ОТЗЫВ

официального оппонента

доктора технических наук, профессора Котенко Игоря Витальевича

на диссертацию Сулавко Алексея Евгеньевича

на тему «Высоконадежная биометрическая аутентификация на основе защищенного исполнения нейросетевых моделей и алгоритмов искусственного интеллекта»

по специальности 2.3.6. Методы и системы защиты информации, информационная безопасность,

представленную на соискание ученой степени доктора технических наук

Актуальность темы исследования

В настоящее время использование биометрических систем для задач аутентификации становится обычной практикой. Биометрия становится широко применимой технологией даже в тех сферах человеческой деятельности, где утечка подобного рода персональных данных может стать непоправимой ошибкой: банковская сфера, государственный сектор, объекты критической информационной инфраструктуры (КИИ). В связи с этим актуальными становятся исследования, посвященные защите биометрических систем от компьютерных атак и биометрических данных от компрометации. Диссертационная работа Сулавко А.Е. дает полное представление о состоянии выше указанного вопроса и представляет собой комплексное исследование научно-технической проблемы, заключающейся в повышении надежности многофакторной биометрической аутентификации и защищенности биометрических систем от компьютерных атак.

Биометрические системы уже сегодня оказываются подвержены множеству потенциальных атак и разрушительных факторов. Среди таковых можно выделить:

- проблемы с защитой биометрических данных от компрометации в результате их неправильного хранения или передачи;
- использование мощных алгоритмов ИИ, способных подделывать биометрические образы и осуществлять хакерские атаки;
- переход к открытым биометрическим образам (голос или лицо), которые в большей степени подвержены изменениям и дрейфу.

Именно поэтому современные биометрические системы должны быть основаны на устойчивых к разрушительным факторам методах искусственного интеллекта (ИИ). В связи с этим, диссертационная работа Сулавко А.Е., направленная на решение важной проблемы повышения надежности многофакторной биометрической аутентификации и защищенности биометрических систем от компьютерных атак на основе технологии защищенного исполнения нейросетевых моделей и алгоритмов искусственного интеллекта, является актуальной.

Оценка структуры и содержания работы

Диссертационная работа состоит из введения, пяти глав, заключения, списка сокращений, списка литературы и приложений. Основной текст диссертации изложен на 391 странице, содержит 108 рисунков и 28 таблиц.

Во введении обосновывается актуальность темы исследования, дается общая характеристика основных научных положений и результатов работы.

В первой главе проведен анализ достигнутых результатов в области биометрической аутентификации и создания систем доверенного ИИ, а также существующих национальных и международных стандартов по защите ИИ от компьютерных атак. Даны определения ключевых понятий, а также описаны основные критерии доверенного ИИ.

Во второй главе разработана концепция защищенного исполнения нейросетевых алгоритмов искусственного интеллекта, а также модель корреляционных нейронов для высоконадежной биометрической аутентификации и нейросетевой преобразователь биометрия-код на их основе. Проведена серия вычислительных экспериментов, чтобы показать эффективность выбранного подхода и определить свойства нового класса нейронов и очертить границы применимости предложенной концепции.

В третьей главе проведён анализ подходов к построению адаптивных моделей искусственного интеллекта (моделей, способных к онлайн-обучению), включая методы глубокого обучения с подкреплением, эволюционный и иммунный подходы. Предложена модель искусственной иммунной сети и модель детекторов, на базе которых формируется эта сеть. Разработаны алгоритмы обучения искусственной иммунной сети (первый алгори́тм оффлайн-обучения с учителем, второй – онлайн-дообучения с подкреплением). Эффективность предложенных технических решений оценивалась путем проведения экспериментов с открытыми наборами данных клавиатурного почерка, а также с набором данных, собранным соискателем в рамках диссертационного исследования.

В четвёртой главе представлены комплексный метод и алгоритм трехфакторной аутентификации с последовательным предоставлением образов слухового канала, голоса и подписи. Помимо обычной подписи (автографа) может использоваться рукописный пароль – тайный биометрический образ, который следует держать в секрете (как и голосовой пароль). Образ слухового канала – это относительно новый тип биометрии. Параметры длины, ширины и других геометрических показателей ушного канала регистрируются с помощью обычного малогабаритного микрофона, встроенного в корпус наушников. В ходе работы проведены эксперименты по оценке надежности алгоритма трехфакторной аутентификации, а также каждого биометрического фактора по отдельности.

В пятой главе обозначены ключевые положения разработанной технологии, которую можно отнести к категории автоматического машинного обучения (AutoML), а также первой редакции разработанного национального стандарта, основанного на данной технологии. Проект стандарта хотелось бы отметить отдельно и отнести к значимым составляющим работы и результатам. Стандарт планируется применять для защиты компонентов ИИ, используемых на объектах КИИ. Приведено описание внедрения результатов работы на примере ряда организаций. Среди прочих внедрений можно отметить разработанную линейку программных продуктов. Результаты работы прошли очень хорошую апробацию и широко задействованы на практике.

В заключении приводятся основные результаты и выводы по проведенной работе.

Степень обоснованности научных положений, выводов и рекомендаций, сформулированных в диссертации

Основные положения и выводы диссертационной работы являются обоснованными. Они базируются на грамотном применении методов теории информации и математической статистики, а также методов спектрального и корреляционного анализа, машинного обучения, кодирования информации и защиты данных от компрометации, а также аппарата искусственных нейронных сетей.

Основные научные положения, выводы и рекомендации, сформулированные в диссертации, подтверждаются широтой их апробации в открытой печати. По проблеме диссертационного исследования опубликовано 80 работ, в том числе 38 статей в журналах из Перечня рецензируемых научных изданий, рекомендованных ВАК, либо в научных изданиях, индексируемых базой RSCI, 21 научная работа в изданиях, включенных в базы Web of Science и Scopus, 11 научных работ в других изданиях и одна коллективная монография. Получен один патент на изобретение и 8 свидетельств о регистрации программ.

Достоверность и новизна полученных результатов

Достоверность основных выводов и результатов диссертационного исследования подтверждается фактами обсуждения результатов исследования на профильных международных и российских конференциях, апробацией полученных результатов в организациях и учреждениях, а также экспериментальными исследованиями с использованием разработанных соискателем моделей и методов.

Все основные положения диссертационной работы и результаты исследований, полученные в рамках разработанной концепции, являются новыми, а именно:

- концепция защищенного исполнения нейросетевых алгоритмов ИИ, которая позволяет сформировать устойчивость модели к извлечению знаний;
- модель корреляционных нейронов и сформированную на их основе модель нейросетевого преобразователя биометрия-код, повышающего защищенность биометрических данных от компрометации и длину ключа, связываемого с биометрическим образом;
- адаптивная нейро-иммунная модель и алгоритмы ее пакетного обучения с учителем и онлайн-обучения с подкреплением, позволяющие предупредить или снизить влияние концептуального дрейфа;
- методы и алгоритм высоконадежной многофакторной аутентификации на основе рукописных и голосовых образов, а также акустических образов уха с обеспечением защиты биометрических данных от компрометации;
- технология автоматического синтеза и обучения нейросетевых моделей доверенного ИИ для высоконадежной биометрической аутентификации и других ответственных приложений ИИ.

О новизне результатов работы также говорит наличие патента на изобретение.

Основные позиции диссертационной работы, выносимые на защиту, подробно сформулированы. Публикации по теме диссертации, включая издания, рекомендованные ВАК РФ, в полной мере отражают основные положения диссертационной работы, обладающие новизной.

Работа поддержана множеством грантов РНФ, РФФИ, Минобрнауки РФ, Фонда содействия инновациям, что также говорит в пользу новизны результатов исследования.

Теоретическая и практическая значимость полученных автором результатов

Результаты работы Сулавко А.Е. выходят немного дальше, чем очерченный круг проблем высоконадежной биометрической аутентификации. Значение результатов для теории биометрической аутентификации, распознавания образов и ИИ заключается в предложенном математическом аппарате, включающем концепцию, модели, методы и алгоритмы машинного обучения на малых выборках. Данный аппарат разрабатывался несколько лет при поддержке РНФ, РФФИ и Минобрнауки РФ.

Практическая ценность результатов диссертации заключается в том, что в ней разработаны проект национального стандарта, который может быть использован не только в области информационной безопасности и биометрической аутентификации в частности, но и в других приложениях ответственного ИИ. Также разработано алгоритмическое, программное и методологическое обеспечение для построения систем высоконадежной биометрической аутентификации, а также систем доверенного ИИ на основе защищенного исполнения нейросетевых моделей и алгоритмов ИИ.

Практическая значимость работы подтверждается восемью актами внедрения (актами использования) результатов работы.

Замечания по диссертационной работе

1. Для обоснования эффективности разработанных алгоритмов обучения нейросетевых моделей ИИ автор приводит исчерпывающие аргументы и доводы, подкрепленные результатами эмпирических исследований и экспериментов с реальными данными. Тем не менее, в работе не приводится строгого математического доказательства сходимости алгоритмов:
   - синтеза и обучения сети корреляционных нейронов;
   - обучения искусственной нейро-иммунной модели с учителем;
   - дообучения искусственной нейро-иммунной модели с подкреплением.

Данный недостаток может быть устранен в рамках будущих исследований, однако автору следовало явно обозначить это, как направление дальнейшего развития темы.

2. Оценки необходимых вычислительных ресурсов и времени, затрачиваемых на процедуру аутентификации с использованием разработанных нейросетевой и нейро-иммунной моделей ИИ, даны весьма поверхностно. В параграфе 5.1 «Границы применимости разработанных методов и технологии» приведены интервальные оценки количества пользователей, которые могут быть обслужены одним процессорным ядром, однако не разъясняется, что означает «при очень высокой плотности запросов».

3. В работе не приводится обоснования выбора архитектур многослойных сверточных нейронных сетей. В частности, автор не приводит аргументов при выборе конфигураций автокодировщиков, применяемых для извлечения признаков из образов наружного уха, а также искусственных нейронных сетей, применяемых для классификации образов наружного уха в главе 4. Хотя эти нейронные сети не являются ключевыми для достижения результатов, автору следовало с большим вниманием отнестись к выбору гиперпараметров нейросетевых моделей.

4. Не все понятия разъяснены в достаточной степени. Например, в работе часто встречаются термины «синтетический образ» и «естественный образ» (пример, образец данных). Однако не приводятся определений этих терминов. Хотя при ознакомлении с работой становится ясен смысл этих понятий, их значение стоило раскрыть в первой главе, где описана терминология. Кроме того, термин модель ИИ (машинного обучения), который употребляется по тексту всей работы, может быть неверно интерпретирован, так как в первой главе приводится определение термина шаблон/модель (template/model), имеющего несколько иное значение.

5. Некоторые промежуточные результаты приводятся недостаточно детально. В частности, в работе отсутствуют результаты проверки гипотезы о нормальном распределении значений используемых признаков. Автор ограничивается констатацией факта, что эти проверки были выполнены. Однако не указано, каким образом это было реализовано.

6. В исследовании присутствуют константы, которые не подкреплены должным объяснением:

- не ясно, исходя из каких соображений, выбраны пороговые значения 0,1, 0,4 и 0,6 на рисунке 2.16.
- не в полной мере понятно, почему во второй главе при проведении вычислительного эксперимента выбраны именно такие показатели информативности генерируемых признаков: $I \approx 1$, $I \approx 0{,}5$, $I \approx 0{,}15$.

Заключение

Несмотря на изложенные замечания, они не влияют на общую положительную оценку диссертационной работы, в которой получены новые теоретические результаты и методологические подходы, основанные на использовании автором современной методологии обеспечения информационной безопасности, а также перспективных методов машинного обучения и искусственного интеллекта, что свидетельствует о широком кругозоре и высокой научной квалификации соискателя.

Диссертация Сулавко А.Е., представленная на соискание ученой степени доктора технических наук, обладает внутренним единством, научной новизной, теоретической и практической значимостью и является завершенной научно-квалификационной работой, в которой получены научно-обоснованные результаты, направленные на решение актуальной проблемы повышения надежности многофакторной биометрической аутентификации и защищенности биометрических систем от компьютерных атак на основе технологии защищенного исполнения нейросетевых моделей и алгоритмов искусственного интеллекта.

Автореферат раскрывает основное содержание диссертации. Количество и содержание публикаций соответствуют требованиям, предъявляемым к докторским диссертациям.

Тема диссертации и область исследования соответствуют специальности 2.3.6. Методы и системы защиты информации, информационная безопасность.

На основании вышеизложенного, считаю, что представленная к защите диссертационная работа Сулавко А.Е. удовлетворяет требованиям п. 9 Положения о порядке присуждения ученых степеней, предъявляемым к диссертациям на соискание ученой степени доктора технических наук, а её автор — Сулавко Алексей Евгеньевич

заслуживает присуждения ему ученой степени доктора технических наук по специальности 2.3.6. Методы и системы защиты информации, информационная безопасность.

Официальный оппонент:

Доктор технических наук, профессор,
главный научный сотрудник
лаборатории проблем компьютерной безопасности,
Федеральное государственное бюджетное учреждение науки
«Санкт-Петербургский Федеральный исследовательский центр
Российской академии наук»

Котенко Игорь Витальевич

Докторская диссертация защищена по специальностям:
20.02.13 - "Информатика и компьютерные технологии в военном деле"
20.01.09 - "Военные системы управления и связи"

Даю согласие на обработку персональных данных.

Адрес места основной работы: 199178, г. Санкт-Петербург, 14-я линия Васильевского острова, 39
Рабочий телефон: +7 (812) 328 34 11
Адрес эл. почты: ivkote@comsec.spb.ru

Подпись руки Котенко И.В. заверяю
Начальник отдела кадров СПб ФИЦ РАН
Д.В.Токарев
20 24 г.