МИНИСТЕРСТВО ОБРАЗОВАНИЯ РЕСПУБЛИКИ БЕЛАРУСЬ

Учреждение образования «БЕЛОРУССКИЙ ГОСУДАРСТВЕННЫЙ

ТЕХНОЛОГИЧЕСКИЙ УНИВЕРСИТЕТ»

Факультет информационных технологий

Кафедра информационных систем и технологий

Специальность Информационные системы и технологии

ОТЧЕТ ПО ЛАБОРАТОРНОЙ РАБОТЕ №13 НА ТЕМУ:

Исследование методов текстовой стенографии

Выполнила:

Студентка 3 курса 1 группы ФИТ

Шимчёнок Елизавета Константиновна

**Цель:** изучение стеганографического метода встраивания/извлечения тайной информации с использованием электронного файла-контейнера текстового формата, приобретение практических навыков программной реализации методов.

**Задачи:**

1. Закрепить теоретические знания из области текстовой стеганографии, классификации, моделирования стеганосистем подобного вида и сущности основных методов.

2. Изучить основные алгоритмы встраивания/извлечения тайной информации на основе методов текстовой стеганографии, получить опыт практической реализации методов.

3. Разработать приложение для реализации алгоритмов встраивания/извлечения тайной информации на основе методов текстовой стеганографии.

4. Познакомиться с методиками оценки стеганографической стойкости методов.

5. Результаты выполнения лабораторной работы оформить в виде описания разработанного приложения, методики выполнения экспериментов с использованием приложения и результатов эксперимент.

**Теоретические сведения**

Для понимания сущности некоторых из методов полезно познакомиться с важнейшими особенностями и параметрами использования стилей (в том числе пространственно-геометрическими параметрами шрифтов), на основе которых строится текстовый файлконтейнер. На рисунке 1 показаны основные из параметров шрифта.

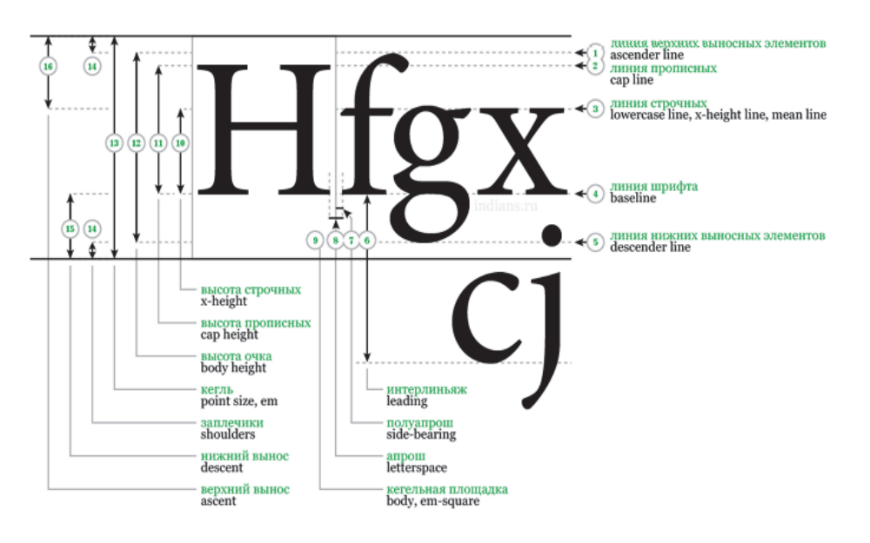


Рисунок 1 – Параметры шрифта

К текстовой стеганографии относятся методы, предусматривающие использование в качестве контейнера файла-документа текстового типа.

Текстовая стеганография:

1) Синтетические методы (затрагивают семантику текстового сообщения):

– изменение расстояния между строками электронного документа;

– изменение расстояния между словами;

– изменение количества пробелов между словами;

– на основе внесения специфических изменения в шрифты;

– изменение интервала табуляции;

– Null Chipper (дословно – несуществующий, нулевай лепет);

– увеличение длины строки;

– использование регистра букв;

– использование невидимых символов.

Достоинства:

– Легко применяются к любому тексту (независимо от содержания, назначения, языка);

– Легко реализуются в программном коде (т.к. они полностью автоматические)

Недостатки:

– Невысокая эффективность (объем встраиваемой информации);

– Перечисленные методы работают успешно до тех пор, пока тексты представлены в коде ASCII;

– Неустойчивы к форматированию текса;

– Неустойчивы к изменению масштаба документа.

2) Лингвистические методы (основаны на эквивалентной трансформации текстовых файлов-контейнеров):

– Метод синонимов;

– Метод переменной длины слова;

– Метод первой буквы;

– Мимикрия.

**Ход работы**

Разработать авторское приложение, реализующее метод текстовой стеганографии на основе модификации изменения длины строки. Данный метод предусматривает искусственное увеличение длины каждой строки за счет пробелов: например, нет пробела – «0», один пробел – «1». Количество строк в контейнере, должно превышать размер сообщения для скрытия.

Для реализации метода необходимо секретное сообщение, которое будет скрываться в текстовом документе: «Elizaveta». Это сообщение будет обрабатываться функцией-перебором всех битов сообщения:

– в случае, если бит равен единице, то к строке добавляется двойной пробел,

– в противном случае одинарный.

На рисунке 2 представлен текстовый файл до сокрытия в нем тайного сообщения.

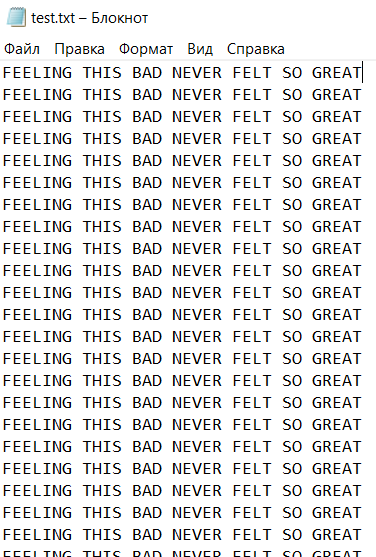


Рисунок 2 – «Пустой» контейнер

Запустим приложение и проанализируем результат выполнения скрытия. Приложения запускаем через консоль передав в качестве параметров название функции, файлы для скрытия, а также само секретное сообщение.

На рисунке 3 приведен контейнер со скрытым сообщением. Однако визуально контейнер не поменялся.

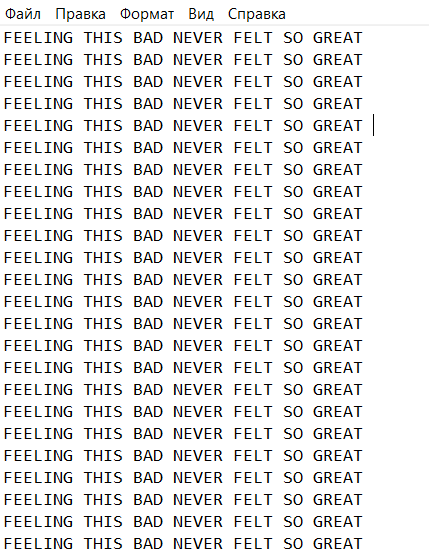


Рисунок 3 – Контейнер после сокрытия

Далее необходимо извлечь скрытую информацию и удостовериться, что контейнер хранит секретную информацию. Выполним извлечение и получим в качестве результата исходное сообщение.

Этот метод является довольно простым способом реализации текстовой стеганографии. Однако довольно непрактичен, так как при минимальном редактировании документа, целостность секретного сообщения будет нарушена.

**Ответы на вопросы**

**1. В чем состоит сущность методов текстовой стеганографии?**

Сущность методов текстовой стеганографии заключается в скрытом передаче информации в текстовых данных без вызывания подозрений у наблюдателей. Основная идея стеганографии заключается в том, чтобы внедрить скрытую информацию в несущественные части текста, называемые носителями. При этом, изменения должны быть незаметны для человеческого восприятия и не вызывать подозрений о наличии скрытых данных.

**2. Охарактеризовать методы синтаксической текстовой стеганографии. Привести примеры конкретной реализации методов.**

Синтетические методы (затрагивают семантику текстового сообщения):

– изменение расстояния между строками электронного документа;

– изменение расстояния между словами;

– изменение количества пробелов между словами;

– на основе внесения специфических изменения в шрифты;

– изменение интервала табуляции;

– Null Chipper (дословно – несуществующий, нулевой лепет);

– увеличение длины строки;

– использование регистра букв;

– использование невидимых символов.

Достоинства:

– Легко применяются к любому тексту (независимо от содержания, назначения, языка);

– Легко реализуются в программном коде (т.к. они полностью автоматические)

Недостатки:

– Невысокая эффективность (объем встраиваемой информации);

– Перечисленные методы работают успешно до тех пор, пока тексты представлены в коде ASCII;

– Неустойчивы к форматированию текса;

– Неустойчивы к изменению масштаба документа.

**3. Охарактеризовать методы лингвистической текстовой стеганографии. Привести примеры конкретной реализации методов**

Лингвистические методы (основаны на эквивалентной трансформации текстовых файлов-контейнеров):

– Метод синонимов;

– Метод переменной длины слова;

– Метод первой буквы;

– Мимикрия.

**4.** Д**ать оценку стеганографической стойкости методов текстовой стеганографии при конвертации текста-контейнера в иной текстовый формат.**

Оценка стеганографической стойкости методов текстовой стеганографии при конвертации текста-контейнера в другой текстовый формат зависит от различных факторов, включая выбранный метод и специфические особенности конвертации.

Замена символов: этот метод может быть относительно простым, но стойкость зависит от способа замены и сложности обнаружения изменений. Если замена символов проводится случайным образом и сложно обнаружить паттерны, это может быть стойким методом.

Внедрение в пробелы и форматирование: стойкость этого метода зависит от того, насколько естественно выглядят изменения в пробелах и форматировании. Если изменения явно выделяются или вызывают подозрения, метод может быть менее стойким.

Внедрение в форматирование документов: этот метод может быть стойким, если изменения в форматировании выглядят естественными и не вызывают подозрений. Однако, если специфическое форматирование может быть обнаружено или удалено при конвертации в другой формат, стойкость может быть низкой.

Лингвистические стеганографические методы: стойкость этих методов зависит от сложности анализа текста и способа внедрения скрытой информации. Если методы лингвистического анализа сложны и обнаружение скрытой информации требует специализированных инструментов, это может быть стойким методом.

**5. Дать оценку стеганографической стойкости методов текстовой стеганографии при визуальном стеганоанализе текста-контейнера.**

Замена символов: этот метод может быть уязвимым при визуальном стеганоанализе, особенно если замена символов приводит к отклонениям от ожидаемого распределения символов или вызывает подозрения. Простые методы замены символов могут быть относительно легко обнаружены.

Внедрение в пробелы и форматирование: если изменения в пробелах и форматировании визуально не выделяются и выглядят естественно, метод может быть стойким при визуальном стеганоанализе. Однако, если есть специфические паттерны или аномалии, которые могут быть обнаружены визуально, стойкость может быть низкой.

Внедрение в форматирование документов: стойкость этого метода при визуальном стеганоанализе зависит от того, насколько изменения в форматировании могут быть обнаружены. Если изменения не вызывают подозрений и не выделяются визуально, метод может быть стойким.

Внедрение в метаданные: при визуальном стеганоанализе метаданные могут быть проверены и анализированы. Если внедрение скрытой информации в метаданные вызывает подозрения или обнаруживается при визуальном анализе, стойкость метода может быть низкой.

Лингвистические стеганографические методы: эти методы могут быть уязвимыми при визуальном стеганоанализе, особенно если изменения в тексте вызывают подозрения или нарушают естественные языковые паттерны. Если визуальный анализ текста выявляет аномалии, стойкость методов лингвистической стеганографии может быть низкой.

Общая оценка стеганографической стойкости методов текстовой стеганографии при визуальном стеганоанализе зависит от специфических особенностей методов, уровня искусства применения и возможностей анализа текста.

**6. Дать общую характеристику стеганоанализу в области текстовой стеганографии на основе метода «χ-квадрат».**

Метод "χ-квадрат" (chi-square) является одним из основных методов стеганоанализа в области текстовой стеганографии. Он используется для статистического анализа текста с целью обнаружения скрытой информации. Общая характеристика стеганоанализа на основе метода "χ-квадрат" включает следующие аспекты.

*Статистический анализ*

Метод основан на сравнении наблюдаемых частот встречаемости символов или других статистических параметров в тексте с ожидаемыми частотами, которые можно вычислить на основе модели статистики текста. При стеганографической внедрении скрытой информации частоты могут измениться, и метод "χ-квадрат" позволяет обнаружить эти аномалии.

*Проверка гипотезы*

Метод используется для проверки статистических гипотез о наличии скрытой информации в тексте. Стеганоанализ заключается в сравнении наблюдаемого значения "χ-квадрат" статистики с критическим значением, определенным на основе выбранного уровня значимости. Если наблюдаемое значение превышает критическое значение, то гипотеза о наличии скрытой информации отвергается.

*Использование моделей статистики текста*

Метод требует знания или создания моделей статистики текста, которые представляют ожидаемые частоты символов или других статистических параметров. Модели могут быть созданы на основе обучающих данных или предварительного анализа текста. Они служат базой для сравнения и вычисления "χ-квадрат" статистики.

*Чувствительность к изменениям*

Метод может быть чувствителен к изменениям в тексте, вызванным внедрением скрытой информации. Если изменения вызывают значительные отклонения от ожидаемых частот или статистических параметров, "χ-квадрат" статистика может показать высокие значения, что указывает на возможное наличие скрытой информации.

**7. Что такое апрош? В чем состоит сущность стеганометода на основе модификации апроша?**

Апрош определяет расстояние между соседними символами текста.

Идея метода: встраивание сообщения в контейнер может быть основано на модификации базового (устанавливаемого текстовым процессором по умолчанию) значения апроша , его изменением от базового до некоторого максимального (или минимального ), которое зрительно не должно отличаться от стандартного. Такое изменение производится с определенным шагом (дискретно) Δ, каждому значению которого присваивается определенный бит или определенная комбинация битов.

**8. Что такое кернинг? В чем состоит сущность стеганометода на основе модификации кернинга?**

Кернинг (kerning) – это процесс настройки промежутков между символами в тексте с целью достижения оптимального визуального расстояния между ними. Кернинг применяется в типографии и дизайне шрифтов для улучшения читаемости и визуального вида текста.

Кернинг, как параметр, будем измерять в тысячных долях круглой шпации () – единицы измерения, которая определяется относительно текущего размера шрифта и равна ширине символа «М»

Сущность стеганометода на основе модификации кернинга заключается в использовании изменений промежутков между символами для внедрения скрытой информации. При этом небольшие изменения в кернинге между символами в тексте-контейнере используются для кодирования скрытого сообщения.

Процесс работы стеганометода на основе модификации кернинга может быть следующим:

1. Выбор текста-контейнера: в тексте-контейнере, который является видимым сообщением, выбираются определенные символы или промежутки между символами, которые будут изменены для внедрения скрытой информации.
2. Кодирование скрытого сообщения: скрытое сообщение разбивается на биты или другие подходящие единицы данных. Затем каждому биту или группе битов присваивается значение, которое будет отражено в изменении кернинга символов.
3. Модификация кернинга: выбранные символы или промежутки между символами в тексте-контейнере изменяются согласно кодированному скрытому сообщению. Эти изменения в кернинге могут быть незначительными и визуально неразличимыми для обычного наблюдателя.
4. Извлечение скрытого сообщения: для извлечения скрытой информации процесс обратный кодированию. Анализируется изменение кернинга в тексте-контейнере, и по этим изменениям восстанавливается скрытое сообщение.

Под кернингом обычно понимается процесс изменения межсимвольного расстояние между отдельными парами символов или кернинговыми парами (именно фактор парности отличает кернинг от апроша).

**Вывод:** в ходе лабораторной работы были изучены стеганографические методы встраивания/извлечения тайной информации с использованием электронного файла-контейнера текстового формата. Более подробно был рассмотрен метод Апрош и разработано приложение для демонстрации шифрования и дешифрования данных.