на правах рукописи

Обыденков Дмитрий Олегович

Методы противодействия анонимности при утечках текстовых документов посредством цифровых водяных знаков

Специальность 2.3.5 – математическое и программное обеспечение вычислительных систем, комплексов и компьютерных сетей

Автореферат

диссертации на соискание ученой степени
кандидата технических наук

Москва – 2024

Работа выполнена в Федеральном государственном бюджетном учреждении науки Институте системного программирования им. В.П. Иванникова Российской Академии Наук.

Научный руководитель: **Маркин Юрий Витальевич,** кандидат технических наук

Официальные оппоненты: **Грибунин Вадим Геннадьевич**, доктор технических наук, доцент, главный научный сотрудник Автономной некоммерческой организации «Институт инженерной физики»

Гамаюнов Денис Юрьевич, кандидат физико-математических наук, доцент кафедры информационной безопасности факультета вычислительной математики и кибернетики Федерального государственного бюджетного образовательного учреждения высшего образования «Московский государственный университет им.М.В.Ломоносова»

Ведущая организация: Федеральное государственное учреждение «Федеральный исследовательский центр «Информатика и управление» Российской академии наук»

Защита диссертации состоится 17 декабря 2024 г. в 16:30 на заседании диссертационного совета 24.1.120.01 при Федеральном государственном бюджетном учреждении науки Институте системного программирования им. В.П. Иванникова Российской Академии Наук по адресу: 109004, г. Москва, ул. Александра Солженицына, д. 25.

С диссертацией можно ознакомиться в библиотеке и на сайте Федерального государственного бюджетного учреждения науки Институте системного программирования им. В.П. Иванникова Российской Академии Наук.

Автореферат разослан «__» _____ 2024 г.

Ученый секретарь
диссертационного совета 24.1.120.01,
кандидат физико-математических наук Зеленов С.В.

Актуальность. Использование информационных систем в коммерческих и государственных организациях приносит значительные выгоды, однако с их внедрением возникают новые угрозы, в частности, угрозы утечки информации. Данные о финансах, технологиях, сотрудниках и клиентах крайне важны для организаций, утечка подобных сведений может нанести серьезный финансовый и репутационный ущерб. Отчеты по инцидентам информационной безопасности подтверждают рост числа утечек. Большинство инцидентов связано с «инсайдерами» — сотрудниками компаний, действующими в сговоре с внешними нарушителями. Согласно исследованию InfoWatch в России доля утечек из-за внутренних нарушителей достигает 79% от общего числа инцидентов.

Для защиты от утечек используются Data Leakage Prevention (DLP) системы, которые могут работать в режиме реального времени или предотвращать возможные утечки превентивно. DLP-системы ориентированы на предотвращение утечек данных через сетевые каналы и не способны эффективно защищать *аналоговые* каналы утечек. К аналоговым сценариям утечек относятся фотографирование выведенного на экран документа или печать документа с последующей оцифровкой за пределами защищаемого контура при помощи сканера или фотоаппарата.

Методы противодействия утечкам информации через фотографирование документов на экране и/или их распечатку делятся на организационные и технические. Технические меры обычно подразумевают нанесение на документы *цифровых водяных знаков* (ЦВЗ) различных типов. ЦВЗ могут содержать информацию, позволяющую деанонимизировать пользователя, ставшего причиной утечки данных. Водяные знаки документов могут быть как явными, так и малозаметными или невидимыми для обнаружения и считывания скрытой информации невооруженным глазом. Методы внедрения малозаметных ЦВЗ представляют особый интерес, поскольку они минимально влияют на удобство работы пользователей с документами.

Разработка методов внедрения ЦВЗ сопровождается поиском баланса между незаметностью изменений в документе, информационной емкостью и устойчивостью к искажениям. Классические подходы, работающие в домене преобразований (например, дискретное косинусное или быстрое преобразование Фурье), либо вносят слишком заметные изменения в изображения документов, либо недостаточно устойчивы к искажениям, возникающим в аналоговых каналах утечек. Вместе с этим, количество внедряемой в ЦВЗ информации как правило уменьшается при изменении параметров метода в сторону повышения незаметности или устойчивости. Разработка методов внедрения ЦВЗ, сочетающих в себе характеристики, позволяющие эффективно решать практические задачи, является актуальной задачей для исследователей в области стеганографии.

Применительно к текстовым документам перспективными являются структурные методы внедрения ЦВЗ. Существующие решения, в том числе EveryTag, используют структурные ЦВЗ для деанонимизации утечек посредством напечатанных документов, однако, не предполагают возможность работы в так называемом *слепом* сценарии – для извлечения внедренной информации требуется наличие оригинала документа. Необходимость хранения оригинальных документов накладывает значительные ограничения (в том числе, создание и поддержку единой базы конфиденциальных документов) на применимость подобных методов, поэтому исследование и разработка методов внедрения ЦВЗ с возможностью слепого извлечения встроенной информации, нацеленных на предотвращение анонимных утечек текстовых документов через распечатанные копии, является актуальной задачей.

Методы нанесения ЦВЗ для защиты документов, отображаемых на экране, делятся на динамические и статические. Динамические методы характеризуются перестроением водяного знака для адаптации под содержимое экрана, что может потребовать значительных вычислительных ресурсов системы, а также вызывать повышенную утомляемость пользователей из-за частых изменений на экране. Статические методы знаки демонстрируют меньшую заметность по метрикам PSNR/SSIM, но обладают низкой устойчивостью к искажениям, возникающим при фотографировании экрана. Для практического применения ЦВЗ также должен быть устойчивым к передаче изображения через мессенджеры, то есть сохранять информацию о распространителе при перекодировании и уменьшении размера изображения. Протестировать существующие на рынке коммерческие системы невозможно, а заявляемые в них характеристики ЦВЗ не подкреплены научными публикациями. Разработка методов нанесения водяных знаков, обеспечивающих низкую заметность для комфортной работы и высокую устойчивость к искажениям при утечках фотографий экрана с выведенным конфиденциальным документом, является актуальной задачей.

Степень разработанности темы. Отечественные и зарубежные исследователи публиковали работы в области внедрения ЦВЗ в контейнеры различных доменов, включая текст, изображения, аудио, видео и прочих. Методы, работающие в домене преобразований, развивали такие ученые, как I. Cox, T. Furon, A. Pramila, P. Dong и др. Внедрение водяных знаков в пространственную область представлено в работах J. Brassil, S. Low, N. Maxemchuk, M. Topkara, Y. Kim, А.А. Грушо, В.О. Писковским, Д.А. Семинихиным и др. За последние несколько лет появились нейросетевые методы, в частности, в работах J. Zhu, M. Tancik, W. Zhang, P. Fernandez и др.

Целью диссертационной работы является разработка 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.

Основные задачи:

1. Разработка архитектуры системы противодействия анонимности при утечках текстовых документов. Система должна обеспечивать внедрение в текстовые документы информации, позволяющей устанавливать виновников публичных утечек;
2. Разработка метода внедрения ЦВЗ в текстовые документы при печати, предполагающего слепое извлечение встроенной информации. Разработанный метод должен обладать устойчивостью к различным искажениям и преобразованиям, сопутствующим печати документа с последующей оцифровкой посредством сканирования или фотографирования. Внедренный в документ ЦВЗ должен быть визуально незаметен. Внедрение ЦВЗ не должно оказывать существенного влияния на скорость печати документов;
3. Разработка метода внедрения ЦВЗ в текстовые документы при выводе на экран, предполагающего слепое извлечение встроенной информации. Разработанный метод должен обладать устойчивостью к различным искажениям и преобразованиям, сопутствующим фотографированию выведенного на экран документа с последующей отправкой фотографии документа через мессенджер. Наличие ЦВЗ не должно вызывать дискомфорт у пользователей при использовании;
4. Реализовать систему противодействия анонимности при утечках текстовых документов с использованием разработанных методов и провести оценку их эффективности.

Научной новизной обладают следующие результаты работы:

1. 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;
2. 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.

Теоретическая значимость

Теоретическая значимость диссертации заключается в разработке и усовершенствовании методов защиты текстовых документов от утечек информации через анонимные каналы с помощью цифровых водяных знаков. В работе предложены новые решения, направленные на предотвращение несанкционированной передачи информации через печатные документы и отображаемые на экране, что расширяет научные представления в области внедрения цифровых водяных знаков. Особую ценность представляют методы, которые обеспечивают эффективную защиту при условии, что документ может быть оцифрован после печати или сфотографирован с экрана, и при этом встроенная в документ информация сохраняется. Важной особенностью работы 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. Предложенные решения являются актуальными в условиях современных угроз информационной безопасности, где важен баланс между эффективностью защиты и удобством использования информационных систем.

Практическая значимость

Разработаны и реализованы методы внедрения ЦВЗ в текстовые документы для защиты от утечек при фотографировании распечатанных или экранных копий. ЦВЗ малозаметны и не создают дискомфорта для пользователей, при этом позволяют деанонимизировать утечки через идентификатор сотрудника и устройства. Тестирование показало, что метод горизонтального смещения слов обеспечивает до 61.7% успешных извлечений для сканированных документов и 56.6% для фотографий (69.7% при ручной обработке), а метод перечеркивания — свыше 80% во всех сценариях. При наложении ЦВЗ на экран точность извлечения достигает 86.67% при непрозрачности цифрового водяного знака 8/255.

Реализованная система противодействия анонимности при утечках текстовых документов внедрена организацией ООО "СиТ" (акт о внедрении №612/0924 от 29.09.24).

Методология и методы исследования. В разработке и при тестировании алгоритмов внедрения ЦВЗ в текстовые документы был использован системный подход, основанный на моделировании угроз и нарушителя. Основные методы исследования включают анализ существующих решений, разработку и экспериментальное тестирование алгоритмов, а также математическое моделирование и статистическую

обработку данных. В совокупности эти методы позволили объективно оценить эффективность созданных решений, а также их устойчивость к возможным угрозам.

Основные положения, выносимые на защиту:

1. Структурный метод внедрения ЦВЗ, предполагающий слепое извлечение внедренной информации, на основе сегментации изображения документа с помощью нейросетевого алгоритма, обладающего визуальной незаметностью и устойчивостью к искажениям, возникающим при распечатывании и последующей оцифровке посредством фотографирования или сканирования, и ориентированный под работу на процессоре общего назначения с минимальным потреблением вычислительных ресурсов;
2. Метод генерации ЦВЗ нейросетевым алгоритмом, предполагающий слепое извлечение внедренной информации и обладающий свойствами визуальной незаметности и устойчивости к искажениям, возникающим при фотографировании экрана и сжатии алгоритмами, применяемым в мессенджерах;
3. На основе предложенных методов реализована система противодействия анонимным утечкам текстовых документов, обеспечивающая внедрение уникальных идентификаторов сотрудников и используемых ими устройств в текстовые документы при печати и выводе на экран.

Апробация работы. Результаты работы обсуждались на следующих конференциях:

- Ежегодная научная конференция «Ломоносовские чтения», Москва, 20 – 29 апреля 2021 г.
- Международная конференция «Иванниковские чтения», Нижний Новгород, 24 – 25 сентября 2021 г.
- Научно-практическая Открытая конференция ИСП РАН им. В.П. Иванникова, Москва, 2 – 3 декабря 2021 г.
- Международная конференция «Иванниковские чтения», Казань, 23 – 24 сентября 2022 г.
- Научно-практическая Открытая конференция ИСП РАН им. В.П. Иванникова, Москва, 1 – 2 декабря 2021 г.
- Всероссийская конференция «Методы и технические средства обеспечения безопасности информации», Санкт-Петербург, 24 – 27 июня 2024 г.

Публикации и личный вклад автора

По теме диссертации опубликовано 8 научных работ. Работы [1, 2, 5, 6, 8] опубликованы в журнале, входящем в список ВАК. Работа [3] опубликована в научном журнале, индексируемом системами Web of

Science и Scopus. Получено 5 свидетельств регистрации программ для ЭВМ.

В работах [1, 7] автором представлена архитектура разработанной системы. В работах [2, 4] автором лично предложены методы внедрения структурных ЦВЗ, применимых к текстовым документам при печати. В статье [3] автору принадлежит методика коррекции ошибок при извлечении информации из ЦВЗ на фотографии текстового документа, выведенного на экран. В работе [5] автором выполнен обзор существующих методов внедрения ЦВЗ в текстовые документы. В статье [6] описана разработанная автором методика тестирования методов нанесения водяных знаков при печати, приближенная к условиям эксплуатации. В работе [8] автором собран набор данных для тестирования и предложен набор преобразований для имитации искажений, возникающих при оцифровке посредством сканирования и фотографирования распечатанных копий текстовых документов.

Содержание работы

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

Первая глава посвящена обзору актуальных методов защиты информации от утечек, используемых специалистами по информационной безопасности. Проводится анализ возможных каналов утечек, слабо защищенных существующими средствами защиты, и формулируются возможные сценарии утечек.

В разделе 1.1 рассматриваются DLP-системы, направленные на предотвращение утечек данных. Активные системы работают в реальном времени, блокируя подозрительную активность, но могут допускать ошибки. Пассивные системы выявляют потенциальные угрозы до их реализации. DLP-системы используют агентов на рабочих станциях (data-in-use) для контроля утечек, а также сетевые компоненты с DPI для анализа трафика и компоненты для сканирования хранилищ (data-at-rest). Системы данного класса наиболее эффективно применяются для противодействия утечкам через сетевой канал.

В разделе 1.2 рассматриваются представленные на отечественном рынке решения, нацеленные на защиту “аналоговых” каналов утечек конфиденциальных документов посредством фотографирования распечатанной копии или выведенных на экран. В разделе описаны решения от шести различных разработчиков решений данного типа.

Информация о данных продуктах взята из публичных источников, таких как сайты разработчиков или партнеров.

Средства защиты информации, внедряющие скрытые ЦВЗ, можно разделить на две категории: одни изменяют структуру документа для создания уникальной копии, другие накладывают водяной знак поверх отображаемого документа. Первая группа решений при расследовании утечек сопоставляет изображение утечки документа со всеми ранее сгенерированными копиями. Этот подход требует централизованного защищенного хранилища оригиналов документов и информации обо всех созданных копиях с водяными знаками. Если данных об уникальной копии документа в хранилище нет, то идентификация субъекта утечки становится невозможной. Решение Docs Security Suite использует изменение межбуквенных интервалов для кодирования информации, однако водяной знак на основе этого метода визуально более заметен, чем подходы, основанные на смещении слов.

Решения, использующие полупрозрачное окно для наложения ЦВЗ поверх содержимого документа, не имеют ограничений на формат документов. Однако для методов этого типа особенно актуальны вопросы точности декодирования внедренной информации и заметности водяного знака. Поскольку эти решения являются коммерческими продуктами, проведение полноценного тестирования для объективной оценки их характеристик затруднено.

В разделе 1.3 рассматриваются методы внедрения ЦВЗ в документы различного типа, представленные в научных публикациях. В подразделе 1.3.1 представлен обзор опубликованных методов встраивания ЦВЗ, подходящих для защиты канала утечек через распечатанные документы. Существует большое разнообразие методов внедрения ЦВЗ, использующих область преобразований (transform domain). В частности, в диссертации Pramila 2018 года, посвященной встраиванию ЦВЗ в изображения, описан подход внедрения битовой последовательности при помощи преобразования Фурье с автокоррекцией перспективы на основе специальных паттернов в пространственной области (spatial domain). Этот и другие методы, оперирующие в домене преобразований, больше подходят для изображений с плавными переходами цветов, тогда как на изображениях текстовых документов подобные ЦВЗ слишком заметны. Методы, ориентированные на работу в пространственной области, предполагает модификацию оригинальных документов с использованием информации о структуре, например, данных о местоположении слов, строк или текстовом содержимом. В данной группе можно выделить подгруппу лингвистических методов, изменяющих семантические и/или синтаксические свойства текстового содержимого документа, например в работах Topkara (2006), Meral (2007) или Kim (2008) и других. Подобные методы имеют ограниченную применимость, поскольку методы данного

типа изменяют исходное содержимое документов. Описание структурного метода на основе смещения текстовых элементов, изменяющего параметры визуального представления документа, но не изменяющие смысл/содержимое текста, публиковались Low, Maxemchuk в 1995 году и развивались Bender (1996), Brassil (1999) и другими. Исследователи столкнулись с проблемой выделения текстовой разметки на изображении документа. Кодирование информации реализовано посредством смещения строк, данный подход имеет низкую емкость водяного знака. В дальнейшем публиковались структурные методы Gutub (2007), Alginahi (2013), основанные на изменении визуального представления символов арабского алфавита.

Подраздел 1.3.2 включает обзор опубликованных методов внедрения ЦВЗ в документы, выводимые на экран монитора и устойчивые к искажениям, возникающим при фотографировании экрана. Метод Gugelman и др., предложенный в 2018 году, основан на изменении яркости областей на экране. Каждому биту цифровой метки ставится в соответствие круговая область, яркость которой повышается или понижается в зависимости от значения бита. В основе методов Kurilin (2011), Ge (2022) и других положена идея поиска таких признаков областей для встраивания ЦВЗ, что эти признаки сохраняются на фотографии экрана. Для одного алгоритма признаки областей задаются положением особых точек алгоритма I-SIFT, для другого поиск осуществляется с помощью детектора Харриса-Лапласа. ЦВЗ встраивается в домен преобразования найденных областей – дискретного косинусного преобразования или дискретного преобразования Фурье. Данные методы относятся к категории динамических – в процессе работы выполняется перестроение ЦВЗ, что менее комфортно для пользователей.

Структурные методы позволяют внедрять в текстовые документы ЦВЗ низкой заметности, данные методы могут использоваться для встраивания информации с возможностью извлечения из изображения распечатанного документа. Структурные методы используют текстовую разметку документа, эффективное вычисление текстовой разметки изображения документа является актуальной проблемой. Статические водяные знаки, накладываемые поверх выводимой на экран информации при помощи полупрозрачного окна-оверлея, позволяют защищать документы любого формата в режиме реального времени. Основной проблемой для исследователей, реализующих данный подход, является обеспечение низкой заметности ЦВЗ и устойчивости к искажениям, возникающим при фотографировании экрана.

Вторая глава посвящена описанию архитектуры 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.

Раздел 2.1 описывает механизмы интеграции в операционную систему (ОС) алгоритмов внедрения ЦВЗ в текстовые документы. Используются алгоритмы двух типов: алгоритмы внедрения ЦВЗ в документы при печати и алгоритмы нанесения ЦВЗ на документы, выводимые на экран. Первые обеспечивают защиту документов от анонимных утечек фотографий или сканов распечатанных конфиденциальных документов, вторые защищают от анонимных утечек фотографий экрана с выведенным конфиденциальным документом. Разработанные алгоритмы внедрения ЦВЗ описаны в Главах 3 и 4 соответственно.

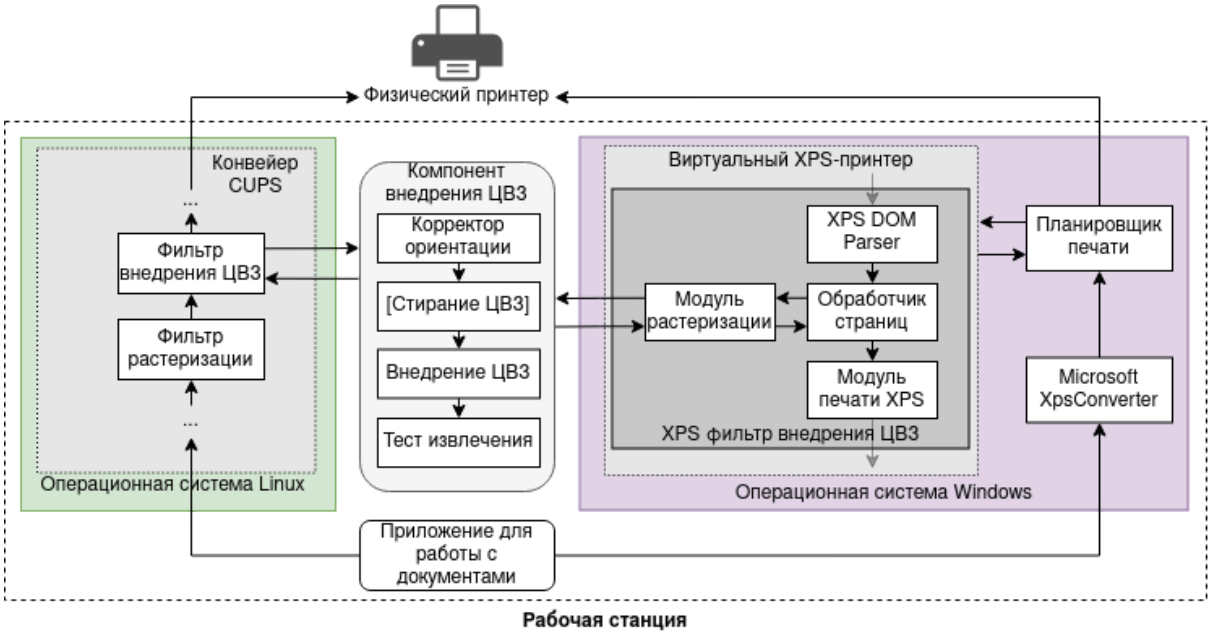


Рисунок 1. Схема внедрения ЦВЗ в текстовые документы при печати.

Алгоритмы внедрения ЦВЗ при печати ориентированы на работу с изображениями текстовых документов, таким образом ЦВЗ внедряются в документы любого формата. Внедрение ЦВЗ выполняется бесшовно для пользователя и приложения, в которым он работает. При установке системы на рабочую станцию под управлением ОС семейства Microsoft Windows создается и конфигурируется виртуальный принтер, осуществляющий внедрение ЦВЗ в документ и последующее перенаправление копии документа с ЦВЗ на физический принтер. В ОС семейства Linux процесс внедрения ЦВЗ реализован через подсистему CUPS (Common UNIX Printing System), в которую при установке системы добавляется фильтр внедрения ЦВЗ в документ. Виртуальный принтер и CUPS-фильтр выполняют функции: растеризация и разбор многостраничного документа на отдельные страницы, корректировка ориентации страницы, внедрение ЦВЗ в каждую страницу документа. При повторной печати документа с уже встроенным ЦВЗ, водяной знак заменяется на новый, содержащий данные о пользователе, инициировавшем печать.

Механизм наложения ЦВЗ на документы при выводе на экран реализован при помощи окна-оверлея. Водяной знак, отображаемый поверх всех окон, статический — накладываемое изображение генерируется при запуске программы и не изменяется в течение сеанса работы пользователя за исключением таких сценариев, как изменение разрешение экрана, подключение дополнительного монитора и других подобных. ЦВЗ отображается с заданным коэффициентом непрозрачности α в отдельном окне, поддерживаемом поверх всех окон.

Раздел 2.2 содержит описание информации, закодированной в ЦВЗ и позволяющей осуществлять деанонимизацию утечек текстовых документов. ЦВЗ содержит информацию о сотруднике и рабочей станции: домен пользователя, имя учетной записи пользователя, серийный номер диска и некоторые дополнительные атрибуты. Для хранения в ЦВЗ полного перечня атрибутов в явном виде требуется емкость, кратно превышающая емкость существующих алгоритмов, поэтому битовая последовательность заданной длины генерируется при помощи хэш-функции. Данная последовательность, именуемая здесь и далее *идентификатором сотрудника и рабочей станции*, не является уникальной, поскольку существует вероятность коллизии идентификаторов. Коллизия идентификаторов означает, что минимум два сотрудника будут иметь один идентификатор, внедряемый в ЦВЗ. Таким образом при инциденте утечки от сотрудников службы безопасности потребуется провести дополнительные изыскания, как например, просмотреть логи обращения к объектам утечки. Включение в битовую последовательность числового идентификатора департамента позволило значительно увеличить доступное число идентификаторов при сохранении низкой вероятности коллизий. 32-битный идентификатор, состоящий из 16-битного идентификатора департамента и 16-битного хэш-кода позволяет кодировать более полумиллиона идентификаторов при вероятности коллизии менее 5% и более 230 тысяч идентификаторов при вероятности коллизии менее 1%. Таким образом, идентификатор длиной 32 бита достаточен для использования системы деанонимизации утечек в большинстве организаций.

В разделе 2.3 рассматривается возможность дополнения идентификатора сотрудника и рабочей станции кодами коррекции для увеличения устойчивости ЦВЗ. Раздел 2.3.1 содержит анализ методов обнаружения и исправления ошибок в битовых последовательностях. Добавление к идентификатору БЧХ-кодов (коды Боуза-Чоудхури-Хоквингема) является оптимальным подходом для исправления ошибок в битовых последовательностях небольшой длины. В разделе 2.3.2 проводится более подробный анализ возможных ошибок: рассматриваются ошибки инверсии, при которых значение одного или нескольких бит инвертируется, и синхронизации, при которых возможно

удаление или вставка произвольных бит. БЧХ-код позволяет исправлять до 3 ошибок инверсии и обнаруживать до 6 ошибок при добавлении 18 бит к последовательности длиной 32 бита. Также описан алгоритм, позволяющий исправлять ошибки любого типа, включая ошибки синхронизации, посредством перебора возможных последовательностей.

Раздел 2.4 содержит описание клиент-серверной архитектуры системы деанонимизации утечек текстовых документов. Компоненты системы, устанавливаемые на рабочие станции сотрудников, передают информацию об идентификаторах сотрудника и рабочей станции и соответствующем наборе атрибутов в централизованное хранилище. При инциденте утечки конфиденциального документа сотрудник службы безопасности проводит извлечение ЦВЗ и внедренного в него идентификатора, далее в хранилище выполняется поиск данного идентификатора и в случае совпадения офицеру службы безопасности предоставляется информация об источнике утечки, в том числе имя учетной записи, IP-адрес, серийный номер диска и другие атрибуты.

Деанонимизация утечек осуществляется посредством идентификаторов, внедряемых в ЦВЗ текстового документа при печати и выводе на экране. ЦВЗ встраивается в изображение отправленного на печать документа любого формата при помощи виртуального принтера или CUPS-фильтра. Окно-оверлей с заданным уровнем непрозрачности отображается поверх всех окон и обеспечивает наложение ЦВЗ на все выводимые на экран документы в режиме реального времени.

Третья глава посвящена структурному методу внедрения ЦВЗ, предполагающему 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 под работу на процессоре общего назначения с минимальным потреблением вычислительных ресурсов.

В разделе 3.1 вводится понятие *разметки текстового документа* – организованной схемы, определяющей взаимное расположение текстовых элементов, таких как символы, слова, строки и других. Разметка документа включает множество *ограничивающих прямоугольников* (ОП) – минимальных прямоугольников, полностью охватывающих текстовый элемент, например слово. В подразделе 3.1.1 описывается нейросетевой алгоритм *сегментации* изображения текстового документа – определения пикселей, принадлежащих ограничивающим прямоугольникам слов. На основе данной информации формируется разметка текстового документа. Для оценки качества текстовой сегментации использовались метрики *IOU* – доля пересечения истинного ОП с предсказанным, а также F_1-мера – гармоническое среднее метрик точности (precision), отражающей долю

истинных ОП среди найденных моделью, и полноты (recall), отражающей долю найденных моделью истинных ОП.

Таблица 1. Сравнение инструментов текстовой сегментации.

Инструмент сегментации	*Производительность*			*Устойчивость разметки*		
	TotalTime, с	*UserTime, с*	*MRS, Мб*	Док-ов	*IOU*	F_1
EasyOCR 1.7.1	123.0433	115.3	3472.45	623	0.8909	**0.9599**
U-Net n5f8	3.3652	6.287	1236.52	623	**0.9085**	0.9572
U-Net n4f8	1.6729	1.781	652.87	622	0.9002	0.9526
U-Net n4f3a0.001h	0.6212	0.854	200.05	618	0.8788	0.9394
U-Net n4f2h	**0.3830**	**0.629**	160.08	620	0.8630	0.9129
Tessarac OCR 4.1.1	0.9834	0.973	**117.44**	537	0.8606	0.9055

В подразделе 3.1.2 описаны эксперименты по оптимизации нейросетевой модели для увеличения производительности при работе на процессоре общего назначения. 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. В таблице 1 представлено сравнение разработанных нейросетевых моделей с открытыми инструментами *Tesseract OCR* и *EasyOCR*. Замеры производительности осуществлялись на компьютере с процессором *Intel Core i5-2390T* и объемом оперативной памяти 4 Гб, метрики были получены при помощи утилиты *GNU time*. Для запуска нейросетевых моделей использовалась среда исполнения *ONNX Runtime* версии 1.12.1. Согласно полученным результатам, разработанные модели *n5f8* и *n4f8* сравнимы по качеству сегментации со значительно более ресурсоемким инструментом EasyOCR, а модели *n4f3a0.001h* и *n4f2h* превосходят по времени работы и качеству инструмент Tesseract OCR.

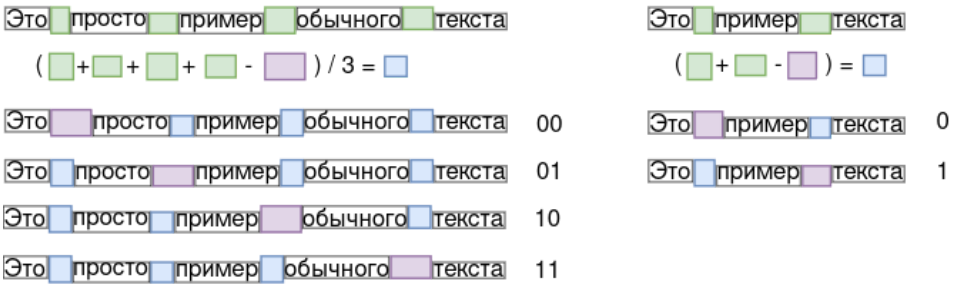
Одной из проблем при текстовой сегментации изображения документа является нестабильная работа на документах с рукописными элементами, такими как инициалы, даты и подписи. В подразделе 3.1.3 описываются эксперименты по разработке нейросетевой модели-фильтра текстовых элементов, не являющихся машинописным текстом. Нейронная сеть на основе архитектуры U-Net вычисляет маску рукописного текста на изображении документа. Такая маска позволяет проверять текстовые элементы на наличие рукописных элементов, и, тем самым, снижать вероятность ошибок вычисления текстовой разметки документа.

Метод нанесения ЦВЗ должен быть устойчив к искажениям, возникающим при печати с последующей оцифровкой документов

посредством фотографирования или сканирования. Для эффективного использования структурных методов внедрения текстовая разметка документа после искажений должна иметь минимальное количество различий в сравнении с исходной разметкой. Методика и результаты оценки устойчивости разметки к искажениям описаны в подразделе 3.1.4. Согласно методике выполняется имитация искажений искажающих и не искажающих геометрию объектов. Искажения применялись с различными параметрами, соответствующих различным сценариям: распечатанная копия документа оцифрована посредством сканера высокого / низкого качества или сфотографирована и отправлена через мессенджер. В таблице 1 представлены усредненные метрики устойчивости разметки F_1 и IOU, вычисленные всех перечисленных сценариев искажения.

оригинальное слово *маска* *изменённое слово*

(a) Кодирование информации перечеркиванием слов.



(b) Кодирование информации горизонтальным смещением слов.

Рисунок 2. Структурные механизмы кодирования информации.

Раздел 3.2 посвящен описанию метода внедрения ЦВЗ при печати. В процессе работы над методом были опробованы механизмы кодирования информации: на основе *перечеркивания слов* и на основе *горизонтального смещения слов*. Первый механизм кодирования изменяет яркость отдельных фрагментов слов (рисунок 2a). Вдоль горизонтальной оси слова между базовой линий и медианой выделяется прямоугольная область, и там, где она пересекает символы текста – буквы и некоторые знаки препинания, например, вопросительный/восклицательный знаки – изменяется яркость. Данный эффект похож на перечеркивание слова осветляющим маркером и визуально имитирует дефекты печати. Основа второго механизма кодирования – горизонтальное смещение слов. Документ разбивается на множество строк, строки посредством жадного алгоритма разбиваются на блоки последовательно расположенных в строке слов, разделенных четырьмя или двумя пробелами (рисунок 2b). При внедрении ЦВЗ слова горизонтально смещаются таким образом, чтобы

величина одного из пробелов в блоке увеличилась, а величина остальных пробелов уменьшилась так, чтобы общая длина блока осталась неизменной. Позиция длинного пробела в блоке позволяет кодировать информацию: для блоков из четырех пробелов – 2 бита, для блоков из двух пробелов – 1 бит.

Метод структурного внедрения ЦВЗ в изображение текстового документа состоит из последовательности шагов:

1. Получение текстовой разметки документа с использованием нейросетевой модели сегментации;
2. Применение фильтров к текстовой разметке, включая фильтр рукописного текста и других элементов немашинописного текста;
3. Определение в текстовой разметке информационных блоков;
4. Смещение или перечеркивание слов, кодирующих информацию.

Извлечение внедренной в текстовый документ информации осуществляется аналогичным образом, но вместо смещения или перечеркивания выполняется считывание закодированной информации.

Разработанный структурный метод внедрения ЦВЗ в текстовые документы использует нейросетевую модель на основе U-Net для вычисления текстовой разметки. Полученная модель превосходит Tesseract OCR и сопоставима с ресурсоемким инструментом EasyOCR по метрикам F_1 и IOU. Для минимизации времени внедрения ЦВЗ проводилась дистилляция модели, что позволило повысить производительность с минимальной потерей качества. Информация кодируется в документе при помощи подходов на основе горизонтального смещения и изменения яркости слов. Метод предполагает слепое извлечение внедренной в ЦВЗ информации, фильтр рукописных элементов на основе нейросетевой модели позволяет повысить точность разметки.

Четвертая глава посвящена методу генерации ЦВЗ нейросетевым алгоритмом, предполагающим 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 разделе 4.1 описывается принцип работы предлагаемого метода. Графическое представление ЦВЗ синтезируется нейронной сетью, сеть кодирует идентификатор сотрудника и устройства, представляющий собой битовую последовательность, в изображение фиксированного размера в оттенках серого цвета. Изображение водяного знака обладает свойствами: высокая степень размытия, благодаря чему отсутствуют резкие переходы цвета между соседними пикселями, а также сохранением первого свойства при объединении изображений водяного знака, при размещении рядом,

либо друг над другом. Данное изображение используется для создания полностью покрывающей экран сетки – ЦВЗ.

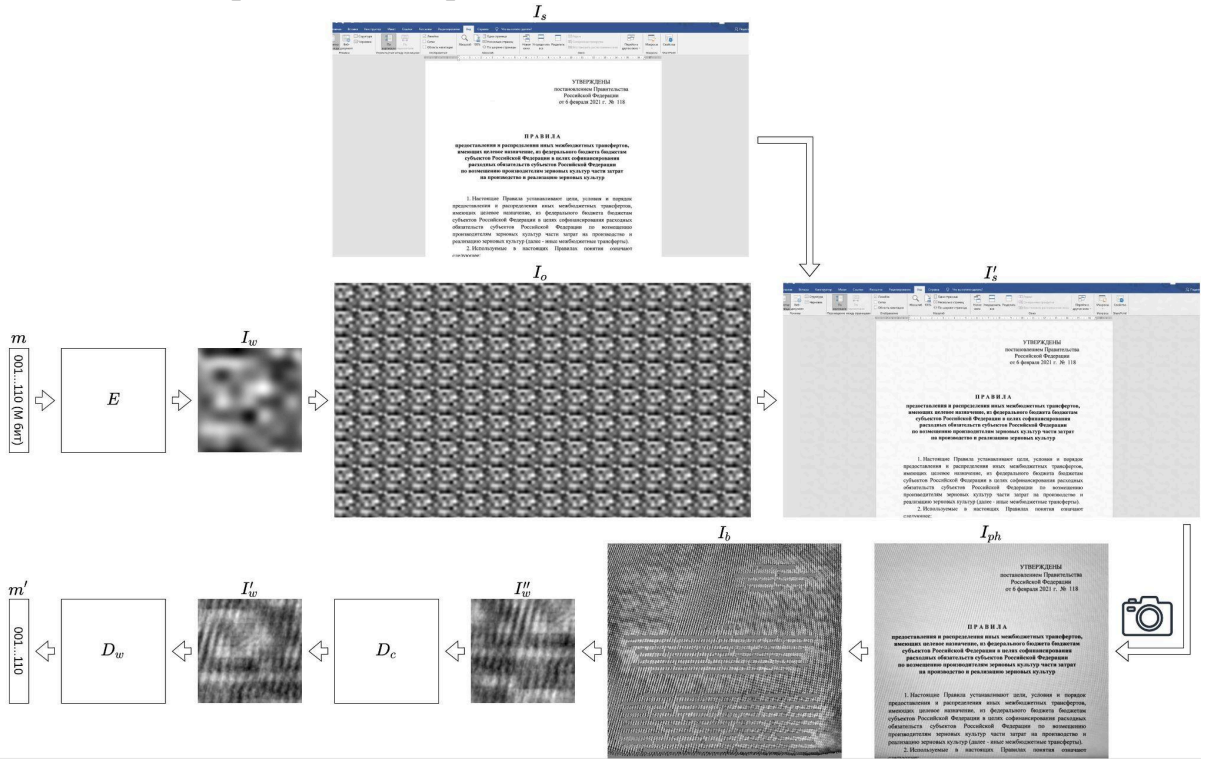


Рисунок 3. Схема работы метода наложения статического ЦВЗ на экран.

Созданный ЦВЗ отображается на экране с определенным уровнем непрозрачности поверх изображений всех окон в ОС. Поскольку для генерации изображения водяного знака не требуется текущее изображение на экране, статический ЦВЗ не требуется перерисовывать в зависимости от действий пользователя устройства. Это позволяет создавать ЦВЗ заранее один раз на основе заданной бинарной последовательности и разрешения экрана. Более того, ЦВЗ всегда присутствует на экране, что позволяет использовать данный метод в режиме реального времени. Статический метод наложения ЦВЗ обеспечивает комфортную работу пользователей при оптимально подобранном уровне непрозрачности.

Внедренная в ЦВЗ информация извлекается из снимка экрана, предварительно подвергнутого процедурам коррекции перспективы и обрезки областей, не относящихся к самому экрану. Алгоритм извлечения информации из ЦВЗ определяет период и смещение сетки на фотографии. Затем изображение водяного знака вычисляется путем усреднения яркости областей на фотографии, с учетом вычисленного периода. Далее, из усредненного изображения ЦВЗ при помощи декодирующей нейронной сети извлекается идентификатор сотрудника и устройства.

Раздел 4.2 содержит описание архитектуры и обучения разработанных нейронных сетей: генерирующей изображение водяного знака (1), определяющей период и смещение (2) и извлекающей внедренную в ЦВЗ битовую последовательность (3). Нейронные сети (1) и

(2) разработаны на основе архитектуры *U-Net*, но имеют отличие от классической реализации данной архитектуры в виде циклического заполнения (circular padding) для достижения упомянутых ранее свойств изображения водяного знака. Задачу извлечения битовой последовательности можно интерпретировать как задачу классификации, поэтому нейронная сеть (3) реализована на основе архитектуры *EfficientNet*, показывающей наибольшую точность в данной задаче на наборе данных *ImageNet*. Обучение нейронных сетей происходит одновременно с использованием слоя аугментаций, имитирующего искажения, возникающие при фотографировании документа на экране. Функция потерь, использованная при обучении, состоит из трех компонент. Первая минимизирует ошибку при определении периода и величины смещения, вторая отвечает за сохранение плавного перехода яркости, а третья обеспечивает корректность извлекаемой битовой последовательности.

В разделе 4.3 описан алгоритм извлечения битовой последовательности, встроенной в ЦВЗ. Алгоритм автоматически выявляет фоновые области на предварительно обработанном изображении утечки, осуществляет поиск периодической структуры на изображении и определяет период, проводит усреднение периодической структуры с изменением масштаба для получения изображения водяного знака, далее извлекает битовую последовательность из изображения водяного знака и исправляет ошибки при помощи БЧХ-кода.

Метод наложения ЦВЗ на текстовые документы состоит из последовательности шагов:

1. Запуск окна-оверлея при старте графической сессии пользователя;
2. Генерация изображения ЦВЗ размером 120×120 пикселей при помощи нейронной сети;
3. Создание сетки, соответствующей размеру экрана, с заданным количеством изображений ЦВЗ по вертикали;
4. Отображение окном-оверлеем сетки с ЦВЗ поверх всех остальных окон с заданным коэффициентом непрозрачности.

Извлечение внедренной в ЦВЗ информации из фотографии экрана состоит из последовательности шагов:

1. Корректировка перспективы и обрезка областей, не относящихся к экрану;
2. Выделение на фотографии документа с ЦВЗ фоновых областей;
3. Поиск периодической структуры на фоновых областях;
4. Формирование усредненного изображения фоновых областей размера ЦВЗ;
5. Определение циклического сдвига при помощи нейронной сети вычисления смещения;
6. Извлечение внедренной в ЦВЗ информации при помощи нейросети.

Предложенный метод позволяет накладывать ЦВЗ на текстовые документы при выводе на экран с возможностью слепого извлечения информации из фотографий экрана. Нейронная сеть на базе архитектуры U-Net обучена генерировать изображение ЦВЗ с плавными переходами яркости для уменьшения его визуальной заметности. Алгоритм извлечения внедренной в ЦВЗ информации корректирует смещение на фотографии при помощи нейросети архитектуры U-Net и декодирует внедренную бинарную последовательность при помощи нейросети архитектуры EfficientNet. Нейросетевые модели обучались с учетом различных искажений, имитирующих фотографирование экрана.

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.

В разделе 5.1 описана методика и результаты тестирования метода внедрения ЦВЗ в текстовые документы при печати на основе структурных механизмов кодирования информации. Экспериментальная оценка осуществлялась на основе количественной и качественной оценок свойств: *емкость*, *незаметность* и *устойчивость*. Исходя из требований к электронным документам, составленным согласно ГОСТ Р 7.0.97–2016, максимальная емкость документа при встраивании ЦВЗ методом на основе горизонтального смещения слов составляет 208 бит для листа формата А4 и 520 бит при использовании метода на основе перечеркивания слов. Метод встраивания ЦВЗ на основе перечеркивания слов показал значения метрик незаметности PSNR 31.81 дБ и SSIM 0.996, что свидетельствует о минимальной визуальной заметности ЦВЗ. Однако для метода, основанного на горизонтальном смещении слов, метрики PSNR и SSIM не применимы из-за особенностей встраивания. В связи с этим для данного метода была проведена оценка стойкости к стеганографическому анализу, которая позволяет определить его устойчивость к попыткам выявления или извлечения скрытой информации. Для оценки устойчивости разработанного метода к искажениям была разработана методика тестирования, приближенная к условиям реальной эксплуатации системы. Тестовый набор документов был сформирован на основе открытых источников и состоял из документов различного форматирования, содержащих в том числе изображения и таблицы. Для печати и сканирования документов использовалось устройство KYOCERA TASKalfa 181, для фотографирования – смартфон Xiaomi Mi A1. Методика предполагает проверку работы в сценариях:

- *Печать и сканирование (Print-Scan)* – в изображение текстового документа внедряется ЦВЗ, содержащий битовую последовательность длиной 32, документ печатается и сканируется, из сканированного изображения извлекается битовая последовательность и сравнивается с внедренной;
- *Повторная печать и сканирование (Double-Print-Scan)* – в изображение текстового документа с ЦВЗ повторно внедряется ЦВЗ, содержащий битовую последовательность длиной 32, документ печатается и сканируется, из сканированного изображения извлекается битовая последовательность и сравнивается с внедренной;
- *Печать и фотографирование (Print-Cam)* – в изображение текстового документа внедряется ЦВЗ, содержащий битовую последовательность длиной 32, документ печатается, фотографируется с фиксированного расстояния 25 см и подвергается обработке для имитации отправки фотографии через мессенджер, из фотографии без обработки извлекается бинарная последовательность и сравнивается с внедренной; в другом сценарии фотографии подвергались *ручной обработке* (обрезка областей, не относящихся к листу бумаги, цветокоррекция, удаление шумов и артефактов съемки) для повышения качества изображения документа и, следовательно, точности извлечения.

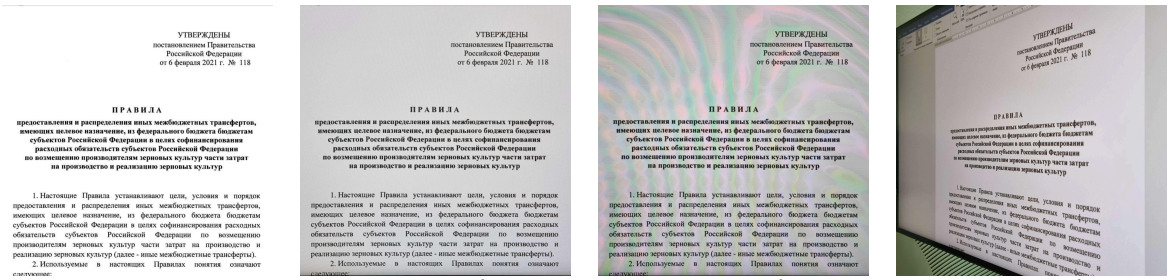
Таблица 2. Результаты тестирования структурного метода внедрения ЦВЗ на основе различных механизмов кодирования информации.

Кодирование	*Сценарий*	$\overline{BER}$	$\frac{E_{BER=0}}{I}$	*E*	*I*	*D*
Горизонтальное смещение слов	*Print-Scan*	0.144	0.617	127	136	160
	Double-Print-Scan	0.189	0.567	128	134	160
	Print-Cam	0.185	0.565	74	76	80
	Print-Cam + ручная обработка	0.123	0.697	74	76	80
Перечеркивание слов	*Print-Scan*	0.001	0.840	112	132	160
	Double-Print-Scan	0.003	0.825	112	132	160
	Print-Cam	0.006	0.803	57	66	80

Сравнение разработанных алгоритмов осуществлялось с использованием следующих метрик: общее количество *D* текстовых документов при тестировании сценария; количество *I* изображений текстовых документов, в которые успешно внедрен ЦВЗ; количество *E* изображений документов с внедренным ЦВЗ, из которых извлечена битовая последовательность; средняя доля ошибок в извлеченной битовой

последовательность $\overline{BER}$ и доля $\frac{E_{BER=0}}{I}$ документов с ЦВЗ, из которых битовая последовательность была извлечена без ошибок.

Результаты проведенных тестов (таблица 2) позволили сделать вывод о практической применимости метода. Разработанный метод внедрения ЦВЗ в текстовые документы при печати позволяет определять источник утечки в различных сценариях. Изменения, вносимые в документ, обладают высокой незаметностью, что подтверждается экспертной оценкой и статистическими методами. Метод внедрения ЦВЗ на основе перечеркивания слов показывает более высокие метрики точности извлечения, однако обладает более высокой заметностью.



Оригинальный документ | *Фотография экрана* | *Фотография экрана с эффектом муара* | *Фотография экрана с искаженной перспективой*

Рисунок 4. Фотографии экрана с различными искажениями.

Раздел 5.2 содержит описание методики и результатов тестирования метода наложения ЦВЗ при выводе текстовых документов на экран. Согласно поставленной задаче ЦВЗ должен обладать визуальной незаметностью и устойчивостью к искажениям, возникающим при фотографировании выведенного на экран документа с последующей оцифровкой посредством фотографирования. Методика тестирования предполагала оценку метода посредством вывода документа на различные модели мониторов и фотографировании на различные модели смартфонов. Тестирование включало четыре этапа:

1. Подбор оптимального уровня непрозрачности α;
2. Оценка влияния расстояния съемки на точность извлечения;
3. Оценка влияния угла съемки на точность извлечения;
4. Оценка влияния качества сжатия JPEG на точность извлечения.

Для определения оптимальной непрозрачности α был проведен эксперимент с 9 парами камер и мониторов на фиксированном расстоянии 40 см. Результаты показали (таблица 3), что при $\alpha > 7/255$ внедренная информация корректно извлекается из большинства фотографий, за исключением случаев с эффектом муара. Последующие этапы тестирования проводились с зафиксированным уровнем непрозрачности $\alpha = 7/255$. Оценка влияния расстояния (таблица 4) и угла съемки (таблица 5) показала, что эффект муара вызывает ошибки при извлечении

меток, особенно на расстояниях 25 и 40 см. Тестирование устойчивости к JPEG-сжатию показало (таблица 6), что внедренная информация корректно извлекается из 99% фотографий при качестве JPEG не менее 40.

Таблица 3. Извлечение внедренной в ЦВЗ битовой последовательности без ошибок при различном уровне непрозрачности.

α	$\frac{3}{255}$	$\frac{4}{255}$	$\frac{5}{255}$	$\frac{6}{255}$	$\frac{7}{255}$	$\frac{8}{255}$	$\frac{9}{255}$	$\frac{10}{255}$
Извлечено	40/90	66/90	71/90	76/90	**78/90**	85/90	86/90	87/90

Таблица 4. Извлечение внедренной в ЦВЗ битовой последовательности без ошибок при фотографировании экрана с различного расстояния.

Расстояние	25 см	40 см	60 см	80 см	100 см
Извлечено	51/90	79/90	85/90	89/90	87/90

Таблица 5. Извлечение внедренной в ЦВЗ битовой последовательности без ошибок при фотографировании экрана под различным углом.

Угол	0°	15°	30°	45°	60°
Извлечено	79/90	83/90	84/90	88/90	82/90

Таблица 6. Извлечение внедренной в ЦВЗ битовой последовательности без ошибок при фотографировании экрана с различным уровнем сжатия.

Качество JPEG	10	15	20	30	40	50	80
Извлечено	19/50	31/50	42/50	47/50	48/50	50/50	50/50

Незаметность ЦВЗ на экране оценивалась с использованием метрик *PSNR* и *SSIM*, усредненных по 50 изображениям документов и 50 изображениям из набора данных *Open Images V6*. SSIM показала высокую степень незаметности для всех значений α, в то время как *PSNR* указала на меньшую заметность ЦВЗ на изображениях не текстовых документов.

Проведенные тесты показали высокую устойчивость метода к искажениям различного рода. Наибольшей проблемой для метода является эффект муара, проявляющийся индивидуально для комбинации камеры, монитора, а также расстояния и угла съемки.

Система противодействия анонимности при утечках текстовых документов посредством ЦВЗ протестирована согласно описанной методике. Система обеспечивает внедрение в текстовые документы 32-битного идентификатора сотрудника и устройства при печати и выводе на экран. Протестированы два алгоритма кодирования информации при печати: на основе горизонтального смещения и перечеркивания слов. Первый механизм показал эффективность до 61.7% для сканированных документов и до 56.5% для фотографий (69.7% при ручной обработке), второй механизм позволил деанонимизировать более 80% утечек документов с ЦВЗ во всех сценариях. При выводе документов на экран

ЦВЗ накладывается с помощью окна-оверлея с непрозрачностью $\alpha = 7/255$, обеспечивая 86.67% успешных извлечений. Увеличение непрозрачности до 8/255 повышает точность до 94.44%, но также повышает заметность водяного знака. Система показала высокую устойчивость к фотографированию под разными углами и с различного расстояния, а также к сжатию фотографий, сохраняя эффективность извлечения более 80% в большинстве сценариев.

В **заключении** приведены основные результаты работы, которые заключаются в следующем:

1. Разработана архитектура системы деанонимизации при утечках изображений текстовых документов, обеспечивающая внедрение уникальных идентификаторов сотрудников и используемых ими устройств в документы при печати и выводе на экран;
2. Разработан обладающий научной новизной структурный метод внедрения ЦВЗ, предполагающий слепое извлечение внедренной информации, на основе сегментации изображения документа с помощью нейросетевого алгоритма. Разработанный метод обладает визуальной незаметностью и устойчивостью к искажениям, возникающим при печати и последующей оцифровке посредством фотографирования или сканирования, и ориентирован для выполнения на процессоре общего назначения с минимальным потреблением вычислительных ресурсов;
3. Разработан обладающий научной новизной метод генерации ЦВЗ нейросетевым алгоритмом, предполагающий слепое извлечение внедренной информации и обладающий свойствами визуальной незаметности и устойчивости к искажениям, возникающим при фотографировании экрана и сжатии алгоритмами, применяемыми в мессенджерах при пересылке изображений;
4. На основе разработанных архитектуры и методов внедрения/извлечения ЦВЗ реализована система противодействия анонимным утечкам текстовых документов. Система протестирована на целевых сценариях утечек – фотографирование документов на экране и сканирование распечатанных документов. Апробация подтвердила эффективность системы в различных условиях, демонстрируя способность успешно извлекать внедренную в ЦВЗ информацию и деанонимизировать утечки.

Также определены направления дальнейшей работы:

- повышение точности и устойчивости к искажениям разработанных методов внедрения ЦВЗ в текстовые документы;
- разработка методов внедрения ЦВЗ в нетекстовые документы (изображения, аудиозаписи, видеозаписи и другие);

- исследование применимости методов внедрения ЦВЗ для решения других практических задач.

Публикации по теме диссертации

1. Система маркирования документов для проведения расследований при их утечке / Д.О. Обыденков, А.Ю. Якушев, Ю.В. Маркин, А.Е. Фролов, С.А. Фомин, С.В. Козлов, Д.Д. Громей, А.В. Козачок, Б.В. Кондратьев // Труды Института системного программирования РАН. — 2021 г. — Т. 33, № 6. — С. 161-174. — (ВАК).
2. Методы маркирования текстовых документов при печати посредством вертикального сдвига и изменения яркости фрагментов слов / Д.О. Обыденков, А.Е. Фролов, Ю.В. Маркин, С.А. Фомин, Б.В. Кондратьев // Труды Института системного программирования РАН. — 2021 г. — Т. 33, №5. — С. 65-82. — (ВАК).
3. Docmarking: Real-Time Screen-Cam Robust Document Image Watermarking / A.Yakushev, Y. Markin, D. Obydenkov, A. Frolov, S. Fomin, M. Akopyan, A. Kozachok, A. Gaynov // 2022 Ivannikov Ispras Open Conference (IVMEM). — 2022 г. — P. 142-150. — (Scopus).
4. Методы маркирования текстовых документов при печати / А.И. Гетьман, Д.О, Обыденков, А.Е. Фролов, Ю.В. Маркин // Ежегодная научная конференция «Ломоносовские чтения» секция «Вычислительной математики и кибернетики». — 2021 г. — P. 55-57.
5. Маркирование текстовых документов на экране монитора посредством изменения яркости фона в областях межстрочных интервалов / А.Ю. Якушев, Ю.В. Маркин, С.А. Фомин, Д.О. Обыденков, Б.В. Кондратьев // Труды Института системного программирования РАН. — 2021 г. — Т. 33, №4. — С. 147-162. — (ВАК).
6. Реализация маркирования в подсистеме печати ОС семейства Windows на основе виртуального XPS-принтера / С.В. Козлов, С.А. Копылов, Б.В. Кондратьев, Д.О. Обыденков // Труды Института системного программирования РАН. — 2020 г. — Т. 32, №5. — С. 95-110. — (ВАК).
7. Предотвращение анонимных утечек конфиденциальных документов / Д.О. Обыденков, А.Ю. Якушев, С.А. Фомин, Ю.В. Маркин, А.В. Козачок, А.Е. Фролов, С.В. Козлов, Д.Д. Громей, А.А. Акименков, А.В. Мякутин, В.А. Челина // Материалы 33-й Научно-технической конференции «Методы и технические средства обеспечения безопасности информации». — 2024 г. — С. 114-115.
8. Экспериментальная оценка алгоритма маркирования текстовых документов на основе изменения интервалов между словами / А.В. Козачок, В.И. Козачок, С.А. Копылов, П.Н. Горбачев, Ю.В. Маркин, Д.О. Обыденков // Труды Института системного программирования РАН. — 2022 г. — Т. 34, №4. — С. 153-1172. — (ВАК).

Свидетельства о государственной регистрации программы для ЭВМ

1. Библиотека маркирования текстовых документов на экране путем изменения яркости в областях межстрочных интервалов / Ю.В. Маркин, А.В. Козачок, С.А. Фомин, Д.О. Обыденков, А.Ю. Якушев, В.А. Падарян ; ФГБУН Институт системного программирования РАН. — № 2020667628 ; заявл. 25.12.2020 (Рос. Федерация).
2. Модуль маркирования текстовых документов при печати для ОС семейства Linux / Ю.В. Маркин, А.В. Козачок, С.А. Фомин, М.С. Акопян, Д.О. Обыденков, П.Н. Горбачев, С.В. Козлов, Д.Д. Громей, С.А. Копылов, В.А. Падарян ; ФГБУН Институт системного программирования РАН. — № 2020667580 ; заявл. 24.12.2020 (Рос. Федерация).
3. Библиотека маркирования текстовых документов при печати за счет горизонтального смещения слов / Ю.В. Маркин, А.В. Козачок, С.А. Фомин, Д.О. Обыденков, П.Н. Горбачев, С.В. Козлов, Д.Д. Громей, С.А. Копылов, Б.В. Кондратьев, В.А. Падарян ; ФГБУН Институт системного программирования РАН. — № 2020667592 ; заявл. 24.12.2020 (Рос. Федерация).
4. Модуль маркирования текстовых документов при печати для ОС семейства Windows / Ю.В. Маркин, А.В. Козачок, С.А. Фомин, М.С. Акопян, Д.О. Обыденков, П.Н. Горбачев, С.В. Козлов, Д.Д. Громей, С.А. Копылов, Б.В. Кондратьев, В.А. Падарян ; ФГБУН Институт системного программирования РАН. — № 2020667579 ; заявл. 24.12.2020 (Рос. Федерация).
5. Модуль маркирования текстовых документов на экране для ОС семейства Windows / Ю.В. Маркин, А.В. Козачок, С.А. Фомин, М.С. Акопян, Д.О. Обыденков, А.Ю. Якушев, В.А. Падарян ; ФГБУН Институт системного программирования РАН. — № 2020667308 ; заявл. 22.12.2020 (Рос. Федерация).

Обыденков Дмитрий Олегович
Методы противодействия анонимности при утечках текстовых документов посредством цифровых водяных знаков

Автореф. дис. на соискание ученой степени канд. тех. наук

Подписано в печать ___.___.____. Заказ № _____
Формат 60×90 / 16. Усл. печ. л. 1. Тираж 100 экз.
Типография _____