**Лабораторная работа №3**

**Алгоритмы пассивного стеганоанализа**

***1. Визуальный стеганоанализ для контейнеров-изображений***

Рассмотрим принцип построения визуальной атаки, позволяющей выявить факт наличия скрываемого сообщения, вложенного в изображение-контейнер. Пусть стеганосистема построена таким образом, что наименее значимые биты (НЗБ) элементов изображения заменяются на биты скрываемого сообщения. Например, биты внедряемого сообщения замещают младшие биты яркостной компоненты каждого пикселя изображения. Ранее считалось, что НЗБ яркостной или цветовой компонент пикселя изображения, равно как и младшие биты отсчетов речевых или аудиосигналов независимы между собой, а также независимы от остальных битов элементов рассматриваемых контейнеров. Однако на самом деле это не так. Младшие биты не являются чисто случайными. Между младшими битами соседних элементов естественных контейнеров имеются существенные корреляционные связи. Также выявлены зависимости между НЗБ и остальными битами элементов естественных контейнеров.

На рисунке 1 показано изображение мельницы, слева рисунок представляет пустой контейнер, справа в каждый НЗБ цветовой компоненты пикселей последовательно бит за битом вложено скрываемое сообщение. Различие между пустым контейнером и заполненным контейнером визуально не проявляется. Но если изображение сформировать только из НЗБ пикселей, то можно легко увидеть следы вложения.

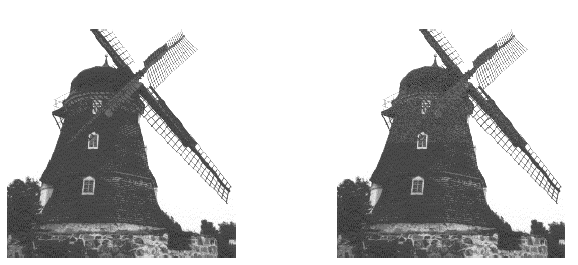


Рисунок 1. Изображение мельницы: слева – пустой контейнер, справа - с вложенным сообщением

На рисунке 2 слева показано изображение, состоящее из НЗБ пустого контейнера. Видно, что характер изображения существенно не изменился. Справа представлено изображение из младших битов наполовину заполненного скрываемым сообщением контейнера. Видно, что верхняя часть изображения, куда внедрено сообщение, представляет собой случайный сигнал.

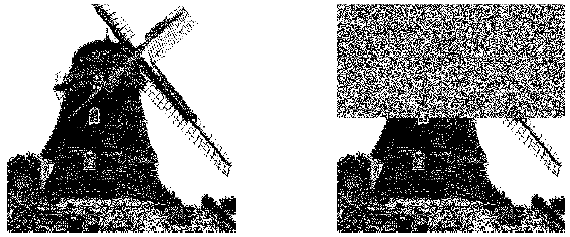


Рисунок 2. Визуальный стеганоанализ: слева – изображение из НЗБ пустого контейнера, справа – изображение из НЗБ наполовину заполненного

Если в стеганосистеме скрываемое сообщение до встраивания зашифровывается, то каждый его бит практически равновероятен и независим от соседних битов, что позволяет легко визуально выявить факт его встраивания, сопоставляя изображения из младших битов заполненного контейнера и пустых естественных контейнеров, соответственно. В некоторых стеганосистемах сообщения до встраивания сжимаются. Это целесообразно как для уменьшения размера скрытно внедряемой информации, так и для затруднения его чтения посторонними лицами. Архиваторы данных преобразуют сжимаемое сообщение в последовательность битов, достаточно близкую к случайной. Чем выше степень сжатия, тем ближе последовательность на выходе архиватора к случайной, и тем проще обнаружить факт существования стеганоканала при визуальной атаке. Однако даже если скрываемое сообщение до встраивания не шифруется и не сжимается, то его вероятностные характеристики не совпадают с вероятностными характеристиками НЗБ используемых пустых контейнеров.

Заметим, что отправитель сообщения может подобрать контейнер с законом распределения, совпадающим с законом распределения конкретного встраиваемого сообщения. В этом случае визуальная атака, как и статистические атаки, неэффективна. Но трудности подбора требуемого контейнера могут сделать такую стеганосистему непрактичной.

***2. Статистический стеганоанализ для контейнеров-изображений***

Визуальная атака целиком основана на способности зрительной системы человека анализировать зрительные образы и выявлять существенные различия в сопоставляемых изображениях. Визуальная атака эффективна при полном заполнении контейнера, но по мере уменьшения степени его заполнения глазу человека все труднее заметить следы вложения среди сохраненных элементов контейнера.

Кроме того, подобные методы не эффективны для методов сокрытия данных в частотной области.

Статистический стеганоанализ представляет введение в контейнер скрываемой информации как нарушение статистических закономерностей естественных контейнеров и является вероятностным. Анализируются статистические характеристики исследуемой последовательности и устанавливается, похожи ли они на характеристики естественных пустых контейнеров (если да, то скрытой передачи информации нет), или они похожи на характеристики заполненных контейнеров (если да, то выявлен факт существования скрытого канала передачи информации).

Класс статистических методов стеганоанализа использует множество статистических характеристик, таких как оценка энтропии, коэффициенты корреляции, вероятности появления и зависимости между элементами последовательностей, условные распределения, различимость распределений по критерию Хи-квадрат и многие другие.

Рассмотрим метод стеганоанализа на основе анализа статистики Хи-квадрат.

Пусть младший бит цветовой компоненты каждого пикселя контейнера-изображения заменяется битом скрываемого сообщения. Исследуем закономерности в вероятностях появления значений цветовой компоненты в естественных контейнерах и в сформированных стеганоконтейнерах. При замене младшего бита цветовой компоненты очередного пикселя контейнера на очередной бит предварительно зашифрованного или сжатого сообщения номер цвета пикселя стеганоконтейнера или равен номеру цвета пикселя контейнера, или изменяется на единицу. Для поиска следов вложения проводится анализ закономерностей в вероятностях появления соседних номеров цвета пикселей. Номер цвета, двоичное представление которого заканчивается нулевым битом, назовем левым (L), а соседний с ним номер цвета, двоичное представление которого заканчивается единичным битом - правым (R).

Пусть цветовая гамма исходного контейнера включает 8 цветов. Следовательно, при встраивании сообщения в НЗБ цветовой компоненты пикселей необходимо исследовать статистические характеристики в 4 парах номеров цвета. На рисунке 3 слева показана одна из типичных гистограмм вероятностей появления левых и правых номеров цвета в естественных контейнерах. Справа показана гистограмма вероятностей появления левых и правых номеров цвета в стеганоконтейнере.

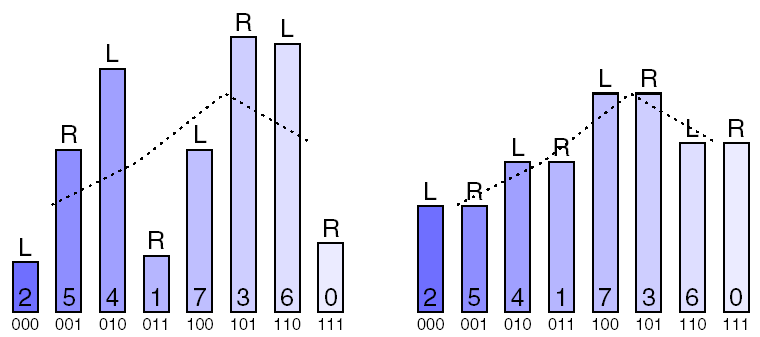


Рисунок 3. Гистограмма частот появления левых и правых номеров цвета, слева – до встраивания, справа – после

Видно, что вероятности появления левых и правых номеров цвета в естественных контейнерах существенно различаются между собой во всех парах, а в стеганоконтейнере эти вероятности выровнялись. Это является явным демаскирующим признаком наличия скрываемой информации. Заметим, что среднее значение вероятностей для каждой пары в стеганоконтейнере не изменилось по сравнению с контейнером (показано на рисунке 3 пунктирной линией).

При замещении битами внедряемого сообщения младших битов яркостной компоненты пикселей контейнера-изображения проявляются аналогичные статистические различия.

Степень различия между вероятностными распределениями элементов естественных контейнеров и полученных из них стеганоконтейнеров может быть использована для оценки вероятности существования стеганоканала. Данную вероятность удобно определить с использованием критерия согласия Хи-квадрат. По критерию Хи-квадрат сравнивается, насколько распределение исследуемой последовательности близко к характерному для стеганограмм распределению.

В исследуемой последовательности подсчитывается сколько раз  ее элемент  принял рассматриваемые значения, где всего  элементов. Например, в гистограмме левых и правых номеров цвета в левой части рисунка 3 номер цвета 000 появился 2 раза (), а номер 001 – 5 раз   
(). При встраивании очередных битов скрываемого сообщения в НЗБ этой пары номер цвета 000 должен появляться в среднем  раз

 (1)

Зная общее число *n* появления всех элементов исследуемой последовательности, легко подсчитать ожидаемую вероятность появления этих элементов в стеганоконтейнере по правилу:

= / (2)

Соответственно, для исследуемой последовательности вероятности равны:

= / (3)

Величина Хи-квадрат для сравниваемых распределения исследуемой последовательности и ожидаемого распределения стеганоконтейнера равна

, (4)

где  есть число степеней свободы. Число степеней свободы равно числу  минус число независимых условий, наложенных на вероятности . Наложим одно условие вида

 (5)

Вероятность *p* того, что два распределения одинаковы, определяется как

*p*, (6)

где *Г* есть гамма-функция Эйлера.

Чем больше значение *p,* тем выше вероятность встраивания скрываемой информации в исследуемую последовательность.

Рассмотрим использование критерия Хи-квадрат для отыскания следов стеганоканала.

Пусть в контейнерное изображение ”Мельница”, показанное в левой части рисунке 1, в НЗБ спектральных коэффициентов изображения, начиная с его верхнего края до его середины, последовательно внедрено 3600 байт скрываемого сообщения.

На рисунке 4 показана вероятность встраивания скрываемой информации в зависимости от размера исследуемой последовательности.

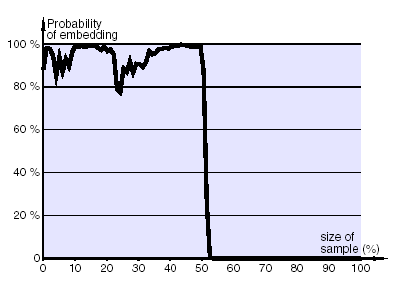


Рисунок 4. Вероятность встраивания по критерию Хи-квадрат

Начало графика получено при анализе первого фрагмента стеганоконтейнера, составляющего одну сотую часть всего стеганоконтейнера. Значение *p* составило 0,8826. Затем к анализируемому фрагменту была добавлена еще одна сотая часть стеганоконтейнера, и так далее. На втором шаге вероятность составила 0,9808 и далее при анализе стеганоконтейнера не опускалась ниже 0,77. При переходе к анализу нижней части изображения, не содержащей скрываемой информации, величина *p* скачком уменьшилась до нуля.

**Контрольные вопросы**

1. Визуальный стеганоанализ для контейнеров-изображений.

2. Статистический стеганоанализ.

3. Метод стеганоанализа на основе анализа статистики Хи-квадрат.

**Практическое задание**

Реализовать метод стеганоанализа на основе анализа статистики Хи-квадрат.