

Вестник НГУЭУ. 2024. № 1. С. 10–29  
Vestnik NSUEM. 2024. No. 1. P. 10–29

Научная статья  
УДК 004.052  
DOI: 10.34020/2073-6495-2024-1-010-029

**ОЦЕНКА НАДЕЖНОСТИ ВЫЧИСЛИТЕЛЬНЫХ СИСТЕМ  
НЕПАРАМЕТРИЧЕСКИМ МЕТОДОМ  
ПО МАЛЫМ ВЫБОРКАМ  
ЭКСПЛУАТАЦИОННЫХ ДАННЫХ**

**Никулин Владимир Сергеевич<sup>1</sup>, Пестунов Андрей Игоревич<sup>2</sup>**

<sup>1,2</sup> *Новосибирский государственный университет  
экономики и управления «НИНХ»*

<sup>1</sup> nikulin-94@inbox.ru

<sup>2</sup> kocheevaalina@gmail.com

**Аннотация.** Предложена новая методика оценки надежности небольших вычислительных систем, позволяющих сформировать только малые выборки эксплуатационных данных. Результатом применения методики является апостериорная плотность распределения отказов, на основе которой могут быть рассчитаны различные показатели надежности. Методика состоит из двух этапов: подготовка эксплуатационных данных, включающая обнаружение отказов методами машинного обучения, и построение плотности распределения отказов адаптированным методом Розенблатта–Парзена. Повышение эффективности оценок с помощью предлагаемой методики достигается за счет учета цензурированных данных, компенсации смещения плотностей распределения отказов и нахождения оптимального параметра сглаживания.

**Ключевые слова:** надежность вычислительных систем, эксплуатационные данные, машинное обучение, метод Розенблатта–Парзена, параметр сглаживания, плотность распределения отказов

**Для цитирования:** Никулин В.С., Пестунов А.И. Оценка надежности вычислительных систем непараметрическим методом по малым выборкам эксплуатационных данных // Вестник НГУЭУ. 2024. № 1. С.10–29. DOI: 10.34020/2073-6495-2024-1-010-029.

Original article

## ASSESSMENT OF COMPUTING SYSTEMS RELIABILITY BY NON-PARAMETRIC METHOD BY SMALL SAMPLES OF OPERATIONAL DATA

Nikulin Vladimir S.<sup>1</sup>, Pestunov Andrey I.<sup>2</sup>

<sup>1,2</sup> *Novosibirsk State University of Economics and Management*

<sup>1</sup> nikulin-94@inbox.ru

<sup>2</sup> kocheevaalina@gmail.com

**Abstract.** A new method for assessing the reliability of small computing systems that allows the generation of only small samples of operational data is proposed in the article. The result of applying the technique is the posterior failure distribution density, on the basis of which various reliability indicators can be calculated. The methodology consists of two stages: the first is the preparation of operational data, including detection of failures using machine learning methods, and the second is the construction of the failure distribution density using the adapted Rosenblatt–Parzen method. Increasing the efficiency of estimates using the proposed method is achieved by taking into account censored data, compensating for the shift of failure distribution densities and finding the optimal smoothing parameter.

**Keywords:** reliability of computing systems, operational data, machine learning, Rosenblatt–Parzen method, smoothing parameter, failure distribution density

**For citation:** Nikulin V.S., Pestunov A.I. Assessment of computing systems reliability by non-parametric method by small samples of operational data. *Vestnik NSUEM*. 2024; (1): 10–29. (In Russ.). DOI: 10.34020/2073-6495-2024-1-010-029.

### Введение

Вычислительные системы – это комплексы программно-аппаратных элементов (узлов), которые образуют единую среду для решения задач, связанных с обработкой информации. Они имеют большое значение для математического моделирования, обработки текстового и аудиовизуального контента, обнаружения закономерностей, а также распознавания образов и речевой информации. Они должны обладать высокой надежностью, поскольку даже единичные отказы могут иметь критические последствия. В работах известного специалиста по теории надежности А.М. Половко [8] и ряде стандартов говорится о том, что вероятность безотказной работы вычислительных систем в течение заданного интервала времени должна быть не менее 0,97.

Оценка надежности вычислительных систем предполагает расчет показателей надежности, среди которых вероятность безотказной работы, коэффициент готовности, средняя наработка на отказ и интенсивность отказов. Оценить данные показатели перед вводом вычислительной системы в эксплуатацию можно тремя основными способами:

- теоретически на этапе проектирования;
- методом биномиальных испытаний;
- посредством ускоренных лабораторных испытаний.

Все эти способы имеют недостатки или ограничения, препятствующие получению адекватных оценок надежности, которые можно было бы использовать для принятия решений в процессе обслуживания вычислительной системы. Так, теоретические расчеты надежности во многих случаях оказываются недостоверны, поскольку на этапе проектирования вычислительной системы отсутствует информация о надежности ее составных элементов, а также планируемых режимах эксплуатации и обслуживания. Теоретические расчеты могут осуществляться параметрическими методами, основанными на априорной информации о распределении отказов. Например, О.О. Шмидт разработала обобщенную модель процесса восстановления после отказов, в рамках которой рассчитываются показатели надежности [14], однако эта и другие подобные модели неприменимы при отсутствии априорных данных.

При получении оценок надежности методом биномиальных испытаний для подтверждения достижения указанной выше вероятности безотказной работы порядка 0,97 потребуется поставить в опытную эксплуатацию не менее 128 элементов. Если во время опытной эксплуатации будет зафиксировано не более одного отказа, то можно сделать вывод о достижении заданной вероятности. Однако при наличии большего количества отказов потребуется значительно увеличить либо время наблюдений, либо количество оцениваемых элементов. Подобные масштабы испытаний могут быть осуществлены только на крупных вычислительных системах, состоящих из больших групп однотипных элементов, таких как, например, «Ломоносов-2» из 1686 узлов в МГУ имени М.В. Ломоносова [24]. Отсюда следует, что для получения достоверных оценок надежности небольших вычислительных систем (менее 100 элементов) требуется длительный период опытной эксплуатации, что в большинстве случаев неприемлемо.

Третьей обозначенной выше альтернативой получения оценок надежности являются ускоренные лабораторные испытания, которые заключаются в искусственном повышении нагрузки на элементы. И здесь, как отмечают А.М. Половко и С.В. Гуров [8], расчеты основываются на различных допущениях, в том числе на предположении о том, что законы распределения отказов при (искусственном) увеличении нагрузки остаются такими же, как и при штатной эксплуатации. Подобные допущения сложно доказать или проверить. Есть вероятность того, что в каких-либо случаях они могут оказаться ошибочными. В итоге лабораторные условия будут значительно отличаться от условий реальной эксплуатации, а модели отказов, используемые при ускоренных лабораторных испытаниях, окажутся неадекватны реальному объекту и приведут к неверным оценкам надежности.

Таким образом, для небольших вычислительных систем без использования *эксплуатационных данных* (данных о фактических отказах оборудования в процессе реальной эксплуатации) осуществить достоверный расчет надежности практически невозможно. В настоящей статье предложена методика получения оценок надежности таких вычислительных систем по малым выборкам эксплуатационных данных. Методика включает подготовку эксплуатационных данных (приведение их к виду, пригодному для

дальнейшей обработки) и построение плотности распределения отказов при помощи адаптированного под особенности вычислительных систем непараметрического метода Розенблатта – Парзена [20, 21].

### **Оценка надежности вычислительных систем по эксплуатационным данным**

Имеются общие методики оценки надежности сложных технических систем по эксплуатационным данным. Вычислительные системы – это их частный случай. Так, А.М. Половко и С.В. Гуров предлагают использовать инженерную методику анализа надежности техники, основанную на аналитическом решении специального интегрального уравнения [8]. Однако, как отмечают сами авторы, ее реализация часто невозможна на практике из-за требований о наличии ряда априорных сведений, многие из которых могут быть недоступны.

А.В. Антонов и М.С. Никулин [1] предложили методику оценки надежности насосных агрегатов АЭС на основе непараметрических методов (эти методы не требуют априорного знания о распределении отказов), среди которых метод Розенблатта – Парзена, метод проекционных оценок и метод корневого оценивания [12, 15, 20, 21]. Применение этих и других общих методик для оценки надежности именно вычислительных систем возможно, но в исходном виде затруднено в силу нескольких причин:

- вычислительные системы состоят из аппаратных и программных элементов, отказы которых могут иметь индивидуальные причины;

- вычислительные системы регулярно модернизируются и масштабируются, что обуславливает постоянное изменение количества оцениваемых элементов и формирует потребность получения оценок в режиме реального времени;

- в качестве основной группы элементов вычислительных систем выступают вычислительные серверы, где за наблюдаемый период времени фиксируется крайне малое число отказов (менее 30);

- на некоторых элементах (например, неконтролируемых источниках бесперебойного питания или сетевых медиаконверторах) отказы фиксируются не точными временными значениями, а интервалами (цензурируются интервалом);

- часть элементов может вообще не иметь отказов за период мониторинга, и их невозможно использовать в расчетах единообразно с данными о фактических отказах; однако учет информации о том, что отказа не произошло, потенциально позволяет повысить эффективность оценки надежности (информация об отсутствии отказа – это цензурированные справа данные).

В упомянутых выше методах отсутствует в явном виде аналитическое представление функции плотности распределения времени до отказа, учитывающее не только полные, но и все виды цензурированных наработок на отказ. Методика А.В. Антонова и М.С. Никулина цензурированные справа данные учитывает, но хорошие результаты достигаются, если их

число не превышает 30 % от общего числа отказов. При увеличении доли цензурированных данных вид функции распределения нарушается. Кроме того, в рассмотренных методиках расчет параметра сглаживания в методе Розенблатта – Парзена приведен в общем виде. Конкретный алгоритм не разработан, хотя сам метод Розенблатта – Парзена в целом подходит для построения оценки надежности небольших вычислительных систем.

Таким образом, актуализируется задача адаптации общих методик оценки надежности под небольшие вычислительные системы и повышение их эффективности за счет учета цензурированных данных и конкретизации метода Розенблатта – Парзена.

### **Проблема формирования массива эксплуатационных данных**

Важным фактором, влияющим на точность оценок надежности, является качество сформированного массива эксплуатационных данных. Несмотря на развитие систем мониторинга, сведения об отказах по-прежнему часто фиксируются обслуживающим персоналом в журналах, рекламационных актах и формулярах. В связи с этим вычислительные системы генерируют статистическую информацию об отказах низкого качества. Она может иметь пропуски, быть зашумлена, противоречива или искажена из-за отсутствия стандартизированного способа выявления отказов. При обобщении таких сведений многие важные события могут быть утрачены. Подготовка данных – важный этап решения задач их анализа и может занимать более 50 % времени реализации проекта. Экспертам и аналитикам требуются подробные методики и автоматизированный инструментарий для осуществления подготовки данных. Эта проблема носит общий характер для всех технических систем и исследована В.С. Викторовой [2], А.С. Степанянц [3], А.Г. Тарасовым [10], М.Б. Успенским [11], I. Foster [16], R. Wolski [23] и др. В работе [7] описана методика подготовки данных на концептуальном уровне без конкретизации отдельных этапов. В [23] рассматривается вопрос о разработке новой методики анализа данных в сфере энергетики из-за недостаточной проработанности существующих решений. В [9] исследуются программные реализации интеллектуального анализа данных в различных сферах и затрагиваются вопросы построения соответствующей базы данных. Однако проанализированные работы носят ограниченный характер и применимы только к своим отраслям или конкретным случаям. А.Г. Тарасов [10] и М.Б. Успенский [11] отмечают, что современный уровень развития систем мониторинга позволяет фиксировать эксплуатационные данные в режиме реального времени для большинства аппаратных, но не программных элементов. Отсюда следует, что актуальной является задача разработки единой архитектуры системы сбора информации как с аппаратных элементов, так и с управляющего программного обеспечения вычислительной системы.

Кроме того, при решении задачи формирования массива эксплуатационных данных возникает проблема обнаружения релевантных (имеющих отношение к отказам и к оценке надежности) событий в общем потоке ин-

формации, фиксируемой системой мониторинга. В большинстве случаев информация от системы мониторинга избыточна. Она может содержать справочные данные, которые не могут быть использованы для оценки надежности. Следовательно, актуализируется задача выявления среди всех эксплуатационных данных именно тех, которые нужны для построения оценок надежности. В настоящей статье предложены модели машинного обучения для поиска отказов среди всей информации.

### Общая схема формирования массива данных об отказах

Предлагаемая в настоящей статье технология подготовки данных для последующего интеллектуального анализа надежности состоит из двух этапов: *общего* и *специального*. Целесообразность такого разделения обусловлена тем, что для различных сложных технических систем (к которым относятся и вычислительные) можно выделить некоторые одинаковые задачи по сбору и подготовке данных. Это дает возможность применять элементы разработанной технологии в других областях, не связанных с вычислительными системами, например, к системам видеонаблюдения и системам сетевого оборудования. В то же время ряд задач остаются уникальными для каждого типа систем и требуют отдельной проработки [13]. Указанные этапы можно декомпозировать на отдельные задачи и представить в виде иерархического списка (рис. 1).

1. Общий набор задач.
  - 1.1. Очистка данных.
    - 1.1.1. Обнаружение несоответствий фиксации контролируемых параметров.
    - 1.1.2. Устранение контролируемых параметров с пустыми значениями.
    - 1.1.3. Исправление аномальных значений контролируемых параметров.
  - 1.2. Выбор данных.
    - 1.2.1. Выбор информативных контролируемых параметров.
    - 1.2.2. Определение интервала фиксации записей.
2. Специальный набор задач.
  - 2.1. Форматирование итогового набора данных.
    - 2.1.1. Стандартизирование типов данных под модели машинного обучения.
    - 2.1.2. Создание структуры данных.
  - 2.2. Формирование итогового набора данных.
    - 2.2.1. Формирование зависимостей атрибутов и элементов вычислительной системы.
    - 2.2.2. Добавление атрибутов (вспомогательной информации) в набор данных.
    - 2.2.3. Получение итогового набора данных для обучения машинных моделей.
  - 2.3. Интеграция итогового набора данных с базой данных.
    - 2.3.1. Выявление способа хранения данных.
    - 2.3.2. Коррекция времени фиксации итогового набора в базе данных.
    - 2.3.3. Соотнесение итогового набора данных с моделями машинного обучения.

*Рис. 1.* Декомпозиция этапов технологии подготовки данных для их дальнейшего использования при получении оценок надежности вычислительных систем  
Decomposition of the stages of data preparation technology for their further use in obtaining estimates of the reliability of computer systems

Рассмотрим основные блоки задач более подробно.

*Очистка данных.* Анализ информации, фиксируемой системой мониторинга вычислительной системы, показал, что она содержит пропуски, которые зачастую связаны с асинхронной фиксацией контролируемых параметров. Это свидетельствует о том, что данные пропущены согласно каким-либо закономерностям. Заполнить пропуски с сохранением этих закономерностей можно при помощи *метода подстановки среднего значения* [6]. Если же характер пропуска случайный, то это может свидетельствовать об отказе или «аномальном» событии, поэтому такой пропуск следует включить в итоговый набор данных в исходном виде.

Простое удаление пропусков может приводить к потере зависимостей, поэтому пропущенные данные необходимо восстанавливать. Для нашего случая подходит гибридный адаптационный метод для сложных динамических систем. Еще одним важным звеном первого этапа является оценка качества данных, включающая проверку данных на наличие ошибок, дубликатов и противоречий.

*Выбор данных.* В данном блоке решаются задачи уменьшения размерности и выбора информативных признаков. Если в наборе данных присутствуют избыточные, неинформативные или слабо информативные признаки, это может негативно сказываться на эффективности модели машинного обучения. Путем уменьшения размерности данных модель становится проще, что влечет уменьшение объема памяти, занимаемого набором данных, и ускоряет работу моделей машинного обучения на нем. Существуют различные методы уменьшения размерности, например, методы выбора признаков и методы выделения признаков. Одним из эффективных методов является L1-регуляризация, заключающаяся в расчете коэффициента средней абсолютной ошибки для каждого признака и штрафования признаков с большими коэффициентами.

*Форматирование итогового набора данных.* В связи с тем что модели машинного обучения (которые используются для распознавания отказов) не могут работать с текстовыми данными в явном виде, их необходимо представить в десятичном виде. При необходимости изменяются типы данных и синтаксическая структура атрибутов и значений. На данном этапе определяется целевая переменная, определяющая состояние устройства «В РАБОТЕ» или «ОТКАЗ» в заданный момент времени на основе показателей контролируемых параметров.

*Формирование итогового набора данных.* Для определенных параметров требуется формирование производных релевантных признаков с дальнейшим преобразованием в векторы специального формата для моделей машинного обучения. В исходном виде признаки использовать не всегда возможно, поскольку их зависимость снижает эффективность машинного обучения. Решение задач этого блока направлено на формирование набора параметров, подходящих для обучения машинных моделей поиска отказов.

*Интеграция итогового набора данных с базой данных.* Цель заключительного блока задач – это соотнесение итогового набора данных с моделями машинного обучения и определение схемы этой интеграции. Здесь выявляются и исправляются конфликты, как всегда возникающие в про-

цессе интеграции, а также устанавливаются соотношения в базе данных с учетом произошедшего обновления. Создаются таблицы со следующими атрибутами: ключ, время события, параметры и состояние контролируемого устройства. Финальная интеграция состоит в формировании итогового набора, куда включены данные обо всех событиях за период подконтрольной эксплуатации.

### **Алгоритм сбора и подготовки данных**

Описанную выше концепцию предлагается реализовать в виде следующего алгоритма.

*Шаг 1.* Запрос в базу данных посредством инструментария Python DB-API и SQL.

*Шаг 2.* Формирование выборок посредством библиотек matplotlib и pandas.

*Шаг 3.* Отбор релевантных контролируемых параметров (L1-регуляризация).

*Шаг 4.* Очистка данных и удаление некорректных значений.

*Шаг 5.* Исключение дублирования данных.

*Шаг 6.* Создание производных признаков (средняя температура по ядрам, загрузка ресурсов ОЗУ и др.).

*Шаг 7.* Форматирование итогового набора данных.

На выходе представленного алгоритма получается массив данных, включающий информацию обо всех событиях, произошедших за период мониторинга вычислительной системы. Среди этой информации присутствуют данные, не относящиеся к отказам, поэтому следующим этапом оценки надежности является их исключение.

### **Обучение машинных моделей для обнаружения отказов**

После того как эксплуатационные данные собраны и подготовлены, необходимо отобрать среди них релевантную информацию, касающуюся именно отказов отдельных элементов вычислительной системы. Остальные данные должны быть исключены из рассмотрения, поскольку не будут использоваться при получении оценок надежности. В большинстве случаев задача распознавания отказов в общем потоке информации, поступающей от системы мониторинга, решается с помощью оценки граничных параметров каждого из контролируемых значений. Однако нередко бывают ситуации, когда контролируемое устройство формирует сигнал работоспособного состояния, но при этом по каким-либо причинам оказывается недоступно. В этом случае целесообразно применение методов машинного обучения.

Общая логика функционирования обученных моделей заключается в следующем. Показатели контролируемых параметров составных элементов вычислительной системы передаются в обученную модель, которая, в свою очередь, проводит комплексную обработку данных. На основе результатов этой обработки модель формирует прогностическую информацию о текущем состоянии оборудования. Если модель распознает отклонение

или потенциально возможный отказ, то она генерирует соответствующее предупреждение и рекомендации для операторов или технического персонала. Результатом обработки является сформированный массив значимых событий по каждому элементу вычислительной системы.

Предлагаемый алгоритм обучения моделей для обнаружения отказов выглядит следующим образом.

*Шаг 1.* Разделить генеральную совокупность на обучающую и тестовую выборку (на основе практических рекомендаций в открытых источниках и собственных предварительных экспериментов выведено оптимальное соотношение 80 и 20 % соответственно).

*Шаг 2.* Выбрать базовую модель и функцию оценки качества обучения модели (деревья решений, регрессионные алгоритмы и др.).

*Шаг 3.* Определить набор управляющих параметров базовых моделей.

*Шаг 4.* Обучить базовые модели с перебором набора управляющих параметров.

*Шаг 5.* Оценить качество обученных моделей; выбрать итоговую модель.

Для поиска оптимальных значений управляющих параметров требуется определить диапазон всевозможных допустимых значений (так называемую «сетку» параметров), которые могут использоваться для обучения машинной модели. Затем применяется метод перебора и на каждой комбинации значений обучается модель. После этого выбирается модель с лучшими показателями.

В разработанных моделях учитываются показатели функционирования программных и аппаратных элементов одновременно. В результате обучения формируются итоговые модели, достигшие в процессе обучения лучших показателей эффективности (минимальных ошибок на тестовых выборках). Модели позволяют определять отказы составных элементов вычислительных систем в процессе их функционирования, тем самым формируя эксплуатационные данные, необходимые для оценки показателей надежности в дальнейшем.

### **Выбор управляющих параметров метода Розенблатта – Парзена**

Исходя из проведенного выше анализа общих методик оценки надежности сложных технических систем, сделан вывод о том, что в качестве базового подхода лучше всего подходит непараметрический метод Розенблатта – Парзена. Для него существуют наработки, учитывающие исторические данные функционирования и цензурированную информацию. Кроме того, этот метод подходит для малых выборок эксплуатационных данных, поскольку даже для единичных событий (отказов) возможно построить частную плотность распределения, которую можно использовать для различных оценок, в том числе показателей надежности.

Основной показатель надежности, который будет рассчитываться в рамках предлагаемой методики, – плотность распределения времени до отказа. На ее основе легко вычисляются остальные показатели [5].

Адаптация непараметрического метода Розенблатта – Парзена осуществлялась в три этапа. На первом этапе проведен анализ влияния управляющих параметров данного метода на точность аппроксимации итоговой плотности распределения и разработан алгоритм вычисления оптимального параметра сглаживания. Пусть  $N$  – количество наблюдаемых элементов вычислительной системы и  $\vec{X} = (x_1, x_2, \dots, x_N)$  – массив значений времени работы между отказами. Тогда базовая оценка плотности распределения методом Розенблатта – Парзена имеет следующий вид:

$$f(t) = \frac{1}{N\sigma} \sum_{i=1}^N V\left(\frac{t-x_i}{\sigma}\right), \quad (1)$$

где  $V$  – функция ядра,  $\sigma$  – параметр сглаживания.

В результате предварительных экспериментов установлено, что выбор функции ядра в основном влияет только на визуальное представление итоговой функции распределения, но от нее практически не зависит аппроксимация, поэтому в качестве  $V(x)$  выбрана простая стандартная функция Гаусса, имеющая следующий вид:

$$V(x) = \frac{1}{\sqrt{2\pi}} \cdot e^{-\frac{x^2}{2}}.$$

Что касается параметра сглаживания ( $\sigma$ ), то его влияние на точность аппроксимации существенно, поэтому в рамках исследования разработан расчет его оптимального значения. Установлено [4], что оценки плотности состоятельны при выполнении следующих условий:

$$\lim_{N \rightarrow \infty} \sigma(N) = 0 \quad \text{и} \quad \lim_{N \rightarrow \infty} \sigma(N) \cdot N = \infty. \quad (2)$$

С учетом формулы (2) параметр сглаживания можно представить в виде [19]

$$\sigma = \frac{\mu}{N^v}, \quad (3)$$

где  $0 < v < 1$ ,  $\mu$  – коэффициент размерности для случайной величины.

Выражение (3) допускает выбор  $\sigma$  в достаточно широких пределах, что приводит к постановке задачи поиска его оптимального значения.

### Вычисление оптимального параметра сглаживания

Предлагаемый алгоритм для расчета его оптимального значения основан на методе Хука – Дживса [17]. Представим его пошагово в виде двух частей: подготовительной и основной.

Пусть  $L$  – логарифмическая функции правдоподобия:

$$L(\sigma) = \sum_{i=1}^N \ln f_{\sigma}(x_i).$$

Подготовительный этап состоит из 4 следующих шагов.

*Подготовительный шаг 1.* Присвоить начальное значение параметра сглаживания согласно приближению Сильвермана [22]:

$$\sigma_b := 0,9 \min \left( \hat{\sigma}, \frac{\mu}{1,34} \right) N^{-\frac{1}{5}},$$

где  $\hat{\sigma}$  – стандартное отклонение элементов выборки,  $\mu$  – медиана выборки.

*Подготовительный шаг 2.* Установить начальное значение шага изменения искомого параметра сглаживания:  $d := \sigma_b / 2$ .

*Подготовительный шаг 3.* Установить минимальное значение ( $e$ ) шага изменения искомого параметра сглаживания, при достижении которого алгоритм останавливается; минимальное значение выбирается согласно требуемой точности оценки искомого параметра.

*Подготовительный шаг 4.* Вычислить начальное значение функции  $L(\sigma_b)$ .

Теперь переходим к основной части, в которой искомый параметр сглаживания вычисляется итеративно до достижения заданной точности ( $e$ ).

*Шаг 1.* Присвоить  $\sigma^+ := \sigma_b + d$  и вычислить значение функции  $L(\sigma^+)$ .

*Шаг 2.* Если  $L(\sigma^+) > L(\sigma_b)$ , то присвоить  $\sigma_b := \sigma^+$  и перейти на шаг 1, иначе – шаг 3.

*Шаг 3.* Присвоить  $\sigma^- := \sigma_b - d$  и перейти на шаг 6.

*Шаг 4.* Если  $\sigma^- < 0$ , то перейти на шаг 5, иначе – шаг 7.

*Шаг 5.* Вычислить значение функции  $L(\sigma^-)$ .

*Шаг 6.* Если  $L(\sigma^-) > L(\sigma_b)$ , то присвоить  $\sigma_b := \sigma^-$  и перейти на шаг 1, иначе – шаг 7.

*Шаг 7.* Если  $d < e$ , то вернуть  $\sigma_b$  и завершить; иначе  $d := d/2$  и перейти на шаг 3.

Таким образом, результатом работы представленного выше алгоритма является вычисленное оптимальное значение параметра сглаживания.

### Компенсация смещения плотности распределения отказов

При использовании базовой формулы (1) для расчета плотности распределения отказов методом Розенблатта – Парзена возникает проблема, состоящая в том, что необходимо компенсировать смещение построенной плотности. Областью определения выбранной в качестве ядра функции Гаусса является интервал  $(-\infty; +\infty)$ , в то время как используемый в теории надежности аргумент  $t$  (время) определен на интервале  $[0; +\infty)$ , и при наличии близких к нулю наработок на отказ оцениваемая плотность распределения смещается, приводя к нарушению условия  $F(0) = 0$ , где  $F$  – функция распределения отказов, соответствующая плотности  $f$ . Данное условие означает, что отказов при  $t \leq 0$  быть не может, а наличие смещения приводит к возможности неравенства  $F(0) > 0$ , означающего, что будто бы отказы фиксировались до ввода вычислительной системы в эксплуатацию (при  $t \leq 0$ ).

Компенсировать это смещение можно методом зеркального отображения [18], что трансформирует базовую оценку (1) к виду

$$f_p(t) = \frac{1}{N\sigma} \sum_{i=1}^N \left[ V \left( \frac{t-x_i}{\sigma} \right) + V \left( \frac{t+x_i}{\sigma} \right) \right]. \quad (4)$$

### Учет эксплуатационных данных, цензурированных интервалом

На третьем этапе выводится аналитическое представление функции плотности распределения с учетом цензурированной информации. Если цензурированные данные отсутствуют, то оценка плотности распределения отказов методом Розенблатта – Парзена с учетом компенсации смещения осуществляется по формуле (4). Напомним, что цензурирование интервалом – это ситуация, когда момент наступления отказа оборудования неизвестен (или не достигнут), но имеется априорная информация об интервале времени, на котором произошел отказ. В этом случае при оценке плотности распределения вводятся дополнительные определения:

$\vec{L} = [(l_1, l_2); (l_2, l_3); \dots; (l_{j-1}, l_j)]$  – массив интервалов цензурирования, на которых зафиксировано  $\vec{V} = (v_1, v_2, \dots, v_j)$  случайных событий,  $\Delta_j$  – длина интервала цензурирования. Функция плотности для таких данных будет иметь вид:

$$f_{\Delta}(t) = \frac{1}{\sigma N} \left[ \sum_{j=1}^J v_j \int_0^1 V \left( \frac{t - l_j - u \Delta_j}{\sigma} \right) du \right]. \quad (5)$$

### Учет эксплуатационных данных, цензурированных справа

Для учета данных, цензурированных справа, в методе ядерных оценок в [1] предложен подход, заключающийся в замене величины  $t$ , распределенной на интервале  $(l, +\infty)$ , случайной величиной  $\tau = \frac{1}{l}$ , распределенной на интервале  $\left(0, \frac{1}{l}\right)$ , в этом случае плотность распределения наработок, цензурированных справа, будет иметь вид

$$f_r(t) = \frac{1}{\sigma N} \left[ \sum_{s=1}^S \frac{v_s}{t^2} \int_0^1 V \left( \frac{t - \frac{1}{l_s} - u \Delta_s}{\sigma} \right) du \right], \quad (6)$$

где  $S$  – количество наработок на отказ, цензурированных справа.

Проведенные исследования [1] показывают, что плотность, построенная только по формуле (4), остается распределенной на интервале  $(-\infty, +\infty)$ , и интеграл плотности  $\int_0^{\infty} f_r(t)$  оказывается меньше 1, что является нарушением условий нормировки (интеграл от плотности должен быть равен единице). Таким образом, возникает необходимость компенсировать это смещение. С помощью метода зеркального отображения данных получено выражение

$$f_r(t) = \frac{1}{\sigma N} \left[ \sum_{s=1}^S \frac{v_s}{t^2} \int_0^1 V \left( \frac{\frac{1}{t} + \frac{u}{l_s}}{\sigma} \right) - V \left( \frac{\frac{1}{t} - \frac{u}{l_s}}{\sigma} \right) du \right]. \quad (7)$$

Особенностью функции (7) является то, что она устанавливает единое поведение для всех цензурированных справа данных от любых источников, так как имеет только два управляющих параметра: значение границы  $t$  и параметр сглаживания  $\sigma$ . Данные параметры не зависят от поведения всей предыдущей статистики и, как показывают исследования [1], малочувствительны к изменению управляющих параметров.

В рамках нашей адаптации метода Розенблатта – Парзена предлагается альтернативный вариант с использованием легкодоступной априорной информации о правой границе цензурированных данных. Источником такой информации должны быть сведения о надежности аналогичного оборудования, в частности, некоторое согласованное экспертами конечное значение  $t_n$  (например, допустимый срок эксплуатации), на которое может быть заменена бесконечность на правой границе. Если это единственная априорная информация и нет сведений о поведении функции распределения на интервале  $(\bar{t}, t_n)$ , то на основании принципа максимума энтропии оценка строится по формуле (3) для цензурированных интервалами данных. В отличие от параметрических методов, данный вариант адаптации не требует в качестве априорной информации сведения о поведении функции распределения на интервале времени, тем самым сохраняет свою универсальность и независимость от априорных законов распределения.

В отсутствие априорной информации о значении  $t_n$  для его определения предлагается воспользоваться приемом, заключающимся в расчете  $t_n$  для каждого из источников, исходя из пропорции относительно числа элементов в источнике и числа отказов, произошедших до момента наступления цензурирования:

$$t_n = \frac{N}{N-r} \cdot \bar{t},$$

где  $r$  – число цензурированных справа наработок,  $N$  – число элементов.

Достоинствами предложенного приема является простота оценки. Кроме того, обеспечивается связь поведения функции распределения за пределами границы цензурирования с ее поведением до границы. Достигается единообразие используемых выражений, обеспечивается гладкость функции в точке перехода, снимается неопределенность относительно цензурированной информации. Все эти условия позволяют в дальнейшем эффективно осуществить прогнозирование.

### Аналитическое представление функции плотности распределения отказов

На основании полученных в предыдущих разделах частных плотностей сформирована следующая итоговая функция плотности, в которой учтены все три вида наработок: полные, цензурированные справа и цензурированные интервалами:

$$f(t) = \frac{1}{\sigma N} \left[ \sum_{i=1}^I f_p(t) + \sum_{j=1}^J f_{\Delta}(t) + \sum_{s=1}^S f_r(t) \right],$$

где  $I$  – количество полных наработок на отказ,  $J$  – количество наработок, цензурированных интервалом,  $S$  – количество наработок, цензурированных справа.

### **Экспериментальная оценка эффективности и достоверности методики**

Проведен ряд экспериментов по исследованию достоверности и точности разработанной методики на модельных данных. Были сгенерированы выборки случайных величин, моделирующих наработки на отказ, имеющие логнормальное распределение, гамма-распределение и распределение Вейбулла. Показано, что даже при существенной доле цензурированных справа данных (порядка 50 %) средние значения ошибок оценивания по предложенной методике лишь незначительно превышают ошибки оценивания, полученные методом максимума правдоподобия с использованием всех элементов выборки и при априорном знании закона распределения.

Обнаружена связь между учетом цензурированных данных и точностью оценки плотности. Установлено, что при высокой доле (70 % и выше) цензурированных данных их учет повышает точность оценки плотности на 25 % при сравнении с оценкой только по полным наработкам. Если процент цензурированных данных снизить до 50–70 %, то их учет снижает ошибку оценивания на 20–25 %. При низкой доле цензурированных данных (30–50 %) ошибка оценивания снижается на 15–20 %. Отметим, что при использовании упомянутого ранее способа учета цензурированной справа информации [1] снижение ошибки наблюдается только в случае, когда количество цензурированных данных не превышает 40 %. В случае увеличения доли цензурированных данных (более 40 %) ошибка остается на прежнем уровне (как и без учета этих данных).

При этом полученная оценка плотности достигает заданного уровня точности уже при небольших выборках (от 10 случайных величин), что и требовалось обеспечить при получении оценок надежности небольших вычислительных систем.

### **Практическое применение предложенной методики**

Предложенная методика апробирована на реально функционирующей вычислительной системе, состоящей из 20 узлов Supermicro Server 1024US-TRT. Период подконтрольной эксплуатации составил 43 800 ч. За этот период зафиксировано 187 полных и 29 цензурированных наработок. По сформированным эксплуатационным данным согласно разработанной методике получены оценки плотности распределения наработок на отказ и вероятности безотказной работы с учетом и без учета цензурированных данных.

Результаты представлены на рис. 2. Исходя из графиков можно заключить, что надежность вычислительной системы выше, чем давали оценки без учета цензурированных данных. Видно, что при их учете плотность смещается вправо, свидетельствуя о более высокой вероятности безотказной работы. Кроме того, учет цензурированных данных повышает точность определения времени доминирующей группы отказов.

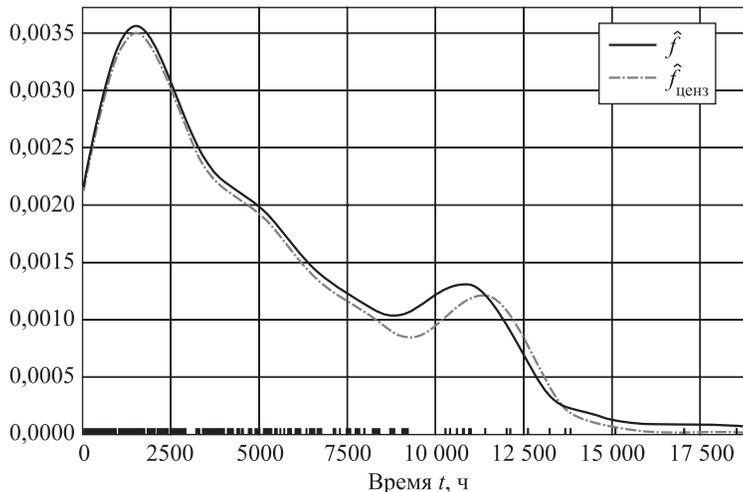


Рис. 2. Сравнение двух оценок плотности распределения отказов реально функционирующей вычислительной системы согласно предложенной методике: с учетом ( $\hat{f}_{\text{ценз}}$ ) и без учета цензурированной информации ( $\hat{f}$ ).

Comparison of two estimates of the failure distribution density of a really functioning computer system according to the proposed methodology: taking into account ( $\hat{f}_{\text{ценз}}$ ) and without taking into account censored information ( $\hat{f}$ ).

### Применение методики для выявления причин отказов

Несмотря на то, что основным результатом применения предлагаемой методики являются апостериорная плотность распределения отказов и расчет показателей надежности вычислительных систем, ее можно использовать для более углубленного анализа надежности, в том числе для выявления причин отказов. По полученным оценкам надежности выявляются временные интервалы с доминирующим количеством отказов и проводится соотнесение этих интервалов и отказавших узлов системы. Полученное соответствие позволяет определить проблемные места эксплуатации. Причинами отказов могут быть как сами узлы, так и факторы, связанные с эксплуатационным обслуживанием.

Так, например, при апробации методики на упомянутой ранее реально функционирующей вычислительной системе определено три временных интервала с доминирующим числом отказов. При анализе удалось установить связь этих отказов с конкретными узлами вычислительной системы.

Первый интервал определен на участке 1500–2000 ч, содержит более 70 % отказов, связанных с выходом из строя драйвера «NVIDIA Driver for Linux x86\_64 v.340.58». Более детальный анализ опыта эксплуатации позволил выявить недостаток запуска пользователями определенного типа задач, связанных с машинным обучением. При выполнении одной из базовых процедур (прерывание обучения машинных моделей) происходил процесс отключения составного элемента (драйвера), требующий перезапуска вычислительной системы. Второй интервал определен на участке 3000–

6000 ч и связан с особенностями технического обслуживания и ремонта. Из-за нерегламентированного процесса обновления программных элементов вычислительной системы происходили сбои в ее функционировании. Третий интервал определен на участке 10 000–12 000 ч и связан с отказами отдельных узлов в период неплановых аварийных отключений. Таким образом, в результате анализа построенной плотности распределения удалось сопоставить отклонения функции с объективными причинами.

Полученные результаты углубленного анализа позволяют рассмотреть различные варианты снижения количества отказов. Так, для рассмотренной выше вычислительной системы предлагается ввести в эксплуатацию программный модуль контейнеризации расчетных задач. Это позволит изолировать программное окружение для каждого пользователя и исключить взаимное негативное воздействие в целом. Такой подход приведет к снижению отказов первого и второго интервала. Для уменьшения третьего отклонения предлагается разработать систему планомерной остановки вычислительных серверов в случае аварийного отключения вычислительной системы, а также добавить в план технического обслуживания проверку работоспособности системы аварийного отключения/включения.

### Заключение

В настоящей статье рассмотрена задача оценки надежности небольших вычислительных систем по малым выборкам эксплуатационных данных, которые могут быть цензурированы справа или цензурированы интервалом. Для решения этой задачи предложена новая методика, результатом применения которой является апостериорная плотность распределения отказов, позволяющая рассчитать различные показатели надежности, среди которых средняя наработка на отказ, интенсивность отказов, вероятность безотказной работы и коэффициент готовности. Особенностью методики является то, что она позволяет осуществить оценку надежности в режиме реального времени без необходимости предварительного формирования больших наборов эксплуатационных данных, что может быть осуществлено только на крупных вычислительных системах или при длительном периоде опытной эксплуатации. Как следствие, предложенная методика позволяет получать адекватные оценки надежности для небольших вычислительных систем практически сразу после их ввода в реальную эксплуатацию.

Предложенная методика состоит из двух этапов: подготовка эксплуатационных данных и построение плотности распределения отказов. Для реализации первого этапа разработана единая архитектура системы сбора информации с аппаратных элементов и управляющего программного обеспечения. Разработана технология подготовки данных, включающая обработку пропусков и выбросов. Предложены модели машинного обучения для распознавания релевантной информации (относящейся именно к отказам элементов системы), которая в дальнейшем будет использоваться для оценки надежности.

Для построения апостериорной плотности распределения отказов предложена адаптация непараметрического метода Розенблатта – Парзена, в рамках которой на основе метода Хука – Дживса с установкой начального приближения Сильвермана разработан алгоритм нахождения оптимального параметра сглаживания. В предложенной адаптации выполнена компенсация смещения построенной плотности, возникшего из-за несоответствия области определения функции ядра Гаусса и области определения временной переменной распределения с учетом цензурированных данных. Учет цензурированных справа данных осуществляется за счет определения правой границы, а данные, цензурированные интервалом, учитываются за счет выбора случайной точки на этом интервале.

На модельных (искусственно сгенерированных) данных показано, что учет цензурированных данных позволяет увеличить точность базового метода Розенблатта – Парзена. Проведена апробация предложенной методики на реально функционирующей вычислительной системе. Представлен способ использования методики не только для расчета показателей надежности, но и для выявления причин отказов.

#### Список источников

1. Антонов А.В., Никулин М.С. Статистические модели в теории надежности: учеб. пособие. М.: Абрис, 2012. 390 с.
2. Викторова В.С., Лубков Н.В., Степаняну А.С. Анализ надежности отказоустойчивых вычислительных систем. М.: ИПУ РАН, 2016. 117 с.
3. Викторова В.С., Степаняну А.С. Динамические деревья отказов // Надежность. 2011. № 3. С. 20–32.
4. Деврой Л., Дьёрфи Л. Непараметрическое оценивание плотности. L1-подход. М.: Мир, 1988. 408 с.
5. Захаров Д.Н., Никулин В.С. Анализ методов статистической оценки эксплуатационной надежности вычислительных комплексов // Научные технологии в космических исследованиях земли. 2020. № 1. С. 64–69.
6. Карлов И.А. Методы восстановления пропущенных значений с использованием инструментария Data Mining // Вестник Сибирского государственного аэрокосмического университета имени академика М.Ф. Решетнева. 2011. № 7. С. 29–33.
7. Матвеевский В.Р. Надежность технических систем: учеб. пособие. М.: Московский государственный институт электроники и математики, 2002. 113 с.
8. Половко А.М., Гуров С.В. Основы теории надежности. СПб.: БХВ-Петербург, 2006. 704 с.
9. Рузанов П.А. Методика создания приложений для работы с базами данных с использованием высокоуровневого языка Python // Цифровые технологии и информационная безопасность бизнес-процессов. Сборник научных статей по итогам научно-практической конференции с международным участием. М., 2022. С. 272–277.
10. Тарасов А.Г. Система мониторинга вычислительного кластера расширенной функциональности: дис. ... канд. техн. наук. Иркутск, 2011. 145 с.
11. Успенский М.Б. Разработка и исследование методов и моделей обработки диагностической информации для обнаружения и локализации неисправностей в системах хранения данных. М., 2020. С. 153.
12. Ченцов Н.Н. Оценка неизвестной плотности распределения по наблюдениям // Доклады АН СССР. 1962. № 1 (147). С. 45–48.

13. Чубукова И.А. Data Mining: учеб. пособие. М.: БИНОМ. Лаборатория знаний, 2008.
14. Шмидт О.О. Обобщенная модель процесса восстановления в теории надежности использования информационных технологий: дис. ... канд. физ.-мат. наук. Красноярск, 2008. 125 с.
15. Эфрон Б. Нетрадиционные методы многомерного статистического анализа. М.: Финансы и статистика, 1988. 261 с.
16. Foster I. Globus Toolkit Version 4: Software for Service-Oriented Systems // Journal of Computer Science and Technology. 2006. No. 4. Pp. 513–520.
17. Hooke R., Jeeves T. “Direct search” solution of numerical and statistical problems // Journal of the ACM. 1961. No. 8. Pp. 212–229.
18. Karunamuni R., Alberts T. A Locally Adaptive Transformation Method of Boundary Correction in Kernel Density Estimation // Journal of Statistical Planning and Inference. 2006. Vol. 136, no. 9. Pp. 2936–2960.
19. Padgett W.J., McNichols D.T. Nonparametric density estimation from censored data // Commun. Statist. Theory and Meth. 1984. Vol. 13, no. 13. Pp. 1581–1611.
20. Parzen E. On estimation of a probability density function and mode // Annals of Mathematical Statistics. 1962. No. 33. Pp. 1065–1076.
21. Rozenblatt M. Remark on some nonparametric estimates of a density function // Annals of Mathematical Statistics. 1956. No. 27. Pp. 832–837.
22. Silverman B. Density estimation for statistics and data analysis. London: Chapman & Hall/CRC. P. 45.
23. Wolski R. Developing a Dynamic Performance Information Infrastructure for Grid Systems. [Электронный ресурс]. URL: <https://www.researchgate.net/publication/2527468> (дата обращения: 04.01.2024).
24. Интеллектуальная система тематического исследования наукометрических данных. Суперкомпьютер «Ломоносов-2». [Электронный ресурс]. URL: <https://istina.msu.ru/equipment/card/19828901/> (дата обращения: 04.01.2024).

## References

1. Antonov A.V., Nikulin M.S. Statisticheskie modeli v teorii nadezhnosti: ucheb. posobie [Statistical models in reliability theory: textbook]. Moscow, Abris, 2012. 390 p.
2. Viktorova V.S., Lubkov N.V., Stepanjanc A.S. Analiz nadezhnosti otkazoustojchivyh vychislitel'nyh system [Reliability analysis of fault-tolerant computing systems]. Moscow, IPU RAN, 2016. 117 p.
3. Viktorova V.S., Stepanjanc A.S. Dinamicheskie derev'ja otkazov [Dynamic fault trees], *Nadezhnost' [Reliability]*, 2011, no. 3, pp. 20–32.
4. Devroj L., D'jorfi L. Neparаметрическое оценивание плотности. L1-podhod [Nonparametric density estimation. L1 approach]. Moscow, Mir, 1988. 408 p.
5. Zaharov D.N., Nikulin V.S. Analiz metodov statisticheskoj ocenki jekspluatacionnoj nadezhnosti vychislitel'nyh kompleksov [Analysis of methods for statistical assessment of the operational reliability of computing systems], *Naukoemkie tehnologii v kosmicheskikh issledovanijah zemli [Science-intensive technologies in space research of the earth]*, 2020, no. 1, pp. 64–69.
6. Karlov I.A. Metody vosstanovlenija propushhennyh znachenij s ispol'zovaniem instrumentarija Data Mining [Methods for recovering missing values using Data Mining tools], *Vestnik Sibirskogo gosudarstvennogo ajerokosmicheskogo universiteta imeni akademika M.F. Reshetneva [Bulletin of the Siberian State Aerospace University named after academician M.F. Reshetneva]*, 2011, no. 7, pp. 29–33.
7. Matveevskij V.R. Nadezhnost' tehniceskikh sistem: ucheb. posobie [Reliability of technical systems: textbook]. Moscow, Moskovskij gosudarstvennyj institut jelektroniki i matematiki, 2002. 113 p.

8. Polovko A.M., Gurov S.V. Osnovy teorii nadezhnosti [Fundamentals of reliability theory]. SPb, BHV-Peterburg. 2006. 704 p.
9. Ruzanov P.A. Metodika sozdaniya prilozhenij dlja raboty s bazami dannyh s ispol'zovaniem vysokourovnevnogo jazyka Python [Methodology for creating applications for working with databases using the high-level Python language]. Cifrovye tehnologii i informacionnaja bezopasnost' biznes-processov [Digital technologies and information security of business processes]. Sbornik nauchnyh statej po itogam nauchno-prakticheskoj konferencii s mezhdunarodnym uchastiem. Moscow, 2022. Pp. 272–277.
10. Tarasov A.G. Sistema monitoringa vychislitel'nogo klastera rasshirenoj funkcional'nosti [Computer cluster monitoring system with extended functionality]: diss. ... kand. tehn. nauk. Irkutsk, 2011. 145 p.
11. Uspenskij M.B. Razrabotka i issledovanie metodov i modelej obrabotki diagnosticheskoj informacii dlja obnaruzhenija i lokalizacii neispravnostej v sistemah hranenija dannyh [Development and research of methods and models for processing diagnostic information for detecting and localizing faults in data storage systems]. Moscow, 2020. P. 153.
12. Chencov N.N. Ocenka neizvestnoj plotnosti raspredelenija po nabljudenijam [Estimation of the unknown distribution density from observations], *Doklady AN SSSR [Reports of the USSR Academy of Sciences]*, 1962, no. 1 (147), pp. 45–48.
13. Chubukova I.A. Data Mining: ucheb. posobie [Data Mining: textbook]. Moscow, BINOM. Laboratorija znanij, 2008.
14. Shmidt O.O. Obobshhennaja model' processa vosstanovlenija v teorii nadezhnosti ispol'zovanija informacionnyh tehnologij [Generalized model of the recovery process in the theory of reliability of the use of information technologies]: diss. ... kand. fiz.-mat. nauk. Krasnojarsk, 2008. 125 p.
15. Jefron B. Netradicionnye metody mnogomernogo statisticheskogo analiza [Non-traditional methods of multivariate statistical analysis]. Moscow, Finansy i statistika, 1988. 261 p.
16. Foster I. Globus Toolkit Version 4: Software for Service-Oriented Systems. *Journal of Computer Science and Technology*, 2006, no. 4, pp. 513–520.
17. Hooke R., Jeeves T. "Direct search" solution of numerical and statistical problems. *Journal of the ACM*, 1961, no. 8, pp. 212–229.
18. Karunamuni R., Alberts T. A Locally Adaptive Transformation Method of Boundary Correction in Kernel Density Estimation. *Journal of Statistical Planning and Inference*, 2006, vol. 136, no. 9, pp. 2936–2960.
19. Padgett W.J., McNichols D.T. Nonparametric density estimation from censored data. *Commun. Statist. Theory and Meth.*, 1984, vol. 13, no. 13, pp. 1581–1611.
20. Parzen E. On estimation of a probability density function and mode. *Annals of Mathematical Statistics*, 1962, no. 33, pp. 1065–1076.
21. Rozenblatt M. Remark on some nonparametric estimates of a density function. *Annals of Mathematical Statistics*, 1956, no. 27, pp. 832–837.
22. Silverman B. Density estimation for statistics and data analysis. London: Chapman & Hall/CRC. P. 45.
23. Wolski R. Developing a Dynamic Performance Information Infrastructure for Grid Systems [Electronic resource]. Available at: <https://www.researchgate.net/publication/2527468> (accessed: 04.01.2024).
24. Intellektual'naja sistema tematicheskogo issledovanija nauko-metricheskikh dannyh. Superkomp'yuter «Lomonosov-2» [Intelligent system for thematic research of scientometric data. Supercomputer "Lomonosov-2"]. [Electronic resource]. Available at: <https://istina.msu.ru/equipment/card/19828901/> (accessed: 04.01.2024).

**Сведения об авторах:**

**В.С. Никулин** – аспирант, Новосибирский государственный университет экономики и управления «НИНХ», Новосибирск, Российская Федерация.

**А.И. Пестунов** – кандидат физико-математических наук, доцент, заведующий кафедрой, Новосибирский государственный университет экономики и управления «НИНХ», Новосибирск, Российская Федерация.

**Information about the authors:**

**V.S. Nikulin** – Graduate Student, Novosibirsk State University of Economics and Management, Novosibirsk, Russian Federation.

**A.I. Pestunov** – Candidate of Physical and Mathematical Sciences, Associate Professor, Head of Department, Novosibirsk State University of Economics and Management, Novosibirsk, Russian Federation.

**Вклад авторов:** все авторы сделали эквивалентный вклад в подготовку публикации. Авторы заявляют об отсутствии конфликта интересов.

**Contribution of the authors:** the authors contributed equally to this article. The authors declare no conflicts of interests.

<i>Статья поступила в редакцию</i>	<i>11.01.2024</i>	<i>The article was submitted</i>	<i>11.01.2024</i>
<i>Одобрена после рецензирования</i>	<i>20.01.2024</i>	<i>Approved after reviewing</i>	<i>20.01.2024</i>
<i>Принята к публикации</i>	<i>22.01.2024</i>	<i>Accepted for publication</i>	<i>22.01.2024</i>