

**ОТЗЫВ официального оппонента**  
**о диссертации на соискание ученой степени кандидата физико-**  
**математических наук Казачук Марии Андреевны**  
**на тему: «Динамическая аутентификация пользователей на основе**  
**анализа работы с клавиатурой компьютера»**  
**по специальности 05.13.11 – «Математическое и программное**  
**обеспечение вычислительных машин, комплексов и компьютерных**  
**сетей»**

**Актуальность выбранной темы**

Диссертация Казачук М.А. посвящена исследованию и разработке алгоритмов динамической аутентификации пользователей на основе анализа их работы с клавиатурой компьютера. Данная задача относится к классу задач поведенческой биометрии и её решение сводится к поиску аномалий в поведении пользователей, воспринимаемого как последовательности нажатий на клавиши с набором характеристик каждого нажатия. Решение задач такого типа является важной составляющей обеспечения компьютерной и общественной безопасности в ряде областей, например, в обеспечении контроля доступа к информации ограниченного пользования. Обширно используемые в настоящее время статические методы аутентификации на основе паролей или аппаратных авторизационных токенов уязвимы к компрометации, ими может завладеть третье лицо, и этот факт неразличим со стороны системы. Поэтому актуальны методы биометрической аутентификации, в том числе методы, основанные на анализе клавиатурного почерка. Основная сложность задачи обнаружения аномалий клавиатурного почерка состоит в том, что работа ведется над данными большой размерности, при этом многие признаки являются нерелевантными с точки зрения выделения целевого класса (поведение легитимного пользователя) либо взаимозависимыми, вследствие чего актуальной является проблема предобработки данных клавиатурного почерка пользователей. Для решения данной проблемы автором предлагается новый

подход к обработке данных, описывающих динамику работы пользователей с клавиатурой, позволяющий строить информативное и стабильное по времени признаковое пространство. Для поиска аномалий в данных в существующих работах для рассматриваемого класса задач используется kernel-подход, связанный с переходом в пространство признаков большей размерности (RKHS) и поиском аномалий непосредственно в нем. Однако, существующие kernel-методы поиска аномалий имеют ряд недостатков, связанных с формой и центром областей, которые они строят в RKHS для выделения аномалий. Для преодоления данных недостатков, в диссертационной работе предлагается нечеткий метод выявления аномалий в данных на основе эллиптической кластеризации в RKHS, строящий эллиптические контуры с оптимальным центром в пространстве признаков высокой размерности, что позволяет более точно описывать обучающую выборку и строить более точную модель соответственно. По результатам экспериментальных исследований, разработанный метод превзошел качество распознавания существующих решений. Поскольку зачастую необходимо определять степень аномальности не отдельного события, а множества последовательных событий, автором также предложен метод оценки степени аномальности целого набора событий на основе t-статистики Уэлша, ранее не применяемый для решения задач данного класса. Данный метод сравнивает отклики классификатора при классификации отдельных событий из валидационной выборки и новых тестовых данных и в качестве результата выдает степень аномальности целого набора событий. Таким образом, тема диссертационной работы Казачук М.А., безусловно, является актуальной.

## **Содержание диссертации**

Диссертация состоит из введения, четырех глав, заключения и списка использованной литературы. Общий объем диссертации составляет 155 страниц. Список литературы содержит 100 наименований.

Во введении приводится краткая характеристика работы: обосновывается актуальность темы исследования, описываются основные цели работы, научная новизна, перечисляются используемые методы

исследования и выносимые на защиту положения, приводятся сведения об аprobации диссертации и практическом использовании полученных результатов.

В первой главе проводится аналитический обзор существующих решений по тематике диссертационной работы. Автором были выявлены основные особенности задачи, а также проанализированы достоинства и недостатки существующих подходов. По результатам проведенного обзора соискателем формулируются направления дальнейших исследований.

Вторая глава посвящена исследованию и разработке алгоритмов обработки данных, описывающих динамику работы пользователей с клавиатурой компьютера. Соискателем предложен собственный подход, включающий способ построения признакового пространства и дальнейшего сокращения его размерности путем выделения наиболее стабильных признаков на основе использования критерия Колмогорова-Смирнова. Для решения проблемы мультимодального распределения признаков, автором предлагается подход на основе дискретизации признаков по квантилям. Использование предложенной комбинации алгоритмов позволило решить проблему падения качества распознавания пользователей с течением времени (в том числе и при смене используемого оборудования).

В третьей главе проводится исследование и разработка методов поиска аномалий в данных клавиатурного почерка пользователей. Автором предлагается новый эффективный нечеткий метод выявления аномалий, строящий эллиптические области с оптимальным центром в пространстве характеристик высокой размерности. Проводится оценка сложности алгоритма, доказывается его сходимость. Для подбора оптимальных значений метапараметров данного метода предлагается разработанный автором алгоритм, работающий с валидационной выборкой легитимного пользователя. Для получения агрегированной оценки степени аномальности поведения пользователя за продолжительный промежуток времени предлагается алгоритм на основе расчета t-статистики Уэлша, сравнивающий аномальности событий на валидационной и тестовой выборках. По результатам экспериментальных исследований, комбинация разработанных

алгоритмов превзошла качество работы существующих решений при классификации как отдельных событий, так и целого их набора.

Четвертая глава посвящена исследованию и разработке экспериментального образца программного комплекса (ЭО ПК), использующего комплекс разработанных в данной диссертации алгоритмов для обнаружения аномалий в поведении пользователей на основе анализа их клавиатурного почерка. Приводится детальное описание архитектуры и программной реализации разработанного ЭО ПК. Проводится экспериментальное исследование данного ЭО ПК, а также исследуется производительность его основных программных модулей.

В заключении подводятся итоги проделанной работы и описываются возможные направления дальнейших исследований.

Текст диссертации написан грамотным научным языком, логично выстроен. Автореферат правильно и достаточно полно отражает содержание диссертационной работы.

### **Научная новизна, обоснование, достоверность и практическая ценность полученных научных положений, выводов и рекомендаций**

Новые научные результаты исследования, предложенные автором, заключаются в:

- Предложенном новом подходе к уменьшению размерности признакового пространства путем выделения наиболее значимых признаков на основе уровня их стабильности;
- Предложенном новом методе выявления аномалий в данных на основе эллиптической кластеризации в RKHS. Данный метод позволяет строить контуры эллиптической формы с оптимальным центром для выделения аномалий в пространстве признаков высокой размерности (RKHS). Также разработан метод подбора оптимальных значений метапараметров алгоритмов одноклассовой классификации, использующий для работы данные только легитимного класса;
- Предложенном новом методе оценки степени аномальности поведения пользователя за длительный период его работы за компьютером,

позволяющем на основе последовательности откликов классификатора получать единое число – степень аномальности поведения пользователя за целую сессию его работы за компьютером.

Все результаты работы являются полностью обоснованными и достоверными. Достоверность результатов диссертации обоснована проработкой литературы по теме диссертационной работы, корректным и достаточно строгим применением методов математической статистики и теории машинного обучения, математическими доказательствами, а также большим количеством проведенных экспериментальных исследований по сравнению качества работы предложенных в диссертации методов с известными (в том числе, с использованием открытого набора тестовых данных). Диссертационная работа представляет собой полное и законченное исследование, выполнена на высоком научном уровне. По теме диссертации автором опубликовано 7 научных работ, три из которых – в изданиях, рецензируемых Web of Science, Scopus, RSCI. Опубликованные работы достаточно полно отражают содержание диссертации. Основные результаты диссертационного исследования докладывались и обсуждались на шести международных и всероссийских конференциях.

Практическая ценность диссертационной работы состоит в разработке и программной реализации экспериментального образца программного комплекса обнаружения аномалий в поведении пользователей на основе анализа работы с клавиатурой компьютера, использующего предложенные в работе алгоритмы. Данный экспериментальный образец программного комплекса был успешно использован в НИР «Разработка технологий биометрической идентификации пользователя по признакам, проявляющимся при использовании устройств ввода данных персональных ЭВМ» (Номер договора №01-04/15 от 08 апреля 2015 г.).

### **Замечания по диссертации**

По содержанию диссертационной работы можно сделать следующие замечания:

- В обзоре методов и средств динамической аутентификации пользователя на основе клавиатурного почерка рассмотрены преимущественно коммерческие системы. Это ограничивает возможности сравнения с существующими методами, так как в документации на коммерческие системы не приводятся математические основы используемых в них методов и описания алгоритмов.
- В исследовании устойчивости реализованного метода обнаружения аномалий клавиатурного почерка приводится описание различий аппаратных средств, используемых в эксперименте, но это описание недостаточно подробное – описаны различия в форме клавиш, но не описана геометрия самой клавиатуры и особенности взаимного расположения клавиш, что, возможно, затруднит воспроизведение эксперимента сторонними исследователями.

## **Заключение**

Вместе с тем, указанные замечания не умаляют значимости диссертационного исследования. Диссертация отвечает требованиям, установленным Московским государственным университетом имени М.В. Ломоносова к работам подобного рода. Содержание диссертации соответствует паспорту специальности 05.13.11 – «Математическое и программное обеспечение вычислительных машин, комплексов и компьютерных сетей» (по физико-математическим наукам), а также критериям, определенным пп. 2.1–2.5 Положения о присуждении ученых степеней в Московском государственном университете имени М.В. Ломоносова, а также оформлена согласно приложениям №5, 6 Положения о диссертационном совете Московского государственного университета имени М.В. Ломоносова.

Таким образом, соискатель Казачук Мария Андреевна заслуживает присуждения ученой степени кандидата физико-математических наук по специальности 05.13.11 – «Математическое и программное обеспечение вычислительных машин, комплексов и компьютерных сетей».

Официальный оппонент:  
кандидат физико-математических наук,  
старший научный сотрудник кафедры Информационной безопасности  
факультета Вычислительной математики и кибернетики  
МГУ им. М.В. Ломоносова  
ГАМАЮНОВ Денис Юрьевич

12.12.2019

Контактные данные:

Тел.: +7 (495) 939-00-91, e-mail: gamajun@seclab.cs.msu.su

Специальность, по которой официальным оппонентом  
зашита диссертация:

05.13.11 «Математическое и программное обеспечение  
вычислительных машин, комплексов и компьютерных сетей»

Адрес места работы:

119991, г. Москва, Ленинские горы,  
МГУ имени М.В. Ломоносова, 2-й учебный корпус  
Кафедра Информационной безопасности факультета  
Вычислительной математики и кибернетики  
МГУ имени М.В. Ломоносова  
Тел: +7 (495) 939-25-96; e-mail: cmc@cs.msu.ru

Подпись сотрудника  
факультета Вычислительной математики и кибернетики  
МГУ имени М.В. Ломоносова  
Д.Ю. Гамаюнова удостоверяю:

