Экз. №___

Отзыв официального оппонента
на диссертационную работу Обыденкова Дмитрия Олеговича
«Методы противодействия анонимности при утечках текстовых документов посредством цифровых водяных знаков», представленной на соискание ученой степени кандидата технических наук по специальности 2.3.5 – «Математическое и программное обеспечение вычислительных систем, комплексов и компьютерных сетей»

Актуальность темы работы. Повсеместное внедрение автоматизированных систем (АС) влечет за собой рост угроз информационной безопасности (ИБ). Эти угрозы создают нарушители, причем большинство угроз создается внутренним нарушителем. Одним из типов средств противодействия угрозам внутреннего нарушителя являются средства защиты от утечек (DLP). Данные средства обычно решают одну из двух задач: 1) предотвращение реализации угрозы ИБ, например, путем пресечения попытки нарушителя передать конфиденциальную информацию за пределы локальной сети; 2) обеспечение расследования состоявшегося инцидента утечки информации.

При решении второй задачи важным является определить источник произошедшей утечки, то есть деанонимизировать нарушителя. С учетом того, что все действия операторов АС должны быть подотчетны, определить источник утечки цифровой информации обычно не составляет труда. Однако для случая утечки информации, представленной в аналоговом виде, это не так. Нарушитель может выполнить внешними по отношению к защищаемой АС средствами фотографирование экрана монитора либо распечатанного конфиденциального документа. В этом случае для обеспечения дальнейшего расследования инцидента стороне защиты было необходимо внести в отображаемое изображение либо распечатываемый бумажный документ дополнительную информацию, позволяющую впоследствии при необходимости идентифицировать нарушителя.

К вносимой дополнительной информации, называемой цифровым водяным знаком (ЦВЗ) обычно предъявляются противоречивые требования по пропускной способности, незаметности для оператора, робастности

к искажениям печати/сканирования и фотографирования. Несмотря на то, что в этой области за последние 25 лет выполнено большое количество исследований, приемлемого для практики результата достичь не удалось. Поэтому тема диссертационной работы Обыденкова Д.О. является актуальной.

Научная новизна и практическая значимость исследований. В диссертации Обыденковым Д.О. представлены результаты, обладающие научной новизной:

- разработан структурный метод внедрения ЦВЗ на основе сегментации текстового документа с помощью нейросетевых алгоритмов с возможностью слепого извлечения встроенной информации, устойчивый к искажениям, возникающим при распечатывании и последующей оцифровке через фотографирование или сканирование, оптимизированный для работы на процессоре общего назначения с минимальным использованием вычислительных ресурсов.

- предложен метод внедрения статических ЦВЗ, сгенерированных нейросетевым алгоритмом, в текстовые документы с возможностью слепого извлечения внедренной информации из фотографии экрана, устойчивый к алгоритмам сжатия изображений, применяемым в мессенджерах.

Новизна этих результатов заключается в том, что для извлечения встроенной информации не требуются оригиналы документов, кроме того, надежное извлечение выполняется в условиях искажений, вызванных сканированием и фотографированием.

Важным **теоретическим результатом** является использование нейросетевых алгоритмов для сегментации и внедрения информации в текстовые документы, что позволяет добиться высокой устойчивости к искажениям, возникающим в процессе передачи изображения, и минимизировать визуальные изменения, что делает методы практически незаметными для пользователя.

Практическим результатом работы является то, что на основе полученных теоретических результатов реализованы программные средства. Тестирование программных средств показало их высокую эффективность. Реализованная система противодействия анонимности при утечках текстовых документов внедрена организацией ООО "СиТ" (акт о внедрении № 612/0924 от 29.09.24).

Обоснованность и достоверность научных положений и выводов. Обоснованность и достоверность полученных результатов обусловлена применением в исследовании системного подхода, методов математического моделирования и статистической обработки данных.

Основные результаты диссертации опубликованы в 8 научных работах, в том числе 5 научных статьях в рецензируемых журналах, включенных ВАК в перечень ведущих периодических изданий, а также 1 публикации Scopus.

Результаты диссертационной работы докладывались и обсуждались на многочисленных конференциях и научных семинарах и достаточно широко опубликованы. Имеется 5 свидетельств о государственной регистрации программы для ЭВМ.

Содержание автореферата соответствует содержанию диссертации.

Краткая характеристика основного содержания диссертации. Диссертация Д.О. Обыденкова состоит из введения, пяти глав, заключения и списка литературы.

Во **введении** обосновывается актуальность исследования, проводимого в рамках диссертационной работы, ставятся цели и задачи, формулируется научная новизна и практическая значимость работы, а также приводятся основные положения, выносимые на защиту.

В **первой главе** выполнен обзор актуальных методов защиты информации от утечек, используемых в области информационной безопасности. Проведен анализ возможных каналов утечек, по отношению к которым противодействие существующих средств защиты недостаточно, сформулированы возможные сценарии утечек.

Во **второй главе** представлена архитектура системы деанонимизации при утечках текстовых документов, выводимых на печать или экран, на основе методов внедрения в них ЦВЗ, содержащих уникальные идентификаторы сотрудников и используемых ими устройств.

В **третьей главе** описан разработанный структурный метод внедрения ЦВЗ, предполагающий слепое извлечение внедренной информации, на основе сегментации изображения документа с помощью нейросетевого алгоритма, обладающего устойчивостью к искажениям, возникающим при распечатывании и последующей оцифровке посредством фотографирования или сканирования, и ориентированный на работу

на процессоре общего назначения с минимальным потреблением вычислительных ресурсов.

В **четвертой главе** представлен метод генерации ЦВЗ с применением нейросетевого алгоритма, предполагающего слепое извлечение внедренной информации. Генерируемый ЦВЗ обладает свойствами визуальной незаметности и устойчивости к искажениям, возникающим при фотографировании экрана и сжатии алгоритмами, применяемыми в мессенджерах при отправке изображений.

В **пятой главе** приведены результаты тестирования реализованных методов противодействия анонимности при утечках текстовых документов посредством ЦВЗ со слепым декодированием, обеспечивающих устойчивость к искажениям, возникающим при печати или фотографировании отображаемых на экране документов с последующей передачей изображения через мессенджеры, а также имеющих визуальную незаметность и не вызывающих дискомфорта у пользователей. Результаты экспериментов показали высокую эффективность разработанных решений.

В **заключении** приведены основные результаты работы, а также намечены основные направления дальнейших исследований.

Несмотря на общую положительную оценку работы, к содержанию работы могут быть сделаны следующие **замечания**:

1) В разделе 2.3 утверждается, что добавление к идентификатору БЧХ-кодов является оптимальным подходом для исправления ошибок в битовых последовательностях небольшой длины. Однако утверждение об оптимальности в работе не доказывается.

2) В подразделе 3.1.2 при описании экспериментов по оптимизации нейросетевой модели наилучшие результаты получены при дистилляции обученной нейросетевой модели архитектуры U-Net на другую модель архитектуры U-Net с меньшим числом слоев и карт признаков, а также при конвертации типа весов модели к 16-битному вещественному типу. Однако, каких-либо объяснений тому, почему именно так получилось, в работе не приведено.

3) Как отмечено в главе 5, наибольшей проблемой для разработанного метода является эффект муара, проявляющийся индивидуально для комбинации камеры, монитора, а также расстояния и угла съемки. Возможно, было бы целесообразно исследовать методы устранения муара.

Указанные замечания не снижают значимости полученных результатов и не влияют на общую положительную оценку диссертационного исследования.

Заключение

В диссертационной работе Д.О. Обыденкова решена актуальная задача противодействия анонимности при утечках текстовых документов на основе ЦВЗ со слепым декодированием, обеспечивающих устойчивость к искажениям, возникающим при печати или фотографировании отображаемых на экране документов с последующей передачей изображения через мессенджеры, а также имеющих визуальную незаметность и не вызывающих дискомфорта у пользователей.

Уровень решаемых задач соответствует требованиям, предъявляемым к кандидатам технических наук. Содержание диссертации соответствует специальности 2.3.5 – Математическое и программное обеспечение вычислительных систем, комплексов и компьютерных сетей.

Диссертация является завершенной научно-квалификационной работой, которая по критериям актуальности, научной новизны, обоснованности и достоверности выводов соответствует требованиям п.7 «Положения о присуждении ученых степеней», а ее автор, Обыденков Дмитрий Олегович, заслуживает присуждения ему ученой степени кандидата технических наук по специальности 2.3.5 – Математическое и программное обеспечение вычислительных систем, комплексов и компьютерных сетей.

Официальный оппонент
Главный научный сотрудник
АНО «Институт инженерной физики»
доктор технических наук, доцент

Грибунин Вадим Геннадьевич

26.11.2024