то было решено заменить состояния системы на числа: «1» - нормальная работа датчика, «0.5» - восстановление работы датчика, «0» - датчик не работает (рис. 4).

Стоит отметить, что такие модели машин-

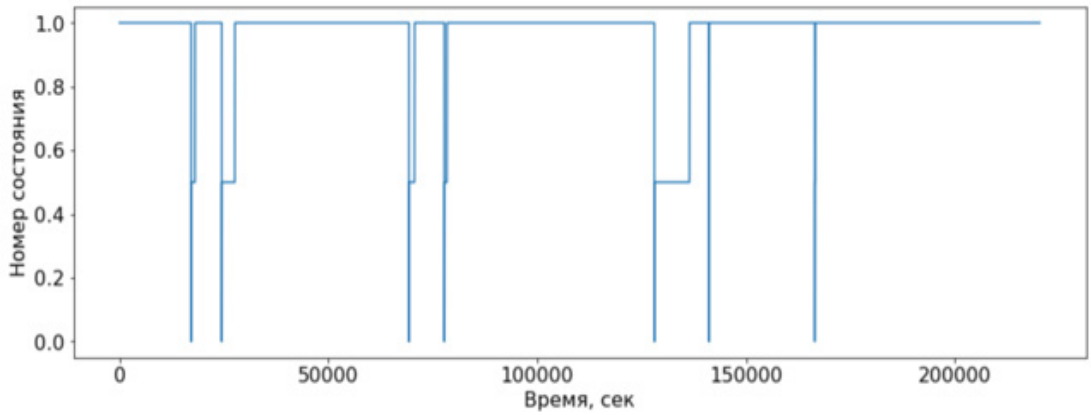

Рис. 4. Временное распределение сенсора с заменой категориальной переменной

ного обучения для работы с временными данными, как модель скользящего среднего (ARIMA) и разновидность данной модели, учитывающая сезонность (SARIMA), не смогли предсказать, через какое время датчик выйдет из строя. Поэтому была построена рекуррентная сеть с долгой краткосрочной памятью (LSTM) (рис. 5) [7]. Для борьбы с переобучением использовалась Lasso регуляризация – Dropout(0.3). Количество эпох для обучения составило 50. В качестве алгоритма оптимизации [9] был выбран Adaptive moment estimation (Adam). А в качестве функции ошибки – среднеквадратичная ошибка (RMSE) [10]:

$$\mathrm{RMSE} = \sqrt{\frac{1}{n}\sum(y_i - \gamma_i)^2}\ , \qquad (1)$$

где n – количество записей, y_i – предсказанный ответ моделью, γ_i – истинный ответ.

На тестовом датасете среднеквадратичная ошибка составила 0.016. На рис. 6 показан процесс обучения модели на тренировочном наборе данных, где после 30-й эпохи модель вышла на плато. Так как на тестовом наборе данных с увеличением эпохи ошибка не увеличивается, сделан вывод о том, что модель не была подвержена переобучению.

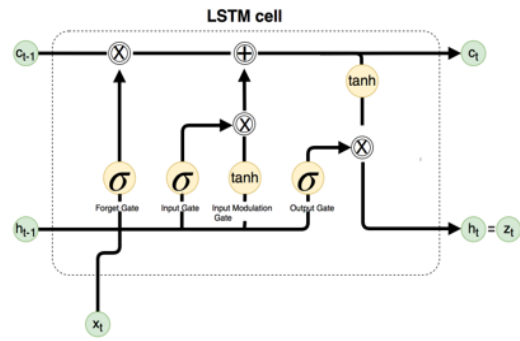

Рис. 5. Архитектура LSTM

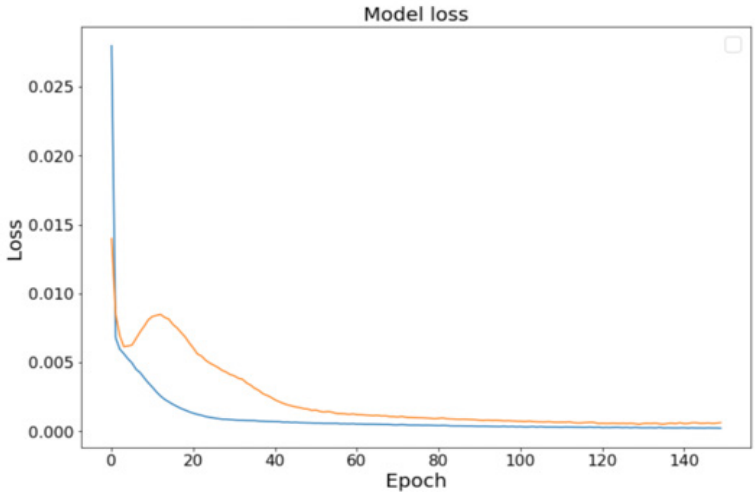

Рис. 6. График обучения модели на тренировочном и тестовом датасетах

На рис. 7 представлены графики значений, предсказанных разработанной моделью, и истинных значений. При сравнении графиков видно, что модель предсказала все

5 отказов в системе и спрогнозировала время ремонта (замены) датчиков, близкое к реальным значениям. Кроме этого, на тестовых данных отсутствуют ошибки второго рода, что является преимуществом данной модели. Стоит также отметить, что разработанная модель позволяет спрогнозировать состояние системы на достаточно длительный период времени.

Полученные результаты позволяют сделать вывод о применимости реализованного метода предиктивной защиты информации на практике для анализа данных АСУ ТП. Разработанные модели могут использоваться для предотвращения нарушений других свойств безопасности.

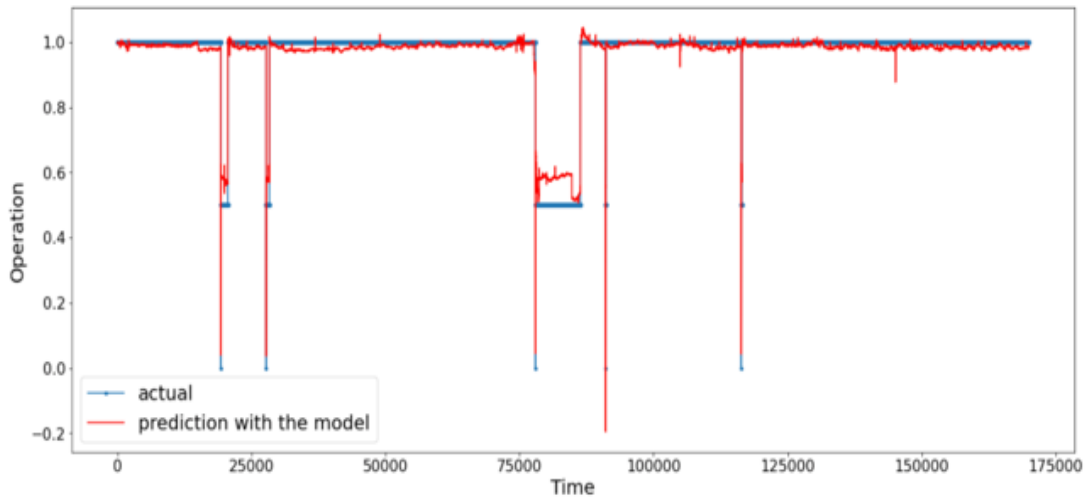

Рис. 7. Графики значений, предсказанных моделью (prediction with the model), и истинных значений (actual)

Литература

1. Гарбук С.В., Правиков Д.И., Полянский А.В., Самарин И.В. Обеспечение информационной безопасности АСУ ТП с использованием метода предиктивной защиты // Вопросы кибербезопасности, 2019. No3(31). С. 30-36.

2. Боровков А.И. «Умные» цифровые двойники – основа новой парадигмы цифрового проектирования и моделирования глобально конкурентоспособной продукции нового поколения. Трамплин к успеху// Журнал АО «ОДК». 2018. No 13. С. 12-18.

3. Правиков Д.И. Об одном подходе к обеспечению информационной безопасности автоматизированных систем // Вопросы защиты информации. 2007. No 3. С. 17-19.

4. Гарбук С.В., Бурцев А.Г. Методические основы исследования уязвимостей компонентов АСУ ТП // Защита информации. Inside. 2012. No 3. С. 34-38.

5. Гарбук С.В. Перспективы применения интеллектуальных технологий для решения задач безопасности // Национальная безопасность / 2016. No 4. С. 451-457.

6. Pump sensor data (2021). Доступ к ресурсу по ссылке https://www.kaggle.com/nphantawee/pump-sensor-data (23 мая 2021).

7. Асяев Г.Д. Обнаружение вторжений на основе анализа аномального поведения локальной сети с использованием алгоритмов машинного обучения с учителем / Г.Д. Асяев, А.Н. Соколов //Вестник УрФО. Безопасность в информационной сфере. – 2020 № 1(35). – С.77-83

8. Luzhnov V.S Simulation of Protected Industrial Control Systems Based on Reference Security Model using Weighted Oriented Graphs / V.S. Luzhnov, A.N. Sokolov, A.E. Barinov //Proceedings - 2019 International Russian Automation Conference, RusAutoCon 2019. – 2019

9. Sokolov A.N. Applying Methods of Machine Learning in the Task of Intrusion Detection Based on the Analysis of Industrial Process State and ICS Networking / A.N. Sokolov, I.A. Pyatnitsky, S.K. Alabugin //FME Transactions. –2019. –Vol. 47 No. 4. – P.782-789

10. Соколов А.Н. Применение методов одноклассовой классификации для обнаружения вторжений / А.Н. Соколов, С.К. Алабугин, И.А. Пятницкий //Вестник УрФО. Безопасность в информационной сфере. – 2018. –Том - № 2(28). – С.43-48

References

1. Garbuk S.V., Pravikov D.I., Polyanskiy A.V., Samarin I.V. Obespecheniye informatsionnoy bezopasnosti ASU TP s ispol'zovaniyem metoda prediktivnoy zashchity // Voprosy kiberbezopasnosti, 2019. No3(31). S. 30-36.

2. Borovkov A.I. «Umnyye» tsifrovyye dvoyniki – osnova novoy paradigmy tsifrovogo proyektirovaniya i modelirovaniya global'no konkurentosposobnoy produktsii novogo pokoleniya. Tramplin k uspekhu// Zhurnal AO «ODK». 2018. No 13. S. 12-18.

3. Pravikov D.I. Ob odnom podkhode k obespecheniyu informatsionnoy bezopasnosti avtomatizirovannykh sistem // Voprosy zashchity informatsii. 2007. No 3. S. 17-19.

4. Garbuk S.V., Burtsev A.G. Metodicheskiye osnovy issledovaniya uyazvimostey komponentov ASU TP // Zashchita informatsii. Inside. 2012. No 3. S. 34-38.

5. Garbuk S.V. Perspektivy primeneniya intellektual'nykh tekhnologiy dlya resheniya zadach bezopasnosti // Natsional'naya bezopasnost' / 2016. No 4. S. 451-457.

6. Pump sensor data (2021). Accessed at https://www.kaggle.com/nphantawee/pump-sensor-data (May 23, 2021).

7. Asyayev G.D. Obnaruzheniye vtorzheniy na osnove analiza anomal'nogo povedeniya lokal'noy seti s ispol'zovaniyem algoritmov mashinnogo obucheniya s uchitelem / G.D. Asyayev, A.N. Sokolov //Vestnik UrFO. Bezopasnost' v informatsionnoy sfere. – 2020 № 1(35). – C.77-83

8. Luzhnov V.S. Simulation of Protected Industrial Control Systems Based on Reference Security Model using Weighted Oriented Graphs / V.S. Luzhnov, A.N. Sokolov, A.E. Barinov //Proceedings - 2019 International Russian Automation Conference, RusAutoCon 2019.-2019

9. Sokolov A.N. Applying Methods of Machine Learning in the Task of Intrusion Detection Based on the Analysis of Industrial Process State and ICS Networking / A.N. Sokolov, I.A. Pyatnitsky, S.K. Alabugin //FME Transactions. -2019.-Vol. 47 No. 4.- P.782-789

10. Sokolov A.N. Primeneniye metodov odnoklassovoy klassifikatsii dlya obnaruzheniya vtorzheniy / A.N. Sokolov, S.K. Alabugin, I.A. Pyatnitskiy //Vestnik UrFO. Bezopasnost' v informatsionnoy sfere. – 2018. – Tom - № 2(28). – C.43-48

АСЯЕВ Григорий Дмитриевич, аспирант кафедры защиты информации высшей школы электроники и компьютерных наук ФГАОУ ВО «Южно-Уральский государственный университет (национальный исследовательский университет)». Россия, 454080, г. Челябинск, проспект Ленина, д. 76. E-mail: asiaevgd@susu.ru.

СОКОЛОВ Александр Николаевич, кандидат технических наук, доцент, заведующий кафедрой защиты информации высшей школы электроники и компьютерных наук ФГАОУ ВО «Южно-Уральский государственный университет (национальный исследовательский университет)». Россия, 454080, г. Челябинск, проспект Ленина, д. 76. E-mail: sokolovan@susu.ru.

ASYAEV Grigorii Dmitrievich, postgraduate student of the department of information security of the school of electrical engineering and computer science in FSAEI HE «South Ural State University (national research university)». 76, Lenin prospect, Chelyabinsk, Russia, 454080. E-mail: asiaevgd@susu.ru.

SOKOLOV Alexander Nikolaevich, Ph.D., Associate professor, Head of the department of information security of the school of electrical engineering and computer science in FSAEI HE «South Ural State University (national research university)». 76, Lenin prospect, Chelyabinsk, Russia, 454080. E-mail: sokolovan@susu.ru.

УДК 004.588

Вестник УрФО № 4(42) / 2021, с. 46–58

Астахова Л.В., Киряев А.И.

DOI: 10.14529/secur210405

ИНТЕГРАЦИЯ АВТОМАТИЗИРОВАННОГО УПРАВЛЕНИЯ ДОКУМЕНТАМИ И ОСВЕДОМЛЕННОСТЬЮ СОТРУДНИКОВ ОБ ИНФОРМАЦИОННОЙ БЕЗОПАСНОСТИ МАЛОГО И СРЕДНЕГО ПРЕДПРИЯТИЯ

В статье определены факторы, влияющие на организацию работы с документированной информацией и управление осведомленностью сотрудников в области информационной безопасности (ИБ) в процессе ее управления на малых и средних предприятиях. К ним относятся требования международных и национальных стандартов управления информационной безопасностью (УИБ) и других нормативных документов; документальный формат проверок состояния организаций ИБ регулирующими органами; недостаточная осведомленность персонала о документально оформленных правилах обеспечения информационной безопасности; недостаточное внимание предприятий к документированию процессов обеспечения ИБ организаций, а также осведомленности сотрудников об ИБ. Особенно низка эффективность работы с документацией по ИБ на малых и средних предприятиях. В процессе анализа мирового потока научной литературы была выявлена необходимость усиления взаимосвязей процессов управления информационной безопасностью, их документирования и автоматизации. Цель исследования - разработать средство оптимизации системы менеджмента информационной безопасностью (СМИБ) организации на основе многофункциональной автоматизированной деятельности с документированной информацией об ИБ путем ее интеграции с управлением осведомленностью персонала об этой документации. Материалы и методы. Для решения проблемы использованы интеграционный подход к процессам построения СМИБ, аналитико-синтетические методы, метод моделирования процессов и программирования. Результаты. В контексте выявленных тенденций и накопленного практического опыта в статье обосновывается необходимость и результат разработки многофункционального программного приложения для организации работы с документацией по управлению

информационной безопасностью малых и средних предприятий. Приложение может выполнять функции создания и обновления; распространения, доступа, поиска и использования; хранения и консервации; управления изменениями; контроля осведомленности сотрудников о документированной системе информационной безопасности организации. Научная новизна заключается в обосновании и реализации интегративного подхода к управлению документацией по ИБ и управлению осведомленностью об ИБ сотрудников организации. Практическая значимость работы заключается в возможности использования разработанного веб-приложения на малых и средних предприятиях для оптимизации процессов защиты информации.

Ключевые слова: *управление информационной безопасностью, документация, осведомленность, интеграция, автоматизация, программное приложение, малые и средние предприятия.*

Astakhova L.V., Kiryaev A.I.

INTEGRATION OF AUTOMATED MANAGEMENT OF DOCUMENTS AND AWARENESS OF EMPLOYEES ABOUT INFORMATION SECURITY OF A SMALL AND MEDIUM ENTERPRISE

The article identifies the factors influencing the organization of work with documented information and the management of the awareness of employees in the field of information security (IS) in the process of its management in small and medium enterprises. These include the requirements of international and national standards for information security management (ISM) and other regulatory documents; the documentary format of inspections of the state of information security organizations by regulatory authorities; insufficient awareness of personnel about the documented rules for ensuring information security; insufficient attention of enterprises to documenting the processes of ensuring information security of organizations, as well as awareness of employees about information security. The efficiency of work with information security documentation is especially low in small and medium-sized enterprises. In the process of analyzing the global flow of scientific literature, the need to strengthen the interconnections of information security management processes, their documentation, and automation was identified. The purpose of the study is to develop a tool for optimizing an organization's information security management system (ISMS) based on multifunctional automated activities with documented information about IS by integrating it with personnel awareness management about this documentation. Materials and methods. To solve the problem, and integrated an approach to the processes of building an ISMS, analytical and synthetic methods, a method of modeling processes and programming were used. Results. In the context of the identified trends and the accumulated practical experience, the article substantiates the need and the result of developing a multifunctional software application for organizing work with documentation on information security management of small and medium-sized enterprises. The

application can perform creation and update functions; distribution, access, search and use; storage and conservation; change management; control of the awareness of employees about the documented information security system of the organization. Scientific novelty lies in the substantiation and implementation of an integrative approach to the management of information security documentation and management of information security awareness of the organization's employees. The practical significance of the work lies in the possibility of using the developed web application in small and medium-sized enterprises to optimize information security processes.

Keywords: *information security management, documentation, awareness, integration, automation, small and medium enterprises.*

1. Введение

Анализ утечек конфиденциальной информации за 9 месяцев 2020 года показал, что в мире было зарегистрировано на 7,4% меньше утечек, чем за аналогичный период прошлого года. Однако в России за тот же период количество утечек увеличилось на 5,6%. Если в мире 52,6% утечек вызваны внешними воздействиями, то в России - в пределах 21%, поскольку более 79% утечек произошли в результате внутренних нарушений. Также в России доля утечек по вине сотрудников вдвое выше, чем в мире - более 72%. [1]. Эти статистические данные свидетельствуют о том, что решение внутри организационных проблем управления информационной безопасностью требует особого внимания как в теоретическом, так и в практическом плане. Необходимы новые подходы к управлению информационной безопасностью организации. Важнейшими направлениями управленческой деятельности в области информационной безопасности являются управление документацией по ИБ и управление осведомленностью персонала об этой документации. Однако эти процессы чаще всего реализуются локально, изолированно друг от друга. Это обусловливает актуальность и цельнастоящегоисследования - разработать средство оптимизации системы менеджмента информационной безопасностью (СМИБ) организации на основе многофункциональной автоматизированной деятельности с документированной информацией об ИБ путем ее интеграции с управлением осведомленностью персонала об этой документации.

2. Вопросы управления документами и осведомленностью об информационной безопасности в теории и практике СМИБ

Управление информационной безопасностью активно изучается в мировой теории и практике, чему способствует стандартизация этой сферы деятельности. Первоначально стандартизация была сосредоточена на технологических аспектах - контроле доступа, сетевой безопасности и т. д. В начале двадцатого века начался процесс выпуска международных оценочных стандартов, определяющих критерии оценки соответствия компьютерных систем требованиям безопасности. Далее стали создаваться документы, в которых внимание сосредоточено исключительно на проблемах СМИБ: построение СМИБ организации, управление рисками, инциденты, аудит ИБ и др. Были разработаны серия международных стандартов ISO / IEC 27000, ISO / IEC TT 33052: 2016, а также библиотека мер контроля из серии стандартов NIST 800 и немецкий стандарт BSI. Наконец, последней тенденцией является разработка стандартов корпоративного управления (управления) информационной безопасностью: O-ISM3 (OpenInformationSecurityMaturity Model); COBIT 19, ISO / IEC 27014-20 и т. д.

Каждый из этих стандартов уделяет внимание управлению документацией по информационной безопасности. Так, согласно стандарта ISO / IEC 27001-2013 [2], эксперты включают в состав документации по ИБ обязательные и необязательные документы и записи. В число обязательных документов они включают: Область действия СМИБ (п. 4.3); Политика и цели информационной безопасности (пункты 5.2 и 6.2); Методология оценки рисков и обработки рисков (п. 6.1.2) и др. [3]. Специалисты также предлагают другие классификации документов, выделяя обязательные (28 документов), рекомендуемые (32 документа) и полезные (15 документов) документы [4].

Большое внимание уделяется контролю документированной информации в практике. Документация должна контролироваться с точки зрения: 1) создания и обновления (идентификации и описания (название, дата,

автор, версия); формата (язык, версия программного обеспечения, графика) и носителя (бумажный, электронный); проверки и утверждения для пригодности и адекватности); 2) распространения, доступа, поиска и использования (информация о распространении (имя получателя, количество копий, местонахождение); контроля доступа, извлечения и использования (уровень доступа, что должно быть доступно, время и продолжительность доступа, используемый метод, срок действия, количество поисков); регулярной проверки предоставленных прав доступа (в том числе новый запрос доступа / поиска, прекращение или отзыв доступа); 3) хранения и консервации (носитель (в сети Интернет, в печатном или электронном виде); прав человека на доступ к хранению документированной информации); 4) контроля изменений (детали модификации (имя человека, дата изменения, журналы истории, номера версий),причины изменений); 5) хранения и порядка утилизации (срок хранения, причины хранения, подробности утилизации (ответственное лицо, дата утилизации, что и почему должно быть удалено) [5, 6]. ISO / IEC TT 33052: 2016 также выделяет группу процессов управления документами и их отношения с группами процессов PRM.

COBIT 2019 усиливает акцент на создании системы управления информационной безопасностью, в задачи которой входят и задачи, связанные с людьми, их компетенциями, навыками икультурой поведения [5]. Представленные задачиможно использовать для подготовки должностных инструкций, программ повышения осведомленности и др. [7].

Особое внимание уделяется документу «Политика информационной безопасности», которая является одним из наиболее важныхформальных документальных средств контроля [8, 9]. Так, была разработана концептуальная основа для создания политики информационной безопасности [10]. В последние годы интерес исследователей прикован к компьютеризированным инструментам, которые поддерживают работу менеджеров по информационной безопасности [9, 11 и др.]. В качестве примера приведем обзор существующих исследований в области управления политикой ИБ, целью которого было изучение соотношения ручной и компьютеризированной поддержки СМИБ и средств их реализации. Авторы пришли к выводу, что существующие исследования сосредоточены в основном на поддержке ручного управления, поэтому разработали программное обеспечение для компьютеризированной поддержки отдельных процессов управления рисками, разработки, соблюдения и мониторинга политики ИБ [12].

В цифровой экономике открываются новые возможности для автоматизации работы с документацией, что немаловажно для сектора информационной безопасности. В условиях цифрового производства должны быть переведены в цифровой формат не только документация, но и методы ее формирования и методологии управления. Автоматизированные системы управления информационной безопасностью решают множество задач, включая обеспечение управления документами. Это позволяет не только улучшить эффективность системы защиты информации, но и выполнить требования большого количества международных и российских стандартов в области информационной безопасности.

Зарубежные специалисты выявили требования к компьютеризированным инструментам и создали программу «TheInformation SecurityGovernanceToolbox (ISGT)». После установления требований безопасности организации, выбора соответствующих мер безопасности и поддерживающих процедур безопасности программа динамически составляет документацию по безопасности, необходимую для обеспечения соблюдения этих мер информационной безопасности. К числу создаваемых документов относятся: корпоративная политика информационной безопасности, краткая и полная версия; политики вторичного уровня в виде различных поддерживающих стандартов компании, которые отражают выявленные меры безопасности; соответствующие процедуры безопасности, связанные с политикой, заявление о применимости. Этот набор документации предлагается пользователю в виде документов Word, которые можно изменять и корректировать в соответствии с конкретными потребностями организации. Программа также хранит копии этих документов, к которым можно получить доступ в любое время, пока не будут разработаны их новые версии [11].

Рассматриваются возможности автоматизации различных видов работ с документацией с использованием программных роботов. Существует понятие «роботизированная автоматизация процессов», и перспективы, связан-

ные с созданием «роботизированных» документов, весьма широки. В настоящее время уже есть опыт роботизации работы с документами, например, с договорами в государственной корпорации «Росатом» на основе решения компании ABBYY [13]. Это позволяет видеть перспективы работы с документацией по ИБ.

В России разработаны и используются в практике продукты для автоматизации работы с документами по ИБ (КИТ-Журнал, АльфаДок). Например, «КИТ-Журнал» облегчает подготовку следующих документов: Список допущенных к работе с ключами СКЗИ / в помещения с ключами СКЗИ; Список лиц, имеющих доступ к содержанию электронного журнала сообщений; Список лиц, допущенных к защищаемой информации; Заявка на предоставление или изменение доступа; Матрица доступа, перечень защищаемых ресурсов и др. Шаблоны документов встроены в программу, но предусмотрена возможность добавления и пользовательских шаблонов документов. Этот функционал позволяет достичь экономии времени за счет мгновенного формирования документов. Программа «АльфаДок» также обеспечивает автоматизированную разработку документации по защите информации и позволяет поддерживать документы в актуальном состоянии в случае изменений в нормативной базе и внутри организационных изменений, таких как кадровые изменения, замена оборудования, программного обеспечения и др.

Повышение осведомленности персонала об ИБ также является актуальной проблемой. Международные стандарты по управлению информационной безопасностью включают не только требования, но и рекомендации по работе ссотрудниками. Особое внимание уделяют обучению сотрудников правилам информационной безопасности и повышению их осведомленности об этих правилах. Серия стандартов ИСО/МЭК 2700 по управлению информационной безопасностью предполагает наличие в организации эффективной программы повышения осведомленности, обучения и подготовки по ИБ, доводящей до сведения всех сотрудников их обязанности по обеспечению ИБ, сформулированные в политиках и стандартах ИБ, и побуждающая их к соответственным действиям. В стандарте ISO/IEC 27001:2013 InformationTechnology — SecurityTechniques — InformationSecurityManagementSystems — Requirements («Информационная технология — Методы и средства обеспечения безопасности — Системы менеджмента ИБ — Требования») в подразделе 7.3. «Осведомленность» (Awareness) приведены требования к сотрудникам организации: они должны быть осведомлены о политике ИБ; своем вкладе в обеспечение эффективности системы менеджмента ИБ, включая выгоды от улучшения функционирования ИБ; последствиях несоблюдения требований системы менеджмента ИБ [2]. Рекомендации по осведомленности, обучению и тренингам в области информационной безопасности, а также перечень мероприятий на разных этапах занятости сотрудника (в период трудоустройства, занятости и увольнения) содержит стандарт ISO/IEC 27002:2013 InformationTechnology — SecurityTechniques — CodeofPracticeforInformationSecurityControls («Информационная технология. Методы и средства обеспечения безопасности. Свод правил по мерам и средствам контроля и управления информационной безопасностью») [14]. Рекомендации по разработке программы информирования, обучения информационной безопасности представлены в п.7.3. «Awareness» стандарта ISO/IEC 27003:2017 [15].

Повышению осведомленности посвящены рекомендации и руководства. Документ «InformationSecurityAwarenessinFinancialOrganisations» (ENISA, 2008) содержит рекомендации по повышению осведомленности по вопросам ИБ в финансовых организациях [16]. «Модель зрелости осведомленности по вопросам безопасности» (SecurityAwareness Maturity Model) (Европейская научно-образовательная организация «The SANS Institute», 2011) дает возможность организациям определить, на каком этапе находится их программа повышения осведомленности о безопасности в настоящее время и в каком направлении следует двигаться в дальнейшем [17]. В ответ на быстрые темпы цифровизации специалисты сформулировалии изучают концепцию цифровой среды как киберпространства, концепции культуры кибербезопасности, цифровой культуры безопасности и т. д. Эти концепции должны лежать в основе культура информационной безопасности любой организации. Агентство Европейского Союза по сетевой и информационной безопасности (ENISA) приложило большие усилия для развития этих концепций. В 2017 году опубликован документ «CyberSecurityCultureinorganisations», в котором описываются рекомендации

и передовой опыт по созданию программ повышения культуры кибербезопасности, восемь этапов и пошаговые инструкции их реализации [18].

Вначале 20 века появилось важное направление в науке - культурология информационной безопасности. Широта и глубина исследований в этой области подтверждается масштабными обзорами публикаций по данной теме:с 2000 по 2013 год [19]; с 2003 по 2016 год [20]; с 2000 г.до 2017 г. [21] и др.

В России также уделяется определенное (хотя и меньшее) внимание сотруднику организации в контексте информационной безопасности. Развитие нормативной правовой базы (Федеральный закон № 152-ФЗ от 27 июля 2006 года «О персональных данных», приказ ФСТЭК России от 14 марта 2014 г. № 31 «Об утверждении Требований к обеспечению защиты информации в автоматизированных системах управления производственными и технологическими процессами на критически важных объектах, а также объектах, представляющих повышенную опасность для жизни и здоровья людей и для окружающей среды», приказ ФСТЭК России от 25 декабря 2017 г. №239 «Об утверждении требований по обеспечениюбезопасности значимых объектов критической информационной инфраструктуры Российской Федерации», российские стандарты по управлению информационной безопасностью, стандарты Банка России по обеспечению информационной безопасности организаций банковской системы Российской Федерации и другие нормативные правовые и методические документы Российской Федерации) способствовало развитию рынка услуг по повышению осведомленности об информационной безопасности (KasperskySecurityAwareness (Лаборатория Касперского), PhishmanИнформационный центр (Phishman), Антифишинг (Антифишинг) Company), SecurityAwarenessPlatform (UBS), SyssoftSecurityAwareness (SyssoftCompany), DeteactAwareness(DeteAct) и т. д.). Однако, несмотря на это, серьезные проблемыповышенияосведомленности персоналаоб информационной безопасности и культуры их информационной безопасности остаются [22].

В связи с изложенным мы считаем, что одним из важных факторовналичия этой проблемы - отсутствие взаимосвязимежду процессами управления документами иосведомленностью. В существующих исследованиях не уделялось особого вниманиявзаимодействию между различными этапами менеджмента ИБ, таких исследований немного. Так, исходя из требований стандарта [23], автор обосновывает связьмежду всеми группами процессов PRM и группойпроцессы управления документами [24].Тем не менее большинство экспертов приходят к выводу, что дальнейшие исследования должныуделять больше внимания взаимодействию между этапами управления политикой информационной безопасности, развивать дополнительные исследования для разработки компьютеризированной поддержки управления ИБ, изучать степень, до которой компьютеризированная поддержка может улучшить интеграцию этапов управления ИБ и упростить управление процессами [8, 12]. Полагаем, что это в полной мере относится и к взаимосвязям между группами процессов, входящих в «Общие интегрированные процессы управления» - Управление документацией (COM 02) и Управление человеческими ресурсами (COM 03).

Необходимость интеграции процессов управления документацией и осведомленностью сотрудников обусловлена еще одним фактором - документационный формат контроля ИБ и УИБ.

Приступая к конкретной проверке, относящейся к информационной безопасности, аудиторы, проводящие проверку мер и средств контроля и управления информационной безопасностью, обычно начинают со сбора предварительной информации из различных источников. Это результаты предыдущих проверок, тестирований и оценок, частично или полностью относящихся к текущей области проверки и так или иначе выполненных аудиторами, проводящими проверку мер и средств контроля и управления информационной безопасностью (например, предварительные тесты безопасности, проведенные специалистами по обеспечению информационной безопасности, могут дать обширные знания по безопасности основных прикладных систем); сведения о соответствующих инцидентах информационной безопасности, ситуациях, близких к инцидентам, вопросах поддержки и изменениях, полученные от службы технической поддержки ИТ, из процессов менеджмента изменений ИТ, процессов менеджмента инцидентов ИТ и из аналогичных источников;

Согласно ГОСТ Р 56045-2014/ISO/IEC TR 27008:2011 «Информационная технология

(ИТ). Методы и средства обеспечения безопасности. Рекомендации для аудиторов в отношении мер и средств контроля и управления информационной безопасностью», к методам проверки, наряду с опросом и тестированием, относится изучение (п. 7.2.). Объекты проверки обычно включают в себя спецификации - документы, устанавливающие требования (см. ГОСТ ИСО 9000-2011, пункт 3.7.3). К ним относятся политики, планы, процедуры, требования к системам, технические инструкции и руководства пользователя/администратора и др. Кроме того, в приложение В включен раздел В1 Кадровые ресурсы и безопасность для проверки мер и средств контроля и управления ИБ, связанной с сотрудниками организации (чувствует ли персонал себя ответственным и/или подотчетным за свои действия; является ли персонал заслуживающим доверия, чтобы обращаться с чувствительной информацией и системами, которые могут подвергать опасности продолжительность существования организации; является ли персонал таким, которому можно полностью доверять; как определяется и измеряется доверие и др.) [25].

Особенно остро стоит проблема интеграции процессов УИБ на малых и средних предприятиях, работающих в условиях ограниченных финансовых, кадровых, временных и др.ресурсов и, следовательно, уделяющих недостаточное внимание этой проблеме. Эксперты единодушны в том, что объем документированной информации для СМИБ может отличаться от одной организации к другой из-заразмера организации и вида ее деятельности; процессов, продуктов и услуг; сложности процессов и их взаимодействия; компетентности сотрудников [3, 4]. Поэтому обнаружено [11], что многие малые и средние предприятия не придерживаются принципов управления безопасностью, в основном, из-за ограниченных ресурсов и опыта. Это особенно верно для тех принципов, которые используются при разработке политик информационной безопасности и мониторинге их соблюдения. Исследования показывают, что эта проблема существует во всем мире и оказывает особенно большое влияние на малый и средний бизнес, которому в настоящее время требуется не теоретическая, а в большей степени практическая помощь в виде процессов, процедур и инструментов, которые они могут использовать в реальных условиях.

3. Программное решение для интеграции управления документами и осведомленностью об ИБ

Чтобы решить обоснованную выше проблему, мы разработали прототип веб-приложения для интегрированного управления документацией по информационной безопасностии осведомленностью сотрудников о ней. Веб-приложение предназначено для модернизации системы управления информационной безопасностью в малом и среднем бизнесе.

Веб-приложение разработано на Python с использованием Фреймворков Django, Bootstrap (интерфейс). SQLite использовался в качестве базы данных, которая используется Django по умолчанию. Python – это универсальный язык программирования высокого уровня, ориентированный на повышение производительности труда разработчиков и читаемость кода. Django – это бесплатный фреймворк для веб-приложений Python, который использует шаблон проектирования MVC. SQLite – это встраиваемая кроссплатформа баз данных, поддерживающая достаточно полный набор SQL команд. К достоинствам Django можно отнести безопасность, масштабируемость, универсальность, скорость развития и др.

Веб-приложение можно разделить на две части: административная панель и панель пользователя. Административная панель позволяет управлять настройками сети (рис.1).

Информация о пользователе выглядит на панели администратора следующим образом (рис.2).

Пользовательская панель предназначена для просмотра документации по информационной безопасности и информации о ней. Панель авторизации изображена на рис.3.

После авторизации открывается панель пользователя (рис. 4), позволяющая работать с документацией по информационной безопасности и проходить тестирование на осведомленность о них.

Для работы программы администратор вводит в базу данных полный пакет документов по информационной безопасности предприятия и заполняет информацию о них. В содержание информации о каждом документе входит: Тип документа; Имя; Дата утверждения; Период действия; Утверждено; Ответственный; Описание; Список ознакомления.

Администратор может редактировать документ (рис.5).

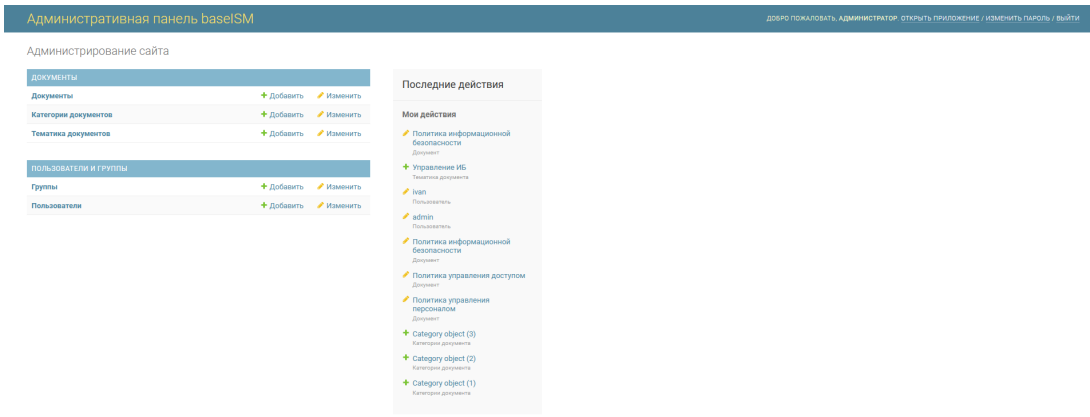

Рис. 1. Панель администратора

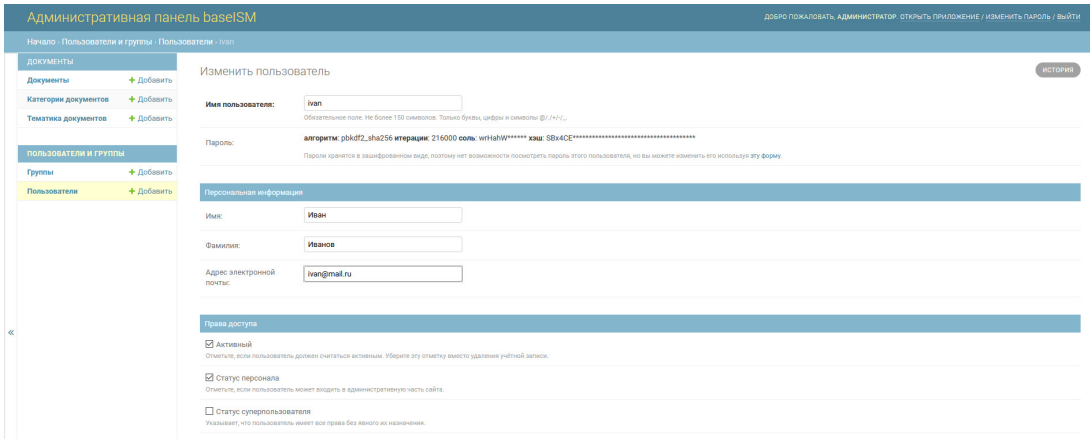

Рис. 2. Информация о пользователе на административной панели

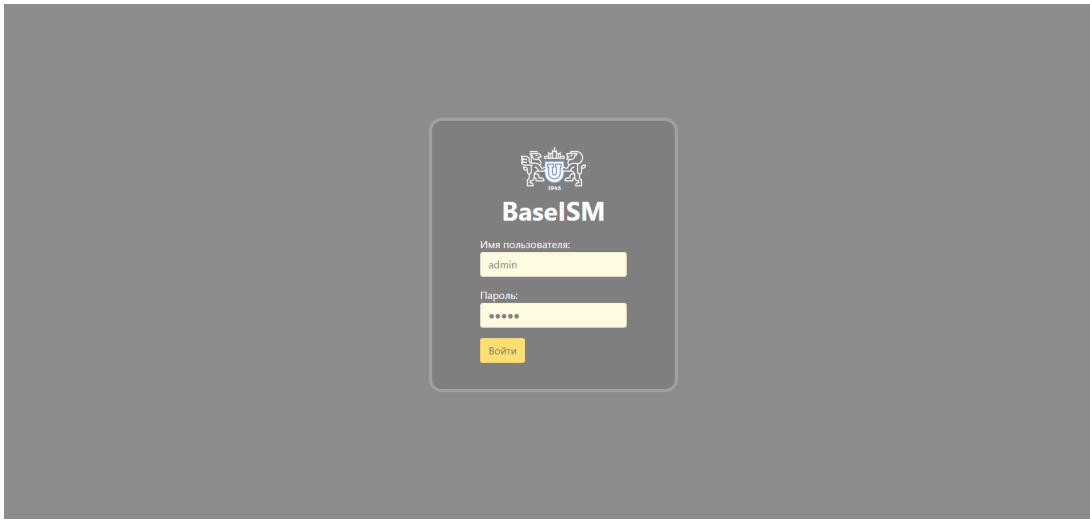

Рис. 3. Панель авторизации пользователя

Сотрудник может работать с документом при наличии права доступа, изучать его, после чего должен сделать отметку об ознакомлении в поле «Ознакомлен» (рис.6).

Администратор следит за процессом ознакомления сотрудников с документацией (рис.6) и процессом тестирования в рамках повышения осведомленности персонала об информационной безопасности (рис.7). Анализ и интерпретация результатов тестирова-

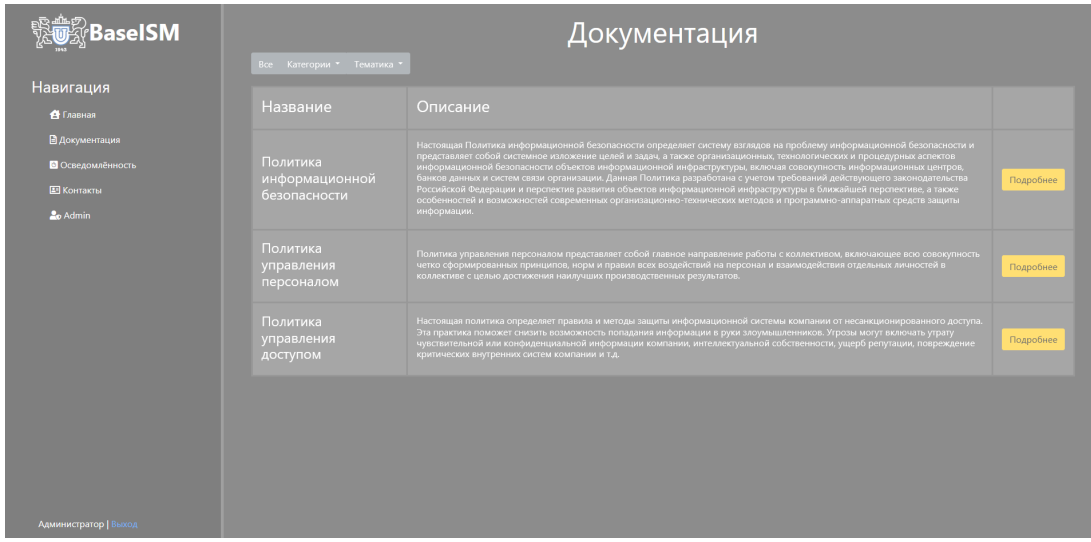

Рис. 4. Панель пользователя

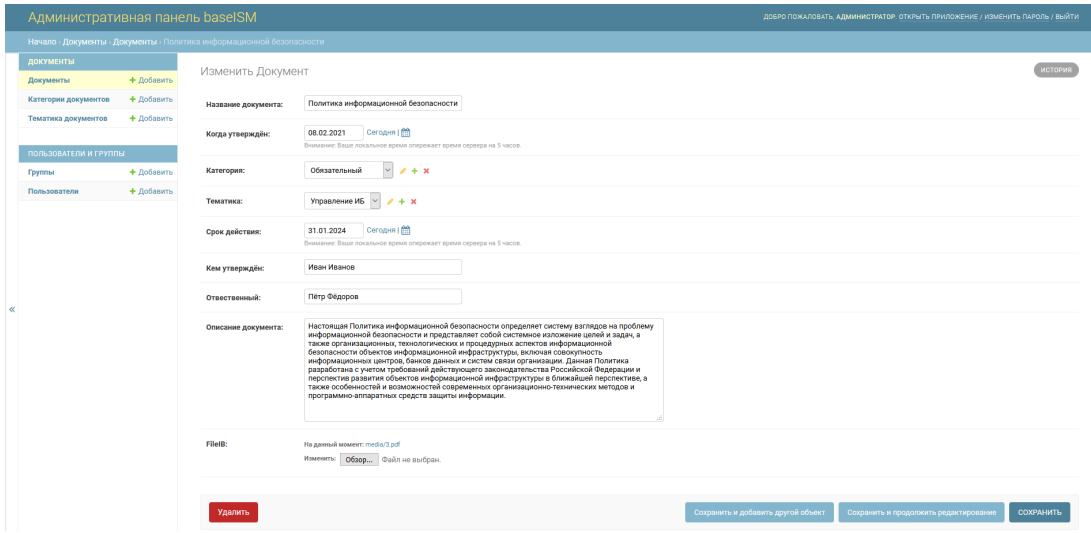

Рис. 5. Панель редактирования информации о документе

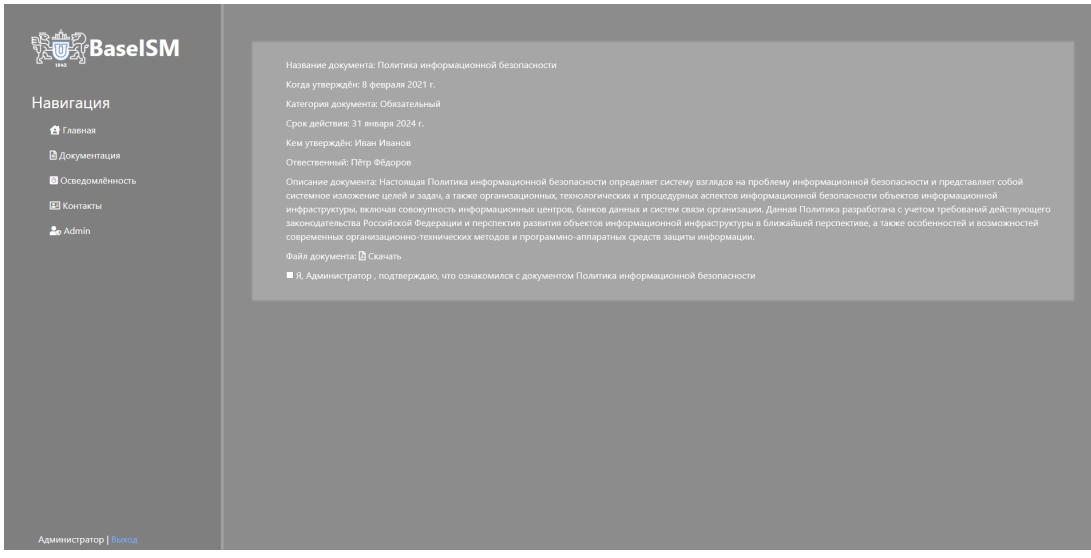

Рис. 6. Панель работы пользователя с документом

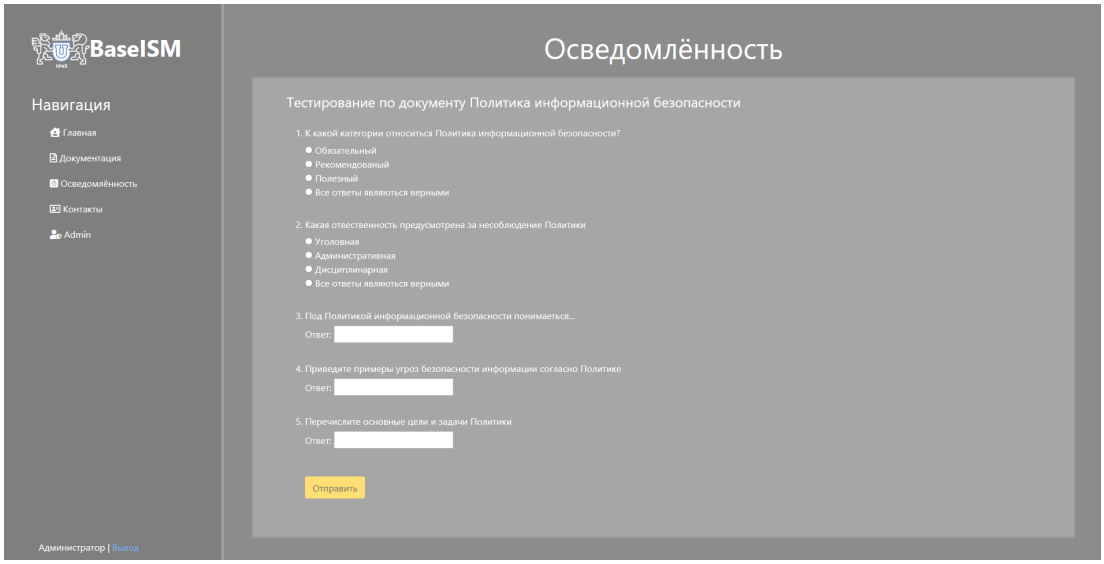

Рис. 7. Панель тестирования сотрудника на знание документа

ния осуществляется только в контексте общей системы управления информационной безопасностью каждой конкретной организации, что является предметом отдельной публикации.

Функциональность представленного веб-приложения не ограничена. В разработке находятся функции управления активами, управления инцидентами информационной безопасности, управления уязвимостями, управления аудитами.

4. Заключение

Анализ публикаций показал проблему отсутствия взаимосвязи между процессами управления информационной безопасностью в организациях. Для решения этой проблемы мы разработали многофункциональный веб-инструмент для интеграции двух процессов управления – управления документами и управления осведомленностью сотрудников о документированных правилах информационной безопасности организации. Продукт может помочь минимизировать затраты на ресурсы для реализации СМИБ-процедур. Научная новизна исследования заключается в обосновании интегративного подхода к управлению системой документации в области ИБ и управлению осведомленностью сотрудников организации об этой системе. Практическое значение работы заключается в возможности использования разработанного веб-приложения на малых и средних предприятиях для оптимизации процессов обеспечения информационной безопасности.

Литература

1. InfoWatch. Restricted Information Leaks: 9 Months 2020 Report. - Available at: https://www.infowatch.ru/analytics/reports/30708 (accessed 31.03.2021).

2. ISO/IEC 27001-2013 Information technology — Security techniques — Information security management systems — Requirements. - Available at: https://www.iso.org/standard/54534.html(датаобращения:19.05.2021).

3. Biswas P. ISO 27001:2013 Information Security Management System. July 20, 2019. - Available at: https://isoconsultantkuwait.com/2019/07/20/2392 (датаобращения: 19.05.2021).

4. ProzorovA. ISMS Documented Information List, v.4.0. - Available at: https://vk.com/isms8020. (датаобращения: 19.05.2021).

5. ISMS Implementation Guideline/ A practical approach. - Available at: file:///C:/Users/1D1D~1/AppData/Local/Temp/1170.pdf (датаобращения: 19.05.2021).

6. ISO 27001: 2013 ISMS Documentation toolkit contents and ISO 27001: 2013 Requirement mapping //Document Control. – 2016. - No. 1. - pp.1-7.

7. COBIT 19 - URL: https://www.isaca.org/resources/cobit(дата обращения: 17.05.2021).

8. Rostami E., Karlsson F., Gao S. Requirements for computerized tools to design information security policies //Computers & Security. – 2020. - Vol. 99. - 102063.

9. Rose A., Okfalisac A. A. Information Security Policy Compliance: Systematic Literature Review // Procedia Computer Science. - 2019. - Vol. 161. - pp. 1216-1224.

10. Flowerday S. V., Tuyikeze T. Information security policy development and implementation: The what, how and who // Computers & Security. – 2016. - Vol. 61. - pp. 169-183.

11. Coertze J., von Solms R. A software gateway to affordable and effective Information Security Governance in SMMEs //Information Security for South Africa. - 2013. - pp. 1-8.

12. Rostami E., Karlsson F., Kolkowska E. The hunt for computerized support in information security policy management. Aliteraturereview //Information&ComputerSecurity. - 2020. - Vol. 28, No. 2. - pp. 215-259.

13. Суровцева Н. Г. Роботизированная документация: проблемы управления // Управление документами в цифровой экономике: Материалы научно-практич. конф. 5 декабря 2018 г. - М, 2019. - С. 23-30.

14. ISO/IEC 27002:2013 Information technology – Security techniques – Code of practice for information security management. – Available at: https://www.iso.org/standard/54533.html (датаобращения: 15.05.2021).

15. ISO/IEC 27003:2017 Information technology — Security techniques — Information security management systems — Guidance. – Available at: https://www.iso.org/obp/ui/#iso:std:iso-iec:27003:ed-2:v1:en. (датаобращения: 15.05.2021).

16. Information Security Awareness in Financial Organisations ENISA. – Available at: https://webcache.googleusercontent.com/search?q=cache:0Us-oaHURyQJ:https://www.enisa.europa.eu/publications/archive/is-in-financial-organisations/at_download/fullReport+&cd=1&hl=ru&ct=clnk&gl=ru(датаобращения: 15.05.2021).

17. Spitzner L. Defining the Security Awareness Maturity Model. – Available at: https://www.sans.org/security-awareness-training/blog/defining-security-awareness-maturity-model (датаобращения: 19.05.2021).

18. ENISA. Cyber Security Culture in organizations. 2017. - Available at: https://doi.org/10.2824/10543 (датаобращения: 19.05.2021).

19. Karlsson F., Åström J., Karlsson M. Information security culture–state-of-the-art review between 2000 and 2013 //Information & Computer Security. - 2015. - No. 23(3). - pp.246-285.

20. Mahfuth A., Yussof S., Baker A.A., Ali N. A systematic literature review: Information security culture // 2017 5th International Conference on Research and Innovation in Information Systems (ICRIIS). DOI:10.1109/ICRIIS.2017.8002442 - Available at: https://www.researchgate.net/publication/319054554_A_systematic_literature_review_Information_security_culture (датаобращения: 19.05.2021).

21. Nasir A., Arshah R.A., Ab Hamid M. R., Fahmy S. An analysis on the dimensions of information security culture concept: A review //Journal of Information Security and Applications. – 2019. - No.44. - pp.12-22.

22. SANS Security Awareness Report. Building Successful Security Awareness Programs. 2018. - Available at: https://www.sans.org/sites/default/files/2018-05/2018%20SANS%20Security%20Awareness%20Report.pdf(датаобращения: 18.05.2021).

23. ISO/IEC TS 33052:2016 Information technology. Process reference model (PRM) for information security management". - Available at: https://www.iso.org/standard/55142.html (датаобращения: 19.05.2021).

24. Sysoeva L.A. Using a reference model of processes in describing information security management processes in an organization. Information security: yesterday, today, tomorrow // Collection of articles based on the materials of the International Scientific and Practical Conference. –M., 2019. - pp. 59-65.

25. ISO / IEC TR 27008: 2011 "Information technology (IT). Methods and means of ensuring security. Recommendations for auditors regarding measures and means of control and management of information security." - URL: https://www.iso.org/ru/standard/45244.html(датаобращения: 19.05.2021).

References

1. InfoWatch. Restricted Information Leaks: 9 Months 2020 Report. - Available at: https://www.infowatch.ru/analytics/reports/30708 (accessed 31.03.2021).

2. ISO / IEC 27001-2013 Information technology - Security techniques - Information security management systems - Requirements. - Available at: https://www.iso.org/standard/54534.html (date accessed: 19.05.2021).

3. Biswas P. ISO 27001: 2013 Information Security Management System. July 20, 2019. - Available at: https://isoconsultantkuwait.com/2019/07/20/2392 (date accessed: 19.05.2021).

4. Prozorov A. ISMS Documented Information List, v.4.0. - Available at: https://vk.com/isms8020. (date of access: 19.05.2021).

5. ISMS Implementation Guideline / A practical approach. - Available at: file: /// C: /Users/1D1D~1/AppData/Local/Temp/1170.pdf (date accessed: 19.05.2021).

6. ISO 27001: 2013 ISMS Documentation toolkit contents and ISO 27001: 2013 Requirement mapping // Document Control. - 2016. - No. 1. - pp. 1-7.

7. COBIT 19 - URL: https://www.isaca.org/resources/cobit (date accessed: 17.05.2021).

8. Rostami E., Karlsson F., Gao S. Requirements for computerized tools to design information security policies // Computers & Security. - 2020. - Vol. 99. - 102063.

9. Rose A., Okfalisac A. A. Information Security Policy Compliance: Systematic Literature Review // Procedia Computer Science. - 2019. - Vol. 161. - pp. 1216-1224.

10. Flowerday S. V., Tuyikeze T. Information security policy development and implementation: The what, how and who // Computers & Security. - 2016. - Vol. 61. - pp. 169-183.

11. Coertze J., von Solms R. A software gateway to affordable and effective Information Security Governance in SMMEs // Information Security for South Africa. - 2013. - pp. 1-8.

12. Rostami E., Karlsson F., Kolkowska E. The hunt for computerized support in information security policy management. A literature review // Information & Computer Security. - 2020. - Vol. 28, No. 2. - pp. 215-259.

13. Surovtseva N. G. Robotic documentation: management problems // Document management in the digital economy: Materials of scientific and practical research. conf. December 5, 2018. - M, 2019.- pp. 23-30.

14. ISO / IEC 27002: 2013 Information technology - Security techniques - Code of practice for information security management. - Available at: https://www.iso.org/standard/54533.html (date accessed: 15.05.2021).

15. ISO / IEC 27003: 2017 Information technology - Security techniques - Information security management systems - Guidance. - Available at: https://www.iso.org/obp/ui/#iso:std:iso-iec:27003:ed-2:v1:en. (date of access: 15.05.2021).

16. Information Security Awareness in Financial Organizations ENISA. - Available at: https://webcache.googleusercontent.com/search?q=cache://Us-oaHURyQJ:https://www.enisa.europa.eu/publications/archive/is-in-financial-organisations/at_download/ fullReport + & cd = 1 & hl = ru&ct = clnk&gl = ru (date accessed: 15.05.2021).

17. Spitzner L. Defining the Security Awareness Maturity Model. - Available at: https://www.sans.org/security-awareness-training/blog/defining-security-awareness-maturity-model (date accessed: 19.05.2021).

18. ENISA. Cyber Security Culture in organizations. 2017. - Available at: https://doi.org/10.2824/10543 (date accessed: 19.05.2021).

19. Karlsson F., Åström J., Karlsson M. Information security culture – state-of-the-art review between 2000 and 2013 // Information & Computer Security. - 2015. - No. 23 (3). - pp. 246-285.

20. Mahfuth A., Yussof S., Baker A.A., Ali N. A systematic literature review: Information security culture // 2017 5th International Conference on Research and Innovation in Information Systems (ICRIIS). DOI: 10.1109 / ICRIIS.2017.8002442 - Available at: https://www.researchgate.net/publication/319054554_A_systematic_literature_review_Information_security_culture (date accessed: 19.05.2021).

21. Nasir A., Arshah R. A., Ab Hamid M. R., Fahmy S. An analysis on the dimensions of information security culture concept: A review // Journal of Information Security and Applications. - 2019. - No.44. - pp.12-22.

22. SANS Security Awareness Report. Building Successful Security Awareness Programs. 2018. - Available at: https://www.sans.org/sites/default/files/2018-05/2018%20SANS%20Security%20Awareness%20Report.pdf (date accessed: 18.05.2021).

23. ISO / IEC TS 33052: 2016 Information technology. Process reference model (PRM) for information security management."- Available at: https://www.iso.org/standard/55142.html (date accessed: 19.05.2021).

24. Sysoeva L.A. Using a reference model of processes in describing information security management processes in an organization. Information security: yesterday, today, tomorrow // Collection of articles based on the materials of the International Scientific and Practical Conference. - M., 2019. - pp. 59-65.

25. ISO / IEC TR 27008: 2011 "Information technology (IT). Methods and means of ensuring security. Recommendations for auditors regarding measures and means of control and management of information security." - URL: https://www.iso.org/ru/standard/45244.html (date of access: 19.05.2021).

АСТАХОВА Людмила Викторовна, доктор педагогических наук, профессор, профессор кафедры защиты информации, Южно-Уральский государственный университет (национальный исследовательский университет). 454080, г. Челябинск, пр. им. В.И. Ленина, 76. E-mail: astakhovalv@susu.ru

КИРЯЕВ Андрей Игорьевич, студент кафедры защиты информации, Южно-Уральский государственный университет (национальный исследовательский университет). 454080, г. Челябинск, пр. им. В.И. Ленина, 76. E-mail: andkir0408@gmail.com

ASTAKHOVA Liudmila Victorovna, Doctor of Pedagogy, Professor, Professor of the Department of Information Security, South Ural State University (National Research University). 454080, Chelyabinsk, Lenin Avenue, 76. E-mail: astakhovalv@susu.ru

KIRYAEV Andrey Igorievich, student of the Department of Information Security, South Ural State University (National Research University). 454080, Chelyabinsk, Lenin Avenue, 76. E-mail: andkir0408@gmail.com

УДК 34.03:044.056.2 Вестник УрФО № 4(42) / 2021, с. 59–63

Минбалеев А. В.

DOI: 10.14529/secur210406

ПРОБЛЕМЫ И ПЕРСПЕКТИВЫ ОБЕСПЕЧЕНИЯ ЗАЩИТЫ ПЕРСОНАЛЬНЫХ ДАННЫХ ГРАЖДАН В ЦИФРОВОМ ПРОФИЛЕ[1]

Активное развитие цифрового профилирования и использование возможностей цифровых профилей гражданина различными коммерческими организациями ставит задачу по обеспечению информационной безопасности данных процессов. Информационные риски и угрозы сегодня отчетливо проявляются как в части несанкционированного сбора информации о гражданах из цифровых профилей, так и в части незаконного их использования, в том числе для мошеннических целей. Государством ставится задача по защите прав граждан в связи с функционированием цифровых профилей, в первую очередь, защите их персональных данных и иной информации, составляющей личную и семейную тайну.

В статье автором анализируются проблемы, складывающиеся в процессы функционирования системы «Цифровой профиль», а также предлагается ряд мер по обеспечению защиты персональных данных граждан в цифровом профиле.

Ключевые слова: *информационная безопасность, правовая охрана, персональные данные, распространение персональных данных в сети Интернет, цифровой профиль.*

Minbaleev A. V.

PROBLEMS AND PROSPECTS OF ENSURING THE PROTECTION OF PERSONAL DATA OF CITIZENS IN THE DIGITAL PROFILE

The active development of digital profiling and the use of the capabilities of digital citizen profiles by various commercial organizations poses the task of ensuring the information security of these processes. Information risks and threats are clearly manifested today both in terms of unauthorized collection of information about citizens from digital profiles, and in terms of

[1] Работа выполнена при финансовой поддержке Совета по грантам Президента Российской Федерации (грант МД-2209.2020.6) «Развитие системы правовых средств обеспечения кибербезопасности в Российской Федерации».

their illegal use, including for fraudulent purposes. The State sets the task of protecting the rights of citizens in connection with the functioning of digital profiles, first of all, the protection of their personal data and other information constituting personal and family secrets.

In the article, the author analyzes the problems developing in the processes of functioning of the "Digital Profile" system, and also proposes a number of measures to ensure the protection of personal data of citizens in the digital profile.

Keywords: *information security, legal protection, personal data, dissemination of personal data on the Internet, digital profile.*

В связи с успешным проведением эксперимента по внедрению и использованию системы «Цифровой профиль» можно однозначно говорить о его переходе в следующем году в режим штатного функционирования и использования государством. Однако создание цифрового профиля, с одной стороны, повышает качество жизни и доступность сферы публичного управления для гражданина, с другой стороны, существенно обостряет риски информационной безопасности, ущерб от которых может быть невосполнимым [1].

Создание единого цифрового профиля без должной его защиты и внесения соответствующих изменений в информационное, гражданское, административное и уголовное законодательство, приведет к увеличению числа преступлений в данной области, поскольку незаконное получение сведений из цифрового профиля будет приобретать еще больший интерес, в связи с большим объемом сведений, содержащихся в нем. Функционирование цифровых профилей предусматривает добровольный характер, в том числе, отказ от сбора и обработки персональных данных. Однако у этого решения будут юридически значимые последствия в связи с невозможностью совершения ряда юридически значимых действий после отзыва согласия на обработку персональных данных, фактически речь идет о «добровольном» ограничении правового статуса. Все эти и другие проблемы, к сожалению, сегодня не решены, что вызывает серьезную обеспокоинность.

Анализ современного состояния правового регулирования цифрового профиля в России позволяет систематизировать и выделить следующие проблемы в данном направлении:

– отсутствие формализованных механизмов обработки данных с целью повышения уровня и качества жизни граждан;

– недостаточный уровень доступности, качества и актуальности государственных данных, необходимых для перехода на цифровое взаимодействие;

– отсутствие инфраструктуры, способной обеспечить унифицированный, безопасный, быстрый и удобный обмен данными между всеми участниками;

– отсутствие стандартов и решений в сфере информационной безопасности, в том числе криптографии, для обеспечения безопасного обмена данными между государственными органами и коммерческими компаниями;

– отсутствие соответствующей нормативной правовой базы [2].

Эти факторы негативно сказываются как для граждан, предпринимателей, так и на органы государственной власти. К числу проблем можно отнести также и достаточно высокий уровень расходов, которые связаны с обработкой бумажных документов, а также необходимостью личного присутствия граждан, ручной проверки и подтверждения предоставления данных, и низкая эффективность внутренних бизнес-процессов организаций, которые связаны с аналитикой данных (скоринг, риск-менеджмент, оценка просроченной задолженности, подготовка отчетности и т.д.), низкое качество клиентского опыта, снижение конверсии, недоступность части услуг для отдельных групп клиентов (например, услуг, требующих личного присутствия (подача документов, идентификация), для граждан, проживающих в труднодоступных районах), сложность персонализации продуктов и услуг на основе данных о физическом лице, полученных или актуализированных из внешних источников, и отсутствие у гражданина возможности управления выданными согласиями в электронном виде и др.

Представляется целесообразным рассмотреть возможность внесения следующих изменений в законодательство Российской Федерации в части развития концепции «цифрового профиля»:

1. Сегодня законодатель предусматрива-

ет создание реестра согласий на обработку персональных данных гражданина, но в рамках них, к сожалению, не происходит реализация технологической возможности для гражданина отслеживать, какие органы государственной власти или организации используют персональные данные гражданина. Не устанавливается взаимодействие Цифрового профиля с первоначальными реестрами данных, в той части, когда запрос на использование данных направляется в первоначальный источник. Таким образом, гражданин не может полностью отследить через систему Цифрового профиля движений своих данных, что может обусловливать определенными злоупотреблениями;

2. К сожалению, на сегодняшний день не выработаны основные принципы внесения изменений в реестры. Сегодня необходимо более четче прописать алгоритм таких изменений и организационно продумать весь процесс;

3. В процессе развития системы Цифрового профиля необходимо проработать и решить вопрос о технологических и функциональных аспектах их реализации. Не в полной мере выработаны технологические и функциональные особенности реализации права на забвение;

4. К сожалению, не решен вопрос и необходимо вносить соответствующие изменения в законодательство об обеспечении должной защиты прав граждан на персональные данные, которые были получены из цифрового профиля коммерческими организациями и обрабатываются ими в коммерческих и иных экономических целях. Явно защита персональных данных в этом случае должна носить особый характер, особенно в части защиты биометрических персональных данных [3-4]. Причем, важно не только предусматривать требования к информационным системам таких персональных данных и требованиям к их функционированию в коммерческих организациях, но и прорабатывать на уровне ФЗ «О персональных данных» [5] и иные меры. Одной меры здесь явно недостаточно, поскольку крупные операторы таких персональных данных легко могут позволить выполнить жесткие требования к информационным системам персональных данных, а предприятия малого и среднего бизнеса не смогут обеспечить такую возможность. С другой стороны, если предусматривать большие штрафные санкции, которые логично могут быть предусмотрены для защиты прав на персональные данные из цифрового профиля, то крупные игроки на рынке легко могут закладывать такие штрафы в риски, связанные с утечкой персональных данных;

5. Важно также предусмотреть в рамках системы Цифрового профиля возможность реализовать на его базе возможности взаимодействия граждан и коммерческих организаций без рассылки рекламных и иных сообщений, если у организации отсутствует согласия на обработку персональных данных;

6. Необходимо решить вопрос, связанный с признанием юридической значимости данных Цифрового профиля, возможность их официального признания и использования в качестве официальных данных в тех или иных случаях, в том числе в качестве доказательств;

7. В связи с тем, что в рамках цифрового профиля гражданин может предоставить право использовать свои персональные данные той или иной организации для получения через нее государственной услуги, важно установить требования к коммерческим и некоммерческим организациям, в которые может обратиться заявитель за организацией предоставления государственных услуг. В рамках предлагаемых требований необходимо установить, как критерии для коммерческих и некоммерческих организаций, так и критерии для государственных услуг, которые могут предоставляться через инфраструктуру коммерческих и некоммерческих организаций гражданам. В качестве них могут быть предложены: количество принятых от заявителей запросов о предоставлении государственной услуги превышает в год (не менее 100-150 тысяч); организация предоставления государственной услуги в организации способствует повышению доли граждан, имеющих доступ к получению такой услуги; для получения услуг и сервисов организации используется результат предоставления государственной услуги либо сфера деятельности организации связана с получением государственной услуги и др.

Полагаем, что данные и другие предложения по развитию системы Цифрового профиля необходимо реализовать на уровне специально разработанных постановлений Правительства Российской Федерации, в которых бы были утверждены требования к организациям, которые получают для своих целей данные о гражданах из Цифрового про-

филя, а также основные требования к коммерческим и некоммерческим организациям, в которые может обратиться заявитель за организацией предоставления государственных услуг, условия предоставления таких услуг и критерии отбора указанных организаций. Ответственными федеральными органами исполнительной власти при этом должны быть Минцифры России, Минэкономразвития России.

При предоставлении персональных данных из цифрового профиля коммерческим и некоммерческим организациям важно закрепить ряд базовых положений:

– организации, предоставляющие государственные услуги и получающие персональные данные о гражданах обязаны обеспечивать защиту информации, доступ к которой ограничен в соответствии с требованиями законодательства Российской Федерации в области персональных данных, а также соблюдать режим обработки и использования персональных данных, в том числе и их работники;

– организации вправе использовать результаты предоставления государственных услуг и информации, ставшей доступной в результате предоставления государственных услуг заявителю, в том числе передавать их третьим лицам, только при наличии информированного согласия заявителя на такое использование, за исключением случаев, когда такие результаты или информация в соответствии с федеральными законами являются общедоступной информацией;

– согласие заявителя оформляется в письменной форме посредством подписания документа на бумажном носителе или с использованием электронных документов, подписанных электронной подписью в соответствии с требованиями Федерального закона «Об электронной подписи», в том числе посредством проставления отметки (галочки) в таком согласии при условии ее проставления заявителем, аутентифицированным в информационной системе организации;

– организация не вправе обуславливать возможность получения государственной услуги заявителем обязанностью предоставить согласие на использование результатов предоставления государственных услуг и информации, ставшей доступной в результате предоставления государственных услуг.

Эти и другие меры будут способствовать развитию системы цифрового профилирования в Российской Федерации, а также обеспечению и защите прав граждан на защиту персональных данных.

Литература

1. Полякова Т. А. Проблемы правового обеспечения информационной безопасности в процессе использования цифровых технологий в глобальной цифровой среде / Т. А. Полякова, А. В. Минбалеев, И. С. Бойченко // Вестник Академии права и управления. – 2018. – № 3(52). – С. 32-36.

2. Полякова Т. А. Концептуальные подходы к правовому регулированию информационной безопасности в условиях цифровизации и трансформации права / Т. А. Полякова, А. В. Минбалеев, И. С. Бойченко // Вестник УрФО. Безопасность в информационной сфере. – 2019. – № 3(33). – С. 64-68.

3. Методические рекомендации Министерства цифрового развития, связи и массовых коммуникаций Российской Федерации «Сценарии использования инфраструктуры Цифрового профиля. Версия 1.2» от 2 апреля 2021 г. URL: https://digital.gov.ru/uploaded/presentations/stsenariiispolzovaniyatspv12.pdf (дата обращения: 20.04.2021 г.).

4. Брызгин, А. А. Правовой режим биометрических персональных данных / А. А. Брызгин, А. В. Минбалеев // Вестник УрФО. Безопасность в информационной сфере. – 2012. – № 2(4). – С. 35-41.

5. Захаров, М. Н. Правовая защита персональных данных / М. Н. Захаров, А. В. Минбалеев // Безопасность информационного пространства: сборник трудов XIII Всероссийской научно-практической конференции студентов, аспирантов и молодых учёных, Челябинск, 26–28 ноября 2014 года / Министерство образования и науки Российской Федерации, Южно-Уральский государственный университет, Кафедра «Безопасность информационных систем». – Челябинск: Издательский центр ЮУрГУ, 2015. – С. 211-215.

6. Федеральный закон от 27.07.2006 № 152-ФЗ «О персональных данных» // Собрание законодательства РФ. 2006. № 31 (1 ч.). Ст. 3451.

References

1. Polyakova T. A. Problemy pravovogo obespecheniya informacionnoj bezopasnosti v processe ispol'zovaniya cifrovyh tekhnologij v global'noj cifrovoj srede / T. A. Polyakova, A. V. Minbaleev, I. S. Bojchenko // Vestnik Akademii prava i upravleniya. – 2018. – № 3(52). – S. 32-36.

2. Polyakova T. A. Konceptual'nye podhody k pravovomu regulirovaniyu informacionnoj bezopasnosti v usloviyah cifrovizacii i transformacii prava / T. A. Polyakova, A. V. Minbaleev, I. S. Bojchenko // Vestnik UrFO. Bezopasnost' v informacionnoj sfere. – 2019. – № 3(33). – S. 64-68.

3. Metodicheskie rekomendacii Ministerstva cifrovogo razvitiya, svyazi i massovyh kommunikacij Rossijskoj Federacii «Scenarii ispol'zovaniya infrastruktury Cifrovogo profilya. Versiya 1.2» ot 2 aprelya 2021 g. URL: https://digital.gov.ru/uploaded/presentations/stsenariiispolzovaniyatspv12.pdf (data obrashcheniya: 20.04.2021 g.).

4. Bryzgin, A. A. Pravovoj rezhim biometricheskih personal'nyh dannyh / A. A. Bryzgin, A. V. Minbaleev // Vestnik UrFO. Bezopasnost' v informacionnoj sfere. – 2012. – № 2(4). – S. 35-41.

5. Zaharov, M. N. Pravovaya zashchita personal'nyh dannyh / M. N. Zaharov, A. V. Minbaleev // Bezopasnost' informacionnogo prostranstva: sbornik trudov XIII Vserossijskoj nauchno-prakticheskoj konferencii studentov, aspirantov i molodyh uchyonyh, CHelyabinsk, 26–28 noyabrya 2014 goda / Ministerstvo obrazovaniya i nauki Rossijskoj Federacii, YUzhno-Ural'skij gosudarstvennyj universitet, Kafedra «Bezopasnost' informacionnyh sistem». – CHelyabinsk: Izdatel'skij centr YUUrGU, 2015. – S. 211-215.

6. Federal'nyj zakon ot 27.07.2006 № 152-FZ «O personal'nyh dannyh» // Sobranie zakonodatel'stva RF. 2006. № 31 (1 ch.). St. 3451.

МИНБАЛЕЕВ Алексей Владимирович, доктор юридических наук, профессор, зав. кафедрой информационного права и цифровых технологий, Московский государственный юридический университет имени О. Е. Кутафина (МГЮА). 123001, г. Москва ул. Садовая-Кудринская, 9.; профессор кафедры теории государства и права, конституционного и административного права, Южно-Уральский государственный университет. 454080, г. Челябинск, пр. Ленина, 76. E-mail: alexmin@bk.ru.

MINBALEEV Aleksey Vladimirovich, Doctor of Sciences (Law), Professor, head Department of Information Law and Digital Technologies of Kutafin Moscow State Law University (MSAL). Sadovaya-Kudrinskaya st., 9, Moscow, 123001.; Professor of the Department of Theory of State and Law, Constitutional and Administrative Law, South Ural State University. 76 Lenin Ave., Chelyabinsk, 454080. E-mail: alexmin@bk.ru.

УДК 532.529

Вестник УрФО № 4(42) / 2021, с. 64–70

Прокудина Л.А., Вихирев М.П.

DOI: 10.14529/secur210407

ЧИСЛЕННЫЙ МЕТОД РАСЧЕТА ВОЛНОВЫХ ХАРАКТЕРИСТИК ТЕЧЕНИЯ ЖИДКИХ ПЛЕНОК ДЛЯ ОБЕСПЕЧЕНИЯ НАДЕЖНОСТИ ТЕХНОЛОГИЧЕСКИХ ПРОЦЕССОВ В АППАРАТАХ ПЛЕНОЧНОГО ТИПА

Тепломассообменные аппараты, в которых реализуется течение тонких слоев вязких жидкостей (жидких пленок), широко применяются в различных областях промышленности (химической, нефтехимической, энергетической, пищевой и др.).

В работе представлены результаты численного моделирования течения вертикальных пленок жидкостей (воды, спирта, молока) со свободной поверхностью при умеренных числах Рейнольдса в рамках дифференциального уравнения в частных производных для отклонения свободной поверхности пленки от невозмущенного состояния. Коэффициенты уравнения включают различные физико-химические факторы, в частности, параметр поверхностного натяжения. Представлена аналитическая зависимость для критических значений параметра Марангони.

Разработан алгоритм численного расчета волновых параметров, обеспечивающий надежность технологических процессов, применяемых в пленочных аппаратах. Проведены вычислительные эксперименты с целью расчета волновых параметров и областей неустойчивости жидких пленок. Рассчитаны критические значения параметра Марангони, при котором достигается эффект разрыва жидкой пленки.

При уменьшении параметра поверхностного натяжения жидкости течение жидкой пленки становится более неустойчивым: наблюдается расширение области неустойчивости, снижение критического значения числа Марангони. Результаты работы могут быть использованы при разработке технологических процессов в жидких пленках.

Ключевые слова: *жидкая пленка, параметр поверхностного натяжения, неустойчивость, параметр Марангони, умеренные числа Рейнольдса, надежность, автоматизированная система управления технологическим процессом.*

NUMERICAL METHOD FOR COMPUTATION OF WAVE CHARACTERISTICS OF LIQUID FILM FLOW FOR ENSURING THE RELIABILITY OF TECHNOLOGICAL PROCESSES IN LIQUID FILM APPARATUSES

Heat and mass transfer devices based on flow of a thin layer of viscous fluid are widely used in numerous industry fields (chemical, petrochemical, energy, food, etc.).

This work presents results of computational modeling of vertical liquid (water, alcohol, milk) film flow at moderate Reynolds numbers in the framework of partial differential equation of the state of the free surface of the liquid film. Equation coefficients include different physicochemical factors such as surface tension parameter. Analytic dependence of Marangoni parameter critical values is presented.

Wave parameters calculation algorithm, ensuring the reliability of technological processes in liquid film apparatuses, is developed. Computational experiments were carried out to calculate wave characteristics and instability regions. Marangoni parameters critical values, at which destruction of film is possible, were calculated.

Increase of surface tension parameter leads to more unstable liquid film flow: increase of instability region and decrease of Marangoni parameter value could be noted. Results could be used in the design of technological processes in liquid films.

Keywords: *liquid film, surface tension parameter, instability, Marangoni parameter, moderate Reynolds numbers, reliability, industrial control systems.*

Введение

Жидкие пленки за счет малого термосопротивления и большой по-верхности контакта эффективны при межфазном тепломассообмене, что обуславливает их широкое применение в промышленности [1–3]. Пленочные аппараты применяются при производстве отдельных видов пластмасс, упаривании пищевых продуктов и т.д.

Нормальная работа пленочных аппаратов возможна при наличии устойчивой однородной пленки на всей рабочей поверхности, так как сухие участки будут подвергаться аномально большому тепловому воздействию, что может привести к аварийному режиму работы аппарата. Безаварийное протекание технологического процесса обеспечивается за счет применения автоматизированных систем управления технологическими процессами (АСУ ТП). Одним из необходимых свойств АСУ ТП является надежность, которую применительно к пленочным аппаратам можно рассматривать как свойство системы сохранять во времени в установленных пределах значения всех параметров, характеризующих способность обеспечивать безаварийное протекание технологического процесса.

При разработке алгоритмов автоматического управления необходимо учитывать не-

линейных характер объектов управления. На режимы и характеристики пленочного течения оказывают влияние различные физико-химические факторы [4–7]. В частности, необходимо учитывать физические свойства жидкости, течение которой осуществляется в пленочном аппарате. К таким свойствам относится параметр поверхностного натяжения жидкости [8]. Для решения этого вопроса применимы методы математического моделирования.

Целью данной работы является разработка алгоритма расчета волновых параметров и численное исследование влияния параметра поверхностного натяжения жидкости на волновые характеристики течения жидкой пленки при умеренных значениях числа Рейнольдса.

Математическая модель состояния свободной поверхности жидкой пленки

Для вертикального течения жидкой пленки по твердой непроницаемой поверхности в системе координат OXY выведено нелинейное дифференциальное уравнение состояния свободной поверхности жидкой пленки [8], линейная часть которого имеет вид:

$$\left(b_1\frac{\partial}{\partial x}-1\right)\frac{\partial\psi}{\partial t}+b_2\frac{\partial^4\psi}{\partial x^4}+b_3\frac{\partial^2\psi}{\partial x^2}+b_4\frac{\partial\psi}{\partial x}=0, (1)$$

где $\psi(x,t)$ – отклонение свободной поверхности жидкой пленки от невозмущенного состояния, t – время, x – координата, $b_1=\frac{5}{24}Re^2F_x$, $b_2=-\frac{1}{3}\sigma Re$, $b_3=-\frac{1}{2}ReM+\frac{3}{40}Re^3F_x^2$, $b_4=-ReF_x$, Re – число Рейнольдса, F_x – число Фруда, σ – параметр поверхностного натяжения, M – число Марангони.

В тепломассообменных процессах высокие градиенты температуры могут привести к разрыву пленки и образованию на поверхности пленочного аппарата «сухого пятна». Критические значения числа Марангони, при которых происходит разрушение пленки, определяются следующим выражением [9]:

$$M_к=\frac{3}{20}Re^2F_x^2-\frac{2}{k^2Re}\left(a_0-\frac{a_1}{a_2}\right),$$

где $a_0=b_2k^4$, $a_1=-b_4k$, $a_2=-b_1k$, k – волновое число.

Вычислительные эксперименты

Подставив в уравнение (1) решение вида $\psi=A\exp(ikx-\omega t)$, получим следующее дисперсионное уравнение:

$$\omega(b_1k+i)+b_2k^4-b_3k^2+b_4ik=0,$$

где $\omega=\omega_r+i\omega_i$ (ω_r – частота, ω_i – инкремент).

Для расчета фазовой скорости используется следующая формула:

$$c_r=\frac{\omega_r}{k}$$

В ходе вычислительных экспериментов исследованы волновые характеристики течения тонкого слоя вязкой жидкости для значений числа Рейнольдса $1 \le Re \le 15$ и волновых чисел $0 \le k \le 0{,}5$. Рассмотрено течение тонких пленок воды, молока и спирта. Параметр поверхностного натяжения воды связан со значением числа Рейнольдса выражением $\sigma_{H_2O}=4887Re^{-\frac{5}{3}}$, параметры поверхностного натяжения молока и спирта приняты примерно равными $0{,}6\sigma_{H_2O}$ и $0{,}33\sigma_{H_2O}$ соответственно.

Для расчета значений волновых характеристик на языке Python реализован следующий алгоритм:

Шаг 1. Рассчитать значения коэффициентов b_1, b_2, b_3 и b_4.

Шаг 2. Для каждого значения $k\in[0;0{,}5]$ (значения выбраны с шагом $\Delta k=0{,}001$) вычислить значения инкремента ω_i и частоты ω_r.

Шаг 3. Используя рассчитанное значение частоты ω_r, вычислить значение фазовой скорости c_r.

Шаг 4. Используя значение волнового числа k, соответствующее максимальному значению инкремента $\omega_{i\,max}$, вычислить критическое значение числа Марангони $M_к$.

Течение жидкой пленки неустойчиво при положительных значениях инкремента. Как показано в таблице 1, при уменьшении значения параметра поверхностного натяжения жидкости наблюдается расширение области неустойчивости.

Результаты расчета значений инкремента и фазовой скорости для неустойчивых режимов течения при $Re = 10$ представлены на рисунках 1 и 2.

В областях неустойчивости жидких пленок величина инкремента достигает максимального значения. В таблице 2 представлены максимальные значения инкремента и соответствующие им значения фазовой скорости для различных значений числа Рейнольдса. Жидкостям с меньшим значением параметра поверхностного натяжения соответствуют большие максимальные значения инкремента и меньшие значения фазовой скорости.

Критические значения числа Марангони, рассчитанные для режима максимального значения инкремента приведены в таблице 3.

Заключение

Для математической модели состояния свободной поверхности при умеренных зна-

Таблица 1

Области неустойчивости

Re	k			Re	k		
	Вода	Молоко	Спирт		Вода	Молоко	Спирт
1	[0; 0,028]	[0; 0,036]	[0; 0,048]	9	[0; 0,17]	[0; 0,219]	[0; 0,294]
2	[0; 0,049]	[0; 0,063]	[0; 0,084]	10	[0; 0,185]	[0; 0,239]	[0; 0,321]
3	[0; 0,068]	[0; 0,088]	[0; 0,118]	11	[0; 0,201]	[0; 0,259]	[0; 0,347]
4	[0; 0,087]	[0; 0,112]	[0; 0,15]	12	[0; 0,216]	[0; 0,278]	[0; 0,373]
5	[0; 0,104]	[0; 0,134]	[0; 0,18]	13	[0; 0,231]	[0; 0,298]	[0; 0,399]
6	[0; 0,121]	[0; 0,156]	[0; 0,21]	14	[0; 0,245]	[0; 0,316]	[0; 0,424]
7	[0; 0,138]	[0; 0,178]	[0; 0,238]	15	[0; 0,26]	[0; 0,335]	[0; 0,45]
8	[0; 0,154]	[0; 0,199]	[0; 0,266]				

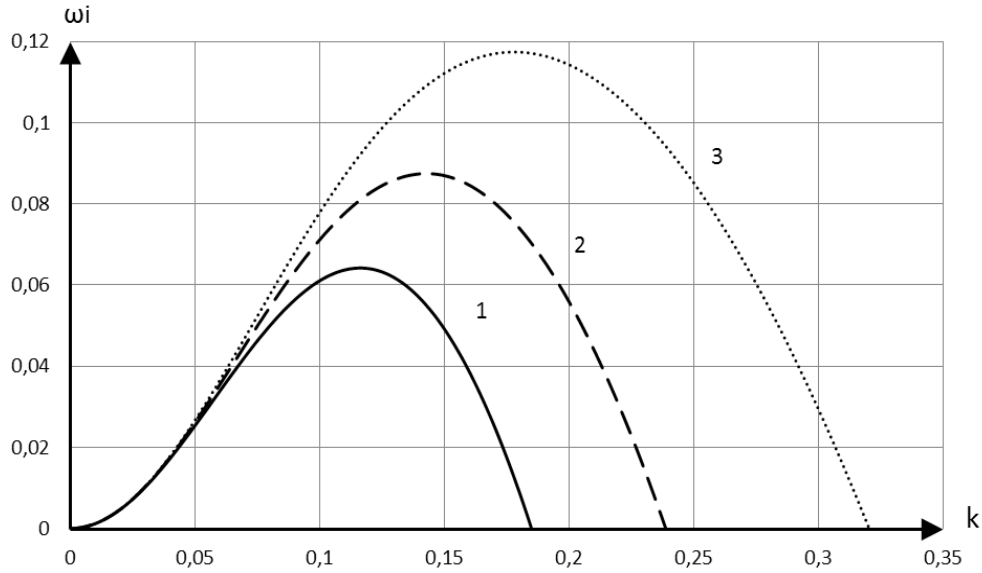

Рис. 1. Инкремент для Re = 10: 1 – вода, 2 – молоко, 3 – спирт

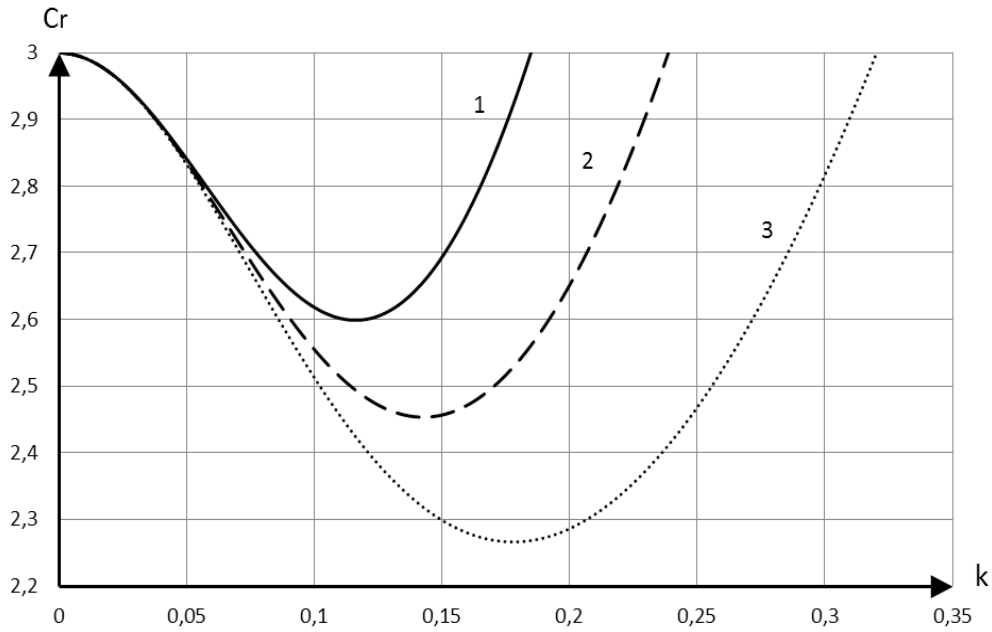

Рис. 2. Фазовая скорость для Re = 10: 1 – вода, 2 – молоко, 3 – спирт

Таблица 2

Инкремент и фазовая скорость

Re	Вода			Молоко			Спирт		
	k	$\omega_{i\,max}$	c_r	*k*	$\omega_{i\,max}$	c_r	*k*	$\omega_{i\,max}$	c_r
1	0,019	0,00022	2,99986	0,025	0,00037	2,99977	0,033	0,00066	2,99959
2	0,034	0,0014	2,99825	0,044	0,00233	2,99709	0,059	0,00419	2,99477
3	0,048	0,0041	2,99231	0,062	0,0068	2,98724	0,083	0,01212	2,97728
4	0,061	0,00871	2,97823	0,078	0,0143	2,96425	0,104	0,02502	2,93746
5	0,072	0,01535	2,95202	0,093	0,02479	2,92252	0,123	0,04206	2,86858
6	0,083	0,02387	2,91047	0,106	0,03759	2,85904	0,139	0,06112	2,77081
7	0,093	0,03376	2,85229	0,118	0,05146	2,77486	0,152	0,07964	2,65156
8	0,102	0,04428	2,7786	0,128	0,06505	2,67475	0,163	0,09568	2,52159
9	0,11	0,05464	2,69266	0,136	0,07726	2,5654	0,171	0,1083	2,39082
10	0,116	0,06419	2,59883	0,143	0,08747	2,45333	0,178	0,11743	2,26606
11	0,122	0,0725	2,50155	0,148	0,09548	2,34361	0,183	0,12348	2,15108
12	0,126	0,07939	2,4046	0,153	0,10139	2,23954	0,187	0,12702	2,04734
13	0,13	0,08484	2,31069	0,156	0,10548	2,14296	0,19	0,12863	1,9549
14	0,133	0,08895	2,22165	0,159	0,10805	2,05459	0,193	0,12879	1,87309
15	0,136	0,0919	2,13841	0,162	0,10939	1,97446	0,195	0,1279	1,8009

Таблица 3

Критические значения числа Марангони

Re	Вода		Молоко		Спирт	
	k	$M_к$	*k*	$M_к$	*k*	$M_к$
1	0,019	26595,3	0,025	15362,6	0,033	8818,0
2	0,034	2078,7	0,044	1242,2	0,059	692,0
3	0,048	465,5	0,062	280,0	0,083	157,4
4	0,061	163,8	0,078	101,1	0,104	58,0
5	0,072	76,6	0,093	46,9	0,123	27,9
6	0,083	41,2	0,106	26,2	0,139	16,2
7	0,093	25,1	0,118	16,5	0,152	10,8
8	0,102	16,8	0,128	11,5	0,163	7,9
9	0,11	12,2	0,136	8,7	0,171	6,2
10	0,116	9,4	0,143	6,9	0,178	5,1
11	0,122	7,6	0,148	5,8	0,183	4,4
12	0,126	6,4	0,153	4,9	0,187	3,9
13	0,13	5,5	0,156	4,3	0,19	3,5
14	0,133	4,8	0,159	3,9	0,193	3,2
15	0,136	4,3	0,162	3,5	0,195	2,9

чениях числа Рейнольдса разработаны вычислительный алгоритм и программа расчета волновых характеристик пленочного течения на языке Python. Представленный алгоритм позволяет рассчитать инкремент, частоту и фазовую скорость и определить границы области неустойчивости пленочного течения, тем самым обеспечивая надежность техноло-

гических процессов, применяемых в пленочных аппаратах.

При снижении параметра поверхностного натяжения жидкости течение жидкой пленки становится более неустойчивым: наблюдается расширение области неустойчивости. Кроме того, наблюдается уменьшение критического значения числа Марангони, что может привести к разрыву жидкой пленки и аварийной ситуации в работе пленочного аппарата.

Результаты работы могут быть использованы при разработке технологических процессов в жидких пленках и проектировании промышленных пленочных аппаратов.

Литература

1. Волновое течение пленок жидкости / С.В. Алексеенко, В.Е. Накоряков, Б.Г. Покусаев. Новосибирск: Наука, 1992. 256 с.

2. Теплообмен в жидкостных пленках / Е.Г. Воронцов, Ю.М. Тананайко. Киев: Техника, 1972. 196 с.

3. Гидродинамика и тепломассообмен с поверхностью раздела / Л.П. Холпанов, В.Я. Шкадов. М.: Наука, 1990. 271 с.

4. Бурмистрова О.А. Устойчивость вертикальной пленки жидкости с учетом эффекта Марангони и теплообмена с окружающей средой // Прикладная механика и техническая физика. 2014. № 3 (55). С. 17–25.

5. Subramaniam V., Garimella S. Numerical Study of Heat and Mass Transfer in Lithium Bromide-Water Falling Films and Droplets // International Journal of Refrigeration, 2014, vol. 40, pp. 211–226.

6. Rahimzadeh A., Ahmadian-Yazdi M.-R., Eslamian M. Experimental Study on the Characteristics of Capillary Surface Waves on a Liquid Film on an Ultrasonically Vibrated Substrate // Fluid Dynamics Research, 2018, vol. 50, no. 6.

7. D'Alessio S.J.D., Seth C.J.M.P., Pascal J.P. The Effects of Variable Fluid Properties on Thin Film Stability // Physics of Fluids, 2014, vol. 26, no. 12.

8. Прокудина Л.А. Влияние неоднородности поверхностного натяжения на волновое течение жидкой пленки // Инженерно-физический журнал. 2014. № 1 (87). С. 158–166.

9. Прокудина Л.А., Вяткин Г.П. Неустойчивость неизотермической жидкой пленки // Доклады Академии наук. 1998. № 6 (362). С. 770–772.

References

1. Alekseenko S.V., Nakoryakov V.E., Pokusaev B.G. Volnovoe techenie plenok zhidkosti [Wave Flow of Liquid Films]. Novosibirsk, Nauka, 1992. 256 p.

2. Vorontsov E.G., Tananayko Yu.M. Teploobmen v zhidkostnykh plenkakh [Heat and Mass Transfer in Liquid Films]. Kiev, Tehnika, 1972. 196 p.

3. Kholpanov L.P., Shkadov V.Ya. Gidrodinamika i teplomassoobmen s poverkhnost'yu razdela [Hydrodynamics and Heat and Mass Transfer with Interface]. Moscow, Nauka, 1990. 271 p.

4. Burmistrova O.A. Ustoychivost' vertikal'noy plenki zhidkosti s uchetom effekta Marangoni i teploobmena s okruzhayushchey sredoy // Pri-kladnaya mekhanika i tekhnicheskaya fizika. 2014. № 3 (55). S. 17–25.

5. Subramaniam V., Garimella S. Numerical Study of Heat and Mass Transfer in Lithium Bromide-Water Falling Films and Droplets // International Journal of Refrigeration, 2014, vol. 40, pp. 211–226.

6. Rahimzadeh A., Ahmadian-Yazdi M.-R., Eslamian M. Experimental Study on the Characteristics of Capillary Surface Waves on a Liquid Film on an Ultrasonically Vibrated Substrate // Fluid Dynamics Research, 2018, vol. 50, no. 6.

7. D'Alessio S.J.D., Seth C.J.M.P., Pascal J.P. The Effects of Variable Fluid Properties on Thin Film Stability // Physics of Fluids, 2014, vol. 26, no. 12.

8. Prokudina L.A. Vliyaniye neodnorodnosti poverkhnostnogo natyazhe-niya na volnovoye techeniye zhidkoy plenki // Inzhenerno-fizicheskiy zhurnal. 2014. № 1 (87). S. 158–166.

9. Prokudina L.A., Vyatkin G.P. Neustoychivost' neizotermicheskoy zhidkoy plenki // Doklady Akademii nauk. 1998. № 6 (362). S. 770–772.

ПРОКУДИНА Людмила Александровна, доктор физико-математических наук, доцент, профессор кафедры вычислительной математики и высокопроизводительных вычислений, Южно-Уральский государственный университет (национальный исследовательский университет). 454080, г. Челябинск, пр. Ленина, 76. E-mail: prokudinala@susu.ru

ВИХИРЕВ Михаил Павлович, аспирант кафедры вычислительной математики и высокопроизводительных вычислений, Южно-Уральский государственный университет (национальный исследовательский университет). 454080, г. Челябинск, пр. Ленина, 76. E-mail: vikhirevmp@susu.ru

PROKUDINA Liudmila Alexandrovna, Doctor of Physical and Mathematical sciences, Associate Professor, Professor of the Department of Computational mathematics and high-performance computing, South Ural State University (National Research University). 454080, Chelyabinsk, Lenina Avenue, 76, E-mail: prokudinala@susu.ru

VIKHIREV Mikhail Pavlovich, postgraduate student of the Department of Computational mathematics and high-performance computing, South Ural State University (National Research University). 454080, Chelyabinsk, Lenina Avenue, 76, E-mail: vikhirevmp@susu.ru

УДК 519.66 + 621.036 Вестник УрФО № 4(42) / 2021, с. 71–79

Гаврилова Т. П.

DOI: 10.14529/secur210408

МЕТОД ОПРЕДЕЛЕНИЯ НЕСТАЦИОНАРНЫХ ТЕМПЕРАТУРНЫХ ПОЛЕЙ ПО РЕЗУЛЬТАТАМ ГРАНИЧНЫХ ИЗМЕРЕНИЙ

При анализе технологических процессов и техническом обслуживании систем широко применяются методы теплового неразрушающего контроля. Применение математических методов для обработки результатов температурных измерений при тепловом мониторинге работающих систем и механизмов позволяет повысить надежность и долговечность оборудования, обеспечить безотказность процесса эксплуатации, оптимизировать параметры теплового воздействия на систему.

В статье рассмотрен метод решения задачи определения внутренних нестационарных температурных полей объекта, сформированных под влиянием внешнего теплового воздействия. Математически процесс теплопереноса представлен обратной граничной задачей с заданным дополнительным условием, сформированным на основании характеристик теплового воздействия на объект. В работе приведен численный метод решения задачи, основанный на применении явной конечно-разностной схемы. Точность и устойчивость представленного метода подтверждена результатами имитационного моделирования.

Ключевые слова: *теплоперенос, обратная задача, численный метод, конечно-разностная схема, целостность информации, вычислительный эксперимент.*

Gavrilova T.P.

METHOD FOR DETERMINING UNSTATIONARY TEMPERATURE FIELDS FROM THE RESULTS OF BOUNDARY MEASUREMENTS

Thermal non-destructive testing is widely used in the analysis of technological processes and maintenance of systems. The application of mathematical methods for processing the results of temperature measurements during thermal monitoring of operating systems and mechanisms allows to increase the reliability and durability of equipment, ensure the safety of the operation process, optimize the parameters of thermal effects on the system.

The article is devoted to a method for determining internal non-stationary temperature fields that are formed in object under the external thermal action. The heat transfer issue is represented by an inverse problem for a one-dimensional parabolic equation with initial, boundary condition and additional condition that are formed on the basis of the characteristics of the thermal effect on the object. We propose a numerical method for solving the inverse problem. This method based on the use of finite-difference inexplicit scheme. The accuracy and stability of the proposed method are confirmed by computational experiment results.

Keywords: *heat transfer, inverse problem, numerical method, finite-difference scheme, information integrity, computational experiment.*

Введение

При диагностике работающих систем и механизмов, управлении технологическими процессами термообработки необходимо гарантировать достоверность информации о внутреннем тепловом состоянии объектов. Источником этой информации являются температурные измерения, полученные вблизи поверхности объекта, подвергающегося термообработке. Актуальной является задача разработки высокоточных и устойчивых методов для определения температуры во внутренних точках объекта по результатам измерений вблизи поверхности.

Математическая модель процесса теплопередачи, при котором информация о внутреннем тепловом состоянии объекта определяется по результатам граничных измерений, относится к классу обратных задач. Теория обратных задач теплопереноса, ориентированная на восстановление неизвестных характеристик нестационарных тепловых процессов, является актуальным направлением современной теплометрии. Обратные задачи, связанные с процессами теплопереноса, изучали О. М. Алифанов, Е. А. Артюхин, Л. А. Коздоба, Ю. М. Мацевитый и многие другие исследователи [1–8]. Среди последних работ следует отметить труды А.Н. Дилигенской [9,10], Н.М. Япаровой [11].

Основной проблемой обратных граничных задач является зависимость погрешности численных решений от уровня шума в исходных данных. Применение к их решению классических методов приводит к неустойчивым решениям и искажению информации о тепловом состоянии объекта. Для повышения устойчивости решений такого рода задач используют методы регуляризации.

Разработке численных методов на основе регуляризующих алгоритмов для обратных задач, представленных интегральными уравнениями, посвящены работы А.С. Апарцина, В.В. Васина, А.Н. Тихонова, А.Г. Яголы и других исследователей [12–16].

Численные методы, основанные на явных и неявных конечно-разностных схемах, представлены в работах П. Н. Вабищевича, А.А. Самарского, В. И. Васильева [17–21].

Постановка задачи

В статье рассматривается задача определения нестационарных температурных полей во внутренних точках изотропного тела, не содержащем внутренних источников тепла. Формирование полей происходит под влиянием внешнего теплового воздействия. Контролировать влияние этого воздействия можно путем проведения тепловых измерений вблизи поверхности. Согласно требованиям к процессу теплопереноса, тепловые потоки в теле распределяются равномерно во всех направлениях и зависят только от времени, также не допускаются резкие скачки температурных градиентов и существенные изменения основных характеристик материала объекта.

Для построения математической модели процесса теплопереноса введем следующие обозначения. Пусть функция $u(x,t)$ определяет значение температуры в точке x в текущий момент времени t. К началу процесса температура во всех точках объекта была одинаковой, поэтому полагаем, что $u(x, 0)=C$. По измеренным температурам на поверхности формируем функцию $\varphi(t)=u(0,t)$, а по измерениям температуры в точке наблюдения x_0 формируем функцию $g(t)=u(x_0,t)$. Пусть L – расстояние от поверхности тела до контрольной точки. Требуется определить температуру тела $u(L,t)=\psi(t)$ в контрольной точке A, а также вдоль линии, соединяющей точки O и A. Рис. 1.

Рассматриваемый процесс теплопереноса математически представим параболическим уравнением

$$u_t=au_{xx},\ x\in(0, L),\ t\geq 0, \qquad (1)$$

граничным и дополнительным условиями

$$u(0,t)=\varphi(t),\ u(0,t)=\varphi(t),\ u(x_0,t)=g(t),\ t\geq 0 \qquad (2)$$

а также начальным условием

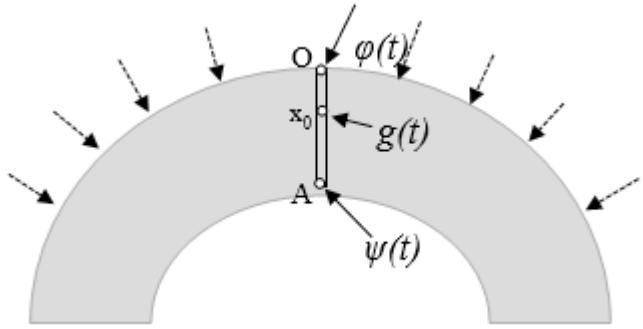

Рис. 1. Измерение температурных функций.
Температурные датчики расположены в точках O и x_0

$$u(x,0) = C,\ x \in [0, L]. \quad (3)$$

В задаче требуется найти значение температуры в контрольной точке

$$u(L,t) = \psi(t) \quad (4)$$

и, на основе полученной функции ψ(*t*), спрогнозировать температурные значения во внутренних точках OA.

При формировании граничного и дополнительного условий задачи необходимо учитывать шум, который неизбежно возникает в результатах измерений. Ситуацию наличия шума представим следующим образом: пусть φ_0 и g_0 – истинные значения температуры на поверхности объекта и в точке наблюдения. Зашумленные значения в рассматриваемых точках обозначим как $\phi\delta$ и $g\delta$ соответственно, при этом $\max\left\{\left\|\varphi_\delta - \varphi_0\right\|, \left\|g_\delta - g_0\right\|\right\} \le \delta$, где δ >0 уровень погрешности датчиков измерений, определяющий допустимый уровень отклонения данных. Таким образом, учитывая погрешности в исходных данных, неизбежно возникающие при измерениях, получаем задачу определения температурной функции $u_\delta(L,t)$ по зашумленным значениям δ, ϕ_δ, g_δ и определения температуры $u_\delta(x,t)$ во внутренних точках объекта.

Вычислительная схема метода

Для построения вычислительной схемы используем идею, предложенную в работе П.Н. Вабищевича [22]. Следуя этому подходу, построим неявную конечно-разностную схему.

Для решения задачи (1)–(4) введем в области [0, L]×[0, T] сетку с равноотстоящими узлами (x_i, t_j):

$$x_i = ih_x,\ \ i = 0,\ldots,N\ ,\ \ N = {}^{L}\!/_{h_x};$$

$$t_j = jh_t,\ \ j = 0,\ldots,M\ ,\ \ M = {}^{T}\!/_{h_t}.$$

Дискретные аналоги функции u(*x*,*t*):

$$u(x_i,t_j) = u_{i,j},$$
$$u(0,t_j) = \varphi_j,$$
$$u(x_i,0) = u_{i,0} = C,$$
$$u(L,t_j) = \psi_j.$$

Пусть точка x_0 является узлом сетки по пространственной переменной с номером *r*, т.е. $x_0 = rh_x$ тогда

$$u(x_0,t_j) = u_{r,j} = g_j. \quad (5)$$

Построим конечно-разностную аппроксимацию уравнения (1), используя четырехточечную разностную схему. При фиксированном j на каждом временном слое получаем следующее уравнение:

$$\frac{u_{i,j+1} - u_{i,j}}{h_t} - a\frac{u_{i+1,j+1} - 2u_{i,j+1} + u_{i-1,j+1}}{h_x^2} = 0$$

, $i = 1, ..., N-1$. (6)

Для проведения дальнейших вычислений введем следующие обозначения:

$$A_i = C_i = \frac{a}{h_x^2},\ B_i = -\frac{2a}{h_x^2} - \frac{1}{h_t},\ D_i = -\frac{1}{h_t}u_{i,j}, \quad (7)$$
$$A_0 = C_0 = 0,\ B_0 = -\frac{1}{h_t}\cdot\frac{\varphi_j}{\varphi_{j+1}},\ D_0 = -\frac{1}{h_t}\cdot\varphi_j.$$

Тогда уравнение (6) примет вид: ,

$A_i u_{i-1,j+1} + B_i u_{i,j+1} + C_i u_{i+1,j+1} = D_i$, $i = 1, ..., N-1$. (8)

Согласно подходу, предложенному в [22], сеточный аналог температурной функции *u*(*x*,*t*) в точке (x_i,t_j) представим с помощью двух вспомогательных функций в следующем виде

$$u_{i,j} = y_{i,j} + \psi_j \cdot z_{i,j},\ i = 0,\ldots,N. \quad (9)$$

При i=N имеем $u_{N,j} = y_{N,j} + \psi_j \cdot z_{N,j}$ и, учитывая граничное условие $u_{N,j} = \psi_j$ задачи (1) – (4), получим $y_{N,j} = 0$ и $z_{N,j} = 1$.

Подставляя выражение (9) в уравнение (8), получим следующие системы уравнений:

$$\begin{cases} A_i y_{i-1,j+1} + B_i y_{i,j+1} + C_i y_{i+1,j+1} = D_i,\ i = 1,\ldots,N-1, \\ y_{N,j} = 0 \end{cases}$$

$$\begin{cases} A_i z_{i-1,j+1} + B_i z_{i,j+1} + C_i z_{i+1,j+1} = 0,\ i = 1,\ldots,N-1, \\ z_{N,j} = 1 \end{cases}$$

где коэффициенты A_i, B_i, C_i, D_i определяются по формулам (7).

Решая системы методом прогонки, определим на каждом j-ом временном слое сеточные функции $y_{i,j}$ и $z_{i,j}$ при i = 1, ..., *N*.

Подставляя в равенство (9) конечно-разностный аналог дополнительного условия в узле сетки x_r, получим равенство $g_j = y_{r,j} + \psi_j \cdot z_{r,j}$. Тогда дискретный аналог температурной функции в контрольной точке определим по формуле $\psi_j = \frac{g_j - y_{r,j}}{z_{r,j}}$ для всех j = 1, ..., *M* при условии $z_{r,j} \neq 0$.

Вычислительный эксперимент

С целью проверки точности и устойчивости предложенного способа определения температурных полей объекта из поверхностных измерений проводился вычислительный эксперимент. В ходе эксперимента выполнялось сопоставление решений, полученных с помощью предложенного алгоритма с тестовыми значениями, сформированными на основе имитационного моделирования, а также были получены экспериментальные оценки погрешностей метода. Вычислительный эксперимент включает следующие этапы:

1. Формируем тестовые значения функции *u(x,t)* в узлах сетки (x_r, t_j) следующим образом. Полагая, что температурная функция ψ(t) нам известна, с помощью неявной схемы находим численное решение прямой задачи:

$$u_t = a u_{xx},\ x \in (0, L),\ t \geq 0, \quad (10)$$

$$u(0,t) = \varphi(t),\ u(L,t) = \psi(t),\ t \geq 0, \quad (11)$$

$$u(x,0) = C,\ x \in [0,L]. \quad (12)$$

Далее моделируем дискретный аналог функции *g(t)* по формуле $g(t_i) = u(x_r, t_i)$.

Затем вводим аддитивный шум в исходные данные:

$$g_\delta(t_i) = g(t_i) + \eta_\delta(t_i),\ \varphi_\delta(t_j) = \varphi(t_j) + \mu_\delta(t_j),$$

где $\eta_\delta(t_i)$ и $\mu_\delta(t_j)$ являются случайными величинами, распределенными равномерно на $[-\delta, \delta]$.

2. На основании предложенного численного метода находим решение $\psi_\delta(t)$ задачи (1)–(3).

3. Оцениваем погрешности определения температуры с помощью функции $\Delta(t) = |\psi_\delta(t) - \psi(t)|$, величин $\Delta_\psi = \max_{t \in [0,T]} \Delta(t)$ и $\varepsilon_\psi = \frac{\Delta_\psi}{\max_{t \in [0,T]} |\psi(t)|}$.

4. Определяем тепловое поле $u_\delta(x, t)$ во внутренних точках линейного объекта на основе температурной функции $\psi_\delta(t)$.

Вычислительный эксперимент проводился при следующих данных: *L*=1, *T*=6000 *c*, *u(x,* 0)=50°C, коэффициент температуропроводности *a*=1.

Результаты вычислительного эксперимента

Приведем результаты вычислительного эксперимента, проведенного для следующих типов тестовых функций:

1) смоделируем распределение температуры в точке на поверхности объекта функцией $\varphi_1(t) = 45 + 1500t(e^{-t} - e^{-3})$, а в контрольной точке – функцией $\psi_1(t) = 45 + 750te^{-t}$;

2) смоделируем распределение температуры в точке на поверхности объекта функцией $\varphi_2(t) = 60 + 2400e^{-2,3t} \ln(t+1)$, а в контрольной точке – функцией $\psi_2(t) = 60 + 2000e^{-2,8t} \ln(t+1)$.

В таблице 1 представлены результаты имитационного моделирования для первой модели теплопереноса с тестовыми функциями $\varphi_1(t)$ и $\psi_1(t)$ при различных способах выбора точки наблюдения x_0 и различных значениях погрешностей исходных данных.

Таблица 1

Экспериментальные оценки погрешностей температурных функций для модели 1

Точка наблюдения	Погрешность исходных данных, δ	Погрешности вычислений	
		Δ_ψ	θ_ψ
$x_0 = 0{,}1 \cdot L$	0,01	118,2624	0,3685
	0,05	397,5678	1,2389
	0,1	560,5154	1,7466
$x_0 = 0{,}8 \cdot L$	0,01	23,1770	0,0722
	0,05	29,1066	0,0907
	0,1	35,6682	0,1111

Функция погрешности определения температуры для модели 1 представлена на рис. 2. Графики тестовой функции $\psi_1(t)$ и численного решения $\psi_\delta(t)$ задачи (1)–(4) приведены на рис. 3. На рис. 4 построены поверхности $u_\delta(x, t)$, соответствующие температурным полям в линейном объекте, сформированным на основе численных решений $\psi_\delta(t)$.

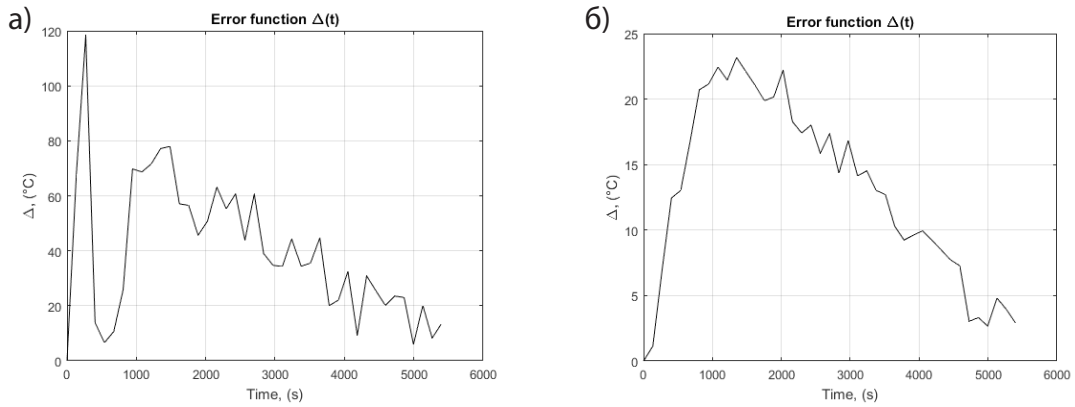

Рис. 2. График функции температурной погрешности $\Delta(t)$ для модели 1
а) при x_0=0,1·L, б) при x_0=0,8·L

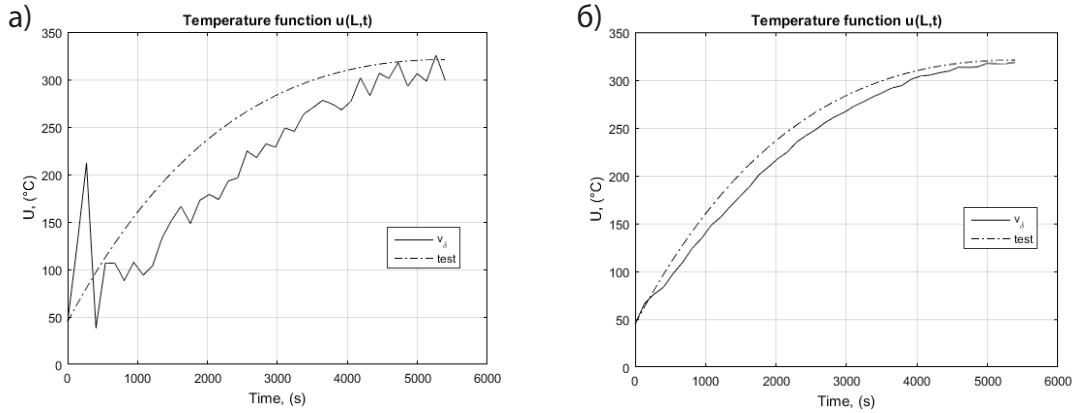

Рис. 3. График численного решения и тестовой функции для модели 1
а) при x_0=0,1·L, б) при x_0=0,8·L

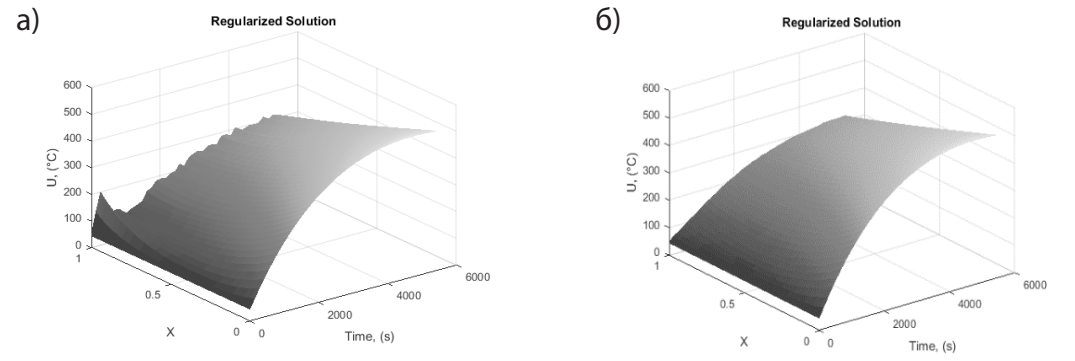

Рис. 4. Распределение температуры внутри линейного объекта для модели 1
а) результаты моделирования температуры при x_0=0,1·L,
б) результаты моделирования температуры при x_0=0,8·L

В таблице 2 представлены результаты вычислительного эксперимента для второй модели теплопереноса с тестовыми функциями $\varphi_2(t)$ и $\psi_2(t)$ при различных способах выбора точки наблюдения x_0.

Функция погрешности определения

Таблица 2

Экспериментальные оценки погрешностей температурных функций для модели 2

Точка наблюдения	Погрешность исходных данных, δ	Погрешности вычислений	
		Δ_ψ	θ_ψ
x_0=0,1·L	0,01	98,5247	0,3438
	0,05	544,2646	1,8994
	0,1	861,1171	3,0051
x_0=0,8·L	0,01	28,7040	0,0935
	0,05	38,9877	0,1360
	0,1	57,7243	0,2014

температуры для модели 2 представлена на рис. 5. Графики тестовой функции $\psi_2(t)$ и численного решения $\psi_\delta(t)$ задачи (1)–(4) приведены на рис. 6. На рис. 7 построены поверхности $u_\delta(x, t)$, соответствующие температурным полям в линейном объекте, определенным на основе граничных условий, найденных предложенным численным методом.

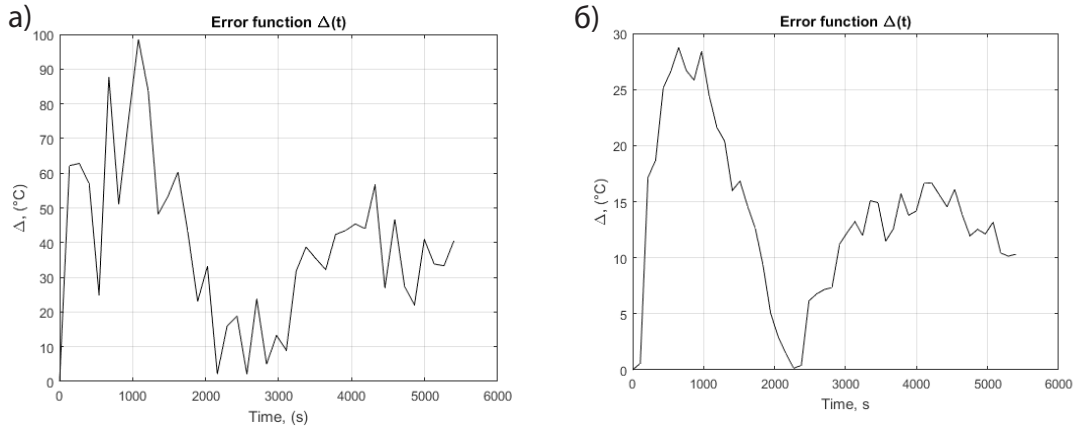

Рис. 5. График функции температурной погрешности $\Delta(t)$ для модели 2
а) при x_0=0,1·L, б) при x_0=0,8·L

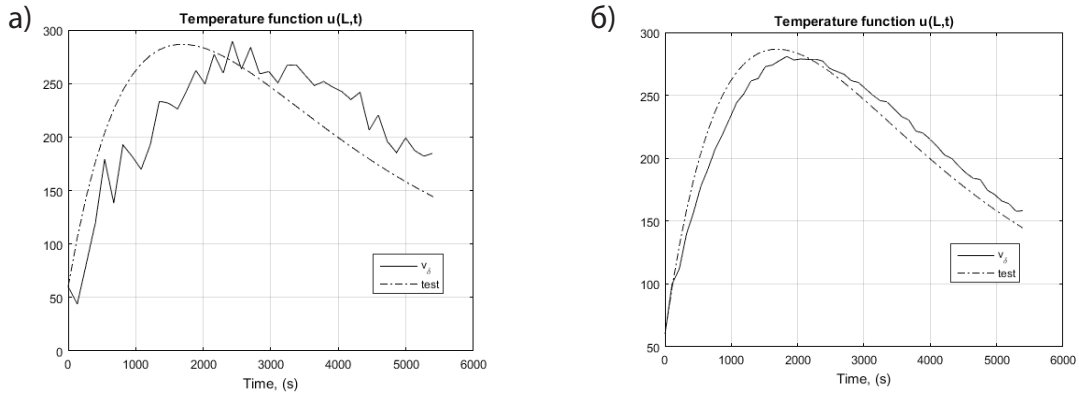

Рис. 6. График численного решения и тестовой функции для модели 2
а) при x_0=0,1·L, б) при x_0=0,8·L

Результаты вычислительного эксперимента свидетельствуют об устойчивости предложенного метода определения температуры. В ходе эксперимента выявлена зави-

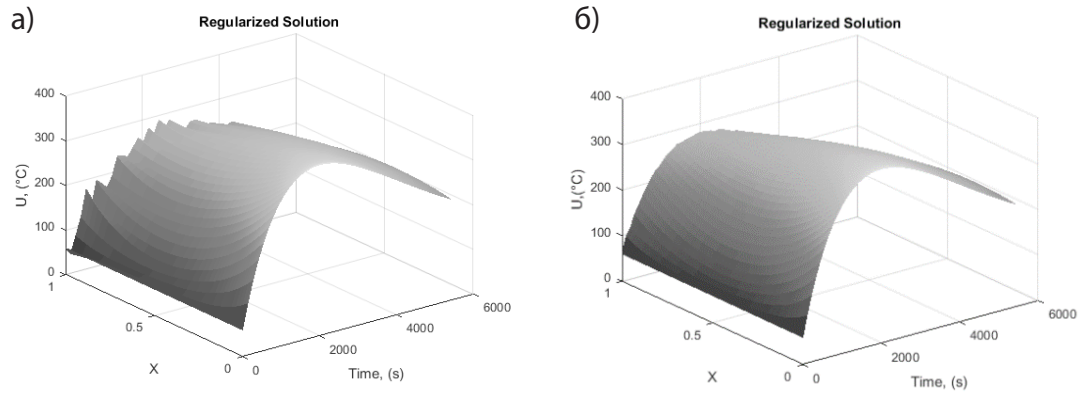

Рис. 7. Распределение температуры внутри линейного объекта для модели 2
а) результаты моделирования температуры при x_0=0,1·L,
б) результаты моделирования температуры при x_0=0,8·L

симость оценки погрешностей температурных функций от выбора точки x_0. Погрешность уменьшается при удалении точки наблюдения от граничной точки O линейного объекта.

Заключение

В данной работе рассмотрена задача определения внутренних нестационарных температурных полей, сформированных под влиянием внешнего теплового воздействия. Учитывая характеристики теплопереноса, математически процесс представлен обратной граничной задачей для параболического уравнения с известными вблизи границы температурными функциями. Численное решение задачи определения температуры основано на применении неявной конечно-разностной схемы.

Результаты вычислительного эксперимента, проведенного на основе имитационного моделирования, подтверждают точность и устойчивость представленного численного метода определения температурных функций внутри объекта. В ходе эксперимента выявлено уменьшение отклонений температурных значений, полученных численным методом от тестовых при удалении точки наблюдения от поверхности объекта.

Литература

1. Алифанов О. М. Идентификация процессов теплообмена летательных аппаратов. – М.: Машиностроение, 1979. – 216 с.

2. Мацевитый Ю. М. Параметрическая и функциональная идентификация тепловых процессов // Мацевитый Ю. М., Гайшун И. В., Борухов В. Т., Костиков А. О. Пробл. машиностроения, 2011. Т. 14, № 3. С. 40–47.

3. Алифанов О. М., Артюхин Е. А., Румянцев С. В. Экстремальные мето

ды решения некорректных задач и их приложения к обратным задачам теплообмена. – М.: Наука. Физматлит, 1988. – 285 с.

4. Коздоба Л. А., Круковский П. Г. Методы решения обратных задач теп

лопереноса. – Киев: Наук. думка, 1982. – 358 с.

5. Мацевитый Ю. М. Обратные задачи теплопроводности. Методология. – Киев: Наук. думка, 2002. – 408 с.

6. Алифанов О.М. Обратные задачи теплообмена / Алифанов О.М. – М: Машиностроение, 1988. – 280 с.

7. Бек Д. Некорректные обратные задачи теплопроводности / Бек Д., Блакуэлл Б., Сент-Клер Ч,мл.– М.: Мир, 1989. – 312 с.

8. Ватульян А. О. Обратные задачи в механике деформируемого твердого тела. / Ватульян А. О. – М.: Физматлит, 2007. – 224 с.

9. Дилигенская А. Н. Метод минимаксной оптимизации в двумерной граничной обратной задаче теплопроводности // Дилигенская А. Н. ТВТ, 2019. Т. 57, выпуск 2. С. 226–233.

10. Дилигенская А. Н. Решение граничных обратных задач теплопроводности на основе методов оптимизации // Дилигенская А. Н. № 3(191), 2016. С. 46–50.

11. Yaparova N. Numerical Methods for Solving a Boundary Value Inverse Heat Conduction Problem / Yaparova N. // Inverse Problems in Science and Engineering. – 2014. – Vol.22, no 5. – P. 832–847.

12. Апарцин А.С. Приближенное решение интегральных уравнений Вольтерра I рода методом квадратурных сумм / Апарцин А.С., Бакушинский А.Б. // Дифференциальные и интегральные уравнения. – Иркутск: Иркут. гос. ун-т. – 1972. – Вып. 1. – С. 248–258.

13. Васин В.В. Регулярный алгоритм аппроксимации негладких решений для интегральных уравнений Фредгольма первого рода / Васин В.В. // Выч. технологии. – 2010. – Т.15, № 2. – С. 15–23.

14. Лаврентьев М.М. Некорректные задачи математической физики и анализа / Лаврентьев М.М., Романов В.Г., Шишатский С.П. – М.: Наука, 1980. – 286 с.

15. Cialkowski M. A sequential and global method of solving an inverse problem of heat conduction equation / Cialkowski M., Grysa K. // Journal of Theoretical and Applied Mechanics. – 2010. – Vol. 48, no. 1. – P. 111–134.

16. Кумицкий Б.М. Математическое моделирование тепловых процессов в условиях промерзания (оттаивания) влажного грунта / Кумицкий Б.М., Саврасова Н.А., Седаев А.А. // Научный журнал строительства и архитектуры. – Воронеж: Воронежский гос. тех. ун-т. – 2018. № 3(51). – С. 31–39.

17. Слепян Л.И. Интегральные преобразования в нестационарных задачах механики / Слепян Л.И., Яковлев Ю.С. – Л.: Судостроение, 1980. – 343 с.

18. Тихонов А.Н. Численные методы решения некорректных задач / Тихонов А.Н., Гончарский А.В., Степанов В.В., Ягола А.Г. – М.: Наука, 1990. – 232 с.

19. Криксин Ю.А. Обратная задача восстановления источника для уравнения конвективной диффузии / Криксин Ю.А., Плющев С.Н., Самарская Е.А., Тишкин В. Ф. // Математическое моделирование –1995, 7:11. – С. 95–108.

20. Самарский А.А. Численные методы решения обратных задач / Самарский А.А., Вабищевич П.Н. – М.: ЛКИ, 2007. – 480 с.

21. Borukhov V. T. Numerical solving an inverse problem of source reconstruction for a parabolic equation // Borukhov V. T., Vabishchevich P. N., Matem. Mod., –1998, 10:11. – P. 93–100.

22. Вабищевич П. Н. Вычислительная идентификация правой части параболического уравнения / Вабищевич П. Н., Васильев В. И., Васильева М. В. // Ж. вычисл. матем. и матем. физ. – 2015, 55:6. – С. 1020–1027.

References

1. Alifanov O. M. Identifikacija processov teploobmena letatel'nyh apparatov [Identification of heat exchange processes of flying apparatuses]. Moskva: Mashinostroenie, 1979. – 216 p.

2. Matsevityi Yu. M., Gaishun I. V., Borukhov V. T., Kostikov A. O. Parametricheskaja i funkcional'naja identifikacija teplovyh processov [Parametric and functional identification of thermal processes]. Probl. mashinostroeniya, 2011, vol. 14, no. 3. – P. 40-47.

3. Alifanov O. M., Artyukhin E. A., Rumyantsev S. V. Jekstremal'nye meto dy reshenija nekorrektnyh zadach i ih prilozhenija k obratnym zadacham teploobmena [Extreme methods methods of solving ill-posed problems and their applications to inverse heat transfer problems]. - Moskva: Nauka. Fizmatlit, 1988. – 285 p.

4. Kozdoba L. A., Krukovsky P. G. Metody reshenija obratnyh zadach teploperenosa [Methods for solving inverse heat transfer problems]. - Kiyev: Nauk. dumka, 1982. – 358 p.

5. Matsevityi Yu. M. Obratnye zadachi teploprovodnosti. Metodologija [Inverse problems of thermal conductivity. Methodology]. Kiyev: Nauk. dumka, 2002. – 408 p.

6. Alifanov O. M. Obratnye zadachi teploobmena [Inverse problems of heat exchange]. Moskva: Mechanical Engineering, 1988. – 280 p.

7. Beck D., Blahwell B., St. Clair Ch, ml. Nekorrektnye obratnye zadachi teploprovodnosti [Incorrect inverse problems of thermal conductivity]. Moskva: Mir, 1989. – 312 p.

8. Vatulyan A. O. Obratnye zadachi v mehanike deformiruemogo tverdogo tela [Inverse problems in the mechanics of a deformable solid]. Moskva: Fizmatlit, 2007. – 224 p.

9. Diligenskaya A. N. Metod minimaksnoj optimizacii v dvumernoj granichnoj obratnoj zadache teploprovodnosti [Minimax optimization method in a two-dimensional partial inverse problem of thermal conductivity]. TVT, 2019, vol. 57, no. 2. – P. 226-233.

10. Diligenskaya A. N. Reshenie granichnyh obratnyh zadach teploprovodnosti na osnove metodov optimizacii [Solution of boundary inverse problems of thermal conductivity based on optimization methods]. No. 3(191), 2016. – P. 46–50.

11. Yaparova N. Numerical Methods for Solving a Boundary Value Inverse Heat Conduction Problem. Inverse Problems in Science and Engineering, 2014, vol.22, no. 5. – P. 832–847.

12. Apartsin A. S., Bakushinsky A. B. Approximate solution of Volterra integral equations of the first generation by the method of quadrature sums Differential and integral equations [Priblizhennoe reshenie integral'nyh uravnenij Vol'terra I roda metodom kvadraturnyh summ]. Irkutsk: Irkut. state un-t, 1972, vol. 1. – P. 248-258.

13. Vasin V. V. Regular algorithm for approximation of nonsmooth solutions for Fredholm integral equations of the first kind [Reguljarnyj algoritm approksimacii negladkih reshenij dlja integral'nyh uravnenij Fredgol'ma pervogo roda]. Vych.tekhnologii, 2010, vol. 15, no. 2. – P. 15–23.

14. Lavrentiev M. M., Romanov V. G., Shishatsky S. P. Nekorrektnye zadachi matematicheskoj fiziki i analiza [Incorrect problems of mathematical physics and analysis]. Moscow: Nauka, 1980. – 286 p.

15. Cialkowski M., Grysa K. A sequential and global method of solving an inverse problem of heat conduction equation. Journal of Theoretical and Applied Mechanics, 2010, vol. 48, no. 1. – P. 111–134.

16. Kumitsky B. M., Savrasova N. A., Sedaev A. A. Mathematical modeling of thermal processes in conditions of freezing (thawing) of wet soil [Matematicheskoe modelirovanie teplovyh processov v uslovijah promerzanija (ottaivanija) vlazhnogo grunta]. Scientific Journal of Construction and Architecture. Voronezh: Voronezh State Technical University. un-t, 2018, no. 3(51). – P. 31–39.

17. Slepyan, L. I., Yakovlev, Yu. P. Integral'nye preobrazovanija v nestacionarnyh zadachah mehaniki [Integral transforms in nonstationary problems mechanical key]. Leningrad: Shipbuilding, 1980. – 343 p.

18. Tikhonov A. N., Goncharsky A.V., Stepanov V. V., Yagola A. G. Chislennye metody reshenija nekorrektnyh zadach [Numerical methods for solving ill-posed problems]. Moskva: Nauka, 1990. – 232 p.

19. Kriksin Ju.A., Pljushhev S.N., Samarskaja E.A., Tishkin V. F. Inverse problem of reconstructing the source for the equation of convective diffusion. [Obratnaja zadacha vosstanovlenija istochnika dlja uravnenija konvektivnoj diffuzii] Mathematical modeling, 1995, 7:11. – P. 95–108.

20. Samarskiy A. A., Vabishchevich P. N. Chislennye metody reshenija obratnyh zadach [Numerical methods for solving inverse problems]. Moskva: LKI, 2007. – 480 p.

21. Borukhov V. T., Vabishchevich P. N. Numerical solving an inverse problem of source reconstruction for a parabolic equation. Matem. Mod., 1998, 10:11. – P. 93–100.

22. Vabishevich P. N., Vasiliev V. I., Vasilyeva M. V. Computational identification of the right part of the parabolic equation [Vychislitel'naja identifikacija pravoj chasti parabolicheskogo uravnenija]. Zh. matem. and math. phys. - 2015, 55:6. – P. 1020-1027.

ГАВРИЛОВА Татьяна Петровна, старший преподаватель кафедры вычислительной математики и высокопроизводительных вычислений ФГАОУ ВО «Южно-Уральский государственный университет (национальный исследовательский университет)». Россия, 454080 Челябинск, проспект Ленина, 76. E-mail: gavrilovatp@susu.ru.

GAVRILOVA Tatiana Petrovna, Senior Lecturer of Deparetment of Computational Mathematics and High-Performance Computing, Federal State Autonomous Educational Institution of Higher Education "South Ural State University (national research university)". 76, Lenin prospekt, Chelyabinsk, Russia, 454080. E-mail: gavrilovatp@susu.ru.

Материалы к публикации отправлять по адресу E-mail: urvest@mail.ru
в редакцию журнала «Вестник УрФО. Безопасность в информационной сфере».

Или по почте по адресу: Россия, 454080, г. Челябинск, пр. им. Ленина, д. 76,
ЮУрГУ, Издательский центр.

ВЕСТНИК УрФО
Безопасность в информационной сфере № 4(42) / 2021
Подписано в печать 30.06.2021.
Дата выхода в свет 12.07.2021. Формат 70×108 1/16. Печать цифровая.
Усл.-печ. л. 6.3. Тираж 100 экз. Заказ 213/248.
Цена свободная.

Отпечатано в типографии Издательского центра ЮУрГУ.
454080, г. Челябинск, пр. им. В. И. Ленина, 76.

Bulletin of the Ural Federal District
Security in the Sphere of Information No. 4(42) / 2021
Signed to print June 30, 2021.
Date of publication of the 12.07.2021. Format 70×108 1/16. Screen printing.
Conventional printed sheet 6.3. Circulation – 100 issues. Order 213/248. Open price.

Printed in the printing house of the Publishing Center of SUSU.
76, Lenina Str., Chelyabinsk, 454080