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 стеганографии

Выполнила студентка 3 курса 1 группы

Пригодич Вера Валерьевна

Минск 2023

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

**Теоретические сведения**

**Классификация, сущность и основные особенности базовых методов текстовой стеганографии**

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

Текстовая стеганография:

1) Синтаксические 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.

**Практическая часть**

Разработать авторское приложение, реализующее один из методов текстовой стеганографии на основе модификации пространственно-геометрических параметров текста-контейнера. Варианты заданий приведены в таблице.

|  |  |
| --- | --- |
| Вариант | Реализуемые методы |
| 8 | Модификация числа пробелов; модификация цвета |

Метод изменения количества пробелов между словами (частный случай метода Word-Shift Coding) основан та том, что, например, чередование одинарного пробела и двойного (хх\_хх\_\_хх) кодирует «1», переход же с двойного пробела на одинарный кодирует «0» (хх\_\_хx\_xx).

Для реализации метода был использован word-документ в формате docx.

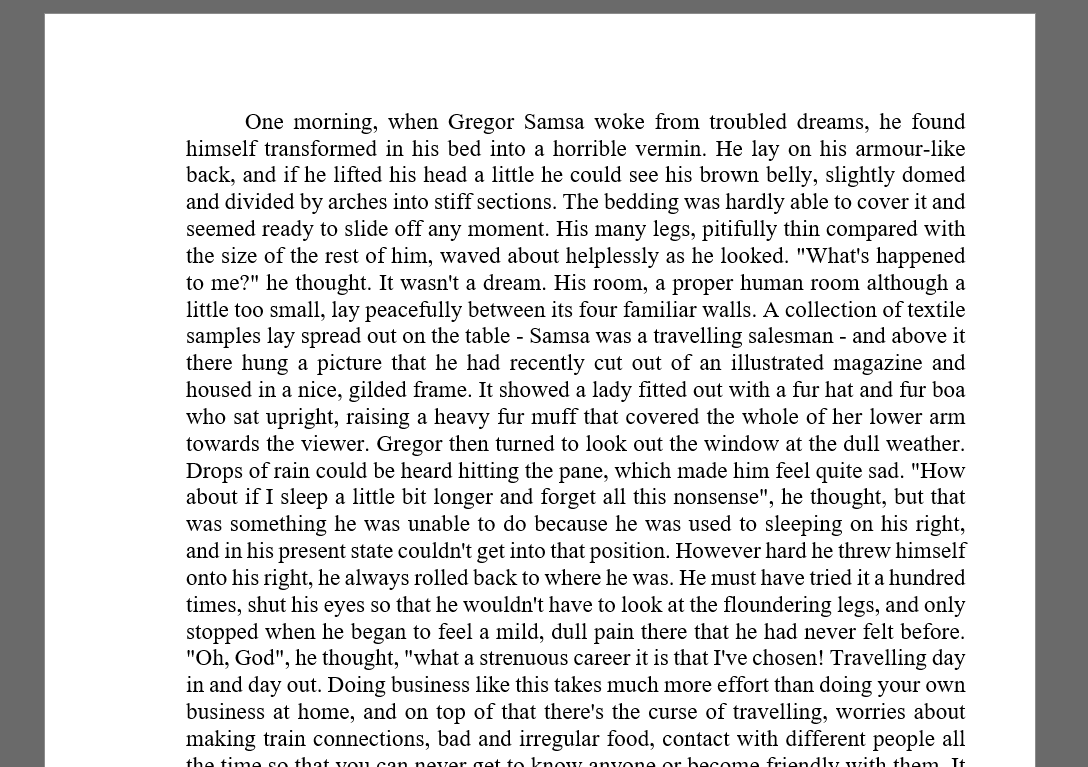


Рисунок 1 – Документ-контейнер

Осадим в данный контейнер сообщение «vera».

Для этого переведем сообщение в бинарный вид. Получаем последовательность «01110110011001010111001001100001».

В соответствии с методом, добавим пробелы в документ.

Так для первого символа «0» необходимо добавить пробел после первого слова, сделав двойной пробел, следующий пробел должен быть одинарным.

Далее следует символ «1», следовательно, следующий пробел должен быть одинарным, после него двойным.

Закодировав всю последовательность получаем документ следующего вида:

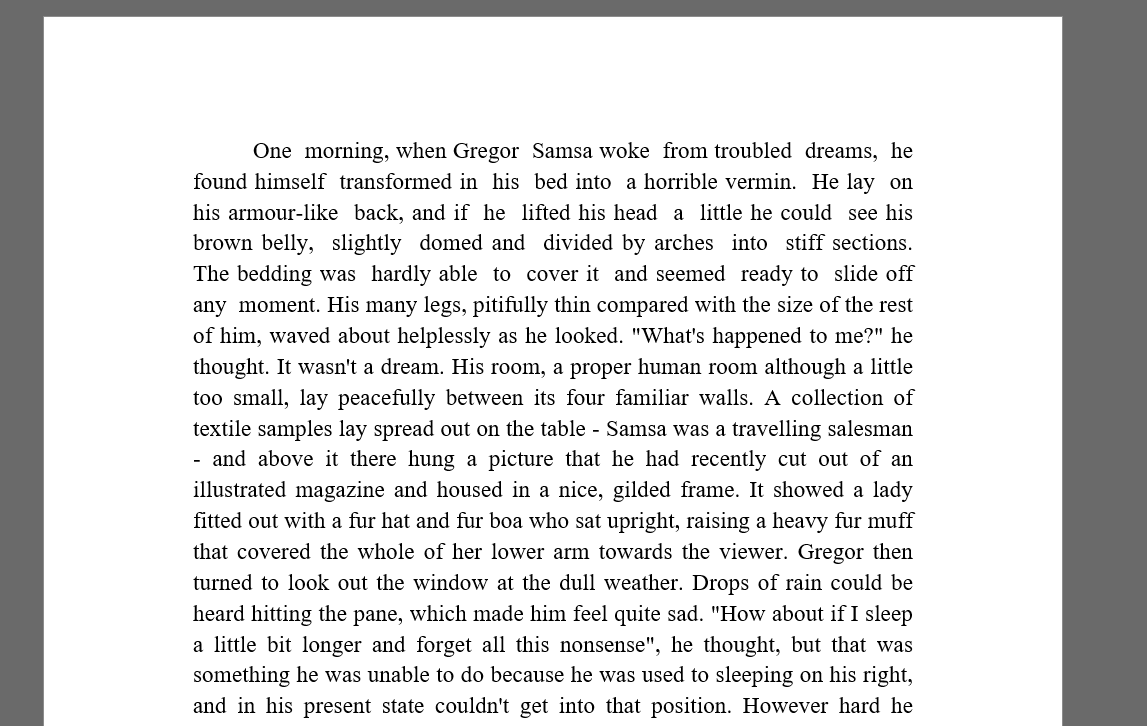


Рисунок 2 – Контейнер с осажденным сообщением

Визуально разница с исходным документом не заметна, однако если отобразить все знаки в документе или проверить документ на ошибки можно заметить следующее:

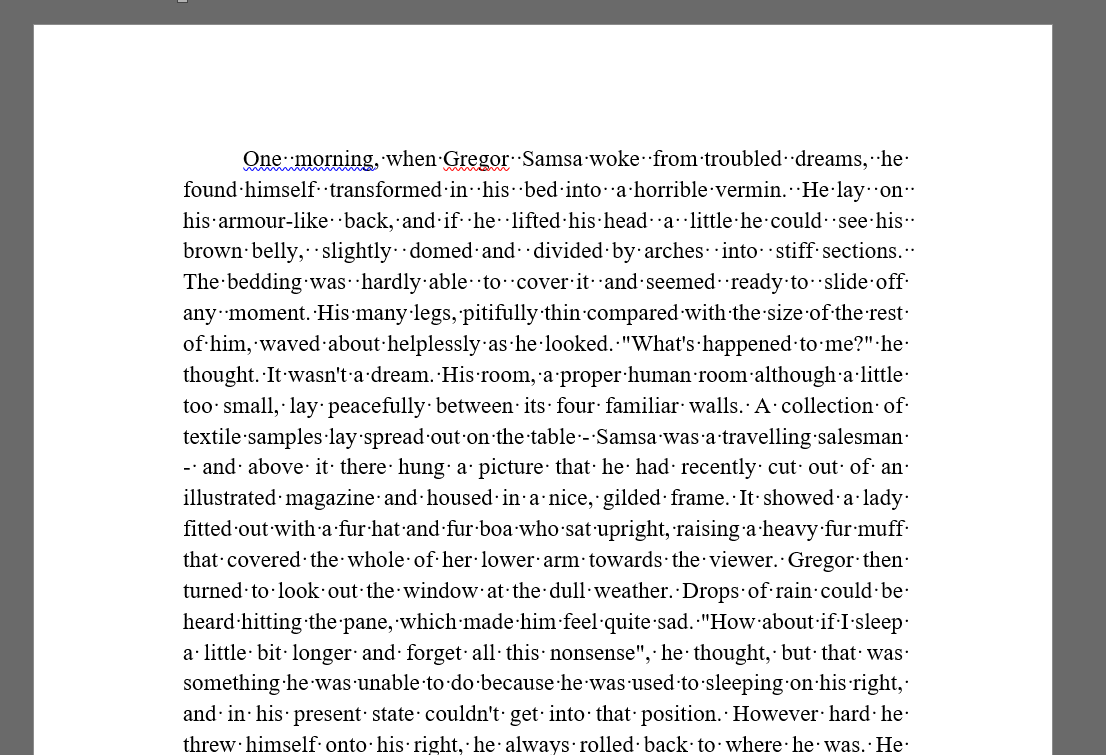


Рисунок 3 – Отображение всех знаков

Проверяя документ таким образом, третье лицо может заметить нетипичную расстановку пробелов.

Получим осажденное сообщение из контейнера.

Подсчитаем количество идущих подряд пробелов. Если в тексте встречается двойной пробел, потом одинарный в буфер для декодирования сообщения записывается «0», если наоборот – «1».

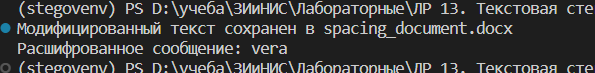


Рисунок 4 – Результат вытягивания осаждённого сообщения

Реализация метода на основе модификации цвета символов текста-контейнера по сути своей схожа с классическим методом наименее значащих битов (см. лабораторную работу № 12) и опирается на использовании цветовой модели RGB.

Для реализации данного метода был использован тот же текст, но в формате html. Также при осаждении сначала сообщение переводится в бинарный вид. Затем для каждого символа в тексте контейнера задаётся значение цвета. В зависимости от выбранного цвета осаждение сообщение может быть более или менее заметным для третьего лица. Так если мы выберем яркие цвета для кодирования битовой последовательности, можно получить следующий результат:

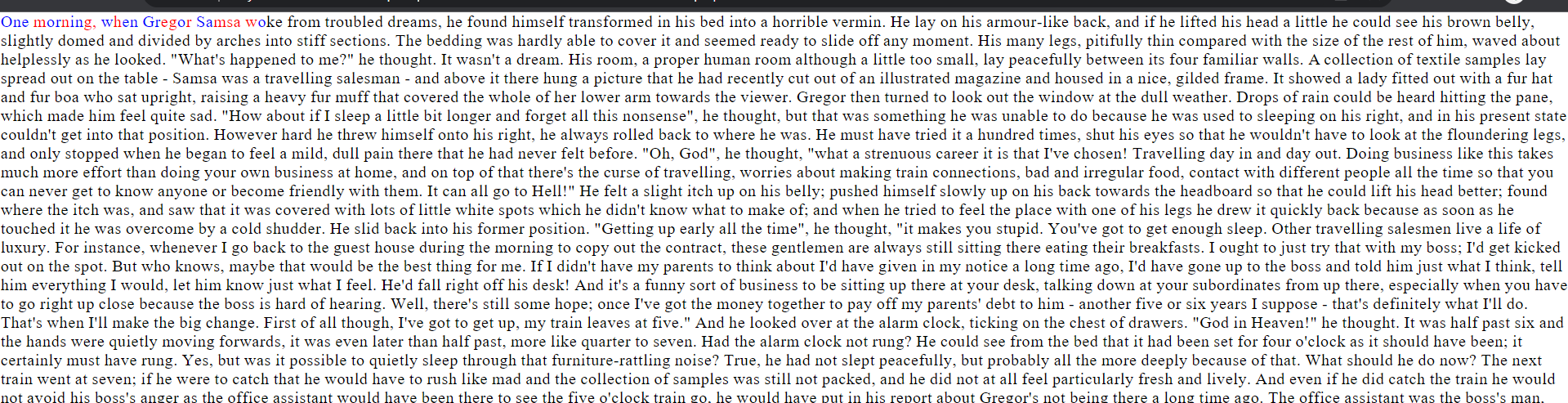


Рисунок 5 – Контейнер с осажденным сообщением

С другой стороны, мы можем выбрать цвета, близкие к цвету исходного текста и получить следующий результат:

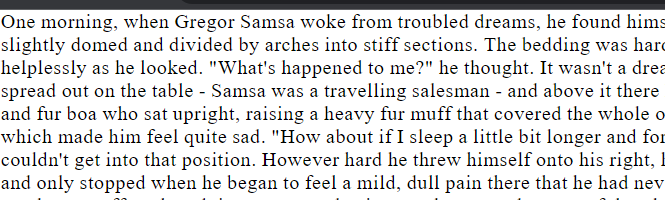


Рисунок 6 – Контейнер с осажденным сообщением

В данном случае изменен цвет тех же символов, что и в первом примере, но цвета, используемые для кодирования бинарной последовательности, являются близкими к черному, поэтому невооруженным взглядом заметить отличия в тексте проблематично.

При вытягивании сообщения анализируется цвет всех элементов текста контейнера. Если цвет соответствует оттенкам, указанным при осаждении, соответствующий бит записывает в буфер для декодированного сообщения.



Рисунок 7 – Результат вытягивания осаждённого сообщения

Вывод: в ходе лабораторной работы были изучены методы текстовой стеганографии и разработаны приложения для использования методов на основе модификации числа пробелов и модификации цвета.

**Контрольные вопросы**

1. В чем состоит сущность методов текстовой стеганографии?

Сущность методов текстовой стеганографии заключается в скрытом передаче информации в текстовых данных, чтобы эта информация оставалась незамеченной для посторонних наблюдателей. Основной целью текстовой стеганографии является сохранение конфиденциальности коммуникации путем внедрения секретных сообщений в текстовые ресурсы без вызова подозрений.

1. Охарактеризовать методы синтаксической текстовой стеганографии. Привести примеры конкретной реализации методов.

Методы синтаксической текстовой стеганографии основаны на изменении синтаксической структуры текста для встраивания секретной информации. Вместо изменения отдельных символов или форматирования, эти методы используют грамматические правила и структуру предложений для скрытой передачи данных. Ниже приведены примеры некоторых методов синтаксической текстовой стеганографии:

– 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.

1. Охарактеризовать методы лингвистической текстовой стеганографии. Привести примеры конкретной реализации методов. \

Методы лингвистической текстовой стеганографии основаны на использовании лингвистических свойств текста для встраивания секретной информации. Эти методы включают анализ семантической структуры текста, использование особенностей языка и контекстуальной информации для скрытой передачи данных. Вот некоторые примеры методов лингвистической текстовой стеганографии:

– Метод синонимов;

– Метод переменной длины слова;

– Метод первой буквы;

– Мимикрия.

1. Дать оценку стеганографической стойкости методов текстовой стеганографии при конвертации текста-контейнера в иной текстовый формат.

При конвертации текста-контейнера в другой текстовый формат следует учитывать следующие факторы:

– Визуальная и статистическая незаметность: Методы текстовой стеганографии должны обеспечивать незаметность визуального и статистического анализа текста. Если скрытые данные вызывают подозрение или приводят к изменениям в статистических свойствах текста, то метод может быть легко обнаружен.

– Устойчивость к изменениям: Методы должны быть устойчивыми к обычным операциям конвертации текста, таким как изменение кодировки, добавление или удаление пробелов, форматирование и другие преобразования. Если скрытые данные не сохраняются после таких операций, метод может быть непрактичным или легко обнаружимым.

– Сопротивление к атакам: Методы должны быть устойчивыми к попыткам атакующего восстановить или удалить скрытую информацию. Если скрытые данные могут быть успешно извлечены или удалены без больших усилий, метод считается небезопасным.

– Зависимость от контекста: Методы, которые сильно зависят от конкретного контекста текста или языка, могут быть менее устойчивыми при конвертации в другой текстовый формат, особенно если изменяется язык или структура предложений.

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

Замена символов: При визуальном стеганоанализе этот метод может быть уязвимым, 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 не выделяются визуально и выглядят естественно, то метод может быть стойким при визуальном стеганоанализе. Однако, если есть специфические паттерны или аномалии, которые могут быть обнаружены визуально, стойкость метода 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.

1. 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 определения наличия скрытых данных в тексте-контейнере и оценки стойкости стеганографического метода.

Характеристики метода "χ-квадрат" в стеганоанализе текстовой стеганографии:

Статистический анализ: Метод "χ-квадрат" основан на статистическом анализе распределения символов или других единиц текста в тексте-контейнере. Он сравнивает наблюдаемое распределение с ожидаемым распределением, чтобы определить наличие скрытых данных.

Оценка стойкости: Метод "χ-квадрат" позволяет оценить стойкость стеганографического метода путем вычисления статистической разницы между наблюдаемым и ожидаемым распределением. Если разница значительна, это может указывать на наличие скрытых данных.

Чувствительность к изменениям: Метод "χ-квадрат" может быть чувствительным к незначительным изменениям в тексте-контейнере, вызванным внедрением скрытой информации. Даже незначительные изменения в распределении символов могут быть обнаружены методом "χ-квадрат".

Ограничения: Метод "χ-квадрат" не является универсальным и может иметь ограничения в определении стеганографической стойкости. Он может быть менее эффективным, если стеганографический метод использует продвинутые техники, которые маскируют скрытые данные или обманывают статистический анализ.

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

Апрош – расстояние между соседними символами текста. Встраивание сообщения в контейнер может быть основано на модификации базового (устанавливаемого текстовым процессором по умолчанию) значения апроша, его изменением от базового до некоторого максимального значения, которое зрительно не должно отличаться от стандартного. Такое изменение производится с определенным шагом, каждому значению которого присваивается определенный бит или определенная комбинация битов.

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

Кернинг – это сдвиг между символами, используемый, как правило, для лучшей читаемости текста. Сущность метода на основе модификации кернинга заключается в использовании изменений промежутков между символами для внедрения скрытой информации. Небольшие изменения в кернинге между символами в тексте-контейнере используются для кодирования скрытого сообщения.

1. Дать сравнительную оценку методов на основе модификации пространственно-геометрических и цветовых параметров символов текста-контейнера (критерий: отношение оправданного объема осаждаемой информации к объему контейнера).

Методы на основе изменения цветовых параметров символов схожи с методом замены наименьших битов. Их эффективности при визуальном анализе напрямую зависит от числа модифицируемых битов цветов. Чем больше используется бит, тем меньше стойкость метода. При этом файловый вес контейнера в таких методах не изменяется.

Изменение же пространственно-геометрических параметров имеет меньшую стойкость, так как легко обнаруживается при визуальном анализе (легко заметить отсутствие засечек у символов).