ОТЗЫВ ОФИЦИАЛЬНОГО ОППОНЕНТА

на диссертацию Обыденкова Дмитрия Олеговича

«Методы противодействия анонимности при утечках текстовых документов посредством цифровых водяных знаков»,

представленную на соискание ученой степени кандидата технических наук по специальности 2.3.5 – «математическое и программное обеспечение вычислительных систем, комплексов и компьютерных сетей»

Использование информационных систем в коммерческих и государственных организациях повышает их эффективность, но также увеличивает риски утечек конфиденциальных данных и коммерческой тайны. Утечка финансовой информации, технической документации и сведений о клиентах может нанести значительный материальный и репутационный ущерб. Более половины инцидентов утечки конфиденциальной информации связаны с внутренними нарушителями, что подчеркивает необходимость разработки эффективных методов защиты от внутреннего нарушителя. Современные системы предотвращения утечек данных (Data Leakage Prevention – DLP) системы направлены на защиту от утечек данных через цифровые каналы (электронная почта, веб-сайты, файлообменники, мессенджеры, съемные носители и др.), однако они малоэффективны при утечках с помощью фотографирования экрана с выведенной конфиденциальной информацией или сканирования / фотографирования напечатанных документов. Перспективным направлением является внедрение в текстовые документы цифровых водяных знаков (ЦВЗ), позволяющих идентифицировать источник утечки. Разработка методов внедрения ЦВЗ требует поиска баланса между информационной емкостью, незаметностью и стойкостью водяных знаков к различным искажениям, включая фотографирование документа с ЦВЗ с экрана и печать документа с внедренной ЦВЗ с последующей оцифровкой

сканированием или фотографированием. Исследователи работают над проблемой более 30 лет, однако существующие решения имеют различные особенности, ограничивающие применение подобных систем защиты информации на практике. Таким образом тема диссертации Д.О. Обыденкова, нацеленная на устранение ограничений подобных систем путем разработки и реализации системы защиты текстовых документов от анонимных утечек с использованием ЦВЗ, является актуальной и востребованной.

Диссертация состоит из введения, пяти глав, заключения и списка литературы из 67 наименований. Общий объем диссертации 164 страницы с 48 рисунками, 24 таблицами и двумя приложениями.

Во **введении** формулируются цели и задачи диссертационной работы, обосновывается актуальность темы исследования, теоретическая и практическая значимость работы, обозначается научная новизна и основные положения, выносимые на защиту.

Глава 1 содержит обзор существующих решений защиты от утечек информации на основе ЦВЗ и опубликованных методов внедрения ЦВЗ в текстовые документы, устойчивых к искажениям при печати с последующим сканированием / фотографированием и фотографировании выведенного на экран документа.

Глава 2 содержит описание архитектуры системы противодействия анонимности при утечках документов, подходе формирования идентификатора сотрудника и устройства, а также механизмов интеграции методов внедрения ЦВЗ в текстовые документы при печати и выводе на экран.

Глава 3 посвящена разработанному методу внедрения ЦВЗ в текстовые документы при печати. Представлен метод получения разметки изображения текстового документа на основе нейросетевого алгоритма и

описаны структурные механизмы кодирования информации, устойчивые к искажениям при печати с последующей оцифровкой фотографированием.

Глава 4 посвящена созданному методу генерации ЦВЗ нейросетевым алгоритмом для защиты документов при выводе на экран. Описаны особенности обучения и использования нейросетевых моделей генерации ЦВЗ для отрисовки в рамках окна-оверлея. Также представлен алгоритм извлечения внедренного ЦВЗ из фотографии экрана с выведенным документом.

Глава 5 посвящена тестированию реализованной системы противодействия анонимности при утечках текстовых документов. Рассмотрены сценарии утечки изображения распечатанного документа и фотографии экрана с выведенным документом.

В **заключении** приведены основные результаты работы.

В работе получены следующие новые научные результаты:

1. Структурный метод внедрения ЦВЗ с использованием нейросетевой сегментации изображения текстового документа, поддерживающий слепое извлечение, устойчивый к искажениям при печати и оцифровке через фотографирование и сканирование, с низким потреблением вычислительных ресурсов;
2. Метод наложения статических ЦВЗ, генерируемых нейросетевой моделью, поверх выводимых на экран текстовых документов, с возможностью слепого извлечения информации из фотографий экрана и устойчивостью к алгоритмам сжатия изображений в мессенджерах.

Все положения и выводы достоверны и научно обоснованы. Достоверность полученных результатов подтверждается результатами экспериментов на открытых наборах данных. Разработанные методы и алгоритмы основываются на корректном применении аппарата

вычислительной математики, методов оптимизации и методов машинного обучения. Результаты, полученные в ходе диссертационной работы, представлены в 8 публикациях: 5 из которых опубликованы в журналах, рекомендованных ВАК, 1 индексируется Scopus и 2 опубликованы в материалах конференции. Получено 5 свидетельств о государственной регистрации программ для ЭВМ.

Практическая значимость состоит в разработке и реализации системы защиты документов от анонимных утечек с использованием цифровых водяных знаков. Тестирование подтвердило высокую точность извлечения ЦВЗ при утечке изображений, полученных фотографированием как распечатанных, так и отображаемых на экране документов. Система защиты документов от анонимных утечек внедрена организацией ООО "СиТ" (акт о внедрении №612/0924 от 29.09.24).

В качестве замечаний к работе, не снижающих ее общего высокого уровня, следует отметить:

1. В разделе 2.2 описан механизм формирования идентификатора сотрудника и устройства, основанный на использовании серийного номера жесткого диска. Однако данный подход недостаточно надежен в условиях виртуализированной среды, где серийные номера могут быть легко изменены или клонированы.
2. В разделе 2.3 подчеркивается важность механизмов обнаружения и исправления ошибок во внедряемой битовой последовательности, однако в последующих разделах отсутствуют расчеты вероятности возникновения ошибок или статистика их частоты при извлечении информации из ЦВЗ.
3. В механизме кодирования информации посредством горизонтального смещения слов, представленном в разделе 3.2.1, не указана формула

для расчета ширины пробелов, используемых при внедрении ЦВЗ в текстовый документ.

4. В разделе 4.2 описан процесс обучения нейросети для внедрения с использованием набора аугментаций, включая циклический сдвиг, аффинные преобразования и добавление шумов. Для повышения эффективности рекомендуется интегрировать нейронную сеть, имитирующую искажения, возникающие при фотографировании экрана, предложенную в исследовании Fang и соавторов (Fang H., Jia Z., Ma Z., Chang E. C., Zhang W. PIMoG: An effective screen-shooting noise-layer simulation for deep-learning-based watermarking network // Proceedings of the 30th ACM international conference on multimedia. – 2022. – С. 2267-2275).
5. В подразделе 5.1.1 проведена оценка максимальной емкости текстовых документов при полной заполненности страницы текстом, однако отсутствует анализ средней емкости документов в реальных условиях эксплуатации.
6. В подразделе 5.1.3 исследуются различные варианты искажений и оценка устойчивости ЦВЗ к этим искажениям, но множество исследованных искажений довольно ограничено – в частности, рассмотрен вариант двукратной печати и сканирования, но не рассмотрена зависимость от числа повторных попыток печати и сканирования с различными настройками качества печати.

Автореферат правильно отражает содержание диссертации и позволяет достаточно точно оценить основные полученные результаты, степень их обоснованности и достоверности.

Диссертационная работа Обыденкова Д.О. по теме «Методы противодействия анонимности при утечках текстовых документов посредством цифровых водяных знаков» является законченным научным

исследованием, основное содержание диссертации отражено в опубликованных статьях и обсуждено на научных конференциях.

Диссертационная работа отвечает требованиях ВАК РФ, предъявляемым к кандидатским диссертация, её содержание соответствует паспорту специальности 2.3.5 «Математическое и программное обеспечение вычислительных систем, комплексов и компьютерных сетей», а её автор, Обыденков Дмитрий Олегович, заслуживает присуждения ученой степени кандидата технических наук по вышеуказанной специальности.

Официальный оппонент:

Кандидат физико-математических наук Гамаюнов Денис Юрьевич

доцент кафедры информационной безопасности факультета вычислительной математики и кибернетики Федерального государственного бюджетного образовательного учреждения высшего образования «Московский государственный университет имени М.В.Ломоносова»

«29» ноября 2024 г.