ФЕДЕРАЛЬНОЕ АГЕНТСТВО СВЯЗИ

Федеральное государственное образовательное бюджетное учреждение
высшего профессионального образования
«Санкт-Петербургский государственный университет телекоммуникаций
им. проф. М.А. Бонч-Бруевича»

На правах рукописи

Герлинг Екатерина Юрьевна

ИССЛЕДОВАНИЕ И РАЗРАБОТКА МЕТОДОВ ОБНАРУЖЕНИЯ СТЕГОВЛОЖЕНИЙ В НЕПОДВИЖНЫХ ИЗОБРАЖЕНИЯХ

05.12.13 – Системы, сети и устройства коммуникаций

Диссертация на соискание ученой степени
кандидата технических наук

Научный руководитель
доктор технических наук , профессор
Коржик Валерий Иванович

Санкт-Петербург – 2014

Оглавление

Введение 4
1 Обзор методов стеганографии и стеганографического анализа 11
1.1 Общие сведения 11
1.2 Современные стеганографические системы 15
1.3 Общие сведения о стеганографическом анализе 27
1.4 Выводы 30
2 Статистические критерии обнаруживаемости стегосистем 32
2.1 Критерий относительной энтропии 32
2.2 Критерий стойкости стегосистем, основанный на вычислении расстояния Бхаттачариа 35
2.3 ROC-кривые 44
2.4 Выводы 47
3 Исследование методов «слепого» стегоанализа 49
3.1 Общие сведения 49
3.2 Элементы МОВ 52
3.3 Функционалы, используемые для обнаружения стегообъекта по МОВ 57
3.4 Выводы 70
4 Целевой стегоанализ для метода вложения СГ-НЗБ 72
4.1 Визуальный метод стегоанализа 72
4.2 Стегоанализ на основе статистики 1-ого порядка (гистограммная атака) 78
4.3 Стегоанализ на основе статистики 2-ого порядка 87
4.4 Выводы 91
5 Целевой стегоанализ для метода вложения СГ-ШПС 94
5.1 Метод, основанный на «раздвоении пиков» 94
5.2 Стегоанализ, основанный на корреляции яркостей смежных пикселей 98
5.3 Стегоанализ, основанный на метод подсчета нулей в гистограмме 102
5.4 Стегоанализ, основанный на сравнении соседних значений гистограммы. 105
5.5 Выводы 108
6 Комплексные методы стегоанализа 110

6.1 Визуальная атака, применительно с СГ-ШПС 110
6.2 Стегоанализ, основанный на статистике 1-ого порядка, применительно к СГ-ШПС 115
6.3 Стегоанализ, основанный на статистике 2-ого порядка, применительно к СГ-ШПС 120
6.4 Стегоанализ, основанный на методе подсчета нулей гистограммы, применительно к СГ-НЗБ 123
6.5Стегоанализ, основанный на сравнении соседних значений гистограммы, применительно в СГ-НЗБ 125
6.6 Метод подсчета локальных максимумов 127
6.7 Комбинированный метод стегоанализа 134
6.8 Выводы 140
Заключение 142
Список используемых сокращений 144
Литература 145
Приложение А 150
Приложение Б 166
Приложение В 209
Приложение Г 210

Введение

Актуальность темы. На протяжении всей истории человечествасуществовала необходимость скрывать секретную информацию – будь то военные тайны или дворцовые секреты. Одновременно с этим появилась необходимость определять эту скрытую информацию. Так появились методы стеганографии (далее СГ), которые в отличие от криптографии позволяют не просто зашифровать, а совсем скрыть от посторонних присутствие секретной информации в некотором покрывающем объекте (далее ПО), которое выглядит абсолютно «невинно» и не вызывает подозрений. В противовес методам СГ появились методы стеганографического анализа (далее СГА), позволяющие выявить наличие скрытой информации в, на первый взгляд, «невинном» объекте.

Вместе с развитием цивилизации меняются и методы СГ и СГА.

Для оценки необнаруживаемости (секретности) СГС разработаны критерии секретности стегосистем. На сегодняшний день большинство широко распространенных критериев секретности довольно сложны для вычислений. Разработка простых и легких в вычислениях критериев секретности, которые можно будет автоматизировать, является важной задачей.

В современном мире широко распространено цифровое представление информации, например, различные компьютерные файлы – медиа или текстовые. Они хранятся на компьютере дома и в офисе, передаются по домашним и офисным сетям, а также через всемирную сеть Интернет, скачиваются на мобильные телефоны. Компьютерные файлы используются как контейнеры (ПО) дляскрытой передачи секретной информации.Современные методы цифровой СГ реализуются в стеганографический системах (СГС).

Современные методы цифровой СГ позволяют использовать в качестве ПО такие широко распространенные типы файлов, как:

- текстовые файлы (форматы DOC и TXT);
- музыкальные файлы (формат MP3);
- видео-файлы (формат AVI);

- файлы с изображениями (форматы BMP и JPEG);
- и другие.

Для выявления наличия скрытой информации в файлах разрабатываются современные методы цифрового СГА[1]. Такие методы позволяют достаточно точно определить, есть ли вложение секретной информации в данном объекте или нет.Некоторые методы цифрового СГА позволяют также оценить объем вложенной информации или определить, какой метод СГ использовался для вложения.

Сейчас широкое распространение получили алгоритмы, позволяющие вкладывать секретную информацию в неподвижные изображения, например, в цифровые фотографии или рисунки, созданные в программе Paint. Одни из наиболее распространенных алгоритмов вложения в неподвижные изображения на сегодняшний день являются алгоритмы:

- СГ-НЗБ – вложение в наименее значащие биты (далее НЗБ),данный алгоритм прост для самостоятельной реализации, но неустойчив к удалению;
- СГ-ШПС – вложение широкополосного сигнала (далее ШПС), данный алгоритм обладает стойкостью к удалению, но более сложен в реализации, чем СГ-НЗБ.

При этом для алгоритма СГ-НЗБ существуют различные методы СГА, широко освещенные в научных статьях. Тогда как для алгоритма СГ-ШПС методыСГА, достаточно надежно выявляющие наличиеили отсутствие вложения, мало описаны в научных статьях.

Целями данной работы являются экспериментальная проверка СГА для СГ-НЗБ,разработка и экспериментальная проверка методов СГА для СГ-ШПС, а также разработка простого и надежного критерия оценки секретности СГС.

Задачи исследования. При написании данной работы ставились следующие частные научные задачи:

1 Исследовать критерии секретности СГС.

2 Экспериментально исследовать эффективность уже имеющихся методов СГА для СГ-НЗБ. Выбрать пороговые значения на основании результатов исследования методов СГА для СГ-НЗБ, при которых соотношение вероятности ложной тревоги P_{fa} и вероятности пропуска P_m будет оптимальным. Экспериментально исследовать методы СГА СГ-НЗБ для СГ-ШПС.

3 Экспериментально исследовать методы СГА специально предназначенные для СГ-ШПС. Экспериментально исследовать методы СГА СГ-ШПС для СГ-НЗБ.

4 Экспериментально исследовать возможность комбинирования методов СГА для СГ-НЗБ и СГ-ШПС.

Методы исследования. В процессе проведения исследований использовались методы теории вероятностей и математической статистики, теории информации, а так же программы, написанные на языке программирования C++.

Достоверность научных результатов. Достоверность результатов исследования подтверждается:

- корректной постановкой задачи;
- результатами исследования и моделирования, не противоречащими друг друга и известным на сегодняшний день результатам исследований других авторов [2,3];
- апробацией основных положений в печатных трудах, в том числе в международных изданиях, и научно-исследовательской работе «Ярус-СГ».

Научная новизна. Основные результаты диссертации, обладающие научной новизной:

1 Предложен и обоснован критерий оценки секретности СГС, основанный на вычислении расстоянияБхаттачариа.

2 Предложены пороговые значения для известных методов СГА СГ-НЗБ, при которых отношение вероятности ложной тревоги P_{fa} и вероятности пропуска P_m будет оптимальным.

3 Предложены методыСГА для СГ-ШПС и СГ-НЗБ, основанные на методе подсчета нулей гистограммы и на сравнении соседних значений гистограммы.

4 Предложен комбинированный метод СГА, позволяющий эффективно обнаруживать вложения даже в тех случаях, когда не известно, какой именно метод вложения СГ-НЗБ или СГ-ШПС был использован.

Объем исследования.В диссертации исследуются ранее известные и предложенныеавтором критерии секретности СГС, а также ранее известные и предложенные автором методы СГА для алгоритмов вложения СГ-НЗБ и СГ-ШПС и предложенный автором комбинированный метод СГА, даны рекомендации по выбору пороговых значений для методов СГА СГ-НЗБ, приводятся результаты компьютерного моделирования алгоритмов вложения СГ-НЗБ и СГ-ШПС и их последующего СГА, а так же значения вероятностей ложной тревоги P_{fa} и пропуска P_m для различных методов СГА.

Область применения.Исследованы ранее известные и разработанный автором критерии секретности СГС, а также ранее известные и разработанные автором методы СГА позволяют производить проверку различных изображений, хранящихся на электронных носителях и передающихся по внутренним (Ethernet) и всемирной (Internet – Интернет) сетям на наличие или отсутствие вложенной секретной (скрытой) информации [4]. Рассмотренные в диссертации методы СГА могут применяться как в государственных, так и в коммерческих структурах.

Теоретическое и практическое значение работы.В современном мире для передачи секретной информации применяются методы стегоалгоритмов (например СГ-НЗБ или СГ-ШПС), позволяющих вкладывать сообщение в компьютерные файлы (например, в файлы с изображением), передающиеся через всемирную сеть Интернет. Выведенные в процессе исследований методов

СГАСГ-НЗБ пороговые значения для атак, основанных на статистике 1-ого и 2-ого порядка и предложенные автором методы СГА могут быть использованы для проверки изображений, передающихся через всемирную сеть Интернет на наличие в них скрытой информации.

Список публикаций.По теме диссертации опубликовано 6 работ, в том числе 2 статьи в журналах из перечня, рекомендованного ВАК, и одна статья в международном журнале.

Апробация и внедрение результатов.Результаты диссертации использовались в научно-исследовательской работе «Ярус-СГ»[5], выполненной фирмой НПП «Дигитон», что подтверждается Актом об использовании результатов диссертационной работы аспирантки Санкт-Петербургского Государственного университета телекоммуникаций им. проф. М.А. Бонч-Бруевича Герлинг Екатерины Юрьевна на тему «Исследование целевых методов обнаружения стегосистем», и в лабораторной работе «Методы обнаружения стегосистем НЗБ и ШПС» в курсе «Основы стеганографии» на кафедре Защищенные системы связи Санкт-Петербургского государственного университета телекоммуникаций им. проф. М.А. Бонч-Бруевича, что подтверждается Актом об использовании результатов диссертационной работы аспирантки Санкт-Петербургского государственного университета телекоммуникаций им. проф. М.А. Бонч-Бруевича Герлинг Екатерина Юрьевнына тему «Исследование и разработка методов обнаружения стеговложений в неподвижных изображениях».

Личный вклад автора.Разработка критерия стойкости СГС, основанного на вычислении расстояния Бхаттачариа, исследование уже имеющихся методов СГА для СГ-НЗБ, выбор пороговых значений для данных методов СГА, разработка методов СГА для СГ-ШПС и комбинированного метода СГА, разработка комплекса компьютерных программ, компьютерное моделирование, сбор статистики, анализ результатов и выводы по диссертации выполнены автором самостоятельно.

Научные положения, выносимые на защиту.

1 Критерий секретности СГС при оптимальных методах обнаружения, основанный на расстоянии Бхаттачариа.
2 Расчет эффективности известных методов СГА для СГ-НЗБ с расчетом оптимальных пороговых значений.
3 Методы СГА для СГ-ШПС, основанные на подсчете нулей гистограммы и на сравнении соседних значений гистограммы.

Структура и объем диссертации.В диссертации исследуются ранее известные и предложенные автором методы СГА для стегоалгоритмов вложения СГ-НЗБ и СГ-ШПС, а также комбинированный метод СГА.

В главе 1 описаны основные цели СГ, приведена классификация различных методов СГ, перечислены типы файлов, которые могут быть использованы в качестве ПО, приведен краткий обзор методов и классификация СГ и СГА, описаны основные методы вложения и извлечения секретной информации.Особое внимание уделено методам вложения СГ-НЗБ и СГ-ШПС, атакже описаны ряд других широко распространенных на сегодняшний день методов вложения, СГА для которых не рассматривается в рамках диссертации, сделаны выводы по главе.

В главе 2 рассмотрены критерии оценки секретности и необнаруживаемости СГС, описаны критерий относительной энтропии,предложен критерий оценки секретности СГС, основанный на вычислении расстояния Бхаттачариа, описан критерий, основанный на ROC-кривых, сделаны выводы по главе.

В главе 3 описаны методы «слепого» стегоанализа (далее ССГА), приведены общие сведения о ССГА, описаны элементы метода опорных векторов (далее МОВ),описаны функционалы, используемые для обнаружения СО по МОВ, сделаны выводы по главе.

В главе4 исследованы широко известные методы СГА алгоритма вложения СГ-НЗБ,подробно описаны принципы действия методов СГАдля СГ-НЗБ, приведены расчетные формулы, показаны результаты компьютерного

моделирования данных атак, собрана статистика по результатам моделирования, показывающая эффективность каждого метода, приведены значения вероятностей ложной тревоги P_{fa} и пропуска P_m, полученные путем анализа собранной статистики, выведены рекомендованные пороговые значений для каждого метода СГА, описаны преимущества и недостатки метода, сделаны выводы по главе.

В главе 5 описаны методы СГА для алгоритма вложения СГ-ШПС, разработанные автором,подробно описаны принципы действия методов СГАдля СГ-ШПС, приведены расчетные формулы, показаны результаты компьютерного моделирования данных атак, собрана статистика по результатам моделирования, показывающая эффективность каждого метода, приведены значения вероятностей ложной тревоги P_{fa} и пропуска P_m, полученные путем анализа собранной статистики, выведены рекомендованные пороговые значений для каждого метода СГА, описаны преимущества и недостатки метода, сделаны выводы по главе.

Глава 6 посвящена комплексным методам, которые могут быть использованы для анализа как алгоритма вложения СГ-НЗБ, так и алгоритма вложения СГ-ШПС, поэтому могут применяться в случае, когда алгоритм вложения не известен. Методы, подходяще для СГ-НЗБ проверены на СГ-ШПС, а методы, подходящие для СГ-ШПС – на СГ-НЗБ, проверка методов осуществлена с помощью компьютерного моделирования. На основании собранной статистике приведены значения вероятностей ложной тревоги P_{fa} и пропуска P_m, показана эффективность каждого метода применительно к алгоритмам СГ-НЗБ и СГ-ШПС,описан метод СГА, основанный на подсчете локальных максимумов и комбинированный метод СГА, сделаны выводы по главе.

В приложении А приведено более полное, чем в [3] доказательствометода СГА, основанного на статистике 2-ого порядка.

1 Обзор методов стеганографии и стеганографического анализа

1.1 Общие сведения

Стенография в переводе с греческого означает «тайнопись». СГ является одной из двух основных частей науки о скрытии информации (Information hiding), второй частью которой является цифровые водяные знаки.

Задача СГ – погрузить секретное (скрываемое) сообщение в ПО так, чтобы сам факт присутствия скрываемой информации нельзя было бы обнаружить нелегитимным пользователям. Основное отличие СГ от криптографии заключается в том, что криптография делает невозможным понимание содержания сообщения, сохраняя при этом возможность обнаружить факт его наличия (шумоподобные сигналы), тогда как СГ утаивает сам факт погружения скрываемой информации в ПО. Таким образом, основная концепция СГ – найти "шумовые компоненты" (области) в ПО и заменить их на зашифрованное (т.е. шумоподобное) секретное сообщение. В современном мире широко распространена цифровая СГ (здесь и далее речь пойдет именно о цифровой СГ, поэтому для краткости будем называть ее просто СГ), когда все ПО представляются в цифровой форме, а вложение и извлечение секретной информации производится на компьютерах.

СГ используется как:

- альтернатива криптографии при ее запрещении или ограничении уровня стойкости;
- скрытие пользователями секретной информации с ее дальнейшим хранением или передачей;
- передача секретной информации через транзитных пользователей;
- передача секретных сигналов и команд определенным пользователям сети интернет;
- отслеживание распространителей информации.

В качестве ПО в современной СГ часто используются:

- файлы с неподвижными изображениями;
- файлы с видео;
- аудио файлы;
- файлы, содержащие речь;
- файлы содержащие печатный смысловой текст;
- файлы с графическими представлениями текста и схем;
- файлы интернет-протоколов;
- файлы формата EXE (программы) и DLL.

В качестве секретной информации, которую необходимо скрыть, часто выступают:

- изображения;
- текстовые сообщения и данные;
- речевые сообщения.

ПО с вложенный в него скрываемым сообщением называется стеганографическим объектом (далее СО). В процессе хранения и передачи СО может быть подвергнут различным преобразованиям:

- естественные преобразования (фильтрация, сжатие, масштабирование. передача по каналам с шумом);
- преднамеренные (атаки, удаление скрытого сообщения).

Основной проблемой при разработке СГС является тот факт, что статистика таких сложных ПО как звук, изображение и смысловой текст известна не полностью и весьма сложна. Поэтому существует опасность, что атакующий знает ее лучше, чем разработчик СГС.

Отметим, что в современном мире СГС нередко используются преступниками для обмена информацией. Например, террористы применяют СГС для координации своих атак, используя при этом открытые каналы связи [6]. Попытка создания скрытого калана связи вызвала бы подозрения, но использование обычных общедоступных каналов связи не вызывает подозрений,

хотя невинное на первый взгляд сообщение может содержать важную информацию.

Например, по мнению специалистов, террористы, связанные с бэн Ладеном и группировкой Аль-Кайда передавали друг другу карты местностей, диаграммы, важные фотографии и текстовые послания, использую в качестве ПО файлы с неподвижными изображениями, сами изображения при этомразмещались, например, на сайтах с фотогалереями или даже [7].

Еще один пример использования СГ в незаконных целях – использование СГС русскими нелегалами (в том числе Анна Чапман), задержанными в США в июне 2010 года [7]. Американские и британские СМИ утверждают, что русские нелегалы использовали СГС для передачи секретной информации, спрятанной в неподвижных изображениях.

Обмен сообщениями через телекоммуникационные каналы между террористическими и криминальными группировками может производиться разными способами [8], например:

- через письма электронной почты;
- использование различных веб-сайтов (таких как E-Bay или сайты с фотогалереями);
- с помощьюсъемных носителей, хранения информации на жестких дисках и на собственных серверах;
- через взлом учетных записей пользователей различных социальных сетей, то есть использование «невинных» пользователей;
- использование телевизионных каналов для передачи сообщений;
- использование радиостанций для передачи сообщений;
- через рассылку СПАМ-сообщений.

Поскольку преступники и террористы все чаще используют методов СГ в незаконных целях, пересылая при этом сообщения по открытым каналам связи, разработка методов выявления СО среди большого потока невинных объектов на сегодняшний день является важно задачей.

Для борьбы с СГ были разработаны различные методы СГА (атаки на СГС).

При разработке методов СГ и СГА принято считать, что атакующий знает:

- модель СГС, то есть алгоритмы вложения и извлечения информации, если они не являются частью ключа;
- общие статистические свойства ПО. Стоит отметить, что атакующий никогда не должен знать в точности ПО, иначе становиться возможным тривиальная атака по обнаружению скрытой информации путем сравнения ПО и исследуемого (проверяемого) объекта.

Для оценки СГ алгоритмов разработаны критерии эффективности СГС:

- вероятность пропуска СО;
- вероятность ложного обнаружения СО;
- вероятность ошибочного бита при извлечении легитимными пользователями вложенного сообщения;
- качество ПО после вложения (отношение сигнал/шум или более сложные, в том числе экспертные, оценки);
- скорость вложения (число бит вложенного сообщения на один отсчет ПО).

Основная классификация СГС:

- для легитимных пользователей:
  - с известным ПО на легитимном декодере (информированный декодер);
  - с неизвестным ПО на легальном декодере ("слепой" декодер);
  - с использованием ПО в легальном кодере (информированный кодер);
- для атакующих:
  - с известном СО (выполняется всегда);
  - с известным скрываемым сообщением;
  - с выбранным скрываемым сообщением;
  - с известным (или выбранным ПО) – для каналов с шумом.

1.2 Современные стеганографические системы

Как было сказано ранее, в современных СГС в качестве ПО используются компьютерные файлы[9]. Для разных алгоритмов СГ используются файлы разных форматов. Программное обеспечение, которое использует современных методов СГ, широко распространены во всемирной сети.

В приведенной ниже таблице 1.1 показаны наиболее распространенные на сегодняшний день СГ приложения [5].

Таблица 1.1 – Перечень СГ приложений (прочерки поставлены в тех местах, где характеристики используемого метода СГ не известны)

Программное обеспечение	**Метод вложения**	**Покрывающие объекты**
Steganos Security Suite 2006	–	Файлы с неподвижным изображением и музыкальные файлы
StegoVideo	–	Видео файлы
StegaNote	Модифицированный алгоритм вложения в наименее значащие биты	Файлы с неподвижным изображением формата BMP 24 бит/пиксель
StegoMagic	–	Текстовые файлы, файлы формата WAV, BMP 24 бита/пиксель и 256 цветов.
Puff	16 различных алгоритмов вложения	Файлы форматов BMP, JPG, PCX, PNG, TGA, AIFF, MP3, NEXT / BC, WAV, и Win PE модули
wbStego4.3open	–	Файлы форматов BMP, TXT, HTML / XML, PDF
Steganography 4.0	–	Файлы формата BMP

Продолжение таблицы 1.1

Программное обеспечение	**Метод вложения**	**Покрывающие объекты**
SecurEngine Professional 1.0	–	ФайлыформатаBMP, GIF, PNG, HTM. Создание архивов и саморасшифровывающихся архивов.
Hermetic Stego v6.5	–	Файлы формата BMP
PhotoCrypt 1.1	–	Файлы формата BMP
Invisible Secrets v4.0	–	Файлы формата JPEG, PNG, BMP, HTML и WAV
CryptArkan	–	Файлы формата WAV и BMP.
Gifshuffle v2.0	Используется перетасовка colourmap (порядок цветов в палитре меняется)	Файлы формата GIF
JPegX	–	Файлы формата JPEG.
The Third Eye	–	Файлы формата BMP, GIF и PCX.
WeavWav	–	Файлы формата WAV
InfoStego	–	Файлы формата BMP
Camouflage	Скремблирует скрываемый файл и добавляет его к концу выбранного ПО	Любой формат файла
BMP Secrets	–	Файлы формата BMP
S-Mail Shareware v1.3	–	Файлы формата EXE и DLL
S-Tools v4	–	Файлы формата BMP, GIF и WAV

Продолжение таблицы 1.1

Программное обеспечение	**Метод вложения**	**Покрывающие объекты**
Encrypt Pic	–	Файлы формата BMP 24 бита/пиксель
Contraband Hell Edition	–	Файлы формата BMP
Steghide 0.4.6.b	–	Файлы формата JPG, BMP, WAV и Au
Hide4PGP v2.0	–	Файлы формата BMP, WAV и VOC
Blindside	–	Файлы формата BMP
TextHide	Замена синонимами слов в тексте	Текстовые файлы
JP Hide and Seek	–	Файлы формата JPG
MP3Stego	–	Файлы формата MP3 и WAV.
Revelation	Алгоритм вложения в наименее значащие биты	Файлы формата BMP 24 бита/пиксель
Stego Machine	Алгоритмы вложения в наименее значащие биты или добавление информации в конец файла	Файлы формата GIF и JPEG.
Stella	Использует два алгоритма вложения	Файлы формата GIF, BMP и JPEG.
SGPO	Используется перетасовка colourmap (порядок цветов в палитре меняется)	Файлы формата GIF
Snow	Вложение осуществляется путем добавления табуляции и пробелов в конце строк	Текстовые файлы

Продолжение таблицы 1.1

Программное обеспечение	**Метод вложения**	**Покрывающие объекты**
Invisible Encryption (IVE)	–	Файлы формата GIF на интернет страницах.
Visual Encryption (VE)	–	Файлы формата GIF на интернет страницах.

Как видно из таблицы 1.1, в качестве ПО чаще всего используются[10]:

- файлы с неподвижными изображениями формата BMP;
- файлы с неподвижными изображениями формата JPEG;
- файлы с неподвижными изображениями формата GIF;
- файлы с оцифрованным аудиопотоком WAV.

Далее в данной работе для проведения экспериментов в качестве ПО и полученных их них СО будут использоваться неподвижные изображения формата BMP и JPEG. Неподвижные изображения в качестве исследуемых объектов выбраны по нескольким причинам:

- неподвижные изображения являются наиболее общим видом для медиафайлов, доступных на сегодняшний день в качестве ПО[11];
- большинство основных принципов построения СГС и методов СГА для неподвижный изображений могут быть легко адаптированы для других медиафайлов, таких как видеоили аудио;
- в данной работе привести примеры экспериментов проще на неподвижных изображениях, поскольку их можно распечатать на бумаге;
- в современных широко распространенных СГС наиболее часто в качестве ПО используются файлы с неподвижными изображениями, а именно файлы форматов BMP и JPEG (таблица 1.1).

Наиболее распространенные СГС(таблица 1.1):

- СГС, использующие СГ-НЗБ;

- СГС, использующие СГ-ШПС;
- лингвистические СГС (Л-СГС);
- графические СГС (Г-СГС);
- интернет СГС (И-СГС).

Рассмотрим более подробно наиболее распространенные СГС.

СГ-НЗБ изменяет наименее значащие биты ПО в соответствии с вкладываемыми битами скрываемого сообщения [12]. В качестве ПО можно использовать, например, файлы с неподвижным изображением формата BMP или JPEG. В качестве секретного сообщения может выступать любой текст или черно-белое изображения, например, чертеж или схема.

Пусть $C(n)$, $n = 1,2,\ldots,N$ это цифровые отсчеты. Эти отсчеты могут выбираться как во временной области (в области пикселей изображений формата BMP), так и в частотной области (например, DCT-коэффициенты формата изображения JPEG). Тогда эти отсчеты могут быть представлены в двоичной базе следующим образом:

$$C(n) = \sum_{i=0}^{L-1} c_i(n) 2^i \qquad (1.1)$$

где $c_i(n)$ – двоичные коэффициенты (биты),

L – количество бит для принятого метода квантования.

Тогда после погружения по методу СГ-НЗБ получаем отсчеты СО $C_w(n)$ в следующем виде:

$$C_w(n) = \sum_{i=1}^{L-1} c_i(n) 2^i + b(n), \qquad (1.2)$$

где $b(n) \in \{0,1\}$ – значение бита информации, вложенной в данный отсчет.

Таким образом в СГ-НЗБ при вложении младший (наименьший значащий) бит заменяется на бит вкладываемой информации.

Преимущества СГ-НЗБ:

- метод прост в реализации;
- дает небольшое искажение ПО;

- дает высокую скорость вложения (1 бит на отсчет);
- дает возможность вкладывать информацию не во все, а лишь в определенные пиксели, задаваемые секретным стегоключом, что понижает скорость вложения, но повышает секретность СО.

Недостатки СГ-НЗБ:

- достаточно легко обнаруживается с использованием специальных визуальных и статистических методов;
- вложенная информация легко удаляется при помощи рандомизации наименьших значащих бит, то есть при помощи их замены на случайно сгенерированные биты.

Одним из основных недостатков СГ-НЗБ считается нестойкость к визуальному и статистическому СГА. Для решения данной проблемы была создана модифицированная СГ-НЗБ(которая в зарубежной литературе называетсяLSB matching) [13]. В модифицированной СГ-НЗБ вложение происходит также в НЗБ, но по измененному алгоритму

$$C_w(n) = \begin{cases} C(n) \text{ если } b(n) = \text{НЗБ}(C(n)) \\ C(n)+1 \text{ с вероятностью вложения } 0{,}5 \\ C(n)-1 \text{ с вероятностью вложения } 0{,}5 \end{cases}.$$

Как видно из алгоритма вложения, данный метод СГ меньше влияет на гистограммуменьше, чем при вложении методом СГ-НЗБ (см. параграф 4.2).Следовательно, модифицированная СГ-НЗБ является более устойчивой к СГА, использующим изменения статистических свойств объекта, возникающие при вложении [14,15].

Основным недостатком модифицированной СГ-НЗБ, также как и у обычной СГ-НЗБ является возможность стегоаналитика легко удалить скрытое сообщение путем рандомизации НЗБ.

Тем не менее, в свободном доступе СГ приложения с использованием СГ-НЗБ встречаются значительно чаще, чем с использованием модифицированной СГ-НЗБ. Поэтому далее в данной работе будет рассматриваться СГ-НЗБ.

СГ-ШПС изменяет числовое значение байта на определенную величину в большую (например, при вложении бита с значением 1) или меньшую (при вложении бита с значением 0) сторону[13]. В качестве ПО можно использовать, так же как и в случае СГ-НЗБ, файлы с неподвижным изображением формата BMP или JPEG. В качестве секретного сообщения может выступать любой текст или черно-белое изображения, например, чертеж или схема

Вложение информации в СГ-ШПС выполняется по правилу:

$$C_w(n) = C(n) + \alpha(-1)^b \pi(n), n = 1,2 \cdots N \qquad (1.3)$$

где α – коэффициент глубины погружения,

b – вкладываемый бит (0 или 1) информации,

$\pi(n)$ – псевдослучайная последовательность, генерируемая по стегоключу.

Выделение вложенных бит производится легальным пользователем, имеющим стеганографический ключ (далее стегоключ)при помощи корреляционного декодера и, следовательно, способным сгенерировать $\pi(n)$. Здесь возможны два случая:

а) Информированный декодер (при известном $C(n)$)

Тогда решение о бите b принимается по правилу:

$$b = \begin{cases} 0, \text{если } \Lambda \geq 0, \\ 1, \text{если } \Lambda < 0, \end{cases} \qquad (1.4)$$

где $\Lambda = \sum_{n=1}^{N} (C'_w(n) - C(n))\pi(n)$.

Атаку по удалению вложенной информации можно описать как

$$C'_w(n) = C_w(n) + \varepsilon(n), n = 1,2,\ldots, N \qquad (1.5)$$

где $\varepsilon(n)$ – аддитивный шум с заданной мощностью σ_ε^2, который добавляет атакующий, пытаясь сделать невозможным извлечение бита b легальным пользователем.

В этом случае вероятность ошибки при извлечении бита b легальным пользователем может быть рассчитана по формуле:

$$p = Q\left(\sqrt{\frac{N}{\eta - 1}}\right), \qquad (1.6)$$

где $\eta = \eta_w / \eta_a$, $\eta_w = \frac{\sigma_c^2}{\alpha^2}$, $\eta_a = \frac{\sigma_c^2}{\alpha^2 + \sigma_\varepsilon^2}$, $Q(x) = \frac{1}{2\pi}\int_x^\infty e^{-t^2/2} dt$,

σ_c^2 – дисперсия ПО.

б) «Слепой» декодер

Используется правило (1.4), но Λ вычисляется по-другому:

$$\Lambda = \sum_{n=1}^{N} (C'_w(n) - m_c)\pi(n), \qquad (1.7)$$

где $m_c = E\{C(n)\}$ – среднее значение ПО.

Вероятность ошибки легального пользователя при извлечении бита b рассчитывается по формуле:

$$p = Q(\sqrt{N / \eta_w}). \qquad (1.8)$$

Из формул (1.6) и (1.8) видно, что как для информированного, так и для «слепого» декодера можно получить любую требуемую малую вероятность ошибки p при любом уровне аддитивного шума σ_ε^2, увеличивая N, однако, при этом уменьшается объем вкладываемой информации. Отметим, что информированный и «слепой» декодер существенно отличаются по помехоустойчивости.

Преимущество «слепого» декодера перед информированным в том обстоятельстве, что у декодера не всегда присутствует абсолютно точная копия ПО $C(n)$.

Таким образом, использование СГ-ШПС позволяет избежать атаки удаления вложенной информации, тем более что стегоаналитик не может выбрать слишком большую мощность шума σ_ε^2, поскольку это приведет к ухудшению качества ПО. (Если бы это условие не накладывалось, атака удаления тривиально выполнялась бы для любых СО при помощи простой недоставки СО получателю). Однако существование атаки удаления требует от разработчика

СГСиспользования СГ-ШПС, что всегда приведет к уменьшению объема вкладываемой информации.

Обнаруживаемость СГ-ШПС определяется глубиной вложения α и видом псевдослучайной последовательности $\pi(n)$.

Л-СГС опираются на лингвистические свойства языка, при этом необходим поиск тех частей ПО, которые могут быть заменены на другие без ущерба для ПО[12]. В качестве ПО можно использовать файлы, содержащие текст, например, файлы формата DOC и TXT. В качестве секретной информации может выступать текст.

Чтобы Л-СГС не вызывала подозрений, она не должна менять структуру языка, сохраняя грамматику, синтаксис, семантику и т.д.

Один из методов построения Л-СГС – использование абсолютных или относительных синонимов.

Абсолютный синоним – это слово или фраза, которые могут быть заменены другими словом или фразой в любом контексте без изменения основного смысла текста, используемого как ПО. Например: взгляд – взор, годный – пригодный, гостиница – отель, грусть – печать, доля – часть, лгун – лжец, незаконный – противозаконный.

Относительные синонимы – это слова или фразы, которые могут или не могут заменить друг друга в зависимости от контекста (окружения этих слов или фраз). Например: дать ход (документу) – направить, дать ход (от преследователей) – уехать.

При использовании Л-СГС важным фактором при использовании относительных синонимов является взаиморасположения – допустимое сочетание соседних слов. Подбор относительных синонимов напрямую зависит от соседних с заменяемым слов и от контекста в целом.

Абсолютные и относительные синонимы для различных языков собраны в специальных словарях (например, «Словарь синонимов русского языка»).

Преимущества Л-СГС:

- идеальная секретность[16];

- широкий выбор ПО, в качестве ПО может выступать любой текст.

Недостатки Л-СГС:

- низкая скорость вложения, а, следовательно, передачи секретной информации;
- вложенная информация легко удаляется при помощи замены всех или части слов на синонимы.

Г-СГС вкладывает информацию в графические представления текста путем модификации различных графических характеристик исходного ПО [12]. В качестве ПО можно использовать файлы, содержащие текст и формула, например, файлы формата DOC, файлы, содержащие картинки и схемы, например. Файлы формата BMP. В качестве секретной информации может выступать текст

Г-СГС для вкладывания информации в ПО с помощью:

- изменение расстояний между словами или предложениями;
- изменение интервалов между строками;
- сдвиги отдельных слов вверх и/или вниз;
- незначительно, незаметное для невооруженного глаза, вращение строк.

Основной недостаток Г-СГС – это то, что СО легко обнаруживается при использовании статистического СГА.

Одно из методов Г-СГС, более секретный, чем описанные выше – метод имитации шумового сканирования.

При использовании метода имитации шумового сканирования сканируется уже напечатанный документ, затем в отсканированный файл вносится скрываемая информация, имитируя шум сканера.

Преимущества Г-СГС:

- извлечение скрытой информации легальным пользователем производиться без ошибок;
- СО устойчиво к визуальным атакам и к простейшим статистическим атакам.

Недостатки Г-СГС:

- низкая скорость вложения, а, следовательно, передачи секретной информации;
- вложенная информация легко удаляется при помощи рандомизации графических параметров текста.

И-СГС вкладывает информацию в различные протоколы передачи данных, при этом возможно вложить информацию на всех уровнях OSI[12]. В качестве ПО можно использовать различные интернет-протоколы, например TCP/IP. В качестве секретной информации может выступать текст.

Способы вложения скрываемой информации на различных уровнях архитектуры приведены в таблице 1.2.

Таблица 1.2 – Способы вложения скрываемой информации на различных уровнях

Уровень архитектуры	**Способ вложения**
Прикладной уровень	Обычные методы СГ
Представительский уровень	Погружение данных в поля системных сообщений
Сессионный уровень	Мониторинг чтения пользователями удаленных дисков
Транспортный уровень	Вложение в неиспользуемые части TCP заголовков
Сетевой уровень	Вложение в свободные поля IP пакетов
Уровень данных	Вложение в заголовки фреймов
Физический уровень	Конфликтные ситуации с пакетами: 0 – посылка пакета после конфликта с задержкой, 1 – посылка пакета без задержки

Еще одна из известных СГ – СГС на основе канала с шумом.

Основная идея построения СГС в каналах с шумами – замаскировать вложенное сообщение под шум канала.

Особенности модели:

- обнаружение ведется по зашумленной версии $C_w(n)$;
- ПО может быть в точности известно атакующему.

Практические приложения:

- передача СО по каналам спутниковой, мобильной и оптико-волоконной связи;
- перехват по побочным каналам;
- имитация каналов с шумом.

Задача атакующего в случае СГС на основе канала с шумом – отличить, присутствует ли только наложение шума канала, или суммы шума канала и стеганографического сигнала.

Разработка СГС на основе канала с шумом проще, чем в «классическом» случае, поскольку описание статистики шума канала известно лучше, чем статистики ПО.

1.3 Общие сведения о стеганографическом анализе

Для того, чтобы понять, если ли скрытая информация в исследуемом контейнере или нет применяются различные методы СГА (атаки на СГС). СГА развивался параллельно с развитием СГ. Для новых СГС придумывались и новые атаки. В современном мире на ряду с цифровой СГ применяется и цифровой СГА.

Классификация методов СГА:

- направленный (targeted) СГА: изучаетСГС и выясняет статистические отличия ПО и СО;
- различающий (distinguishing) СГА: СО обрабатывается таким образом, чтобы получить аппроксимацию ПО, которое использовалось в данной СГС, далее сравниваются СО и аппроксимация;
- ССГА: производится «обучение» идентификатора СО по большому количеству СО и ПО, что позволяет выработать в некотором смысле оптимальный алгоритм, принимающий решение о том, является ли представленный образец ПО или СО.

К основным атакам на СГС относят:

- обнаружение СО;
- нахождение объема (доли) скрытой информации;
- определения типа использованного алгоритма вложения;
- чтение скрытой информации;
- удаление скрытой информации без значительного ухудшения качества ПО и даже при необнаружении СО.

Для оценки эффективности различных методов СГА введены специальные критерии:

- P_m – вероятность пропуска – вероятность того, что при применении данного метода СГА результат будет показывать отсутствие скрытого сообщения, хотя на самом деле исследуемый объект будет содержать вложение;
- P_{fa} – вероятность ложной тревоги (ложного обнаружения) – вероятность того, что при применении данного метода СГА результат будет показывать наличие скрытого сообщения, хотя на самом деле исследуемый объект не будет содержать вложения.

Информационный канал в котором создается (используя СГС) и по которому передается СО называют стегоканалом.

Стегоканал состоит из пяти основных элементов:

- канал передачи информации;
- источник ПО;
- скрываемое сообщение;
- алгоритмы вложения и извлечения скрываемого сообщения;
- стегоключ, позволяющий вкладывать скрываемое сообщение, и стегоключ, позволяющий извлекать вложенное скрываемое сообщение (в некоторых СГС эти два стегоключа могут совпадать).

При разработке СГА для принято считать, что стегоканал соответствует принципу Кирхгоффа, то есть алгоритмы вложения и извлечения срываемой информации, а также стегоключ для вложения (если он отличается от стегоключа для извлечения) и общие статистические свойства ПО доступны атакующему и не являются секретными; секретным является только стегоключ для извлечения. На практике принцип Кирхгоффа не всегда выполняется и у атакующего крайне мало информации о стегоканале. Отметим, что для создания эффективного СГА атакующему важна любая информация о СГС и стегоканале.

Атакующий может обладать разным объемом информации об элементах стегоканала. Чем большим объемом информации о стегоканале обладает атакующий, тем более эффективным будет метод СГА, используемый атакующим для анализа подозрительных объектов.

Поскольку в данной работе рассматривается цифровая СГ, то под подозрительными объектами подразумеваются медиафайлы. Отметим, что в современном мире, где большинство информации представлено (или продублировано) в цифровом виде, объем медиафайлов, подозрительных для атакующего, может оказаться очень большим[13].

Существует два способа проанализировать большой поток медиафайлов:

- мониторинг траффика узла коммуникационной сети;
- анализ содержимого «захваченного» компьютера.

Мониторинг траффика можно организовать в автоматическом режиме, с помощью программы, которая проверяет все медиафайлы, проходящие через этот контролируемый узел, и разделяет их на две группы:

- группа медиафайлов, являющихся СО;
- группа «чистых» медиафайлов, не содержащих скрытой информации.

Основной недостаток мониторинга трафика коммуникационной сети – у атакующего очень мало информации о ПО, скрываемом сообщении или СГС. Это создает сложности при СГА, поскольку в данном случае принцип Кирхгоффа выполняется не полностью. В этом случае атакующему надо использовать такой СГА или набор СГА, который позволяет обнаружить максимально широкий спектр СГС.

В случае анализа содержимого «захваченного» компьютера атакующему проводить СГА значительно проще.

На жестком диске могло сохраниться программное обеспечение, которое было использовано для вложения скрываемой информации. И даже если оно было удалено с компьютера, следы его присутствия все равно могли сохраниться. Эта информация дает атакующему возможность определить, какой алгоритм вложения был использован.

Также атакующий может найти на жестком диске несколько почти одинаковых версий одного и того же файла. В этом случае можно исследовать разные версии файла независимо друг о друга как с использованием различных методов СГА, так и с помощью других процессов, например таких, как компрессия JPEG. Такое исследование позволит понять, является ли какой-то из этих файлов СО, а если является, то можно сделать важные выводы о том, какие изменения происходят с ПО после вложения в него срываемой информации.

Для анализа «захваченного» компьютера также можно воспользоваться специальным программным обеспечением, которое позволит проводить исследование файлов на компьютере в автоматическом режиме. При этом методы СГА для одних и тех же файлов при разных способах анализа будут одинаковые.

Какой бы способы анализа большого потока схожих по своим свойствам медиафайлов не использовался, они осуществляется с помощью одних и тех же методов СГА. Другими словами, для эффективного анализа объектов важны используемые методы СГА и их эффективность. Поэтому так важно разрабатывать новые методы СГА и совершенствовать уже существующие. Использовать один и тот же метод СГА для совершенно разных по статистический свойствам объектов неэффективно, поэтому для анализа большого объема медиафайлов необходимо разрабатывать комбинированные методы СГА, включающие в себя различные методы, созданные для анализа медиафайлов, обладающих разными статистическими свойствами.

1.4 Выводы

С развитием цифровых форм коммуникации СГполучила новый виток развития – цифровая СГ, использующая в качестве ПО цифровые медиа файлы. СГС, использующие цифровое представление информации и методы СГА, созданные для этих СГС, с каждым годом совершенствуются. Интерес с СГ и СГА постепенно увеличивается, особенно после того, как стало понятно, что СГ может быть использована в криминальных целях. Отметим, что СГ перестала

быть уделом профессионалов, широкий выбор специализированных программ, размещенных в свободном доступе с сети Интернет, как платных, так и бесплатных, предоставляют любому пользователю возможность использовать цифровую стеганографию в своих целях, в том числе и незаконных. Для борьбы с незаконным использованием цифровой СГ разрабатываются методы цифрового СГА.

Большинство программного обеспечения для создания СО, размещенного в Интернете, использует в качестве ПО наиболее распространенные форматы файлов – видео, музыку и изображения. Это программное обеспечение находится в свободном доступе и все, у кого есть подключение к Интернету могут использовать его. Как показано в параграфе 1.1 и в [7,8], СГС часто используются в террористических и криминальных группировках для обмена информацией. Разработка методов пресечения террористический и криминальных стегоканалов связи, то есть разработка методов СГА, позволяющих выявить криминальные стегоканалы, является важной задачей. Известные методы СГА не позволяют в полной мере решать задачи обнаружения вложений, поэтому задача исследования существующих и разработка новым методов СГА является актуальной, и ей посвящены последние главы диссертации.

Одним из наиболее распространенных форматов, используемых в различных СГС, является формат BMP. Для данного формата разработано множество методов вложения, из которых наиболее распространенные – СГ-НЗБ и СГ-ШПС. Разработка и усовершенствование методов СГА для данных методов вложения является важной и актуальной задачи, и именно этой задаче посвящена данная диссертация.

Для проверки эффективности рассматриваемых методов СГА проведены эксперименты, с использованием в качестве ПО файлов с неподвижными изображениями формата BMPGrayScale 8 бит/пиксель (256 оттенков серого).

2 Статистические критерии обнаруживаемости стегосистем

Атаки на СГС направлены на обнаружение факта присутствия скрытой информации. Для этого тестируется две гипотезы:

- H_0 – в тестируемом объекте нет скрытой информации;
- H_1 – в тестируемом объекте есть скрытая информация.

Атакующий на основании результатов СГА решает, какая из данных гипотез верна. Некоторые СГА также позволяют оценить объем (долю) скрытой информации, прочитать (расшифровать) скрытую информацию или удалить ее из СО даже если гипотеза H_1 не подтвердилась, без ухудшения качества ПС.

Одним из основных критериев оценки эффективности СГС является ее необнаруживаемость (стойкость). СГС называется необнаруживаемой (стойкой) если атакующий не может отличить одну гипотезу от другой, другими словами, даже при использовании наилучшего метода СГА выбор между гипотезами равносилен случайному угадыванию.

2.1 Критерий относительной энтропии

Если известно вероятностное распределение ПО P_S и вероятностное распределение СО P_X, то для определения степени необнаруживаемости СГС в современной научно-технической литературе используется относительная энтропия (расстояние Кульбака-Лейблера[17]) $D(P_S \| P_X)$, $D(P_X \| P_S)$. Относительная энтропия – это мера удаленности друг от друга двух вероятностных распределений. Она, вообще говоря, несимметрична, то есть $D(P_S \| P_X) \neq D(P_X \| P_S)$.

В случае двух непрерывных распределений P и Q относительная энтропия может быть рассчитана, как показано в [18]:

$$D(P\|Q) = \int_{\Omega} P(w)\log\left(\frac{P(w)}{Q(w)}\right)dw, \qquad (2.1)$$

где Ω – область допустимых значений параметров тестируемых объектов.

Эффективность проверки гипотез можно оценить при помощи двух вероятностей:

- вероятности пропуска P_m, когда атакующий решает в пользу гипотезы H_0, хотя в исследуемом объекте содержится скрытая информация;
- вероятности ложной тревоги P_{fa}, когда стегоаналитик решает в пользу гипотезы H_1,хотя в исследуемом объекте не содержится скрытой информации.

Нижнюю границу [19,20] для байсовой вероятности ошибки $P_e = \pi_0 P_{fa} + \pi_1 P_m$ при использовании наилучшего статистического обнаружителя можно представить через относительную энтропию как

$$P_e > \pi_0\pi_1 \exp\left(\frac{-J}{2}\right),$$

где $J = D(P_X\|P_S) + D(P_S\|P_X)$,

π_0 и π_1 – априорные вероятности отсутствия или присутствия СО соответственно.

Согласно теории информации [18] при тестировании любых двух гипотез оптимальным методом справедливы неравенства:

$$P_{fa}\log\frac{P_{fa}}{1-P_m} + (1-P_{fa})\log\frac{1-P_{fa}}{P_m} \le D(P_S\|P_X), \qquad (2.2)$$

$$P_m\log\frac{P_m}{1-P_{fa}} + (1-P_m)\log\frac{1-P_m}{P_{fa}} \le D(P_X\|P_S). \qquad (2.3)$$

При $P_{fa} = 0$ из неравенства (2.2) получаем $P_m \ge 2^{-D(P_S\|P_X)}$.

Из (2.2) и (2.3) видно, что СГС будет абсолютна секретна (необнаруживаема), если $D(P_S\|P_X) = 0 = D(P_X\|P_S)$. Как доказано в

[20],относительные энтропии $D(P_S\|P_X)$ и $D(P_X\|P_S)$, взятые из (2.1) для N-мерных гауссовских распределений с нулевым средним значением P_S и P_X будут иметь вид:

$$D\left(P_{X^N}\middle\|P_{S^N}\right)=-\frac{1}{2}\ln\det\left(I_N+\delta R\right)+\frac{1}{2}tr\left(\delta R\right),$$

$$D\left(P_{S^N}\middle\|P_{X^N}\right)=-\frac{1}{2}\ln\det\left(I_N+\delta\widetilde{R}\right)+\frac{1}{2}tr\left(\delta\widetilde{R}\right), \qquad (2.4)$$

где I_N – единичная матрица размером $N\times N$,

$tr(\bullet)$ – след матрицы,

R_{X^N} и R_{S^N} – матрицы ковариаций гауссовских N-мерных случайных векторов X^N (для СО) и S^N (для ПО) соответственно, а

$$\delta R = R_{X^N}R_{S^N}^{-1}-I_N,$$

$$\delta\widetilde{R} = R_{S^N}R_{X^N}^{-1}-I_N.$$

Длядостаточно секретной СГС, где R_{X^N} близко к R_{S^N}, параметр J в этом случае может быть аппроксимирован соотношением

$$J\approx\frac{1}{2}tr\left(\delta R^2\right). \qquad (2.5)$$

Как видно из формулы (2.5), даже в случае достаточно секретной СГС с гауссовским распределением с нулевым средним значением ПО мы сталкиваемся с серьезной проблемой при вычислении J по формуле (2.5). Более того, гауссовское распределение с нулевым средним значением не типично для большинства ПО, представленных в цифровом виде, например таких, как изображения (наиболее близкая к реальной модель цифрового изображения – это нестационарная модель с корреляцией между соседними пикселями [13]), и тогда вычисление относительной энтропии по (2.1) представляет собой проблему.

2.2 Критерий стойкости стегосистем, основанный на вычислении расстояния Бхаттачариа

Как было отмечено в параграфе 2.1, в некоторых случаях вычисление относительно энтропии, как меры необнаруживаемости СГС, представляет собой сложную задачу. Поэтому в [21] был предложен критерий стойкости СГС, основанный на вычислении расстояния Бхаттачариа.

В соответствии с [22], расстояниеБхаттачариа между двумя вероятностями распределения P_S и P_X в пространстве Ω определено как

$$D_B(P_S, P_X) = -\ln \rho_B(P_S, P_X), \tag{2.6}$$

где

$$\rho_B(P_S, P_X) = \int_{\Omega} \sqrt{P_X(w), P_S(w)}\, dw. \tag{2.7}$$

Из формул (2.6) и (2.7) следует, что $D_B(P_S, P_X) = D_B(P_X, P_S)$. Отметим, что D_B – это симметричная функция (в отличие от относительной энтропии), и, хотя она не удовлетворяет неравенству треугольника [22], ее небольшая модификация в виде $\sqrt{1-\rho_B^2}$ будет удовлетворять данному неравенству. Следовательно, при таком критерии нет необходимости вычислять две величины $D(P_S \| P_X)$ и $D(P_X \| P_S)$, как в случае с критерием, основанном на относительной энтропии.

Если априорные вероятности π_0 и π_1 равны (то есть $\pi_0 = \pi_1 = 1/2$), то байсовая вероятность ошибки P_e при проведении оптимального тестирования гипотезы будет ограничена двухсторонним неравенством

$$\frac{1}{4}\rho_B^2(P_X, P_S) \le P_e \le \frac{1}{2}\rho_B(P_X, P_S). \tag{2.8}$$

Для особого случая, когда $P_X = P_S$ получаем $\rho_B(P_S, P_X) = 1$ и $D_B(P_X, P_S) = 0$. Это означает, что СГС абсолютно секретна (необнаруживаемая). Если $D_B(P_X, P_S) = \varepsilon$, будем называть эту СГС ε -секретной. В этом случае

$$P_e \geq \frac{1}{4}\exp(-2\varepsilon)\,. \qquad (2.9)$$

Неравенство (2.9) показывает, что если $D_B\left(P_X, P_S\right)$ очень близко к 0 или $\rho_B\left(P_X, P_S\right)$ очень близко к 1, то СГС считается почти секретной.

Если X^N и S^N – N-мерные гауссовские случайные векторы, тогда [22]

$$D_B\left(P_{X^N}, P_{S^N}\right) = \frac{1}{8}\left(v_{X^N} - v_{S^N}\right)^{\mathrm{T}} R^{-1}\left(v_{X^N} - v_{S^N}\right) + \frac{1}{2}\ln\frac{\det R}{\sqrt{\det R_{X^N}\det R_{S^N}}}\,, \qquad (2.10)$$

где v_{X^N} и v_{S^N} – среднее значения N-мерных случайных векторов X^N и S^N соответственно,

R_{X^N} и R_{S^N} – матрицы ковариаций N-мерных случайных векторов X^N и S^N соответственно,

T – символ транспонирования матриц,

$$R = \frac{1}{2}\left(R_{X^N} + R_{S^N}\right).$$

В частном случае когда $v_{X^N} = v_{S^N}$ (т.е. когда средние значения векторов X^N и S^N равны 0) из (2.10) получаем

$$D_B\left(P_{X^N}, P_{S^N}\right) = \frac{1}{2}\ln\frac{\det R}{\sqrt{\det R_{X^N}\det R_{S^N}}}\,. \qquad (2.11)$$

В другом частном случае, когда $R_{X^N} = R_{S^N} = R$, из (2.10) получаем

$$D_B\left(P_{X^N}, P_{S^N}\right) = \frac{1}{8}\left(v_{X^N} - v_{S^N}\right)^{\mathrm{T}} R^{-1}\left(v_{X^N} - v_{S^N}\right).$$

Если пространство Ω дискретно, можно изменить (2.7) на

$$\rho_B\left(P_{X^N}, P_{S^N}\right) = \sum_{w\in\Omega}\sqrt{P_{X^N}(w)P_{S^N}(w)}\,.$$

Верхняя граница в (2.8) может быть улучшена [23] если в(2.6) использовать модифицированный коэффициент Бхаттачариа $\widetilde{\rho}_B\left(P_{X^N}, P_{S^N}\right)$ где

$$\widetilde{\rho}_B\left(P_{X^N}, P_{S^N}\right) = \min_{0\leq s\leq 1}\int_{\Omega} P_{X^N}^{s}(w)P_{S^N}^{1-s}(w)dw\,.$$

Для пояснения эффективности использования нового критерия рассмотрим широкополосное вложение аддитивным методом

$$X^N = \alpha S^N + W^N, \qquad (2.12)$$

где покрывающий объект S^N – гауссовский случайный вектор с i.i.d. (independent and identically distributed – независимое равновероятное распределение) значениями, т.е. $S^N \sim N(0, \sigma_S I_N)$ масштабируется с помощью коэффициента α, и W^N также выбран как гауссовский случайный вектор с нулевым средним значением и с дисперсией σ_w^2. Если коэффициент α выбран как $\alpha = 1 - \frac{D}{2\sigma_S^2}$ и $\sigma_w^2 = D\left(1 - \frac{D}{4\sigma_S^2}\right)$, тогда искажение после вложения равно D и $D(P_{X^N} \| P_{S^N}) = 0$ [20]. Это означает, что такая СГС абсолютно секретна.

Стоит отметить, что модель ПО с вектором независимого гауссовского распределения мало приемлема. Если принять, что S^N коррелированный источник, тогда, как доказано в [20], абсолютно секретная СГС может быть достигнута после применения к S^N преобразования Карунена-Лоева (ПКЛ), выполнения вложения (как в (2.12)) в ПКЛ-коэффициенты и восстановления СО путем обратного преобразования Карунена-Лоева.

Однако, применение ПКЛ требует знаний матрицы ковариации S^N. Ясно, что количество реальных медиафайлов, например изображений, которые могут быть использованы как ПО, огромно, и даже при том, что известна средняя матрица ковариации типичного изображения (или другого типичного медиафайла), ее нельзя использовать для оценки результатов ПКЛ для определенного изображения.

Рассмотрим функцию вложения из (2.12) для СО с гауссовской вероятностью распределения, где D и N выбраны таким образом, что бы, по

крайней мере, была обеспечена ε -секретность СГС, т.е. если $D(P_{X^N}\|P_{S^N})=\varepsilon$ или $D_B\left(P_{X^N},P_{S^N}\right)=\varepsilon$.

В этом случае атакующий должен уметь отличать гауссовский шум, вызванный вложением от того шума, который появляется самостоятельно в процессе преобразования (пересохранения, редактирования и т.д.). Обычно, шум имеет не гауссовское распределение и это требует использования вложения с таким шумом, который приближен к шуму самого устройства (устройства хранения, преобразования и т.д.) настолько, насколько это возможно.

Рассмотрим модель с окрашенный гауссовским ПО и вложением с i.i.d. гауссовским шумом. Главная цель исследования – найти, как корреляция стеготекста влияет на секретность СГС.

Предположим, что матрица ковариации гауссовского ПО S^N имеет вид

$$R_{S^N}=\sigma_S^2\begin{bmatrix}1 & r_{12} & \cdots & r_{1,N-1} & r_{1N}\\ r_{21} & 1 & \cdots & r_{2,N-1} & r_{2N}\\ \vdots & \vdots & \ddots & \vdots & \vdots\\ r_{N-1,1} & r_{N-1,2} & \cdots & 1 & r_{N-1,N}\\ r_{N1} & r_{N2} & \cdots & r_{N,N-1} & 1\end{bmatrix},$$

где для каждого элемента $i,j\le N$, $r_{ij}=\dfrac{E\left(S(i)S(j)\right)}{\sigma_S^2}$,

$S(i)$ – i -ые выборочное значение покрывающего сообщения.

Видно, что матрица ковариации стеготекста X^N после вложения по (2.12) будет выглядеть так

$$R_{X^N}=\sigma_S^2 d^{-1}\begin{bmatrix}d & r_{12} & \cdots & r_{1,N-1} & r_{1N}\\ r_{21} & d & \cdots & r_{2,N-1} & r_{2N}\\ \vdots & \vdots & \ddots & \vdots & \vdots\\ r_{N-1,1} & r_{N-1,2} & \cdots & d & r_{N-1,N}\\ r_{N1} & r_{N2} & \cdots & r_{N,N-1} & d\end{bmatrix},$$

где $d = \left(1 - \frac{D}{2\sigma_S^2}\right)^{-2}$.

В частном случае экспоненциальной корреляции, где S^N – это AR(1) последовательность с нулевым средним, появляется матрица Теплица R_{S^N} и «почти» матрица Теплица R_{X^N}

$$R_{S^N} = \sigma_S^2 \begin{bmatrix} 1 & r & \cdots & r^{N-2} & r^{N-1} \\ r & 1 & \cdots & r^{N-3} & r^{N-2} \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ r^{N-2} & r^{N-3} & \cdots & 1 & r \\ r^{N-1} & r^{N-2} & \cdots & r & 1 \end{bmatrix},$$

$$R_{X^N} = \sigma_S^2 d^{-1} \begin{bmatrix} d & r & \cdots & r^{N-2} & r^{N-1} \\ r & d & \cdots & r^{N-3} & r^{N-2} \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ r^{N-2} & r^{N-3} & \cdots & d & r \\ r^{N-1} & r^{N-2} & \cdots & r & d \end{bmatrix}. \qquad (2.13)$$

В то же время вычислить относительную энтропию по формуле (2.4) (и даже по формуле (2.5)) для данных матриц довольно трудно.

Однако, если мы примем расстояние Бхаттачариа как статистическую разницу между двумя гауссовскими окрашенными распределениями с нулевыми средними P_{X^N} и P_{S^N}, то из (2.6), (2.8) и (2.11) получаем

$$P_e \geq \frac{1}{4}\rho^2 = \frac{1}{4}\frac{\sqrt{\det R_{X^N} \det R_{S^N}}}{\det\left(\frac{1}{2}\left(R_{X^N} + R_{S^N}\right)\right)}, \qquad (2.14)$$

где R_{S^N} и R_{X^N} взяты из (2.13). После простого преобразования и использования формулы (2.14) будем иметь

$$R = \frac{1}{2}\left(R_{X^N} + R_{S^N}\right) = \sigma_S^2 t \begin{bmatrix} t^{-1} & r & \cdots & r^{N-2} & r^{N-1} \\ r & t^{-1} & \cdots & r^{N-3} & r^{N-2} \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ r^{N-2} & r^{N-3} & \cdots & t^{-1} & r \\ r^{N-1} & r^{N-2} & \cdots & r & t^{-1} \end{bmatrix},$$

где $t = \dfrac{1+d^{-1}}{2}$. Детерминант матрицы Теплица R_{S^N} можно легко вычислить [24] как

$$\det R_{S^N} = \sigma_S^{2N}\left(1-r^2\right)^{N-1}. \tag{2.15}$$

Для «почти» матриц Теплица R_{S^N} и R их детерминанты могут быть найдены [25] как

$$\det R_{X^N} = \left(d^{-1}\sigma_S^2\right)^N a_N,\ \det R = \left(t\sigma_S^2\right)^N b_N, \tag{2.16}$$

где параметры a_N и b_N вычисляются через рекуррентные формулы

$$\begin{aligned} &\alpha_0 = 1 \\ &\alpha_1 = d + (d-2)r^2 \\ \forall n = 2,\ldots,N-1\wedge: \quad &\alpha_n = \left(d + (d-2)r^2\right)\alpha_{n-1} - \left((d-1)r\right)^2 \alpha_{n-2} \\ &a_N = d\alpha_{N-1} - \left((d-1)r\right)^2 \alpha_{N-2} \end{aligned} \tag{2.17}$$

и

$$\begin{aligned} &\beta_0 = 1 \\ &\beta_1 = d_1 + (d_1-2)r^2 \\ \forall n = 2,\ldots,N-1\wedge: \quad &\beta_n = \left(d_1 + (d_1-2)r^2\right)\beta_{n-1} - \left((d_1-1)r\right)^2 \beta_{n-2}, \\ &b_N = d_1\beta_{N-1} - \left((d_1-1)r\right)^2 \beta_{N-2} \end{aligned} \tag{2.18}$$

где $d_1 = t^{-1}$.

Приведем пример. Возьмем $r = 0{,}5$ и $N = 200$. Тогда из (2.14), (2.15) и (2.16) получаем $\rho^2 = 0{,}856$.

Хотя нижняя граница для P_e приведена здесь в замкнутой форме, ее вычисления для больших N по рекуррентным формулам (2.17) и (2.18) все еще остается трудной задачей.

Можно заменить [26] рекуррентное вычисление (2.17) на явную форму

$$\alpha_n = c_1 x_1^n + c_2 x_2^n , \tag{2.19}$$

где x_1 и x_2 – корни квадратного уравнения

$$x^2 - \left(d + (d-2)r^2\right)x + \left((d-1)r\right)^2 = 0 , \tag{2.20}$$

и коэффициенты

$$c_1 = \frac{\left(d + (d-2)r^2\right) - x_2}{x_1 - x_2},\ c_2 = 1 - c_1 .$$

Упрощенная формула для второго рекуррентного отношения (2.18) может быть найдена аналогично. Но, к сожалению, для больших N вычисления по (2.19) и (2.20) все еще достаточно сложны.

Для «хорошей» СГС ожидается $d \approx 1$, следовательно $\left((d-1)r\right)^2 << 1$, и тогда аппроксимация рекуррентного отношения (2.17) принимает вид:

$$\forall n = 2, \ldots, N-1 : \begin{matrix} \alpha_n \approx \left(d + (d-2)r^2\right)\alpha_{n-1} \\ \alpha_N \approx d\,\alpha_{N-1} \end{matrix} .$$

Отсюда получаем упрощенную формулу

$$a_N = d\left(d + (d-2)r^2\right)^{N-1} . \tag{2.21}$$

Вместо рекуррентного отношения (2.18) получаем

$$b_N = d_1\left(d_1 + (d_1-2)r^2\right)^{N-1} . \tag{2.22}$$

Подставляя формулы (2.15), (2.21) и (2.22) в (2.14) будем иметь

$$\rho^2 = \left(\frac{\sqrt{\left(1-r^2\right)\left(1+\left(1-2d^{-1}\right)r^2\right)}}{\left(1+(1-2t)r^2\right)} \right)^{N-1} . \tag{2.23}$$

В таблице 2.1 приведены величины ρ^2, вычисленные по формуле (2.23), для различных r и N, при $d = 1{,}01$.

Таблица 2.1 Значение ρ^2 по формуле (2.23) для $d = 1{,}01$ и различных r и N

r	N			
	50	100	200	500
0,5	1	0,999	0,999	0,997
0,9	0,961	0,922	0,849	0,664
0,95	0,841	0,705	0,495	0,171
0,99	0,062	04	1,23×10-6	4,82×10-13

Видно, что при выборе типичного параметра $d = \left(1 - \frac{D}{2\sigma_S^2}\right)^{-2} \approx 1{,}01$, СГС с вложением по (2.12) оказывается достаточно стойкой, если элементы покрывающего сообщения имеют малую корреляцию ($r \le 0{,}9$), но при большой корреляции (этот факт подтверждается точными вычислениями по рекуррентным формулам (2.17) и (2.18)) СГС становиться нестойкой. Заметим, что можно уменьшить корреляцию между элементами, если разнести элементы, подвергнутые процессу вложения.

Выведем формулу для вероятности ошибки P_e для легального пользователя СГС, т.е. для пользователя, который знает секретную последовательность W^N, при атаке аддитивным шумом и использовании информированного декодера. Процесс вложения (2.12) может быть представлен как

$$X^{N_0} = \alpha S^{N_0} + (-1)^b W^{N_0},$$

где $N_0 \le N$ – блок, в который вкладывается один информационный бит $b \in \{0,1\}$.

После атаки аддитивным шумом Z^{N_0} с нулевым средним и с дисперсией σ_Z^2 получаем

$$\widetilde{X}^{N_0} = X^{N_0} + Z^{N_0}.$$

Корреляционный информированный приемник легального пользователя принимает решение о информационном бите $\widetilde{b}$ по следующему правилу:

$$\Lambda = \sum_{n=1}^{N_0} \left(\widetilde{X}(n) - aS(n)\right) W(n) \Rightarrow \widetilde{b} = \begin{cases} 0 \text{ если } \Lambda \geq 0 \\ 1 \text{ если } \Lambda < 0 \end{cases},$$

где $\left(\widetilde{X}(n)\right)_{n \leq N_0}$, $\left(S(n)\right)_{n \leq N_0}$ и $\left(W(n)\right)_{n \leq N_0}$ – элементы $\widetilde{X}^{N_0}$, S^{N_0} и W^{N_0} соответственно. Тогда согласно центральной предельной теореме теории вероятностей, вероятность приема P ошибочного бита для любого вероятностного распределения Z^{N_0} может быть представлена [27] как

$$P \approx Q\left(\sqrt{N_0 \frac{D\left(1 - \frac{D}{4\sigma_S^2}\right)}{\sigma_Z^2}}\right), \tag{2.24}$$

где $Q : x \mapsto \frac{1}{\sqrt{2\pi}} \int_x^{+\infty} e^{-\frac{t^2}{2}} dt$.

Определим искажение ПО после атаки как

$$\mathrm{var}\left(\widetilde{X}^{N_0} - S^{N_0}\right) = D + \sigma_Z^2,$$

и отношение сигнал-шум после вложения η_w и после атаки η_a как

$$\eta_w = \frac{\sigma_S^2}{D}, \ \eta_a = \frac{\sigma_S^2}{D + \sigma_Z^2}.$$

Теперь, используя (2.24) вероятность ошибочного бита P можно легко выразить через η_w и η_a как

$$P \approx Q\left(\sqrt{N_0 \frac{(4\eta_w - 1)\eta_a}{4(\eta_w - \eta_a)\eta_w}}\right). \tag{2.25}$$

Если $\eta_w >> 1$, эта вероятность может быть аппроксимирована как

$$P \approx Q\left(\sqrt{\frac{N_0}{\eta - 1}}\right),$$

где $\eta = \frac{\eta_w}{\eta_a}$.

Теперь можно выразить параметр d, представленный в (2.23) через η_w как

$$d = \left(1 - \frac{D}{2\sigma_S^2}\right)^{-2} = \left(1 - \frac{1}{2\eta_w}\right)^{-2} = \frac{4\eta_w^2}{(2\eta_w - 1)^2}. \qquad (2.26)$$

Приведем пример. Возьмем $\eta_w = 100$, $\eta_a = 70$ и $N = 500$. Тогда согласно (2.26) получаем $d = 1{,}01$ и $t = 0{,}995$. Вычисление по формуле (2.23) для $r = 0{,}9$ дает $\rho^2 = 0{,}664$. С другой стороны, из (2.25) получаем $P \approx Q\sqrt{11{,}6} \approx 3 \times 10^{-4}$ для $N_0 = 5$. Это означает, что можно вложить 100 бит информации секретно (поскольку ρ^2 близко к 1) и в то же время достаточно надежно для дальнейшего извлечения легальным пользователем.

Для слепого декодера решение об информационном бите $\hat{b}$ принимает вид

$$\widetilde{\Lambda} = \sum_{n=1}^{N_0} \widetilde{X}(n) W(n) \Rightarrow \hat{b} = \begin{cases} 0 \text{ если } \widetilde{\Lambda} \geq 0 \\ 1 \text{ если } \widetilde{\Lambda} < 0 \end{cases}.$$

Вероятность ошибки для слепого декодера можно аппроксимировать как

$$\widetilde{P} \approx Q\left(\sqrt{N_0 \frac{(4\eta_w - 1)\eta_a}{4(\eta_w - \eta_a + \eta_w\eta_a)\eta_w}}\right). \qquad (2.27)$$

Если $\eta_w >> 1$, то из (2.27) получаем $P \approx Q\left(\sqrt{\frac{N_0}{\eta_w}}\right)$.

При использовании исходных данных из предыдущего примера получаем $\widetilde{P} \approx 1{,}2 \times 10^{-2}$, если $N_0 = 500$. Это означает, что только 1 бит информации может быть вложен «почти» секретно (поскольку ρ^2 близко к 1).

2.3 ROC-кривые

Рассмотренный в параграфе 2.2 критерий позволяет определить потенциальную необнаруживаемость СГС, но не позволяет построить такую систему. Задать параметры необнаруживаемой СГС для последующей реализации, а также оценить секретность СГС можно с помощью ROC-кривых

(Receiver Operating Characteristic – операционная характеристика приемника)[13]. Данная характеристика оценивает стойкость СГС при использовании различных методовСГА. Одна и также СГС будет иметь разную стойкость к разным методамСГА. Следовательно, наиболее эффективной оценкой стойкости СГС будет стойкость при использовании наилучшего метода СГА. Отметим, что с помощью ROC-кривых можно оценивать не только стойкость СГС, но и эффективность методов СГА (чем менее стойка СГС к определенному методу СГА, тем выше эффективность данного метода применительно к данной СГС). На рисунке 2.1 приведены примеры ROC-кривых, где по оси абсцисс отложена вероятность ошибочного обнаружения (False Positive) СО P_{FP}, а по оси ординат – вероятность правильного обнаружения (True Positive) СО P_{TP}. Очевидно, что $P_{FP} = P_{fa}$, а $P_{TP} = 1 - P_m$.

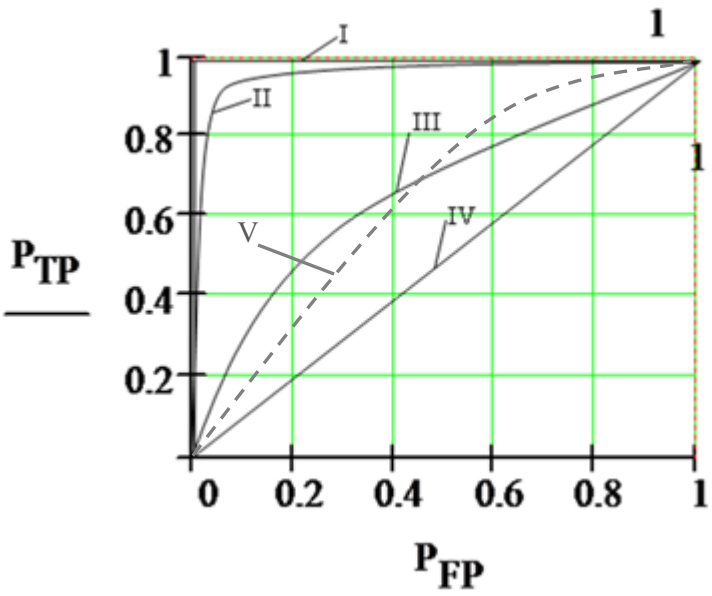


Рисунок 2.1 – Примеры ROC-кривых:
I – абсолютно нестойкая СГС,
II – посредственная СГС,
III и V – хорошие СГС,
IV – идеально стойкая СГС

У идеально стойкой СГС, чей результат равносилен случайному угадыванию, ROC-кривая будет совпадать с диагональю, примером такой кривой является кривая IV на рисунке 2.1. У абсолютно нестойкой СГС ROC-кривая

будет аналогична кривой I, показанной на рисунке 2.1. ROC-кривая II соответствует «слабой» СГС, у которого одновременно P_{FP} достаточно мала, а P_{TP} достаточно велика. ROC-кривые III и V соответствуют «хорошим» СГС, когда одновременно невозможно сделать малой P_{FP} и большой P_{TP}. Можно сравнить стойкость двух СГС сравнивая их ROC-кривые. При этом если $P_{TP}^{\mathrm{I}}(P_{FP}) > P_{TP}^{\mathrm{II}}(P_{FP}) > P_{TP}^{\mathrm{III}}(P_{FP}) > P_{TP}^{\mathrm{IV}}(P_{FP})$ для $P_{FP} \in [0,1]$, то стегосистема I менее стойкая, чем II, которая менее стойкая, чем III, которая менее стойкая, чем IV для всех P_{FP}.

Данный метод сравнения сложно применить, если ROC-кривые двух сравниваемых СГС пересекаются. Тогда для удобства сравнения ROC-кривых можно используют числовой параметр ρ [13]

$$\rho = 2A - 1, \tag{2.28}$$

где A – площадь под ROC-кривой.

Из формулы (2.28) (и из рисунка 2.1) видно, что $\rho_{\mathrm{I}} = 1$, $\rho_{\mathrm{IV}} = 0$, $\rho_{\mathrm{I}} > \rho_{\mathrm{II}} > \rho_{\mathrm{III}} > \rho_{\mathrm{IV}}$ и $\rho_{\mathrm{I}} > \rho_{\mathrm{II}} > \rho_{\mathrm{V}} > \rho_{\mathrm{IV}}$. При этом для сравнения III и V ROC-кривых графика на рисунке 2.1 недостаточно, необходимо рассчитать ρ_{III} и ρ_{V}.

Еще один числовой параметр, удобный для сравнения стойкости СГС,– минимальная средняя вероятность ошибки

$$P_e = \min_{P_m} \frac{1}{2}\left(P_{fa} + P_m\right). \tag{2.29}$$

Из формулы (2.29) видно, чем ближе P_e к 1/2, тем выше стойкость СГС.

Отметим, что все критерии, описанные в данном параграфе, требуют сбора статистической информации не только по самой СГС, но и сбор статистических параметров СГС (P_m и P_{fa}) после атаки на СО наилучшим методом СГА.

2.4 Выводы

Рассмотренные в данной главе методы позволяют оценить стойкость СГС.

Предложенный автором критерий оценки необнаруживаемости СГС, основанный на расстоянии Бхаттачариа, в отличие от критерия относительной энтропии, рассмотренного в параграфе 2.1, может оказаться более простым в вычислении. Вычисления по критерию относительной энтропии представляет собой сложную задачу, тогда как для вычисления критерия, основанного на расстоянии Бхаттачариа, иногда требуется меньше аппаратных и временных ресурсов.

Экспериментальные результаты показали, что предложенный критерий необнаруживаемости СГС прост в вычислении и надежен для оценки секретности СГС, но при этом он не позволяет определить методы СГА и их эффективность. Поэтому в параграфе 2.3 рассмотрен критерий ROC-кривых, позволяющий построить СГС с заданными параметрами секретности.

Возможность оценить необнаруживаемость метода СГ важна как для создателя СГС, так и для разработчика методов СГА, поскольку проверять эффективность нового метода лучше на наиболее стойкой СГС, тогда на менее стойкий СГС метод тоже будет работать эффективно. Другими словами, критерии стойкости методов СГ можно применять для оценки эффективности методов СГА.

Для абсолютно секретной СГС $P_m = P_{fa} = 1/2$. Тогда выбор между гипотезами H_0 и H_1 равносилен случайному угадыванию. Если при оценки стойкости с помощью критерия, основанного на расстоянии Бхаттачариа, исследуемая СГС оказалась абсолютно секретной, то есть $\rho_B(P_S, P_X) = 1$ и $D_B(P_X, P_S) = 0$, то дальнейшая разработка новых методов СГА для такой СГС не имеет смысла. Чем больше величина $D_B(P_X, P_S)$, тем менее секретна данная СГС.

Если $D_B(P_X, P_S) = \varepsilon$, то СГС является ε-секретной. В этом случае разработка новых или усовершенствование уже известных методов СГА имеет смысл [28].

Чем выше стойкость используемой СГС, тем сложнее создать метод СГА, позволяющий раскрыть секретный канал связи. Для построения необнаруживаемой СГС надо точно знать распределение ПО и использовать метод СГ, который создает СО с распределением, близким к исходному распределению ПО. И чем ближе распределения ПО и СО, тем ближе результат методов СГА к случайному выбору между гипотезами H_0 и H_1.

Однако, поскольку точное распределение ПО знать невозможно, и исходное распределение после вложения изменить свои статистические свойства, построить абсолютно секретную СГС на данный момент не представляется возможным. Следовательно, имеет смысл разрабатывать и использовать методы СГА (как целевые, основанные на сравнении статистических параметрах объекта, так и ССГА, основанные на использовании метода опорных векторов и различных функционалов), позволяющие найти статистические различия между ПО и СО.

В последующих главах описаны методы СГА, позволяющие находить статистические различия в распределениях ПО и СО, а также даны рекомендации по выбору параметров данных методов, которые позволяют оптимизировать отношение P_{fa} и P_m.

3 Исследование методов «слепого» стегоанализа

3.1 Общие сведения

При анализе объекта, проверяемого на наличие или отсутствие скрытой информации, алгоритм вложения далеко не всегда известен (см. таблицу 1.1). Когда стегоаналитик не может с достаточной долей уверенности предположить, какой именно метод вложения был использован передающим, для анализа используется «слепой» метод СГА (далее ССГА)[29].

Анализ объекта при использовании ССГА производится в два этапа:

- обучение распознавателя на достаточно большой выборке СО определённого типа и «чистых» ПО;
- применение распознавателя (классификатора), который, используя результат обучения и представленный объект, принимает решение о принадлежности его к СО или к ПО, при этом также возможно определения использованного метода вложения.

Чтобы использовать метод ССГА, атакующему необходимо иметь в наличии достаточное количество СО. Для этого достаточно иметь в своем распоряжении алгоритм вложения в виде «черного ящика», который позволяет получить СО из заданного ПО с использованием произвольного стегоключа, даже если точное описание алгоритма вложения неизвестно. Тогда атакующий может сам получить необходимое количество СО и соответствующих им ПО.

Важным этапом ССГА является выбор некоторых характеристик (функционалов) ПО и СО, используемых как на этапе обучения, так и на этапе принятия решения. От правильности выбора этих функционалов зависит эффективность обнаружения. Хотя предполагается, что в точности алгоритм вложения может быть неизвестен, однако, желательно иметь некоторые знания об этом алгоритме, например, в какую область производится вложение – в область пикселей или в область частот (DCT-коэффициентов). Отметим, что часто оказывается возможным высказать предположение, на какие именно

характеристики (функционалы) ПО должно повлиять вложение при любом возможном алгоритме.

Общая структурная схема ССГА показана на рисунке 3.1.

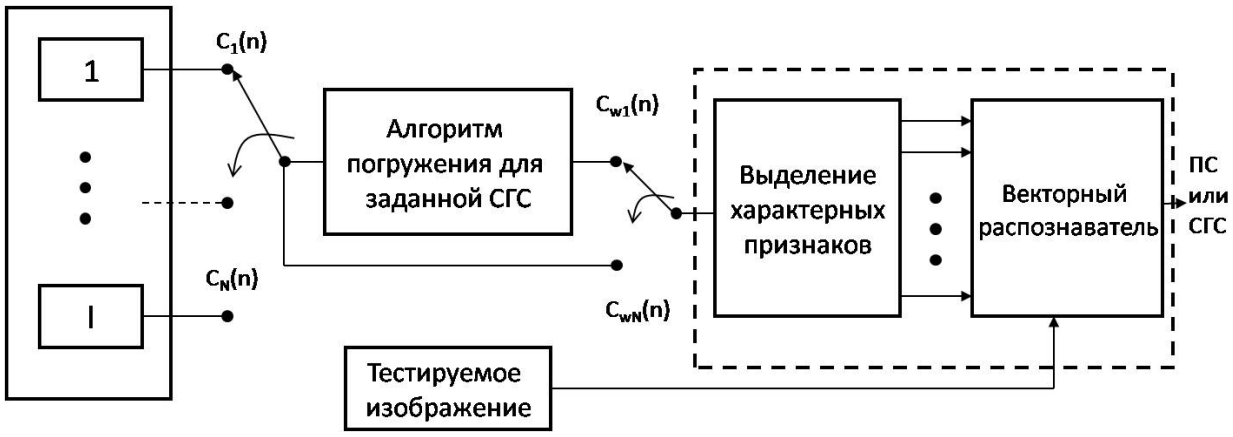


Рисунок 3.1 – Структурная схема ССГА

Рассмотрим рисунок 3.1 более подробно. В процессе обучения с помощью первого переключателя выбираются ПО. Далее часть ПО проходят через блок алгоритма погружения для заданной СГС, при этом получаются СО, которые подводятся к второму переключателю. Другая часть ПО в своем исходном виде приходит на тот же второй переключатель. Далее ПО и СО попадают в блок выделения характерных признаков, где из них выделяют n-мерные векторы-признаки. На вход векторного распознавателя в процессе обучения для обработки и анализа поступают n-мерные векторы-признаки. Результаты обработки сохраняются в памяти, в процессе тестирования выделяются векторы признаки из тестируемого объекта, который поступает на вход распознавателя, где они сравниваются с векторами признаками, полученными в результате обучения, и на выходе появляется признак ПО или СО, а так же, возможно, признак класса СГС (например Outguess или F5).

Заметим, что для векторного распознавателя в англоязычной литературе существует специальный термин «Support Vector Machine», который в русскоязычной литературе переведён как «метод опорных векторов»[30]. Далее

мы будем использовать для этого распознавателя сокращение – МОВ. В качестве исследуемых объектов используются изображения формата JPEG.

Ещё одна техника выделения функционалов по калибровочному изображению показана на рисунке 3.2.

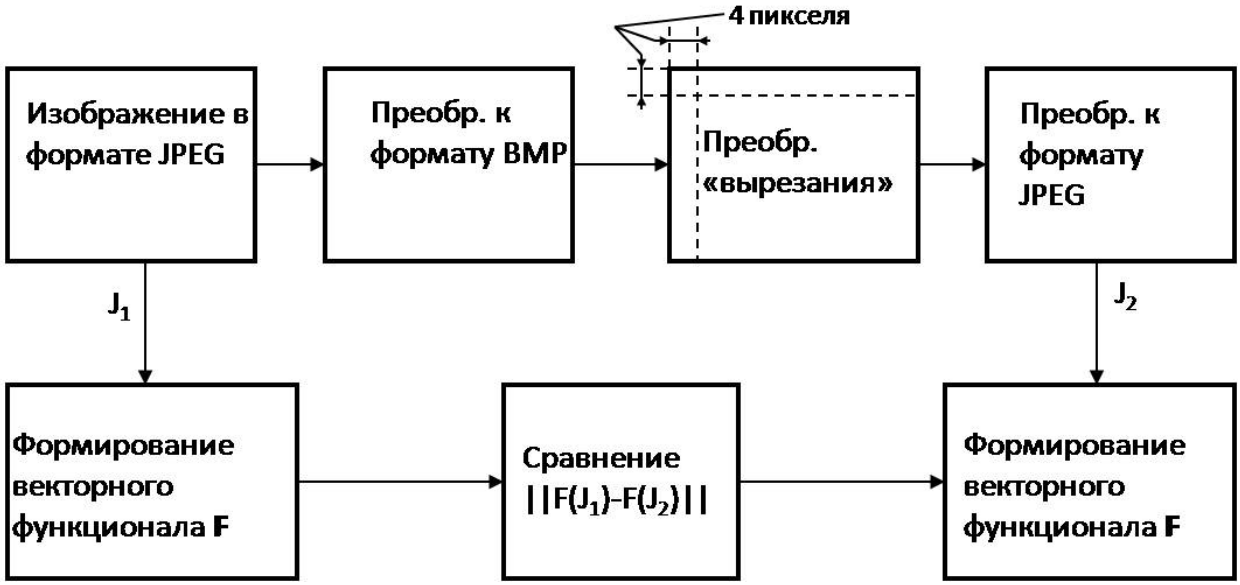


Рисунок 3.2 – Выделения функционалов по калиброванному изображению

Изображение J_2 называют калиброванным по отношению к исходному изображению J_1. Данная калибровка представляет собой нарушение решетки 8x8 (стандартная решетка файлов формата JPEG). Для этого исходное изображение формата JPEG преобразуется к формату BMP, далее вырезаются полосы из 4 пикселей по горизонтали и по вертикали, полученное изображение формата BMP путем сжатия преобразуется в формат JPEG. При этом даже если исходное изображение J_1 являлось СО, в новом, полученном после калибровки, изображении J_2 эффект погружения скрытой информации будет проявляться значительно меньше. Таким образом, J_2 является аппроксимациейПО, который был использован для создания СО J_1, или аппроксимацией самого изображения J_1, если оно не содержит вложения. Следовательно, значение выражения $\|F(J_1)-F(J_2)\|$, где $F(\ldots)$ – некоторый функционал, будет меняться в

зависимости от того, является J_1 ПО или СО.Если значения J_1 и J_2 близки, следовательно, $\|F(J_1) - F(J_2)\|$ близко к 0, то исследуемый объект J_1 является ПО.

3.2 Элементы МОВ

В главе 2 описаны две тестируемые гипотезы H_0 и H_1. Именно эти гипотезы будем проверять и в данной главе.

На этапе обучения имеется множество векторов

$$\{\mathbf{x}_i, y_i\}_{i=1}^{N},\ y_i = \begin{cases} 1, \text{если } \mathbf{x}_\text{i} \in \text{стегообъект}, \\ -1, \text{если } \mathbf{x}_\text{i} \in \text{покрывающий объект}, \end{cases}$$

где $\{\mathbf{x}_i\}_{i=1}^{N} \in R_d$ – вектор характерных признаков (функционалов).

На этапе распознавания вектор $\mathbf{x}_i$ через МОВ отображается в $y_i = \pm 1$ в зависимости от того, в пользу какой тестируемой гипотезе МОВ принял решение.

Различают три основных типа МОВ[5]:

- линейный сепарабельный МОВ (тип 1);
- линейный несепарабельный МОВ (тип 2);
- нелинейный МОВ (тип 3).

На рисунке 3.3 приведены все три типа МОВ для двумерного пространства признаков (двухмерное пространство взято для наглядности, поскольку геометрическая иллюстрация с многомерными векторами нереализуема).

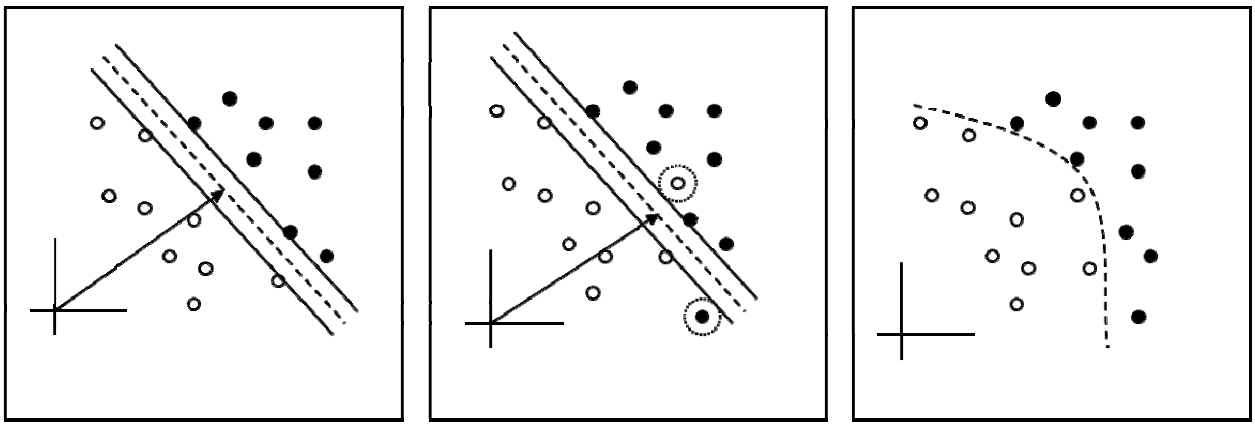


Рисунок 3.3 – Иллюстрация типов МОВ для двумерного пространства признаков

Для объяснения принципов работы МОВ рассмотрим вначале наиболее простой из всех МОВ – линейный.

Правило классификации в случае линейного МОВ имеет следующий вид:

$$y = \mathrm{sgn}\left(\langle \boldsymbol{\omega}, \mathbf{x} \rangle - \omega_0\right), \tag{3.1}$$

где $\mathrm{sgn}(a) = \begin{cases} +1, \text{если } a \geq 0 \\ -1, \text{если } a < 0 \end{cases}$, $\omega_0 \in R_1$ – некоторый числовой порог,

$\langle \boldsymbol{\omega}, \mathbf{x} \rangle$ – скалярное произведение векторов $\boldsymbol{\omega}$ и $\mathbf{x}$, при этом вектор $\boldsymbol{\omega}$ определяет правило решения или некоторую гиперплоскость, разделяющую область ПО от области СО.

Если выборка $\{\mathbf{x}_i\}_{i=1}^{N}$ сепарабельна, то оказывается справедливым соотношение:

$$y_i\left(\langle \boldsymbol{\omega}, \mathbf{x}_i \rangle - \omega_0 \geq 0\right), i = 1,2,...N, \tag{3.2}$$

где $\{\mathbf{x}_i\}_{i=1}^{N}$ – полное множество векторов характерных признаков, полученных на этапе обучения.

Неравенство (3.2) показывает, что существует такая гиперплоскость, определяемая вектором $\boldsymbol{\omega}$ и параметром ω_0, что для всех векторов $\{\mathbf{x}_i\}_{i=1}^{N}$, полученных на этапе тестирования, мы будем иметь однозначную правильную классификацию. Однако, такая идеально разделяющая выборки плоскость может

оказаться не единственной. Естественно, что для решения (3.1) должна быть выбрана такая гиперплоскость, которая бы максимально далеко отстояла от ближайших к ней точек обоих классов. Тогда можно ожидать, что при тестировании ошибка распознавания окажется минимальной [30].

Из правила (3.1) следует, что алгоритм классификации не изменится, если $\boldsymbol{\omega}$ и ω_0 одновременно умножить на одну и ту же положительную постоянную величину. Удобно выбирать эту постоянную таким образом, чтобы для всех пограничных векторов $\mathbf{x}_i$ (т.е. ближайших к разделяющей гиперплоскости) выполнялось бы условие

$$\langle \boldsymbol{\omega}, \mathbf{x}_i \rangle - \omega_0 = y_i. \qquad (3.3)$$

Другими словами, чтобы все пограничные объекты находились от гиперплоскости на одинаковом расстоянии. Остальные объекты будут находиться от гиперплоскости на большем расстоянии (рисунок 3.3, а):

$$\langle \boldsymbol{\omega}, \mathbf{x}_i \rangle - \omega_0 \begin{cases} \geq 1, \text{ если } y_i = +1 \\ \leq -1, \text{ если } y_i = -1 \end{cases}. \qquad (3.4)$$

Условие $-1 < \langle \boldsymbol{\omega}, \mathbf{x}_i \rangle - \omega_0 < 1$ задаёт полосу, разделяющую два класса, причём ни одна из точек $\mathbf{x}_i$ обучающей выборки не будет лежать внутри этой полосы, а сама разделяющая гиперплоскость проходит точно посередине этой полосы. Для лучшей различимости классов гиперплоскость должна проходить так, чтобы ширина этой полосы была бы максимальной.

Рассмотрим две произвольные точки $\mathbf{x}_+$ и $\mathbf{x}_-$, соответствующих $y = +1$ и $y = -1$ соответственно, лежащие на границе полосы, тогда ширина полосы будет равна

$$\left\langle (\mathbf{x}_+ - \mathbf{x}_-) \frac{\boldsymbol{\omega}}{\|\boldsymbol{\omega}\|} \right\rangle = \frac{\langle \boldsymbol{\omega}, \mathbf{x}_+ \rangle - \langle \boldsymbol{\omega}, \mathbf{x}_i \rangle}{\|\boldsymbol{\omega}\|} = \frac{(\omega_0 + 1) - (\omega_0 - 1)}{\|\boldsymbol{\omega}\|} = \frac{2}{\|\boldsymbol{\omega}\|}. \qquad (3.5)$$

Таким образом, если выборка разделима, то оптимальная гиперплоскость, задаваемая вектором $\boldsymbol{\omega}$ и параметром ω_0 должна минимизировать $\|\boldsymbol{\omega}\|$ при условии выполнения (3.4).

Задача условной минимизации сводится к задаче условной максимизации следующего Лагранжиана

$$L = \sum_{i=1}^{N} \lambda_i - \frac{1}{2} \sum_{i=1}^{N} \sum_{j=1}^{N} \lambda_i \lambda_j \left\langle \mathbf{x}_i, \mathbf{x}_j \right\rangle y_i \cdot y_j , \tag{3.6}$$

где $\sum_{i=1}^{N} \alpha_i y_i = 0$, $\alpha_i \geq 0, i = 1,2,...N$.

Тогда оптимальная гиперплоскость, определяемая вектором $\boldsymbol{\omega}$ и параметром ω_0, находится следующим образом

$$\begin{aligned} \boldsymbol{\omega} &= \sum_{i=1}^{N} \lambda_i - y_i \mathbf{x}_i , \\ \omega_0 &= \frac{1}{N} \sum_{i=1}^{N} \left(y_i - \left\langle \boldsymbol{\omega}, \mathbf{x}_i \right\rangle \right). \end{aligned} \tag{3.7}$$

После нахождения множителей $\lambda_i > 0$, обеспечивающих максимизацию L по (3.6), алгоритм классификации может быть представлен в следующем виде

$$y = \operatorname{sgn}\left(\sum_{i=1}^{N} \left\langle \lambda_i y_i \mathbf{x}_i, \mathbf{x} \right\rangle - \omega_0 \right). \tag{3.8}$$

Теперь рассмотрим линейный несепарабельный МОВ. Из рисунка 3.3, б видно, что не существует прямой линии (в многомерном случае гиперплоскости), которая бы разделила выборки, полученные в результате обучения, на два непересекающихся подмножества. Следовательно, можно позволить МОВ допускать ошибки, но количество этих ошибок целесообразно минимизировать. Введём набор дополнительных переменных $\{\xi_i\}_{i=1}^{N}$, характеризующих величину ошибки на тренируемых изображениях.

Тогда вместо неравенств (3.4) получим

$$\langle \boldsymbol{\omega}, \mathbf{x}_i \rangle - \omega_0 \begin{cases} \geq 1 - \xi_i, \text{ если } y_i = 1, \\ < -1 + \xi_i, \text{ если } y_i = -1. \end{cases} \quad (3.9)$$

Тренировочные пары $(\mathbf{x}_i, y_i)$, оказавшиеся в «неправильном» множестве, будут иметь $\xi_i \geq 1$.

Задача оптимизации выбора разделяющей гиперплоскости в этом случае требует минимизации общей ошибки классификации $\sum_{i=1}^{N} \xi_i$ при максимизации интервала между областями $\frac{1}{2}\|\boldsymbol{\omega}\|^2 + C\sum_{i=1}^{N} \xi_i$, где C – постоянная, которая контролирует ошибки идентификации. Решение этой задачи сводится к (3.6), но на переменную α_i накладываются другие ограничения

$0 \leq \alpha_i \leq C$.

Для алгоритма классификации в этом случае остается справедливой формула (3.8).

Теперь рассмотрим нелинейный МОВ (рисунок 3.3, в). В случае нелинейного МОВ невозможно построить линию (в общем случае гиперплоскости), которая бы разделила результаты тестирования на два непересекающихся подмножества без значительных ошибок.

Тогда предлагается сначала отобразить пространство векторов $\mathbf{x}$, в котором производится обучение, в некоторое новое Евклидово пространство $\mathbf{H}$ большей (а возможно и бесконечной) размерности, в котором может оказаться допустимым применение сепарабельного или несепарабельного линейных МОВ.

Такое Евклидово (т. е. имеющее скалярное произведение) пространство $\mathbf{H}$ называется спрямляющим пространством.

Правило классификации (3.8) зависит только от скалярных произведений $(\mathbf{x}', \mathbf{x})$. Следовательно, если выполняется отображение $\mathbf{Y} : \mathbf{X} \to \mathbf{H}$, то после такого отображения МОВ будет использовать то же самое правило, но с заменой $(\mathbf{x}', \mathbf{x})$ на $K(\mathbf{x}, \mathbf{x}') = (K(\mathbf{x}'), K(\mathbf{x}))$, где $K(\mathbf{x}, \mathbf{x}')$ называется ядром преобразования.

Задача построения нелинейного МОВ сводится к выбору ядра. Заметим, что на сегодняшний день выбор ядра, оптимального для решения конкретной задачи, является открытой проблемой.

Часто ограничиваются перебором конечного числа функций, о которых известно, что они являются ядрами, среди которых и выбирается лучшая.

Широко используемым видами ядер являются:

- полиномиальные ядра $K(\mathbf{x},\mathbf{x}') = \left(\langle \gamma\ \mathbf{x},\mathbf{x}'\rangle + r\right)^d, \gamma > 0$,
- гауссовские ядра $K(\mathbf{x},\mathbf{x}') = \exp\left(-\gamma\|\mathbf{x}-\mathbf{x}'\|^2\right), \gamma > 0$,
- сигмоидные ядра $K(\mathbf{x},\mathbf{x}') = th\left(K_0 + K_1\langle \mathbf{x},\mathbf{x}'\rangle\right)$,

где γ, d, K_0, K_1 – параметры ядра.

В качестве иллюстрирующего примера с использованием спрямляющего пространства рассмотрим двумерное исходное пространство $\mathbf{X} = \mathbf{R}^2$ и квадратичное ядро $K(\mathbf{x},\mathbf{x}') = \langle \mathbf{x},\mathbf{x}'\rangle^2$, где $\mathbf{x} = (x_1, x_2), \mathbf{x}' = (x_1', x_2')$.

Преобразуем квадрат скалярного произведения:

$$K(\mathbf{x},\mathbf{x}') = \langle x_1, x_1' + x_2, x_2'\rangle^2 = {x_1}^2, {x'}^2 + {x_2}^2 {x_2'}^2 + 2\cdot x_1 x_1' x_2 x_2' =$$

$$\left\langle \left({x_1}^2, {x_2}^2\sqrt{2}x_1x_2\right), \left({x_1'}^2, {x_2'}^2\sqrt{2}x_1x_2\right)\right\rangle. \qquad (3.10)$$

Из соотношения (3.10) видно, что ядро $K(\mathbf{x},\mathbf{x}') = \langle \mathbf{x},\mathbf{x}'\rangle^2$ представляется в виде скалярного произведения в пространстве $\mathbf{R}^3$, а преобразование $\mathbf{Y}: \mathbf{R}^2 \to \mathbf{R}^3$, соответствующее данному ядру, имеет вид $(x_1x_2) \to \left({x_1}^2, {x_2}^2, \sqrt{2}x_1x_2\right)$.

Линейной поверхности в пространстве $\mathbf{H}$ будет соответствовать квадратичная поверхность в исходном пространстве $\mathbf{X}$.

3.3 Функционалы, используемые для обнаружения стегообъекта по МОВ

Как было отмечено в параграфе 3.2, правильный выбор функционалов (т. е. характеристик, определяющих статистические свойства изображений) оказывает

решающее влияние на эффективность обнаружения даже при использовании такого мощного классификатора, как МОВ. В настоящее время не существует метода выбора оптимальных функционалов. Необходимым условием выбора определённого функционала (т. е. многомерного вектора, являющегося функцией изображения) является его чувствительность к вложению скрываемой информации. Если при замене ПО на СО, выбранный вектор не изменится для значительной части различных ПО, то его нецелесообразно использовать в качестве функционала.

Для МОВ могут использоваться различные функционалы. Например, в работе [31] предполагается сначала выполнить декомпозицию изображений с использованием низкочастотных и высокочастотных фильтров, а затем рассчитать статистические характеристики, такие как среднее значение, дисперсия, перекос (коэффициент асимметрии) и эксцесс полосовых коэффициентов. Кроме того, используется статистика ошибок при оптимальном линейном предсказании амплитуд коэффициентов. В результате формируются многомерные векторы $\mathbf{x}_i$, которые используются для тренировки МОВ и последующего различия с его помощью ПО и СО.

В работе [32] используется более полный набор функционалов в предположении, что при построении СО вложение производится в частотную область, т. е. фактически в DCT-коэффициенты, что типично для ПО и СО в формате JPEG.

Рассмотрим эти функционалы более подробно.

Предположим, что JPEG файл представлен своими DCT-коэффициентами $d_{\mathrm{k}}(i,j)$, где k – номер 8×8 блока, $k=1,2,\ldots,B$, а (i,j) квантованные коэффициенты в k-ом блоке. Тогда в качестве первой характеристики выбирается глобальная гистограмма

$$h_r=\left\{\#(i,j,k):d_{\mathrm{k}}(i,j)=r\right\}, \qquad (3.11)$$

где $r\in\left(L,\ldots,R\right)$ – диапазон квантования DCT-коэффициентов.

Далее выбираются гистограммы индивидуальных DCT-коэффициентов

$$h_r^{ij} = \left\langle \#k : d_{\mathrm{k}}(i,j) = r \right\rangle . \qquad (3.12)$$

Наиболее важными являются низкочастотные коэффициенты (т. е. малые величины i и j), поскольку средние и высокие частоты статистически неэффективны из-за малого количества ненулевых коэффициентов.

В качестве последней характеристики первого порядка выбирается дуальная гистограмма. Для заданной величины d дуальная гистограмма – это 8×8 матрица

$$\left\{ g_{ij}^{d} \right\}_{(i,j)=1}^{8} = \sum_{\mathrm{k}=1}^{B} \delta\left(d, d_{\mathrm{k}}(i,j)\right), \qquad (3.13)$$

где $\delta(u,v) = \begin{cases} 1, \text{ если } u = v \\ 0, \text{ если } u \neq v \end{cases}$.

Другими словами, g_{ij}^{d} – это количество блоков 8×8 бит, в которых величина d совпадает с DCT-коэффициентом с индексами (i,j).

Теперь рассмотрим характеристики второго порядка. Необходимость использования таких статистик объясняется тем, что в реальных изображениях корреляция присутствует не только внутри блоков 8×8 DCT-коэффициентов, но и между соседними блоками, которые, следовательно, нельзя считать независимыми. Поэтому необходимо выбрать характеристики, которые были бы чувствительны к изменению такой зависимости при вложении секретной информации.

Пусть I_r и I_c – векторы блоковых индексов при сканировании изображения по строкам и по столбцам соответственно. Тогда первый функционал второго порядка, оценивающий межблочную зависимость и называемый вариацией V, определяется следующим соотношением

$$V = \frac{\sum_{(i,j)=1}^{8} \sum_{\mathrm{k}=1}^{|I_c|-1} \left| d_{I_r(\mathrm{k})}(i,j) - d_{I_r(\mathrm{k}+1)}(i,j) \right| + \sum_{(i,j)=1}^{8} \sum_{\mathrm{k}=1}^{|I_c|-1} \left| d_{I_c(\mathrm{k})}(i,j) - d_{I_c(\mathrm{k}+1)}(i,j) \right|}{\left|I_r\right| + \left|I_c\right|}. \qquad (3.14)$$

Большинство СГ алгоритмов увеличивают энтропию в области квантованных DCT-коэффициентов, поэтому можно ожидать, что значение величины V скорее будет увеличиваться, чем уменьшаться.

Можно также ожидать, что погружение дополнительной информации в СГС уменьшить корреляцию в пределах границ блока 8×8. Блочность вычисляется по декомпрессированному (т. е преобразованному к формату BMP) изображению и выражает интегральную меру межблоковой зависимости по всем DCT-коэффициентам и по всему изображению

$$B_\alpha = \frac{\sum_{i=1}^{L(M-1)/8}\sum_{j=1}^{N}\left|x_{8i,j}-x_{i+1,j}\right|^2 + \left|\sum_{j=1}^{L(N-1)/8}\sum_{i=1}^{M}\left|x_{1,8j}-x_{i,8j+1}\right|^2\right|}{N\left[(M-1)/8\right]+M\left[(N-1)/8\right]}, \qquad (3.15)$$

где M и N – размеры изображения в пикселях,

$x_{i,j}$ – значение яркости (оттенка) пикселя с индексом (i,j) у изображения, преобразованного из формата JPEG в формат BMP GrayScale.

Последние три функционала вычисляются по матрицам взаимного расположения соседних DCT-коэффициентов. Дадим определение такой матрице $\mathbf{C}_{st}$, размером $D\times D$, $D=R-L+1$

$$\mathbf{C}_{st} = \frac{\sum_{K=1}^{|I_c|-1}\sum_{i,j=1}^{8}\delta\left(s,d_{I_r(k)}(i,j)\right)\cdot\delta\left(t,d_{I_r(k+1)}(i,j)\right)+}{\left|I_c+I_r\right|}$$

$$\frac{+\sum_{K=1}^{|I_c|-1}\sum_{i,j=1}^{8}\delta\left(s,d_{I_c(k)}(i,j)\right)\cdot\delta\left(t,d_{I_c(k+1)}(i,j)\right)}{}. \qquad (3.16)$$

Матрица $\mathbf{C}_{st}$ описывает вероятностное распределение пар соседних DCT-коэффициентов. Обычно существует острый пик для $s=0$, $t=0$, далее идёт быстрое убывание. Пусть $\mathbf{C}_{st}(J_1)$ и $\mathbf{C}_{st}(J_2)$ будут матрицы $\mathbf{C}_{st}$ для исходного изображения в формате JPEG и для его калиброванной версии соответственно.

В работе [32] предлагается использовать следующие функционалы, связанные с матрицами $\mathbf{C}_{st}(J_1)$ и $\mathbf{C}_{st}(J_2)$

$$N_{00} = C_{0,0}(J_1) - C_{0,0}(J_2),$$

$$N_{01} = C_{0,1}(J_1) - C_{0,1}(J_2) + C_{1,0}(J_1) - C_{1,0}(J_2) + C_{-1,0}(J_1) - C_{-1,0}(J_2) +$$

$$+ C_{0,-1}(J_1) - C_{0,-1}(J_2),$$

$$N_{11} = C_{1,1}(J_1) - C_{1,1}(J_2) + C_{1,-1}(J_1) - C_{1,-1}(J_2) + C_{-1,1}(J_1) - C_{-1,1}(J_2) +$$

$$+ C_{-1,-1}(J_1) - C_{-1,-1}(J_2).$$

В качестве функционалов для МОВ могут также эффективно использоваться Марковские функционалы.Вектор Марковских функционалов показывает различия между абсолютными значениями соседних DCT-коэффициентов как Марковский процесс. Расчет функционала начинается с формирования матрицы $F(u,v)$ абсолютных значений DCT-коэффициентов изображения. Далее рассчитываются четыре массива по четырем направлениям – горизонталь, вертикаль, диагональ и минорная диагональ, – обозначенные как, $F_v(u,v)$, $F_d(u,v)$ и $F_m(u,v)$соответственно:

$$F_h(u,v) = F_h(u,v) - F_h(u+1,v),$$

$$F_v(u,v) = F_v(u,v) - F_v(u,v+1),$$

$$F_d(u,v) = F_d(u,v) - F_d(u+1,v+1),$$

$$F_m(u,v) = F_m(u+1,v) - F_m(u,v+1).$$

Далее вычисляются вероятностные матрицы перехода:

$$M_h(i,j) = \frac{\sum_{u=1}^{S_u-2}\sum_{v=1}^{S_v}\delta\left(F_h(u,v)=i, F_h(u+1,v)=j\right)}{\sum_{u=1}^{S_u-1}\sum_{v=1}^{S_v}\delta\left(F_h(u,v)=i\right)},$$

$$M_v(i,j) = \frac{\sum_{u=1}^{S_u}\sum_{v=1}^{S_v-2}\delta\left(F_v(u,v)=i, F_v(u,v+1)=j\right)}{\sum_{u=1}^{S_u}\sum_{v=1}^{S_v-1}\delta\left(F_v(u,v)=i\right)},$$

$$M_d(i,j)=\frac{\sum_{u=1}^{S_u-2}\sum_{v=1}^{S_v-2}\delta\left(F_d(u,v)=i,F_d(u+1,v+1)=j\right)}{\sum_{u=1}^{S_u-1}\sum_{v=1}^{S_v-1}\delta\left(F_d(u,v)=i\right)},$$

$$M_m(i,j)=\frac{\sum_{u=1}^{S_u-2}\sum_{v=1}^{S_v-2}\delta\left(F_m(u+1,v)=i,F_m(u,v+1)=j\right)}{\sum_{u=1}^{S_u-1}\sum_{v=1}^{S_v-1}\delta\left(F_m(u,v)=i\right)},$$

где S_u и S_v – размеры изображения,

$\delta=1$ если аргумент функции выполняется.

В таблице 3.1 приведены примеры функционалов, которые могут быть использованы для обнаружения СО с помощью МОВ.

Таблица 3.1 – Примеры функционалы для МОВ

Название функционала	**Вид функционала**
Глобальная нормированная гистограмма	$\frac{H}{\|H\|_{L1}}$
Индивидуальная нормированная гистограмма для 5 DCT-коэффициента	$\frac{h^{21}}{\|h^{21}\|_{L1}},\frac{h^{31}}{\|h^{31}\|_{L1}},\frac{h^{12}}{\|h^{12}\|_{L1}},\frac{h^{22}}{\|h^{22}\|_{L1}},\frac{h^{13}}{\|h^{13}\|_{L1}}$
Дуальная нормированная гистограмма для 11ти DCT-коэффициентов (-5,…+5)	$\frac{g^{-5}}{\|g^{-5}\|_{L1}},\frac{g^{-4}}{\|g^{-4}\|_{L1}},\cdots\frac{g^{4}}{\|g^{4}\|_{L1}},\frac{g^{5}}{\|g^{5}\|_{L1}}$
Марковские функционалы	M_h, M_v, M_d, M_m
Вариация	V

Продолжение таблицы 3.1

Название функционала	**Вид функционала**
L_1 и L_2 блочность	B_1, B_2
Взаимное расположение	N_{00}, N_{01}, N_{11}

Напомним, что здесь под $\|x\|_{L1}$ понимается $\sum_{i=1}^{n}|x_i|$, если $\mathbf{x}=(x_1,\ldots,x_i,\ldots,x_n)$.

Для оценки эффективности МОВ было проведено тестирование ПО и СО[5,34], созданных с помощью алгоритмов вложения F5 и Outguess, при чем задание состоит не только в определении, какая теория верна – H_0 и H_1, но и в классификации используемых методов вложения в случае СО – F5, Outguess или ModelBased (далее МВ). В тестировании метод МВведен для того, чтобы у МОВ был выбор из нескольких методов вложения, при этом в тестировании не применялись СО, созданные методом МВ. Более подробно метод МВ описанв [33].

Для выполнения этапа обучения стегоанализа по МОВ было использовано множество изображений (более 5000 изображений)[5,34]. Все изображения из множества для упрощения решения были приведены к одинаковому размеру и перекодированы алгоритмом JPEG с одинаковым качеством.

На рисунках 3.4 и 3.5 показаны ПО и СО, полученный из ПО путем вложения по алгоритму F5 и при максимальной доле вложения, соответственно.

Рисунок 3.4 – Покрывающий объект

Рисунок 3.5 – Стегообъект с максимальной долей вложения
по алгоритму F5

По представленным изображениям невозможно визуально отличить СО от ПО.

В таблице3.2 представлены результаты тестирования ПО и СО, созданных по алгоритму F5.Цифры в таблицах показывают процент принятия верного решения (алгоритма вложения).

Таблица 3.2 – Результат МОВ при тестировании ПО и СО с максимальным вложением по алгоритму F5

Используемые функционалы	**ПО**				**СО (F5, доля вложения – 1)**			
	ПО	**F5**	**outguess**	**MB**	**ПО**	**F5**	**outguess**	**MB**
Глобальная гистограмма	0,9318	0,0341	0,0341	0	0	0,9821	0,0179	0
Индивидуальная гистограмма для положения 12	0,9181	0,0128	0,0691	0	0	0,433	0,5967	0
Индивидуальная гистограмма для положения 21	0,9936	0,0051	0,0013	0	0	1	0	0
Индивидуальная гистограмма для положения 13	0,9531	0,0026	0,0443	0	0	0,4524	0,5476	0
Индивидуальная гистограмма для положения 31	0,9659	0,0281	0,0060	0	0	1	0	0
Индивидуальная гистограмма для положения 22	0,9753	0,0013	0,0235	0	0	0,8387	0,1613	0
Дуальная гистограмма -5	1	0	0	0	0	0,9996	0,0004	0
Дуальная гистограмма -4	1	0	0	0	0	1	0	0
Дуальная гистограмма -3	1	0	0	0	0	0,4776	0,5224	0
Дуальная гистограмма -2	1	0	0	0	0	1	0	0
Дуальная гистограмма -1	0,9970	0,0030	0	0	0,0009	0,9087	0,0905	0

Продолжения таблицы 3.2

Используемые функционалы	**ПО**				**СО (F5, доля вложения – 1)**			
	ПО	**F5**	**outguess**	**MB**	**ПО**	**F5**	**outguess**	**MB**
Дуальная гистограмма 0	0,7633	0,2367	0	0	0,1605	0,8395	0	0
Дуальная гистограмма 1	1	0	0	0	0	1	0	0
Дуальная гистограмма 2	1	0	0	0	0	1	0	0
Дуальная гистограмма 3	1	0	0	0	0	0,4545	0,5455	0
Дуальная гистограмма 4	1	0	0	0	0	1	0	0
Дуальная гистограмма 5	1	0	0	0	0	1	0	0

В таблице 3.3 представлены результаты тестирования СО по алгоритмам Outguess и F5. Цифры в таблицах показывают процент принятия решения по видам алгоритма вложения.

Таблица 3.3 – Результат МОВ при тестировании СО с максимальным вложением по алгоритму Outguess и СО с вложением одного байта по алгоритму F5

Используемые функционалы	**Outguess (доля вложения – 1)**				**F5 (вложение – 1 байт)**			
	ПО	**F5**	**outguess**	**MB**	**ПО**	**F5**	**outguess**	**MB**
Глобальная гистограмма	0	0	1	0	0	0,9916	0,0084	0
Индивидуальная гистограмма для положения 12	0	0,0538	0,9462	0	0	0,4385	0,5615	0

Продолжение таблицы 3.3

Используемые функционалы	**Outguess (доля вложения – 1)**				**F5 (вложение – 1 байт)**			
	ПО	**F5**	**outguess**	**MB**	**ПО**	**F5**	**outguess**	**MB**
Индивидуальная гистограмма для положения 21	0	0	1	0	0	1	0	0
Индивидуальная гистограмма для положения 13	0	0,2267	0,7733	0	0	0,4870	0,5130	0
Индивидуальная гистограмма для положения 31	0	0	1	0	0	1	0	0
Индивидуальная гистограмма для положения 22	0,0018	0	0,9982	0	0	0,8621	0,1379	0
Дуальная гистограмма -5	0	0	1	0	0	1	0	0
Дуальная гистограмма -4	0	1	0	0	0	1	0	0
Дуальная гистограмма -3	0	0,2234	0,7766	0	0	0,4711	0,5289	0
Дуальная гистограмма -2	0	1	0	0	0	1	0	0
Дуальная гистограмма -1	0,0009	0,1839	0,8152	0	0,0004	0,9145	0,0851	0
Дуальная гистограмма 0	0,0782	0,9218	0	0	0,1635	0,8365	0	0
Дуальная гистограмма 1	0	1	0	0	0	1	0	0
Дуальная гистограмма 2	0	1	0	0	0	1	0	0

Продолжение таблицы 3.3

Используемые функционалы	**Outguess (доля вложения – 1)**				**F5 (вложение – 1 байт)**			
	ПО	**F5**	**outguess**	**MB**	**ПО**	**F5**	**outguess**	**MB**
Дуальная гистограмма 3	0	0,2110	0,7890	0	0	0,4680	0,5320	0
Дуальная гистограмма 4	0	1	0	0	0	1	0	0
Дуальная гистограмма 5	0	0	1	0	0	1	0	0

В таблице 3.4 представлены результаты тестирования ПО и СО, созданных по алгоритму F5 при различных долях вложения. Цифры в таблицах показывают процент принятия верного решения (выбора верной гипотезы).

Таблица 3.4 – Результат МОВ с применением функционалов DCT-коэффициентов функционалов Маркова

Доля вложенной информации	**Функционалы DCT-коэффициентов**	**Функционалы Маркова**
1	0,9949	0,9980
0,5	0,9880	0,9920
0,25	0,8454	0,8694
0	0,9980	0,9153

Как видно из полученных результатов тестирования, различные функционалы, дают различную надежность распознавания. Отметим, что МОВ достаточно надежен, поскольку верный выбор гипотез происходит с вероятностью не менее 0,84.

Описанный выше МОВ опирается на формирование двух множества, получаемых на этапе обучения – множества, полученное из ПО, и множество, полученного из СО. Для данного МОВ необходимо иметь алгоритмы вложения,

которые могут быть применены для создания СО, в виде черных ящиков. На практике атакующий часто не располагает информацией о том, какие методы СГ будет применять передающий, и не имеет СГС даже в виде черного ящика. Таким образом создание множества СО, необходимых для обучения МОВ, становиться невозможным.

Решить эту проблему можно с помощью альтернативного метода использования МОВ.

При альтернативном методе использования МОВ формируется только одного множество [13]. В этом случае при обучении формируется только множество, полученное из ПО. При тестировании, если исследуемый объект попадает в полученное на этапе обучение множество, то принимается гипотеза H_1, если не попадает в полученное множество – гипотеза H_0.

Основное преимущество альтернативного метода использования МОВ – атакующему не надо знать, какие методы СГ использует передающим, и не надо иметь возможные алгоритмы вложений в виде черных ящиков. Достаточно провести обучение МОВ на достаточно большом объеме ПО.

При применении альтернативного метода использования МОВ вероятность ложной тревоги будет близка к 0, но вероятность пропуска наоборот будет довольно большой. На практике, при использовании МОВ с двумя множествами, множества СО и ПО пересекаются, давая либо несепарабельную линейную МОВ, либо нелинейную МОВ. При использовании только одного множества для обучения МОВ точность СГА зависит от того, как будут очерчены границы данного множества. При аппроксимации, близкой к линейной, часть СО попадут в множество ПО. Уменьшить вероятность пропуска можно, если как можно точнее очертить границы множества ПО. Но, поскольку множество СО неизвестно, точное разграничение этих множеств все равно не осуществимо. Как следствие, возникнет большая вероятность пропуска. Для детального описание альтернативного подхода использования МОВ требуется проведение дополнительного исследования, которое выходит за рамки данной диссертационной работы.

3.4 Выводы

Как показано в параграфе 3.3, ССГА достаточно эффективен и может успешно применяться для атак на подозрительные объекты. При этом данный СГА позволяет не только определить наличие или отсутствие скрытого вложения в исследуемом объекте, но и определить с некоторой долей точности, какой именно метод вложения был применен при создании СО. Экспериментальные результаты по исследования эффективности ССГА были использованы в научно-исследовательской работе «Ярус -СГ» [5].

Эффективность ССГА зависит от того, насколько «качественно» прошел этап обучения МОВ. Для обучения надежного МОВ необходимо прежде всего знать СГС, которые могут быть применены для создания СО. При этом не обязательно знать непосредственно алгоритмы вложения, достаточно иметь СГС в виде черных ящиков.

Также для обучения необходимо иметь большую базу ПО. При этом, чем больше ПО и созданных из них СО будет обработано на этапе обучения, тем эффективней будет МОВ на этапе тестирования.

Для надежного определения наличия или отсутствия вложения и метода СГ объекты, исследуемые на этапе тестирования, должны быть схожи по своим статистическим свойствам с объектами, на которых было произведено обучения МОВ. Если исследуемый объект будет сильно отличаться по статистическим свойствам от объектов, используемых на этапе обучения, эффективность и надежность МОВ будут уменьшаться, а вероятность ложной тревоги и вероятность пропуска будут увеличиваться. Отметим также, что если статистические характеристики ПО, используемых на этапе обучения, будут сильно отличаться друг от друга, то и в этом случае надежность МОВ будет уменьшаться.

Из вышесказанного видно, что для эффективного и надежного СГА МОВ должен применяться к тем исследуемым объектам, чьи статистические свойства схожи со свойствами объектов, используемых на этапе обучения. К сожалению,

предсказать статистические свойства исследуемых объектов нельзя, и это является основным недостатком МОВ.

Хотя для ССГА не нужно заранее знать алгоритм вложения, который мог быть применен к исследуемому объекту, но список СГС (хотя бы в виде черных ящиков), которые могли быть использованы, знать нужно, поскольку без этого не провести обучение МОВ. Следовательно, МОВ будет иметь большие вероятность пропуска, если применять его к СО, созданным СГС, не использованными на этапе обучения.

Отметим, что хотя ССГА имеет преимущество перед методами целевого СГА, а именно, не требует знания алгоритма вложения,тем не менее, целевые методы СГА остаются актуальными по ряду причин:

- многие СГС (в особенности «любительские») применяют алгоритм вложения СГ-НЗБ без модификации, такие СГС быстрее обнаруживаются целевым СГА, чем ССГА;
- многие стойкие СГС лучше обнаруживаются подобранными специально для них целевыми методами СГА;
- методы целевого СГА требуют меньше аппаратного временного ресурса для их реализации.

Поэтому в следующих главах будут рассмотрены целевые методы СГА.

4 Целевой стегоанализ для метода вложения СГ-НЗБ

В параграфе 1.2 был описан алгоритм вложения СГ-НЗБ, который может применяться как для вложения в область пикселей, например, в файлы формата BMP, так и для вложения в частотную область (DCT-коэффициенты файлов формата JPEG). В данной главе будут описаны методы целевого СГА для метода СГ-НЗБ, использующего вложение в область пикселей[35,36].

4.1 Визуальный метод стегоанализа

Пусть $C(n)$, $n = 1, 2..N$ – значения оттенков пикселей изображения формата Grayscale (оттенки серого) BMP (8-битное представление оттенка для каждого пикселя, максимально возможное количество оттенков серого в изображении – 256), где n – порядковый номер пикселя, N – общее количество пикселей в изображении.

Преобразуем исходное изображение $\left(C(n)\right)_{n=1}^{N}$ к черно-белому изображению $\left(\widetilde{C}(n)\right)_{n=1}^{N}$ [37]по правилу:

$$\widetilde{C}(n) = \begin{cases} \text{белое, если } C_0(n) = 1, \\ \text{черное, если } C_0(n) = 0, \end{cases} \tag{4.1}$$

где $C_0(n)$ – значение НЗБ n-ого пикселя.

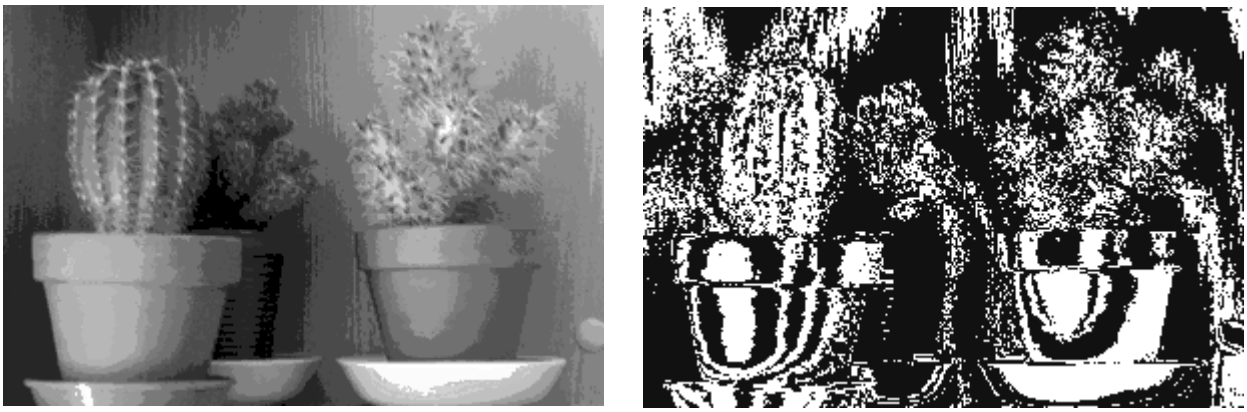
Если в исследуемом изображении нет вложения, то на вновь полученном черно-белом изображении будут видны некоторые контуры исходного изображения. Если же исследуемое изображение содержит вложение, то новое черно-белое изображение будет выглядеть как шумовое поле.

Вложение может быть произведено не во все пиксели, а лишь в некоторые. Тогда доля вложенной секретной информации относительно общего числа пикселей равна P. Следовательно, измениться только часть пикселей, и на изображении $\left(\widetilde{C}(n)\right)_{n=1}^{N}$, полученном из исследуемого изображения, будут

просматриваться контуры первоначального изображения, даже если исследуемое изображение содержит скрытую информацию.

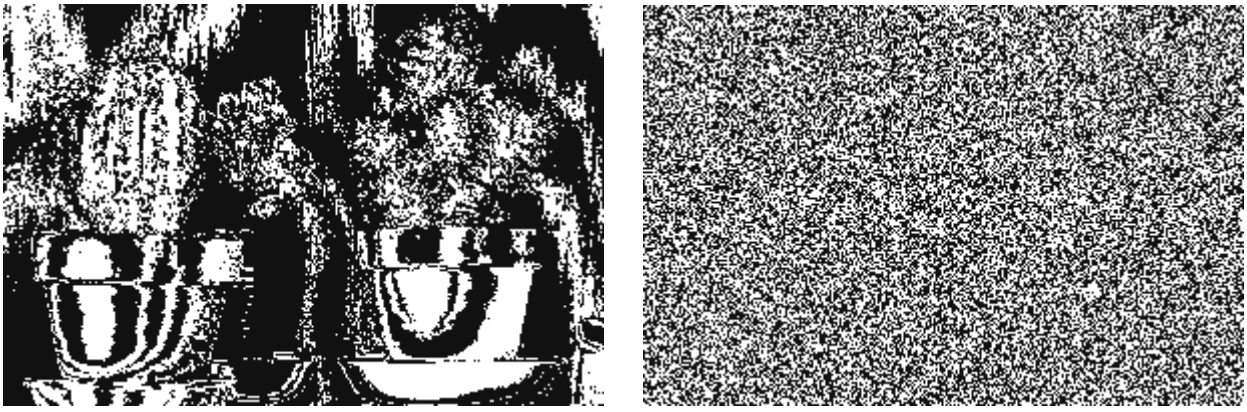
Использование данной атаки возможно только в присутствии оператора-стегоаналитика, поскольку наличие или отсутствие вложения скрытой информации в исследуемом изображении определятся по результатам атаки «на глаз» непосредственно человеком.

На рисунках 4.1 и 4.6 показаны исходные ПО и результаты визуального СГА. На рисунках 4.2-4.5 и 4.7-4.9 показаны результаты СГА на ПО и СО с различными долями вложения P.



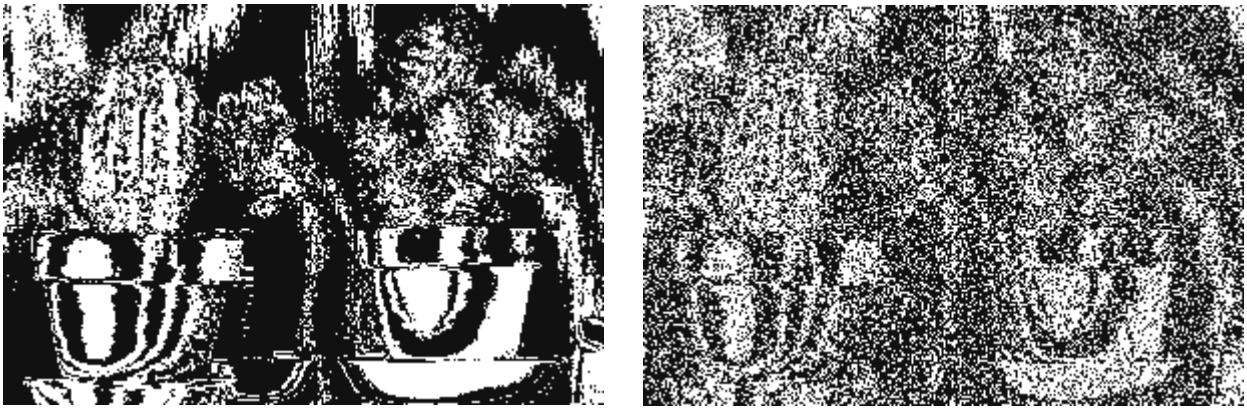
а б

Рисунок 4.1 – а – покрывающее сообщение,
б – покрывающее сообщение после визуальной атаки



а б

Рисунок 4.2 – а – покрывающее сообщение после атаки,
б – изображение с вложением $P = 1$ после атаки



а б

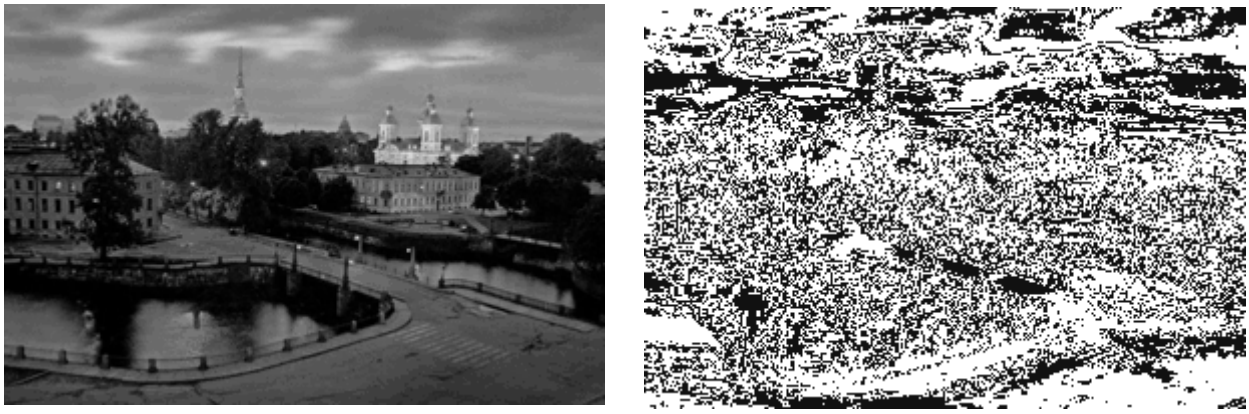
Рисунок 4.3 – а – покрывающее сообщение после атаки,
б – изображение с вложением $P = 0{,}5$ после атаки

а б

Рисунок 4.4 – а – покрывающее сообщение после атаки,
б – изображение с вложением $P = 0{,}1$ после атаки

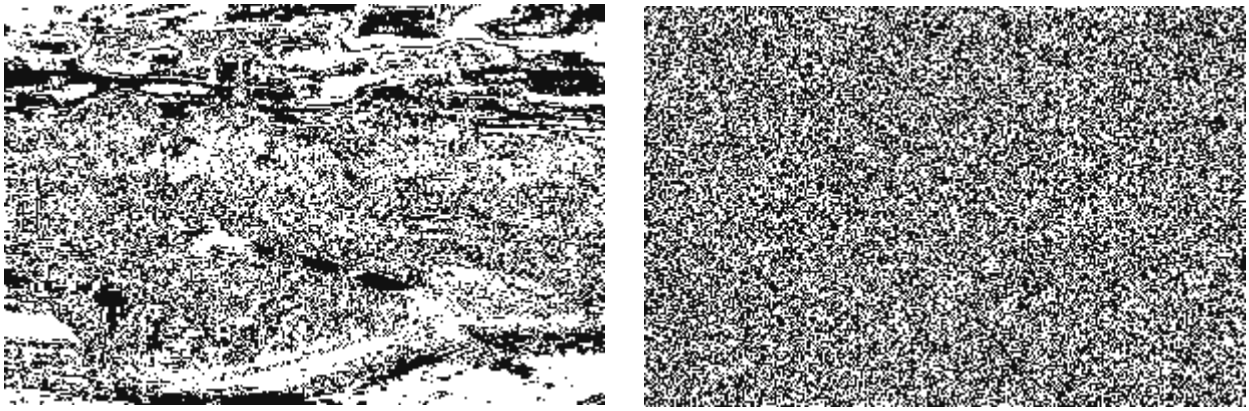
а б

Рисунок 4.5 – а – покрывающее сообщение после атаки,
б – изображение с вложением $P = 0{,}05$ после атаки



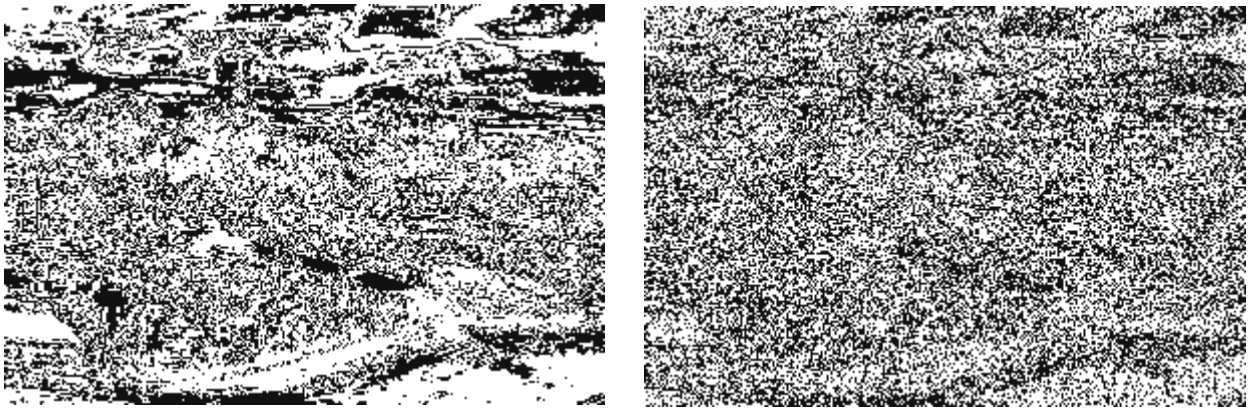
а б

Рисунок 4.6 – а – покрывающее сообщение,
б – покрывающее сообщение после атаки



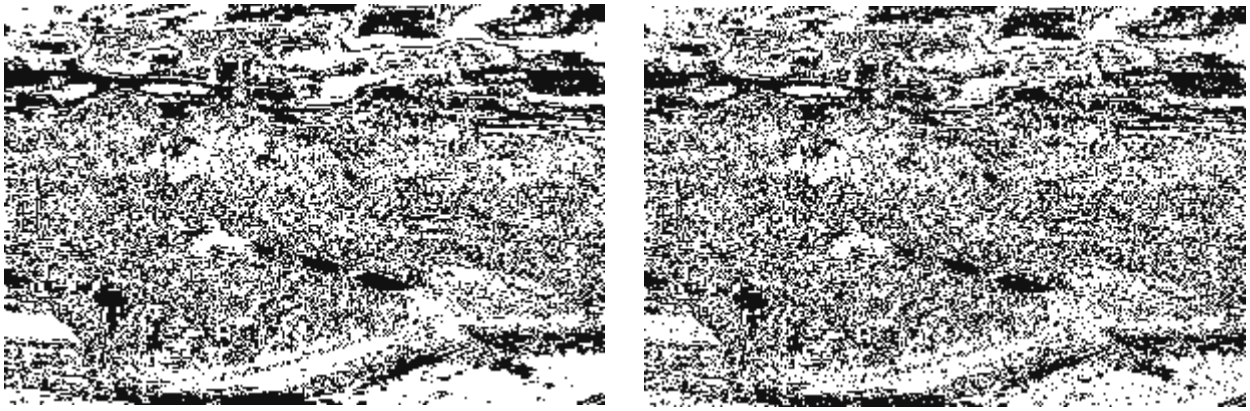
а б

Рисунок 4.7 – а – покрывающее сообщение после атаки,
б – изображение с вложением $P = 1$ после атаки



а б

Рисунок 4.8 – а – покрывающее сообщение после атаки,
б – изображение с вложением $P = 0{,}5$ после атаки



а б

Рисунок 4.9 – а – покрывающее сообщение после атаки,
б – изображение с вложением $P = 0{,}1$ после атаки

Отметим, что $\widetilde{C}(n)$ до вложения не известно стегоаналитику, поэтому он может принять решение о наличии или отсутствии скрытой информации в изображении основываясь лишь на степени зашумленности полученного после атаки изображения.

Сравнение рисунка 4.1б и рисунков 4.2-4.5 б показывает, что для данного изображения можно обнаружить вложение методом визуального СГА, если доля вложения больше или равна 0,1.

Однако, для другого вида изображения (рисунки 4.6-4.9) применение данного метода оказывается затруднительным, поскольку отличить визуально преобразованныйСО от преобразованного ПО можно в том случае, если доля вложения больше или равна 0,5.

Как видно из рисунков 4.1-4.9, визуальный СГА для большинства изображений достаточно надежен при $P \geq 0{,}5$, в ином случае стегоаналитику трудно определить, чем именно вызван шум на изображении $\left(\widetilde{C}(n)\right)_{n=1}^{N}$ – естественным шумом изображения $\left(C(n)\right)_{n=1}^{N}$ или наличием скрытой информации.

Можно выработать единые визуальные «пороговые параметры» для анализа различный изображений. Но такие параметры будут иметь существенные ограничения:

- если считать, что вложение есть только в тех случаях, когда контуры исходного изображения практически не различимы, то есть те случаи, когда доля вложения скрытой информации больше 0,5, то вероятность пропуска P_m будет большой;
- если за «порог» примем результаты атаки с хорошо просматриваемыми контурам, но при этом с небольшим шумом, возрастает вероятность ложной тревоги P_{fa} .

Из вышеперечисленного можно сделать следующие выводы относительно данного метода СГА:

а) Для реализации визуальной атаки необходимо присутствие оператора, который будет принимать решение о наличии или отсутствии вложения. Автоматически распознавать СО методом визуальной атаки в настоящее время не представляется возможным.

б) Наилучшие результаты визуальная атака дает в случае использования в качестве покрывающего сообщения изображений, имеющих малое количество шумовых полей и большое количество однородных текстур, контуров и линий с резкими перепадами яркостей.

в) Данный метод можно рекомендовать для первичного анализа (например, в случае «малошумного» изображения), а также в сочетании с другими методами.

г) Данный метод невозможно применить для анализа изображений формата JPEG, поскольку замена DCT-коэффициентов не приводит исходное изображение к черно-белому.

В следующих главах будет показано, что существуют более эффективные (к тому же полностью автоматические) методы СГА при вложении в НЗБ, поэтому развивать далее визуальный метод не имеет смысла.

4.2 Стегоанализ на основе статистики 1-ого порядка (гистограммная атака)

Данный метод описан в [2]. Гистограммой $V(i)$, $i = 1, 2..L$ изображения $C(n)$, $n = 1, 2..N$ с L градациями яркости, называется относительная частость появления каждого из уровня яркости, или формально

$$V(i) = \frac{\#\{n \in (1, 2..N) : C(n) = i\}}{N},\ i = 1, 2..L. \qquad (4.2)$$

При вложении в НЗБ выполняются условия, приведенные в таблице 4.1.

Таблица 4.1 – Свойства СГ-НЗБ

		$C_w(n) = i$	
		НЗБ=0	НЗБ=1
$C(n) = i$	Четный уровень $i = 2j$	Четный уровень $i = 2j$	Нечетный уровень $i = 2j + 1$
	Нечетный уровень $i = 2j + 1$	Четный уровень $i = 2j$	Нечетный уровень $i = 2j + 1$

Используемые в таблице 4.1 обозначения:

$C(n)$ – яркость n-ого пикселя до вложения,

$C_w(n)$ – яркость n-ого пикселя после вложения,

i, j – целые числа.

Заметим, что $2j \not\to 2i - 1$, $2j + 1 \not\to 2(i + 1)$ при вложении любого бита («0» или «1»), где $\not\to$ – означает невозможность перехода.

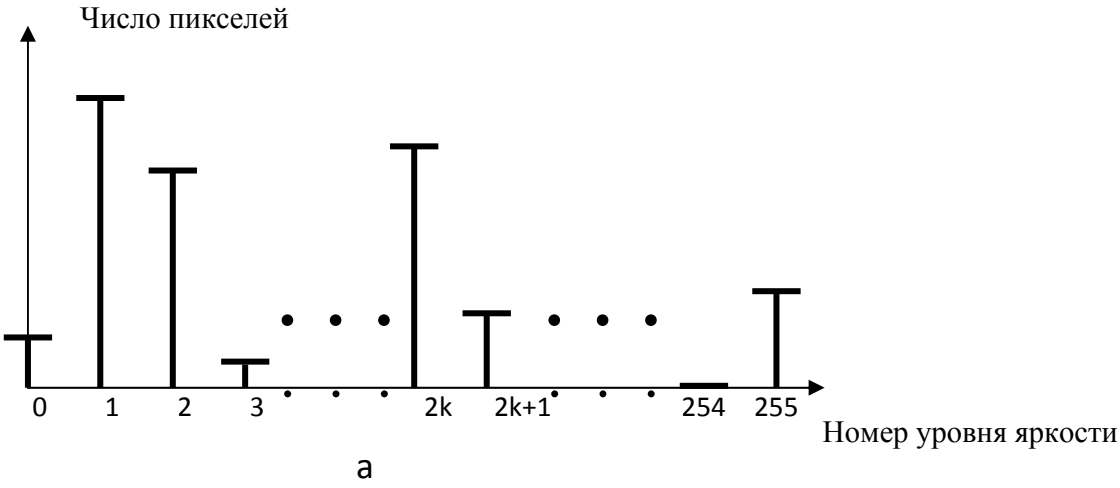
Полагаем, что секретное сообщение зашифровано стойким шрифтом, тогда можно считать, что $P\{w(n) = НЗБ = 0\} = P\{w(n) = НЗБ = 1\} = 1/2$.

Легко видеть, что при любой гистограмме $V(i)$ исходного изображения $C(n)$, в гистограмме $V_w(i)$ изображения после вложения $C_w(n)$ при доле вложения $P = 1$ (вложение производиться в каждый пиксель) будет выполняться условие:

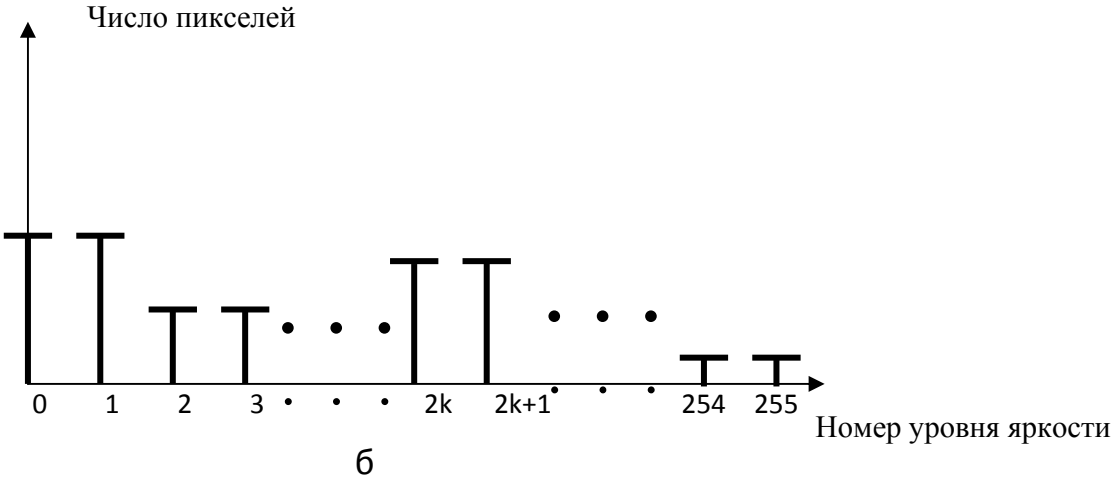
$$E\{V_w(2j)\} = E\{V_w(2j+1)\}_{j=0,1,..(L-1)/2} = \frac{1}{2}\left(V(2j) + V(2j+1)\right), \qquad (4.3)$$

где $E\{...\}$ – символ математического ожидания.

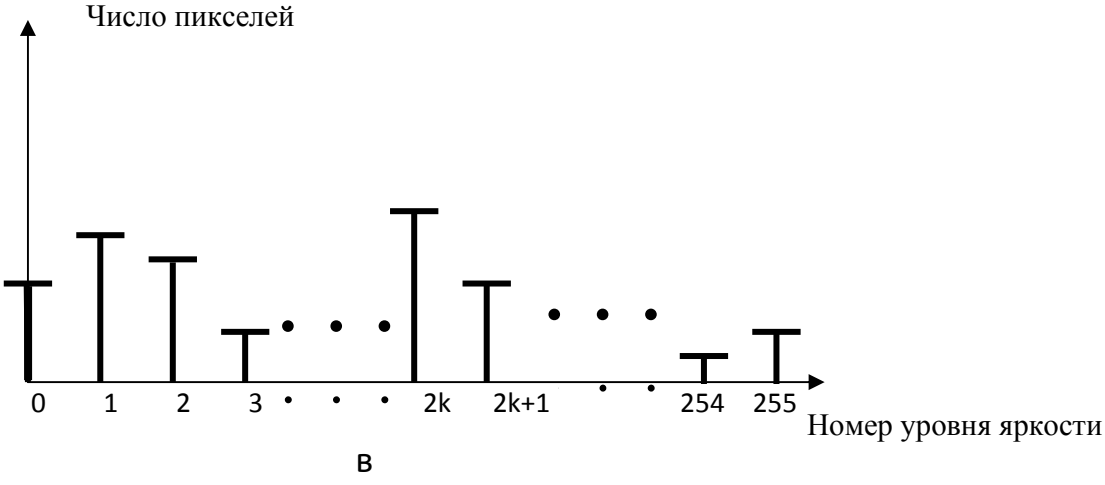
Следовательно, для достаточно больших изображений (содержащих большое количество пикселей N) гистограммы $V(i)$ и $V_w(i)$ будут «качественно» иметь вид, показанный на рисунке 4.10.



а



б



в

Рисунок 4.10 – а – гистограмма ПО $V(i)$,
б – гистограмма СО $V_w(i)$ с вероятностью 1,
в – гистограмма СО $V_w(i)$ с вероятностью меньше 1

Если вероятность вложения секретной информации в каждый пиксель равна $P \neq 1$, а вероятность невложения $1-P$, то есть вложение в НЗБ производиться по секретному ключу, получаем следующее соотношение:

$$\begin{aligned} E\{V_w(2j)\} &= \frac{P}{2}\big(V(2j)+V(2j+1)\big)+(1-P)V(2j), \\ E\{V_w(2j+1)\} &= \frac{P}{2}\big(V(2j)+V(2j+1)\big)+(1-P)V(2j+1). \end{aligned} \tag{4.4}$$

Из соотношения (4.4) видно, что если $P << 1$, то $E\{V_w(2j)\} \neq E\{V_w(2j+1)\}$.

Однако, факт приближенного равенства значений гистограммы для уровней $2j$ и $2j+1$ можно использовать и при $P << 1$.

Из теории вероятностей известно, что если случайный вектор частот ν_i соответствует вероятности P_i, $i = 1, 2..K$, то случайная величина

$$\chi^2 = \sum_{j=1}^{K} \frac{(\nu_j - P_j)^2}{P_j}. \tag{4.5}$$

Пусть $N = \nu_1 + \nu_2 + ... + \nu_K$, тогда в пределе $(N \to \infty)$ будем иметь так называемое χ^2-распределение с $K-1$ степенями свободы:

$$P\{\chi^2 \le x\} = \int_0^x \frac{1}{2^{\frac{K-1}{2}}\Gamma\left(\frac{K-1}{2}\right)} x^{\frac{K-1}{2}-1} e^{-\frac{x}{2}} dx, \tag{4.6}$$

где $\Gamma(\cdot)$ – гамма функция.

Положим $\nu_j = V(2j)$, $j = 0, 1..\left[(L-1)/2\right]$, $[x]$ – целая часть x. При вложении по методу СГ-НЗБ при $P = 1$ получаем из (4.3):

$$P_j = \frac{1}{2}\big(V(2j)+V(2j+1)\big). \tag{4.7}$$

Следовательно, (4.5) принимает вид:

$$\chi^2 = \sum_{j=0}^{\left[(L-1)/2\right]} \frac{\big(V(2j)-V(2j+1)\big)^2}{2(V(2j)+V(2j+1))}. \tag{4.8}$$

Из (4.2) следует:

$$\begin{cases} V(2j) = \dfrac{n_{2j}}{N} \\ V(2j+1) = \dfrac{n_{2j+1}}{N} \end{cases}, \tag{4.9}$$

где n_{2j} – количество пикселей оттенка $2j$,

n_{2j+1} – количество пикселей оттенка $2j+1$,

N – общее число пикселей в изображении.

Подставим (4.9) в (4.8), получаем:

$$\chi^2 = \sum_{j=0}^{\left[(L-1)/2\right]} \frac{\left(\dfrac{n_{2j}}{N} - \dfrac{n_{2j+1}}{N}\right)^2}{2\left(\dfrac{n_{2j}}{N} + \dfrac{n_{2j+1}}{N}\right)},$$

$$\chi^2 = \sum_{j=0}^{\left[(L-1)/2\right]} \frac{\left(n_{2j} - n_{2j+1}\right)^2}{2\left(n_{2j} + n_{2j+1}\right)}. \tag{4.10}$$

Обозначим пороговое значение через χ_0^2. Тогда критерий обнаружения СГ-НЗБ будет иметь следующий вид:

$$\begin{cases} \chi^2 < \chi_0^2 \text{ - принимаем гипотезу } H_1, \\ \chi^2 \geq \chi_0^2 \text{ - принимаем гипотезу } H_0. \end{cases} \tag{4.11}$$

Вероятность необнаружения (пропуска) СО P_m можно вычислить при больших N, используя асимптотическое распределение (4.6):

$$P_m \leq \int_{\alpha}^{\infty} \frac{1}{2^{\frac{K-1}{2}}\Gamma\left(\dfrac{K-1}{2}\right)} x^{-\frac{K-1}{2}-1} e^{-\frac{x}{2}} dx. \tag{4.12}$$

При заданном P_m по (4.12) можно выбрать пороговое значение χ_0^2. Вероятность ложной тревоги P_{fa} приходитсявычислять только экспериментально на массиве ПО. При этом, чем больше разнородных изображений содержится в массиве, тем точнее можно вычислить P_{fa}.

Критерий обнаружения СГ-НЗБ (4.11) выведен для $P = 1$. Однако, как показывают эксперименты [37], на практике данную атаку можно использовать и при $P < 1$.

В таблице 4.2 приведены экспериментальные значения χ^2 для 5 различных изображений при долях вложения 1; 0,5; 0,1; 0,05; 0,01; 0.

Таблица 4.2 – Значение χ^2 при различных долях вложения

№изображения	P	χ^2
1	1	38
	0,5	14338
	0,1	46908
	0,05	52520
	0,01	57010
	0	58211
2	1	61
	0,5	14948
	0,1	48102
	0,05	54069
	0,01	58476
	0	60000
3	1	36
	0,5	14286
	0,1	45917
	0,05	51192
	0,01	55853
	0	56936
4	1	56
	0,5	15129
	0,1	47792
	0,05	53642
	0,01	58449
	0	59589

Продолжение таблицы 4.2

№изображения	P	χ^2
5	1	44
	0,5	15004
	0,1	48409
	0,05	54386
	0,01	58552
	0	59928

Из экспериментальных данных таблицы 4.2 видно, что при вложении информации в НЗБ с уменьшением вероятности вложения P величина χ^2 увеличивается. Отметим, что величина χ^2 для СО при $P < 0{,}05$ мало отличается от величины χ^2 для ПО ($P = 0$).

При $P = 1$ пороговое значение χ_0^2 можно выбрать, используя P_m и неравенство (4.12).

Пример расчета порогового значения χ_0^2 по (4.12) при заданных значениях $\widetilde{P}_m$ 0,05; 0,01; 0,001 и при $K = 128$ (параметр соответствует изображению формата BMP GrayScale 8 бит/пиксель) приведен в таблице 4.3. Для расчета использовались таблицы χ^2-распределения с $K - 1 = 127$ степенями свобод [38].

Таблица 4.3 – Значение χ_0^2 при заданных значениях $\widetilde{P}_m$ – 0,05; 0,01 и 0,001

$\widetilde{P}_m$	0,05	0,01	0,001
χ_0^2	124,3	135,8	149,4

Результат экспериментального вычисления вероятностей P_m и P_{fa} для более чем 100 изображений при выборе порогового значений в соответствии с таблицей 4.3 при $\widetilde{P}_m$ 0,05; 0,01; 0,001 приведены в таблице 4.4.

Таблица 4.4 – Экспериментальные результаты вычисления вероятностей P_m и P_{fa} при заданных значениях $\widetilde{P}_m$ – 0,05; 0,01 и 0,001

	$\widetilde{P}_m$		
	0,05	0,01	0,001
	$\lambda = 124{,}3$	$\lambda = 135{,}8$	$\lambda = 149{,}4$
P_m	0,03	0,01	0
P_{fa}	0	0	0

Из таблиц 4.3 и 4.4 видно, что $\widetilde{P}_m \approx P_m$, следовательно, неравенство (4.12) можно использовать для расчета порогового значения χ_0^2 при $P = 1$.

При $P < 1$ выбор порогового значения χ_0^2 представляет собой значительные трудности, поскольку в данном случае отсутствует теоретическая возможность расчета χ_0^2 и подбирать оптимальное пороговое значение надо экспериментальным путем.

Экспериментальные результаты расчета вероятности пропуска P_m и вероятности ложной тревоги P_{fa} для различных значений P 1; 0,5; 0,1; 0,05; 0,01 приведены в таблице 4.5. Статистика, приведенная в таблице, собрана при анализе массива из более чем 100 изображений. Заметим, что пороговые значения χ_0^2 в данной таблице значительно выше пороговых значений, приведенных в таблице 4.4, это объясняется тем, что для анализа P_m и P_{fa} при малых P требуется высокий порог χ_0^2.

Таблица 4.5 – Экспериментальный расчет величин P_{fa} и P_m

χ_0^2		59000	58000	56000	54000	51000	45000
P_{fa}		0,33	0,30	0,2	0,15	0,09	0,07
$P = 1$	P_m	0	0	0	0	0	0
$P = 0{,}5$	P_m	0	0	0	0	0	0
$P = 0{,}1$	P_m	0	0	0	0	0	0,62
$P = 0{,}05$	P_m	0	0	0	0,54	0,75	0,85
$P = 0{,}01$	P_m	0	0,65	0,73	0,82	0,86	0,90

Из таблицы 4.5 видно, что наиболее приемлемое пороговое значение – 54000, при этом отметим, что пороговое значение можно менять, подбирая его под необходимые значения вероятностей ложной тревоги и пропуска. Дополнительные исследования показали, что по критерию χ^2 лучше обнаруживается СГ-НЗБ для качественных изображений, снятых при правильных настройках (при низком разрешении может возникнуть зернистость (пикселизация) изображение), очищенных от цифрового шума (цифровой шум можно легко увидеть – на изображение как будто наложено множество точек различной яркости и цвета). При этом для изображений с сильным шумом (например, передаваемых по каналу с большим уровнем шума) или имеющих высокую зернистость (например, снятых при очень высоком значении светочувствительности) СОс вложением по методу СГ-НЗБ обнаруживаются ненадежно, поскольку для подобных изображений возрастает вероятность ложной тревоги P_{fa}.

На основании проведенных экспериментов можно сделать следующие выводы относительно данного метода СГА:

а) При использовании критерия χ^2 для проверки изображений с малым шумом и зернистостью СГ-НЗБ обнаруживается достаточно надежно

при $P \geq 0{,}05$, если изображение сохранено в Paint, и при $P \geq 0{,}1$ – в Adobe Photoshop.

б) При $P \leq 0{,}1$ оказывается проблематичным обнаружение СГ-НЗБ для изображений, поврежденных шумом или имеющих высокую зернистость.

Поскольку вся теория χ^2-распределение остается верной и для гистограмм DCT-коэффициентов, можно полагать, что данный метод универсален как при вложении скрываемой информации в область пикселей, так и при вложении в частотную область, то есть при вложении информации в изображение формата JPEG.

Отметим, что данная атака может быть использована и для модифицированной СГ-НЗБ. При анализе 10000 изображений, если принять вероятность пропуска равной 0,5, то вероятность ложной тревоги для метода χ^2-распределение и метода с использованием двухмерного преобразования Фурье от матрицы смежности при доле вложения информации 1 равна 0,14 и 0,13, при доле вложения 0,5 – 0,30 и 0,21, при доле 0,1 – 0,46 и 0,44, – соответственно.

4.3 Стегоанализ на основе статистики 2-ого порядка

Рассмотренные метод предложен в работе [3] и назван там sample pair analysis – парно-выборочный анализ (далее ПВА).

Более подробно, чем в [3], доказательство законности данного метода представлено в Приложении А.

Ведем следующие обозначения для исследуемого изображения:

- C_0 – количество последовательных пар яркостей пикселей, совпадающих в первых 7-ми битах;
- C_1 – количество последовательных пар яркостей пикселей, которые, если откинуть наименее значащий бит, в первых 7 битах будут отличаться друг от друга на величину 1;

- D_0 – количество последовательных пар яркостей пикселей, совпадающих во всех битах;
- D_2 – количество последовательных пар яркостей пикселей, которые отличаются друг от друга на величину 2;
- X_1 – количество последовательных пар яркостей пикселей вида $(2k,2k-1)$ или $(2k-1,2k)$, где k – целое число;
- Y_1 – количество последовательных пар яркостей пикселей вида $(2k,2k+1)$ или $(2k+1,2k)$, где k – целое число.

Как доказано в [3] и в Приложении А, результатом данной атаки $\widetilde{P}$ в анализируемом изображении является наименьшим вещественным корнем квадратного уравнения

$$\frac{\widetilde{P}^2}{4}(2C_0 - C_1) - \frac{\widetilde{P}}{2}(2D_0 + 2Y_1 - D_2 - 2X_1) + Y_1 - X_1 = 0. \qquad (4.13)$$

Решение уравнения (4.13) имеет вид:

$$\widetilde{P} = \frac{\frac{2D_0 + 2Y_1 - D_2 - 2X_1}{2} - \sqrt{\left(-\frac{2D_0 + 2Y_1 - D_2 - 2X_1}{2}\right)^2 - \frac{2C_0 - C_1}{4}(Y_1 - X_1)}}{\frac{2C_0 - C_1}{2}}. \qquad (4.14)$$

Для атаки методом ПВА критерий обнаружения СГ-НЗБ имеет следующий вид:

$$\begin{cases} \widetilde{P} > \widetilde{P}_0 \text{ - принимаем гипотезу } H_1, \\ \widetilde{P} \le \widetilde{P}_0 \text{ - принимаем гипотезу } H_0, \end{cases}$$

где $\widetilde{P}_0$ – пороговое значение.

Метод ПВА позволяет не только определить наличие или отсутствие вложения в исследуемом объекте, значение $\widetilde{P}$ является оценкой доли вложенной скрытой информации.

Экспериментальные результаты исследований [37] для пяти изображений при различных долях вложения P показаны в таблице 4.6.

Таблица 4.6 – Экспериментальные результаты расчета $\widetilde{P}$ по методу ПВА

P	**№ изображения**				
	1	**2**	**3**	**4**	**5**
1	0,5856	0,5273	0,5731	0,5861	0,5806
0,5	0,4074	0,3786	0,3967	0,4044	0,4034
0,1	0,0973	0,0866	0,0800	0,0987	0,0966
0,05	0,0500	0,0392	0,0335	0,0504	0,0456
0,01	0,0107	0,0104	0	0,0093	0,0121
0,005	0,0052	0,0056	0	0,051	0,0049
0,001	0,0009	0,0009	0	0,0007	0,0012
0,0005	0,0005	0	0	0	0,0005
0,0001	0	0	0	0	0,0001
0	0	0	0	0	0,00005

Для расчета вероятности ложной тревоги P_{fa} и вероятности пропуска P_m при атаке исследуемого изображения методом ПВА был проведен эксперимент [37] на массивеиз более чем 100 изображений при заданных пороговых значениях $\widetilde{P}_0 = 0$, как рекомендовано в [3], и $\widetilde{P}_0 = 0{,}001$. Результаты эксперимента приведены в таблице 4.7.

Таблица 4.7 – Экспериментальный расчет величин P_{fa} и P_m при пороговых значениях $\widetilde{P}_0 = 0$ и $\widetilde{P}_0 = 0{,}001$

P	$\widetilde{P}_0 = 0$		$\widetilde{P}_0 = 0{,}001$	
	P_{fa}	P_m	P_{fa}	P_m
1	-	0	-	0
0,5	-	0	-	0
0,1	-	0,02	-	0,02
0,05	-	0,06	-	0,06
0,01	-	0,10	-	0,10
0,005	-	0,23	-	0,23
0,001	-	0,30	-	0,30
0,0005	-	0,43	-	0,45
0,0001	-	0,65	-	0,70
0	0,10	-	0,05	-

Как видно из таблиц 4.6 и 4.7, ПВА позволяет достаточно надежно обнаружить наличие скрытой информации при доле вложения $P \geq 0{,}01$. Таким образом, в изображении формата BMP GrayScale 8 бит/пиксель размером 300x200 пикселей можно вложить не более 600 бит необнаруживаемой информации.

Опираясь на результаты, отображенные в таблице 4.7, в данной работе рекомендуется в качестве порогового значения использовать $\widetilde{P}_0 = 0{,}001$, поскольку при $P \geq 0{,}01$ вероятность пропуска не увеличивается, а при дальнейшем уменьшении доли вложения увеличивается незначительно, но вероятность ложной тревоги уменьшается в 2 раза, по сравнении с $\widetilde{P}_0 = 0$.

Метод ПВА позволяет достаточно точно вычислить долю вложенной скрытой информации $\widetilde{P}$ для $0{,}01 \leq P \leq 0{,}1$. Если $0{,}1 < P \leq 1$, то $\widetilde{P}$ методом ПВА оказывается неточной, при этом, как правило, $\widetilde{P} < P$.

Метод ПВА может быть использован для обнаружения скрытой информации, вложенной в частотную область (DCT-коэффициенты в

изображении формата JPEG), тогда вместо значений оттенков пикселей анализируются уровни DCT-коэффициентов.

Чтобы оценить эффективность метода ПВА при анализе изображений формата JPEG был проведен эксперимент на массиве из более чем 100 изображений по расчету вероятностей пропуска P_m и ложной тревоги P_{fa} при различных долях вложения P. В качестве порогового значения выбрано $\widetilde{P}_0 = 0{,}45$. Результат эксперимента приведены в таблице 4.8.

Таблица 4.8 – Экспериментальное исследование ПВА при вложении в JPEG

P	P_{fa}	P_m
1	-	0
0,5	-	0,05
0,1	-	0,15
0,05	-	0,4
0	0,05	-

Из таблице 4.8 видно, что при вложении в DCT-коэффициенты СО при заданном пороге $\widetilde{P}_0 = 0{,}45$ достаточно хорошо обнаруживается при $P \geq 0{,}1$, имея при этом малую вероятность ложной тревоги P_{fa}.

4.4 Выводы

Рассмотренные в данной главе методы СГА широко описаны в научных статьях. Они достаточно точно распознают изображения с малым количеством шума (или очищенные от шума).

СГА, основанный на визуальной атаке имеет ряд недостатков по сравнению с СГА, основанными на статистике 1-ого и 2-ого порядка. Данный метод позволяет надежно обнаруживать вложение при $P \geq 0{,}5$.

Для метода визуальной атаки нельзя полностью реализовать автоматический алгоритм принятия решения о наличии или отсутствии скрытой

информации в изображении, для принятия данного решения требуется присутствие квалифицированного человека-оператора. Как показали эксперименты, проведенные в диссертации, метод визуальной атаки не достаточно эффективен для анализа изображений формата BMPи абсолютно неприменим для анализа изображений формата JPEG.

Методы, основанные на χ^2-распределении и ПВА, позволяют автоматизировать алгоритм СГА, поскольку решение о наличии или отсутствии вложения принимается с помощью заданного порогового значения.

Эффективность использования данных методов СГА зависит в том числе и от выбора порогового значения. В главе даны рекомендации по выбору пороговых значения для данных методов. Для метода, основанного на χ^2-распределении рекомендуется использовать пороговое значение $\chi_0^2 = 54000$, а для метода ПВА – $\widetilde{P}_0 = 0{,}001$.

Исследования, проведенные в диссертации, показали, что метод χ^2-распределения достаточно надежно обнаруживает вложение для $P \geq 0{,}1$ при $P_{fa} = 0{,}15$ и $P_m = 0$, а метод ПВА для $P \geq 0{,}01$ при $P_{fa} = 0{,}1$ и $P_m = 0{,}1$.

Для большинства изображения метод ПВА более эффективен, чем метод χ^2-распределения. Основные преимущества ПВА: надежность и наличие «дополнительной функции» – оценки доли вложенной информации.

Однако, для некоторых изображений, нарисованных и сохранены через редактор Paint, метод χ^2-распределения оказывается более надежным, чем ПВА. Для рисунков, обработанных и сохраненных в редакторе Adobe Photoshop, или цифровых фотографий более надежным оказывается метода ПВА.

Эксперименты показали, что методы, основанные на χ^2-распределения и ПВА, можно применять и при СГА изображений формата JPEG.

Отметим, что для выявления СО необходимо знать, в какую область было произведено вложение – в частотную область DCT-коэффициентов или в область пикселей.

5 Целевой стегоанализ для метода вложения СГ-ШПС

В параграфе 1.2 был описан метод вложения скрываемой информации методом СГ-ШПС. В главе 5 будут исследованы известные и вновь предложенные методы целевого СГА для СГ-ШПС.

5.1 Метод, основанный на «раздвоении пиков»

Предположим, что вложение скрытой информации производилось во все пиксели, то есть $P = 1$. Пусть ПО имеет пик в некоторой области гистограммы при $i = C_0$ [39], как показано на рисунке 5.1 а.

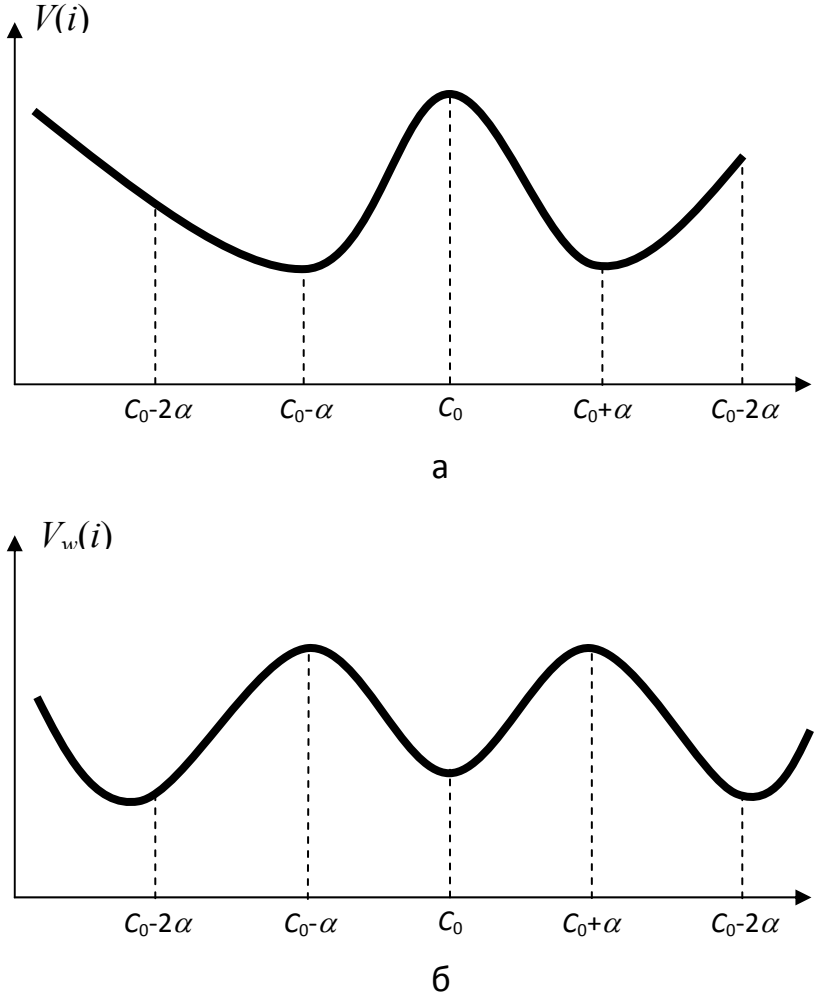


Рисунок 5.1 –Иллюстрация свойства раздвоения пиков
а – гистограмма ПО,
б – гистограмма СО

Если во вкладываемой последовательности символы 0 и 1 равновероятны, то для СГ-ШПС будет справедливо равенство:

$$V_w(i) = \frac{1}{2}\big(V(i-\alpha) + V(i+\alpha)\big),\ i = 1, 2..L, \tag{5.1}$$

где $V(i)$ – гистограмма ПО,

$V_w(i)$ – гистограмма СО.

Если $V(C_0 - \alpha) = V(C_0 + \alpha)$ после вложения, то для ситуации, показанной на рисунке 5.1, где $V(C_0) >> V(C_0 - \alpha)$ и $V(C_0) >> V(C_0 + \alpha)$, совместно с (5.1), получаем:

$$\begin{aligned} V_w(C_0) &= \frac{1}{2}\big(V(C_0 - \alpha) + V(C_0 + \alpha)\big) = V(C_0 - \alpha), \\ V_w(C_0 - \alpha) &= \frac{1}{2}\big(V(C_0 - 2\alpha) + V(C_0)\big) > V_w(C_0), \\ V_w(C_0 + \alpha) &= \frac{1}{2}\big(V(C_0) + V(C_0 + 2\alpha)\big) > V_w(C_0). \end{aligned} \tag{5.2}$$

Из соотношения (5.2) следует факт раздвоения пиков, при этом интервал между вершинами появившихся пиков будет равно 2α (рисунок 5.1 б). Высота вновь появившихся пиков окажется тем больше, чем больше первоначальный пик. Гистограммаизображения до и после вложения информации по методу СГ-ШПС при $\alpha = 3$ приведена на рисунке 5.2. На рисунке четко виден раздвоенный пик.

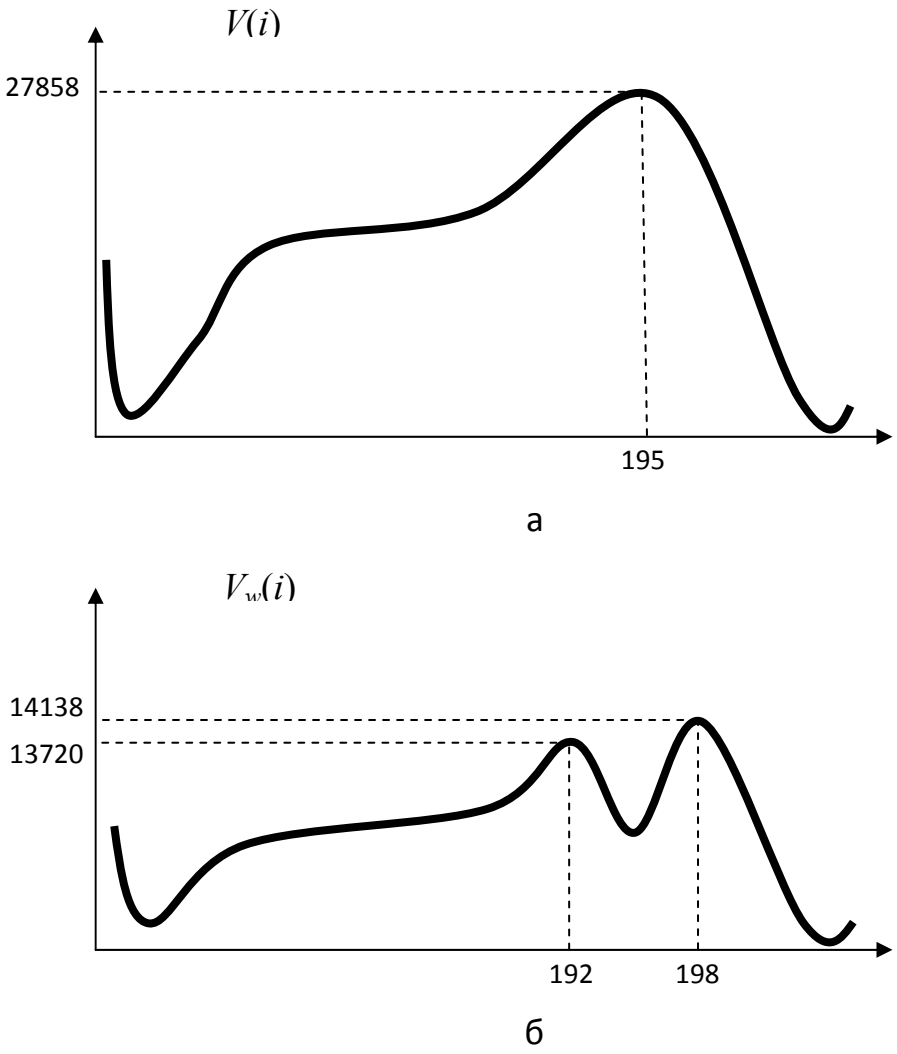
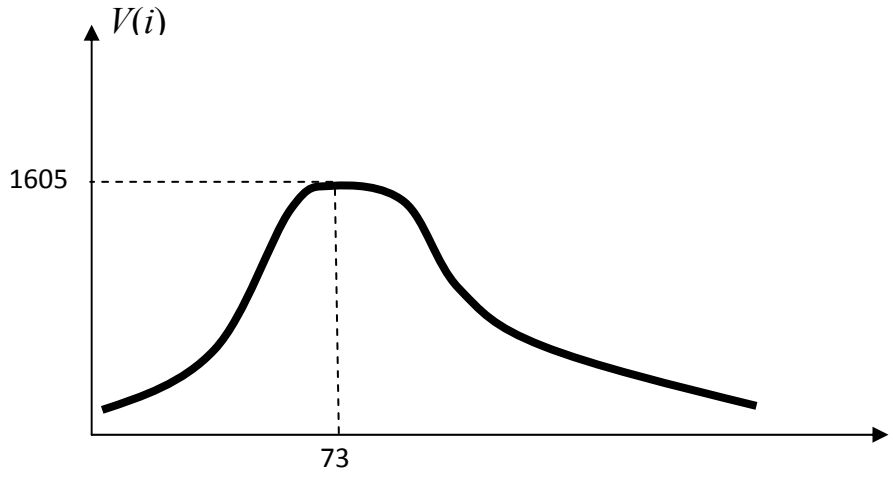
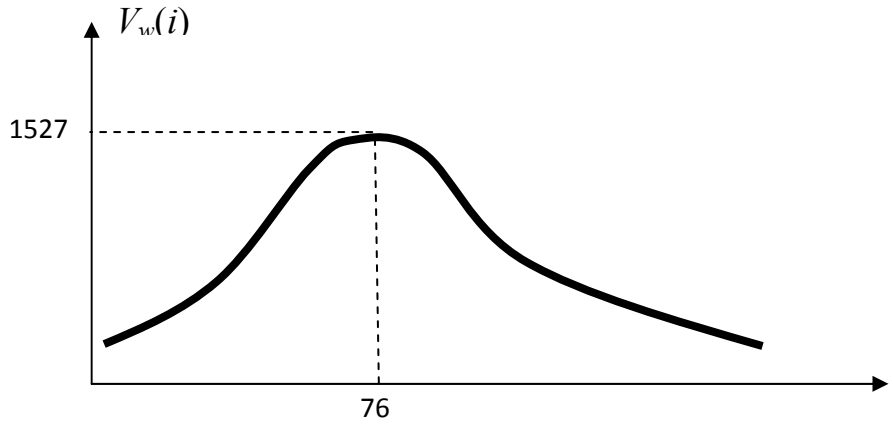


Рисунок 5.2а – гистограмма ПО,
б – гистограмма СО

Однако, даже визуально не всегда можно обнаружить раздвоение пиков. Гистограмма другого исследуемого изображения до и после вложения информации по методу СГ-ШПС при $\alpha = 3$ приведена на рисунке 5.3, и на данном изображении не просматривается раздвоение пиков.



а



б

Рисунок 5.3–а – гистограмма ПО,
б – гистограмма СО

Для более четкого выделения раздвоенных пиков в некоторых работах [39] предлагается обработка изображения при помощи фильтрации. Однако, автоматизация поиска раздвоенных пиков представляется достаточно сложной задачей, и, следовательно, данная атака требует участия человека оператора.

Можно, например, задать критерий обнаружения СГ-ШПС по количеству раздвоенных пиков:

$$\begin{cases} \gamma > \gamma_0 & \text{- принимаем гипотезу } H_1, \\ \gamma \le \gamma_0 & \text{- принимаем гипотезу } H_0, \end{cases}$$

где $\gamma_0 > 0$ – некоторое пороговое значение,

$$\gamma = \#\left\{i \in (1..255) : \left(V(i) < V(i-\alpha) + \Delta\right) \cup \left(V(i) < V(i+\alpha) + \Delta\right)\right\},$$

где $\Delta > 0$ – некоторое пороговое значение.

Выбор порогов γ_0 и Δ для различных изображений затруднителен. Значения вероятности пропуска P_m и вероятности ложной тревоги P_{fa} будут большими, поэтому вряд ли можно ожидать хорошей надежности данной атаки. В следующих параграфах будут подробно исследованы более эффективные методы СГА для СГ-ШПС.

5.2 Стегоанализ, основанный на корреляции яркостей смежных пикселей

Общее свойство СГ-ШПС состоит в дополнительной «рандомизации» изображения за счет вложения. Это свойство проявляется в уменьшении корреляции яркостей смежных пикселей.

При вложении по методу СГ-ШПС (1.3) появляется возможность обнаружить факт вложения скрытой информации, если справедлив принцип Кирхгоффа, когда все параметры СГС, кроме стегоключа, известны атакующему.

Непосредственно из (1.3) следует, что при $Var\left(\pi(n)\right) = 1$ будет справедливо неравенство:

$$Var\left(C_w(n)\right) = \sigma_C^2 + \alpha^2 > \sigma_C^2, \tag{5.3}$$

где $\sigma_C^2 = Var\left(C(n)\right)$ – дисперсия $C(n)$.

Тогда атака по обнаружению СГ-ШПС принимает вид:

$$\begin{cases} \Lambda > \lambda \text{ - принимаем гипотезу } H_1, \\ \Lambda \le \lambda \text{ - принимаем гипотезу } H_0, \end{cases}, \tag{5.4}$$

где $\Lambda = \frac{1}{N}\sum_{n=1}^{N} C_w^2(n) - \left(\frac{1}{N}\sum_{n=1}^{N} C_w(n)\right)^2$,

N – общее количество пикселей $C_w(n)$,

λ – некоторое пороговое значение.

Для исключения тривиальной атаки (5.4) модифицируем метод вложения (1.3). Произведем вложение по следующему правилу:

$$C_w(n) = \beta C(n) + \alpha(-1)^b \pi(n),\ n = 1, 2..N, \qquad (5.5)$$

где $\beta = \sqrt{1 - \frac{\alpha^2}{\sigma_C^2}}$.

Легко проверить, что:

$$Var\left(C_w(n)\right) = \beta^2\sigma_C^2 + \alpha^2 = \sigma_C^2 - \alpha^2 + \alpha^2 = \sigma_C^2 = Var\left(C(n)\right).$$

Тогда атака (5.4) становиться невозможной.

Для модифицированного по (5.5) СО рассмотрим следующую атаку обнаружения:

$$\begin{cases} \Gamma > \lambda \text{ - принимаем гипотезу } H_1, \\ \Gamma \le \lambda \text{ - принимаем гипотезу } H_0, \end{cases} \qquad (5.6)$$

где

$$\Gamma = \frac{1}{2N\widetilde{\sigma}_C^2}\sum_{n=1}^{N}\left(C_w(n+1) - C_w(n)\right)^2, \qquad (5.7)$$

где $\widetilde{\sigma}_C^2 = \frac{1}{N}\sum_{n=1}^{N} C_w^2(n)$.

Для подтверждения возможности атаки по (5.6) найдем математическое ожидание $E(\Gamma)$ для случая, когда $C_w(n) \in$ СО и $C_w(n) \in$ ПО.

Из выражения (5.7) получаем для случая $C_w(n) \in$ ПО(если ПО$= C(n)$ является стационарным случайным процессом):

$$E\{\Gamma\} = \frac{N}{2N\widetilde{\sigma}_C^2} E\left\{C^2(n+1) + C^2(n) - 2C(n)C(n+1)\right\} = 1 - R_C(n, n+1) \qquad (5.8)$$

где $R_C(n, n+1) = \frac{E\{C(n)C(n+1)\}}{\widetilde{\sigma}_C^2}$ – коэффициент корреляции яркостей смежных пикселей ПО.

Для случая, когда $C_w(n) \in$ СОс использованием правила вложения (5.5) получаем в предположении, что $\pi(n) \in \pm 1$:

$$E\{\widetilde{\Gamma}\}=\frac{N}{2N\widetilde{\sigma}_C^2}E\left\{\left[\beta C(n+1)+\alpha(-1)^b\pi(n+1)-\beta C(n)+\alpha(-1)^{\tilde{b}}\pi(n)\right]^2\right\}=$$

$$\frac{1}{2\widetilde{\sigma}_C^2}E\left\{\beta^2C^2(n+1)+2\beta C(n+1)\alpha(-1)^b\pi(n+1)+\left(\alpha(-1)^{\tilde{b}}\right)^2+\beta^2C^2(n)+\right.$$

$$+2\beta C(n)\alpha(-1)^{\tilde{b}}\pi(n+1)+\left(\alpha(-1)^b\right)^2-2\beta^2C(n+1)C(n)-$$

$$-2\beta C(n+1)\alpha(-1)^{\tilde{b}}\pi(n)-2\beta C(n)\alpha(-1)^b\pi(n+1)$$

$$\left.-2\alpha(-1)^b\alpha(-1)^{\tilde{b}}\pi(n+1)\pi(n)\right\} \qquad (5.9)$$

По предположения $\pi(n)\in$i.i.d, и значит, будет справедливо равенство:

$$E\{\pi(n)\}=E\{\pi(n+1)\}=E\{\pi(n)\pi(n+1)\}=0. \qquad (5.10)$$

Подставляя (5.10) в (5.9) получаем:

$$E\{\widetilde{\Gamma}\}=\frac{1}{2\widetilde{\sigma}_C^2}E\left\{\beta^2C^2(n+1)+\left(\alpha(-1)^b\right)^2-2\beta^2C(n+1)C(n)\right\}=$$

$$=\beta^2+\frac{\left(\alpha(-1)^b\right)^2}{\widetilde{\sigma}_C^2}-\beta^2R_C(n,n+1)=1-\beta^2R_C(n,n+1). \qquad (5.11)$$

Из сравнение выражений (5.8) и (5.11) видно, что поскольку для разумных СО $\beta^2<1$, то $E\{\widetilde{\Gamma}\}>E\{\Gamma\}$ и правило обнаружения СО (5.6) является достаточно обоснованным. Следует отметить, что при выводе соотношения (5.8) и (5.11) предполагается стационарность ПО как случайного процесса, то есть выполнение условий:

$$\begin{cases}E\{C^2(n)\}=\text{in var}(n),\ n=1,2..N,\\ R_C(n,n+1)=\text{in var}(n),\ n=1,2..N.\end{cases} \qquad (5.12)$$

Для реальных ПОв виде изображений условия (5.12) выполняться, как правило, не будет, и поэтому можно ожидать трудностей при выборе порога λ при обнаружении СГ-ШПС по правилу (5.6) для некоторых видов изображений.

Типичное значение $\widetilde{\sigma}_C^2$ для покрывающего сообщения в виде изображения является величина порядка 20000. Тогда даже при $\alpha=5$ по (5.5) получаем $\beta=0{,}999$, следовательно, разница между $E\{\Gamma\}$ и $E\{\widetilde{\Gamma}\}$ оказывается небольшой,

что приводит к сложностям обнаружения СО, созданного методом СГ-ШПС. ПО и вкладываемая информация имеют цифровой вид, поэтому при округление малого β получаем $\beta = 1$.

Тогда выражение (5.5) принимает вид:

$C_w(n) = C(n) + \alpha(-1)^b \pi(n)$.

Проводя выкладки, аналогичные доказательству (5.11), получаем для вложения по правилу (1.3):

$$E\{\widetilde{\Gamma}\} = 1 + \frac{\alpha^2}{\sigma_C^2} - R_C(n, n+1).$$

Однако, неравенство $E\{\widetilde{\Gamma}\} > E\{\Gamma\}$ сохраняется, и обнаружение по правилу (5.6) остается возможным.

В таблице 5.1 приводятся результаты экспериментальных вычислений величины Γ для 10 различных изображений с глубиной вложения $\alpha = 1$ и долями вложения $P = 1$.

Таблица 5.1 – Экспериментальные результаты вычислений величины Γ

№ изображения	Γ	$\widetilde{\Gamma}$
1	0,004394	0,004433
2	0,039907	0,040459
3	0,021924	0,022316
4	0,021864	0,022040
5	0,105029	0,141615
6	0,042248	0,042728
7	0,033281	0,033327
8	0,013435	0,013484
9	0,004362	0,004406
10	0,058273	0,059046

Как видно из таблицы 5.1, значения Γ и $\widetilde{\Gamma}$ очень близки приразличных α и P, при этом Γ различных изображений отличаются на большую величину, чем

Γ и $\widetilde{\Gamma}$ одного и того же изображения. Следовательно, выбор какого-либо порога невозможен и применение данного метода на практике представляется затруднительным.

5.3 Стегоанализ, основанный на метод подсчета нулей в гистограмме

Данная атака[28] является гистограммной и основывается на правдоподобном утверждении:

Количество нулей в гистограмме СО при вложениимметодом СГ-ШПС всегда меньше, чем количество нулей в гистограмме соответствующего ему ПО[28].

Если в гистограмме ПО имеется ноль при некотором $i = i_0$, то есть $V(i_0) = 0$, то в гистограмме СО $V_w(i_0) \neq 0$, если $V(i_0 - \alpha) \neq 0$ или $V(i_0 + \alpha) \neq 0$. Следовательно, по аномально малому количеству нулей гистограммы можно обнаружить факт присутствия вложения.

Критерий обнаружения СГ-ШПС имеет следующий вид::

$$\begin{cases} K < K_0 \text{ - принимаем гипотезу } H_1, \\ K \geq K_0 \text{ - принимаем гипотезу } H_0, \text{'} \end{cases} \qquad (5.13)$$

где K –количество нулей в гистограмме исследуемого изображения,

K_0 – некоторое пороговое значение.

Экспериментальные результаты по подсчету количества нулей гистограммы для 5 различных изображений с глубиной вложения $\alpha = 1; 2; 3$ и вероятностью вложения $P = 1; 0{,}5; 0{,}1; 0$ приведены в таблице 5.2.

Таблица 5.2 – Результаты подсчета количества нулей гистограммы для 5 различных изображений

α	P	**№ изображения**				
		1	**2**	**3**	**4**	**5**
1	**1**	165	95	154	138	161
	0,5	139	5	128	93	127
	0,1	141	6	134	100	131
2	**1**	153	68	133	131	157
	0,5	134	36	120	109	130
	0,1	138	36	127	120	136
3	**1**	165	95	154	138	161
	0,5	139	5	127	93	127
	0,1	140	5	132	95	132
0		201	162	193	174	200

Как показали эксперименты, наиболее эффективна данная атака при доле вложения $P = 0{,}5$. Стоит отметить, что существуют изображения с весьма малым количеством нулей гистограммы, даже если в них нет вложенной информации. Применение данной атаки для таких изображений представляется проблематичным.

В таблице 5.3 приведены значения вероятностей ложной тревоги P_{fa} и пропуска P_m для данного метода СГА при анализе 100 различных изображений и использовании различных пороговых значениях K_0 .

Таблица 5.3 – Значения P_{fa} и P_m для СГА, основанного на методе подсчета нулей гистограммы

K_0			**200**	**170**	**160**	**140**	**120**	**100**	**70**
P_{fa}			0,85	0,75	0,25	0,18	0,14	0,12	0,03
$\alpha = 1$	$P = 1$	P_m	0,06	0,12	0,14	0,19	0,27	0,54	0,87
	$P = 0,5$	P_m	0,06	0,06	0,06	0,12	0,18	0,22	0,30
	$P = 0,1$	P_m	0,06	0,06	0,06	0,13	0,20	0,25	0,32
$\alpha = 2$	$P = 1$	P_m	0,06	0,11	0,12	0,15	0,24	0,30	0,53
	$P = 0,5$	P_m	0,06	0,06	0,06	0,12	0,16	0,24	0,31
	$P = 0,1$	P_m	0,06	0,06	0,06	0,13	0,21	0,26	0,34
$\alpha = 3$	$P = 1$	P_m	0,06	0,11	0,12	0,18	0,27	0,52	0,88
	$P = 0,5$	P_m	0,06	0,07	0,08	0,16	0,20	0,46	0,83
	$P = 0,1$	P_m	0,07	0,09	0,10	0,16	0,23	0,58	0,84

Как видно из таблицы 5.3, P_{fa} уменьшается с уменьшением K_0. Наименьшее значение P_m при $\alpha = \text{const}$ и $K_0 = \text{const}$ достигается при $P = 0,5$. Представим более подробно зависимость P_m от α, P и K_0.

При $\alpha = 1$ и $70 \le K_0 \le 200$ – $P_m(P = 0,5) \le P_m(P = 0,1) \le P_m(P = 1)$.

При $\alpha = 2$ и $70 \le K_0 \le 200$ – $P_m(P = 0,5) \le P_m(P = 0,1) \le P_m(P = 1)$.

При $\alpha = 3$ и $120 \le K_0 \le 200$ и $K_0 = 70$ – $P_m(P = 0,5) \le P_m(P = 0,1) \le P_m(P = 1)$.

Только при $\alpha = 3$ и $K_0 = 100$ – $P_m(P = 0,5) \le P_m(P = 1) \le P_m(P = 0,1)$.

Как видно из таблицы 5.3, наименьшее P_m достигается при $P = 0,5$. Также стоит отметить, что почти всегда $P_m(P = 0,1) \le P_m(P = 1)$. Следовательно, данный СГА лучше всего определяет наличие вложения при P, близкой к 0,5. При $K_0 = 140$ вероятность ложной тревоги $P_{fa} = 0,18$, а вероятность пропуска $P_m(0,1 \le P \le 1) < 0,2$. Следовательно, при $0,1 \le P \le 1$ и пороговом значении

$K_0 = 140$ данныйСГА позволяет определить наличие или отсутствие скрытой информации в изображении с вероятностью более 0,8.

5.4 Стегоанализ, основанный на сравнении соседних значений гистограммы

Еще одна атака на СГ-ШПС похожа на атаку (4.10) для СГ-НЗБ. ДанныйСГА так же основан на разнице уровней отсчетов гистограмм[28], но не попарно, а всех соседних отчетов подряд:

$$R = \frac{\sum_{j=0}^{L-2}\left(V(j) - V(j+1)\right)^2 + V^2(0) + V^2(L-1)}{2\sigma_V^2}, \qquad (5.14)$$

где $\sigma_V^2 = \sum_{j=0}^{L-1} V^2(j)$ – дисперсия,

L – количество отсчетов гистограммы.

При вложении методом СГ-ШПС гистограмма «сглаживается», уровни отсчетов «выравниваются». Следовательно, вообще говоря:

$$V_w(j) - V_w(j+1) < V(j) - V(j+1), \qquad (5.15)$$

где $V(j)$ и $V(j+1)$ – уровни отчетов j и $j+1$ гистограммы ПО,

$V_w(j)$ и $V_w(j+1)$ – уровни отчетов j и $j+1$ гистограммы СО.

Из (5.13) следует, что $R_w < R$, тогда критерий обнаружения СГ-ШПС принимает вид:

$$\begin{cases} R^{/} < R_0^{/} \text{ - принимаем гипотезу } H_1, \\ R^{/} \geq R_0^{/} \text{ - принимаем гипотезу } H_0, \end{cases} \qquad (5.16)$$

где $R_0^{/}$ – некоторое пороговое значение,

$R^{/}$ – значение R по (5.12) для тестируемого изображения.

Экспериментальные результаты подсчета R, основанного на сравнении соседних значений гистограммы, для 5 различных изображений с глубиной вложения $\alpha = 1; 2; 3$ и долей вложения $P = 1; 0{,}5; 0{,}1; 0$ приведены в таблице 5.4.

Таблица 5.4 – Расчет величины R по формуле (5.14)

α	P	**№ изображения**				
		1	**2**	**3**	**4**	**5**
1	**1**	0,93559	0,75254	0,78694	0,80936	0,98685
	0,5	0,29622	0,11062	0,34065	0,26763	0,33946
	0,1	0,86273	0,85146	0,78603	0,08442	0,88791
2	**1**	0,94427	0,34849	0,89075	0,81062	0,99248
	0,5	0,91498	0,54254	0,86002	0,71135	0,98853
	0,1	0,97022	0,92904	0,88683	0,92510	0,99833
3	**1**	0,90780	0,76303	0,89106	0,77811	0,99437
	0,5	0,91876	0,83708	0,81983	0,86573	0,99317
	0,1	0,97336	0,96783	0,87973	0,94819	0,99899
0		0,97978	0,99999	0,89268	0,96016	0,99981

Исследования показали, что наиболее эффективна данный СГА при доле вложения $P = 0{,}5$. Для расчета эффективности данного метода СГА проанализируем вероятности ложной тревоги P_{fa} и пропуска P_m для 100 ПО и полученных из них СО при различных пороговых значениях $R_0^{/}$. Результаты эксперимента приведены в таблице 5.5.

Таблица 5.5 – Значения P_{fa} и P_m для СГА, основанного на методе сравнения соседних значений гистограммы

$R_0^/$			**0,9999**	**0,97**	**0,93**	**0,90**	**0,88**	**0,85**	**0,80**
P_{fa}			0,39	0,31	0,21	0,18	0,15	0,10	0,03
$\alpha = 1$	$P = 1$	P_m	0,02	0,10	0,15	0,18	0,20	0,24	0,35
	$P = 0,5$	P_m	0	0,01	0,01	0,01	0,01	0,01	0,01
	$P = 0,1$	P_m	0	0	0	0,02	0,14	0,65	0,78
$\alpha = 2$	$P = 1$	P_m	0,05	0,10	0,18	0,20	0,23	0,25	0,45
	$P = 0,5$	P_m	0,05	0,11	0,17	0,20	0,20	0,23	0,26
	$P = 0,1$	P_m	0,05	0,16	0,55	0,78	0,84	0,86	0,94
$\alpha = 3$	$P = 1$	P_m	0,05	0,14	0,19	0,23	0,24	0,30	0,44
	$P = 0,5$	P_m	0,05	0,12	0,15	0,21	0,31	0,47	0,59
	$P = 0,1$	P_m	0,05	0,25	0,75	0,78	0,80	0,86	0,95

Как видно из таблицы 5.5, при уменьшении $R_0^/$ уменьшается P_{fa}, а P_m увеличивается. Наибольший интерес при рассмотрении статистики, приведенной в таблице 5.5, представляет зависимость P_m от P при $\alpha = \text{const}$ и $R_0^/ = \text{const}$. Наименьшая P_m в большинстве случаев достигается при $P = 0,5$. Рассмотрим результаты таблицы 5.5 более подробно:

При $\alpha = 1$ и $0,93 \le R_0^/ \le 0,9999$ – $P_m(P = 0,1) \le P_m(P = 0,5) \le P_m(P = 1)$.

При $\alpha = 1$ и $0,90 \le R_0^/ \le 0,88$ – $P_m(P = 0,5) \le P_m(P = 0,1) \le P_m(P = 1)$.

При $\alpha = 1$ и $0,80 \le R_0^/ \le 0,85$ – $P_m(P = 0,5) \le P_m(P = 1) \le P_m(P = 0,1)$.

При $\alpha = 2$ и $0,97 \le R_0^/ \le 0,9999$ – $P_m(P = 1) \le P_m(P = 0,5) \le P_m(P = 0,1)$.

При $\alpha = 2$ и $0,80 \le R_0^/ \le 0,93$ – $P_m(P = 0,5) \le P_m(P = 1) \le P_m(P = 0,1)$.

При $\alpha = 3$ и $0,90 \le R_0^/ \le 0,9999$ – $P_m(P = 0,5) \le P_m(P = 1) \le P_m(P = 0,1)$.

При $\alpha = 3$ и $0,80 \le R_0^/ \le 0,88$ – $P_m(P = 1) \le P_m(P = 0,5) \le P_m(P = 0,1)$.

К сожалению, найти однозначную зависимость P_m от α или от P не представляется возможным. Из таблицы 5.5 видно, что данный метод СГА имеет большие величины P_{fa} и P_m. Стоит отметить, что при вложении методом СГ-ШПС, как правило, не выбирают $P > 0{,}9$, поскольку в этом случае даже при $\alpha = 1$ исходное изображение заметно теряет в качестве, и $P < 0{,}3$, поскольку в данном случае объем вкладываемой информации очень мал, следовательно, мала и скорость передачи информации.Данный СГА эффективен при анализе СО с наиболее часто встречаемыми P, а именно, как видно из таблицы 5.5, при $0{,}90 \le R_0^{/} \le 0{,}93$ и $P \ge 0{,}5$ (а для $\alpha = 1$ при $P \ge 0{,}1$) вероятность ложной тревоги $P_{fa} \le 0{,}21$ и вероятность пропуска $P_m \le 0{,}23$. При использовании порогового значений $0{,}90 \le R_0^{/} \le 0{,}93$ данный СГА позволяет определить наличие или отсутствие скрытой информации в изображении с вероятность более 0,80.

Отметим, что в литературе встречаются и другие методы СГА, для СГ-ШПС. Например, в [15] описан метод с использованием двумерного преобразования Фурье от матрицы смежности пикселей. Но данный метод не достаточно надежный, так при вероятности пропуска равной 0,5 вероятность ложной тревоги равна 0,37 для глубины вложения 1 вероятность, 0,31 для глубины вложения 2 и 0,32 для глубины вложения 3.

5.5 Выводы

Как видно из параграфов 5.1 и 5.2, методы СГА, основанные на уже известных алгоритмах анализа, таких как выявление раздвоенных пиков и корреляция смежных пикселей не достаточно эффективны для выявления скрытой информации, вложенной методом СГ-ШПС. Основная проблема, выявленная при анализе результатов моделирования данных атак, – это невозможность выбора универсального порогового значения. Метод СГА, основанный на корреляции яркостей смежных пикселей, может применяться

лишь для сравнения пар изображений, визуально похожих, и выяснения, какое из этих изображений содержит скрытую информацию.

Предложенные автором диссертации методы СГА, основанные на сравнении соседних пикселей (параграф 5.4) и на подсчете нулей гистограммы (параграф 5.3) обладают примерно одинаковой эффективностью. Однако стоит отметить, что эффективность первой атаки немного выше, чем у второй (поскольку P_{fa} у первой атаки ниже, чем у второй), но вторая дает более стабильные результаты при различных долях и различной глубине вложения. Обе атаки дают наилучшие результаты при $P = 0{,}5$.

При анализе изображения рекомендуется вначале применить СГА, основанный на подсчете разности яркостей соседних пикселей. Если данный СГА дал отрицательный результат (вложения нет), то стоит применить СГА, основанный на подсчете нулей гистограммы.

Отметим, что СГ-ШПС является более необнаруживаемая СГС (более секретная СГС), чем СГ-НЗБ. При ограничении искажений, вызванных вложением секретной информации, она становиться абсолютно секретной, но при этом уменьшается скорость передачи информации [13]. Отметим, что как было показано в параграфе 2.2, при малой корреляции между яркостями пикселей в ПО обнаружение вложения СГ-ШПС даже оптимальным методом СГА становится невозможным.

Разработанные автором методы СГА – метод,основанный на сравнении соседних пикселей, и метод, основанный на подсчете нулей гистограммы, – используются в лабораторной работе «Методы обнаружения стегосистем НЗБ и ШПС» в курсе «Основы стеганографии» на кафедре «Защищенные системы связи».

6 Комплексные методы стегоанализа

В главах 4 и 5 были рассмотрены методы СГА, разработанные специально для СГ-НЗБ и СГ-ШПС соответственно.

Как правило, стегоаналитику заранее неизвестно, какой именно метод СГ был применен для создания СО, как следствие, невозможно подобрать максимально эффективным метод целевого СГА для конкретного объекта. Наиболее эффективное решение этой проблемы – создание единого метода СГА, позволяющего обнаруживать вложения как СГ-НЗБ, так и СГ-ШПС.

В данной главе будут приведены экспериментальные результаты описанных выше методов, примененных к «не своим» методам вложения, для проверки их эффективности при исследовании «незнакомого» алгоритма СГ. К СГ-ШПС будут применены методы стегоанализа 1-ого и 2-ого порядка, а так же визуальный метод стегоанализа. К СГ-НЗБ – методы, основанные на подсчете нулей гистограммы и на сравнениизначений соседних отсчетов гистограммы.

Однако, как будет показано ниже, эти методы СГА имеют свои «слабости». Поэтому, помимо уже рассмотренных в главах 4 и 5 методов СГА будет рассмотрен еще один метод – метод локальных максимумов. Также будет рассмотрен алгоритм комбинирования ранее рассмотренных методов и метода локальных максимумов, что позволяет достаточно надежно определить наличие вложения в изображения с различными статистическими характеристиками при использовании разных алгоритмов СГ.

6.1 Визуальная атака, применительно с СГ-ШПС

Алгоритм данного метода СГА подробно описан в 4.1

Как было показано ранее, при отсутствии вложения на полученном после атаки изображении будут просматриваться контуры первоначального изображения. Если в исследуемом изображении присутствует скрытое вложение, то в зависимости от четности или нечетности значения яркости пикселя

исследуемого изображения и четности или нечетности величины α получим значения НЗБ после вложения в соответствии с таблицей 6.1.

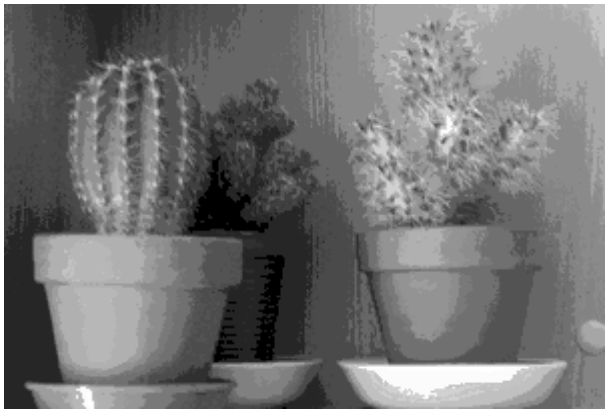
Таблица 6.1 - Значение НЗБ после вложения

		$C(n)$	
		четный	**нечетный**
α	**четный**	0 (черный)	1 (белый)
	нечетный	1 (белый)	0 (черный)

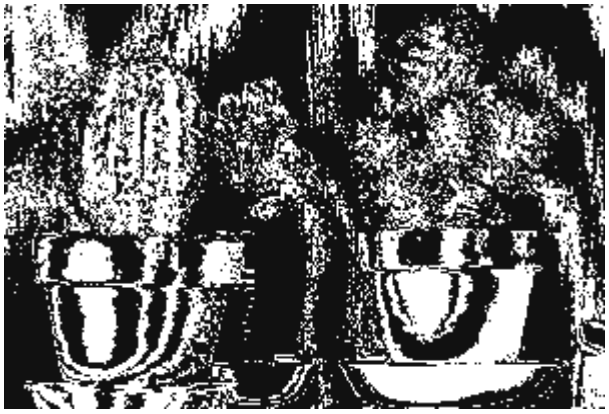
Как видно из таблицы 6.1, на изображениях, полученных после визуальной атаки на СО, при вложении методом СГ-ШПС с нечетной α при $P = 1$ (вложение в каждый пиксель) происходит инвертирование черного и белого цветов (по сравнению с изображениями, полученными после визуальной атаки на ПО). При нечетной α и $P = 0{,}5$ в изображении после атаки пиксели, в которых происходит инвертирование цветов, чередуются с пикселями, вложение в которых не производится. Следовательно, только половина пикселей меняет цвет по сравнению с изображением после атаки на ПО. Это приводит к появлению на изображении после визуальной атаки шумового поля. Если при нечетной α начать уменьшать долю вложения ($P < 0{,}5$), то на изображениях, полученных после визуальной атаки, будут проявляться контуры исходного изображения. При этом, чем меньше P, тем отчетливей будут контуры.

При вложении методом СГ-ШПС с четной α НЗБ пикселей, подвергшихся вложению, не поменяют своих значений. Следовательно, изображение, полученное после визуальной атаки на СО, при любом P будет точно такое же, как изображение, полученное после визуальной атаки на соответствующее ПО.

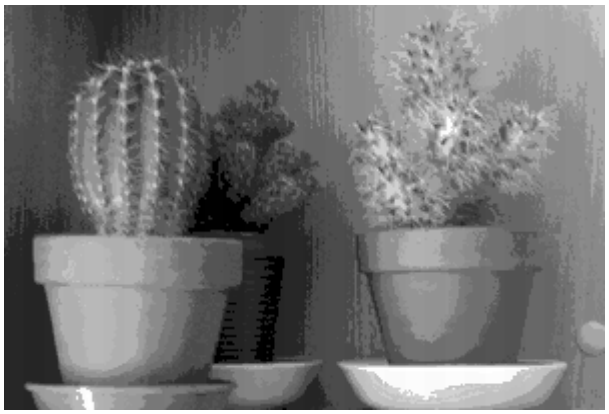
Данные свойства подтверждаются результатами экспериментов, представленных на рисунках 6.1-6.3.



а

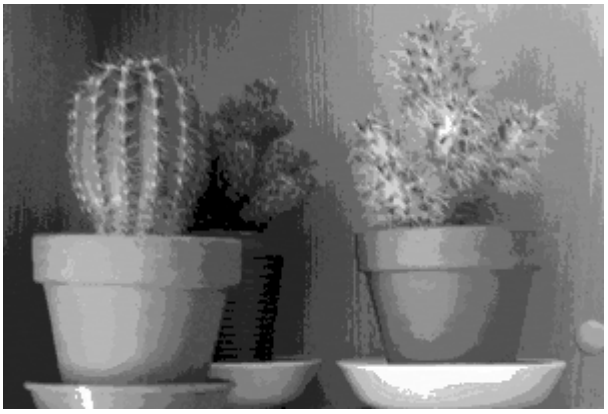


б

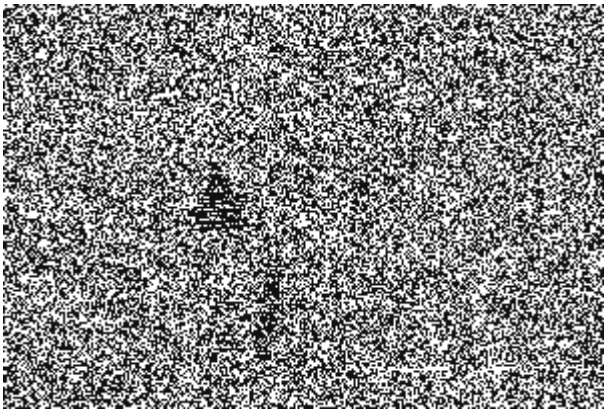


в

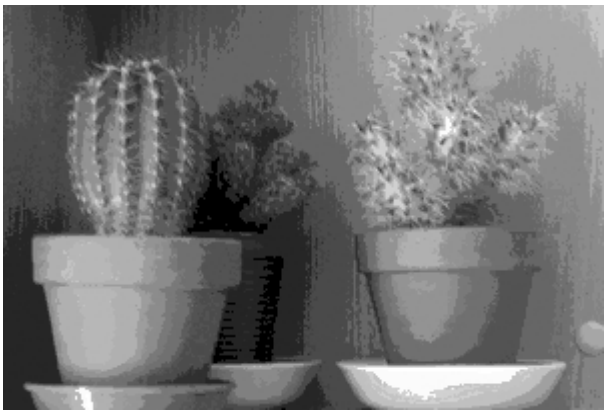
г



д



е



ж

з

Рисунок 6.1 – а, в, д, ж – изображение ПО и СО с вложением α=1 и $P = 1$, $P = 0{,}5$, $P = 0{,}1$ соответственно, б, г, е, з – изображение ПО и СО с вложением α=1 и $P = 1$, $P = 0{,}5$, $P = 0{,}1$ после атаки соответственно

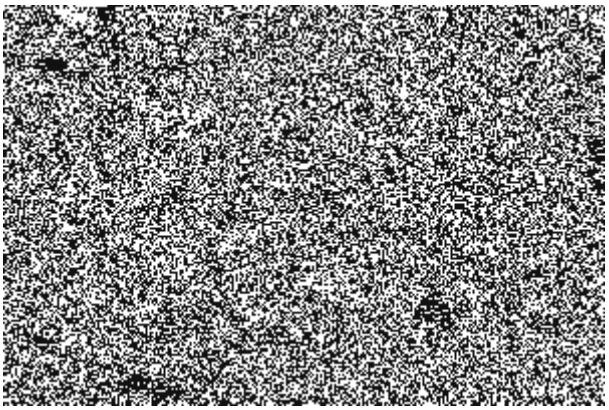
а

б

в

г

д

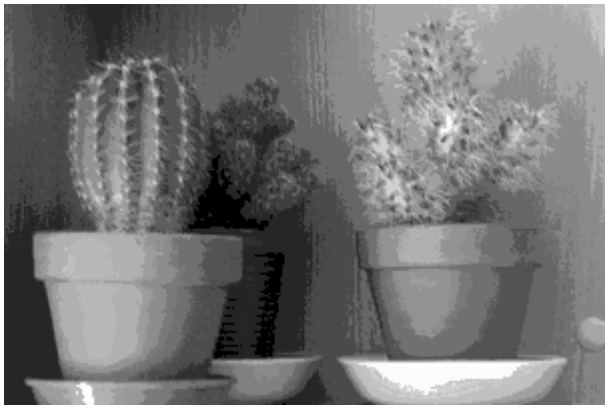


е

ж

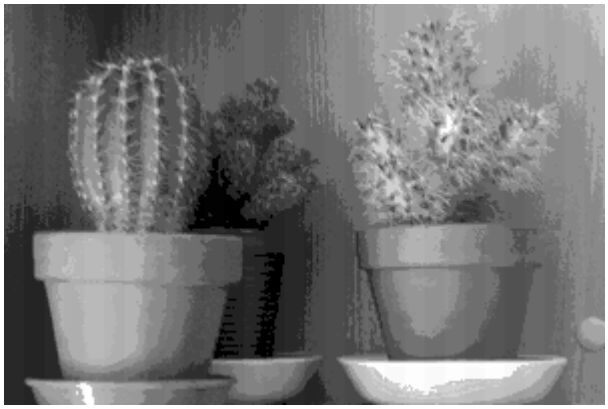
з

Рисунок 6.2 – а, в, д, ж – изображение ПО и СО с вложением α=1 и $P = 1$, $P = 0{,}5$, $P = 0{,}5$ соответственно, б, г, е, з – изображение ПО и СО с вложением α=1 и $P = 1$, $P = 0{,}5$, $P = 0{,}5$ после атаки соответственно



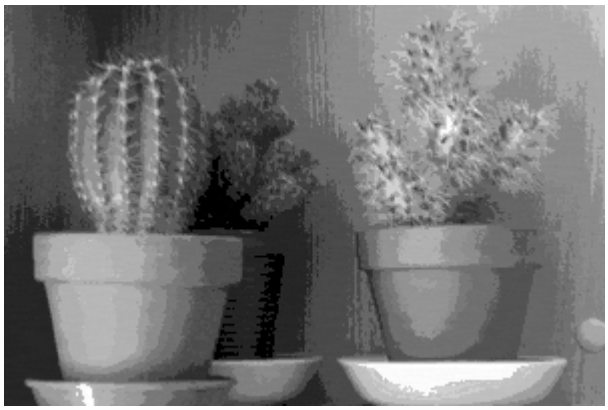
а

б



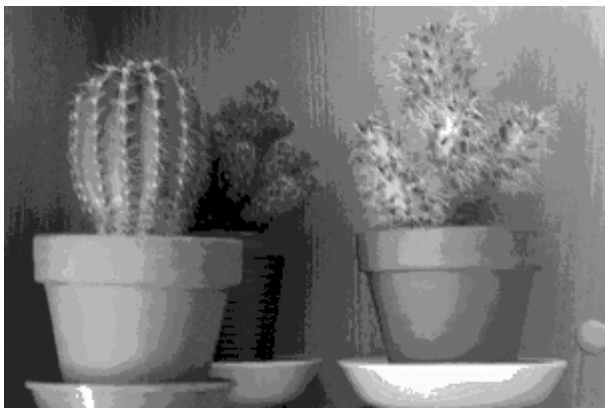
в

г

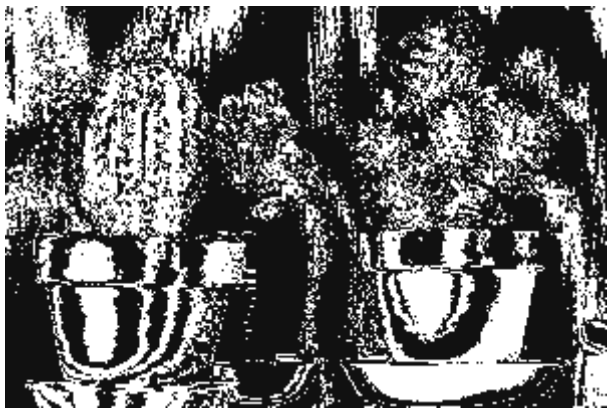


д

е



ж



з

Рисунок 6.3 – а, в, д, ж – изображение ПО и СО с вложением α =2 и $P = 1$, $P = 0{,}5$, $P = 0{,}5$ соответственно, б, г, е, з – изображение ПО и СО с вложением α =2 и $P = 1$, $P = 0{,}5$, $P = 0{,}5$ после атака соответственно

Как видно из результатов экспериментов, эффективность визуальной атаки не убывает монотонно, как это происходило при вложении методом СГ-НЗБ. При методе вложения СГ-ШПС данная атака наиболее эффективна при $P = 0{,}5$ и наименее эффективна при $P = 1$ и $P < 0{,}1$. Стоит отметить, что приведенные выше выводы оценки эффективности данной атаки справедливы только при нечетной α. Если же α четная, то визуальная атака становиться абсолютно неэффективной при любой P. По сравнению с уже рассмотренными в главе 5 методами СГА, метод визуальной атаки значительно менее эффективен и неудобен в использовании, поскольку, как было сказано в параграфе 4.1, при его использовании невозможно установить пороговое значение, следовательно, требуется постоянное присутствие оператора.

6.2 Стегоанализ, основанный на статистике 1-ого порядка, применительно к СГ-ШПС

СГА, основанный на статистике 1-ого порядка, подробно описан в параграфе 4.2.

Если в вкладываемой последовательности символы 0 и 1 распределены равномерно, то для СГ-ШПС будет справедливо равенство:

$$V_w(i) = \frac{1}{2}\left(V(i-\alpha) + V(i+\alpha)\right),\ i = 1, 2..L, \qquad (6.1)$$

где $V(i)$ – гистограмма ПО,

$V_w(i)$ – гистограммаСО, созданного методом СГ-ШПС.

Следовательно, для СГ-ШПС, даже при $P = 1$, не выполняется равенство $V_w(2j) = V_w(2j+1)$, $j = 1, 2..\frac{(L-1)}{2}$,которое справедливо для СГ-НЗБ при $P = 1$. Следовательно, применение метода СГА, основанного на статистике 1-ого порядка для СГ-ШПС оказывается необоснованным. Однако, при вложении методом СГ-ШПС происходит перераспределение уровней гистограммы, и некоторое «выравнивание» значений соседних отсчетов. При этом вид

гистограммы после вложения методом СГ-ШПС при $P = 0{,}5$ приближается к виду гистограммы после вложения методом СГ-НЗБ при $P = 1$ (рисунок 4.10 б).

В таблице 6.2 приведены значений χ^2 для пяти различных изображений с глубиной вложения $\alpha = 1; 2; 3$ при различных долях вложения $P = 1; 0{,}5; 0{,}1; 0$.

Таблица 6.2 – Значение χ^2 при различных долях вложения

α	**№изображения**	P	χ^2
1	1	1	51248
		0,5	15885
		0,1	45781
		0	58211
	2	1	40026
		0,5	3629
		0,1	43308
		0	60000
	3	1	41215
		0,5	14170
		0,1	43152
		0	56936
	4	1	48827
		0,5	15048
		0,1	46793
		0	59589
	5	1	54579
		0,5	17852
		0,1	47715
		0	59928

Продолжение таблицы 6.2

α	**№изображения**	P	χ^2
2	1	1	54824
		0,5	47791
		0,1	54739
		0	58211
	2	1	20386
		0,5	22046
		0,1	50453
		0	60000
	3	1	56609
		0,5	48792
		0,1	54379
		0	56936
	4	1	51950
		0,5	38243
		0,1	52950
		0	59589
	5	1	58085
		0,5	55079
		0,1	57540
		0	59928

Продолжение таблицы 6.2

α	**№изображения**	P	χ^2
3	1	1	49463
		0,5	50267
		0,1	55932
		0	58211
	2	1	39683
		0,5	44964
		0,1	55701
		0	60000
	3	1	47860
		0,5	45163
		0,1	53000
		0	56936
	4	1	39539
		0,5	43668
		0,1	55149
		0	59589
	5	1	56996
		0,5	56408
		0,1	58899
		0	59928

Значения вероятности ложной тревоги P_{fa} и вероятности пропуска P_m для данной атаки при пороговом значении $\chi_0^2 = 54000$ приведены в таблице 6.3

Таблица 6.3 – Экспериментальный расчет величин P_{fa} и P_m

α	P	P_{fa}	P_m
1	1	-	0,08
	0,5	-	0,02
	0,1	-	0,03
	0	0,15	-
2	1	-	0,17
	0,5	-	0,09
	0,1	-	0,2
	0	0,15	-
3	1	-	0,10
	0,5	-	0,07
	0,1	-	0,3
	0	0,15	-

Как видно из таблиц 6.2 и 6.3, при вложении методом СГ-ШПС критерий χ^2 лучше всего обнаруживает вложения в изображениях высокого качества, без цифрового шума, без повышенной зернистости. Для изображений с сильным шумом или высокой зернистостью наличие или отсутствие вложения по данному методу выявляется плохо.

Данная атака может применяться для выявления вложений,сделанных методом СГ-ШПС. Как показали исследования, наилучшее обнаружение происходит при $P = 0{,}5$.Стоит отметить, что при нечетных α результат атаки значительно лучше, чем при четных, а с увеличением α вероятность пропуска P_m увеличивается. Данное свойство СГ-ШПС может быть использовано при разработке СГС, устойчивых к СГА.

6.3 Стегоанализ, основанный на статистике 2-ого порядка, применительно к СГ-ШПС

СГА, основанный на статистике 2-ого порядка подробно описан в параграфе4.3. Как показывает анализ вывода уравнения (4.14), приведенного в Приложение А), данное уравнение не может быть распространено на СГ-ШПС. В Приложении А описаны различные множества, состоящие из разных пар пикселей. Переход из одного множества в другое происходит посредством шаблонов {00;01;10;11}. Данные шаблоны были выведены для СГ-НЗБ и не подходят для аналогичной атаки на СГ-ШПС.

Тем не менее, были проведены эксперименты по расчету величины $\widetilde{P}$ для методеа вложения СГ-ШПС по уравнению (4.14). Результаты эксперимента при $P = 1; 0{,}5; 0{,}1; 0$ и $\alpha = 1; 2; 3$ приведены в таблице 6.4.

Таблица 6.4 – Экспериментальные результаты расчета $\widetilde{P}$ по методу ПВА

α	P	**№ изображения**				
		1	**2**	**3**	**4**	**5**
1	**1**	0,00306	0	0,03965	0,00525	0,00603
	0,5	0,03307	0	0,10544	0,01854	0,02475
	0,1	0	0	0,00580	0,00482	0,00468
2	**1**	0,00030	0	0	0,00005	0
	0,5	0	0,05735	0	0	0,00397
	0,1	0	0,00211	0	0	0,00050
3	**1**	0,00050	0,00335	0,01936	0,00215	0
	0,5	0,00061	0	0	0,00183	0
	0,1	0,00009	0,00397	0	0,00012	0
0		0	0	0	0	0,00005

Как видно из таблицы 6.4, при применении данного метода СГА к СГ-ШПС величина $\widetilde{P}$ не является оценкой доли вложения, то есть в случае СГ-ШПС данный метод не позволяет оценить долю вложенной информации P.

В таблице 6.5 приведены вероятности ложной тревоги P_{fa} и пропуска P_m при пороговом значении $\widetilde{P}_0 = 0{,}0001$.

Таблица 6.5 – Экспериментальный расчет величин P_{fa} и P_m

α	P	P_{fa}	P_m
1	1	-	0,27
	0,5	-	0,25
	0,1	-	0,35
	0	0,05	-
2	1	-	0,7
	0,5	-	0,52
	0,1	-	0,57
	0	0,05	-
3	1	-	0,21
	0,5	-	0,33
	0,1	-	0,42
	0	0,05	-

Из таблиц 6.4 и 6.5 видно, что данный метод СГА имеет большую вероятность пропуска P_m, поэтому не рекомендуется использовать данный метод для выявления СО, созданного по методу СГ-ШПС. Отметим, что экспериментальные исследования показали, что наилучший результат данная атака показывает при доле вложения $P = 0{,}5$.

Рассмотренные ранее экспериментальные результаты СГА, примененных к СГ-ШПС также показали наилучшие результаты при доле вложения $P = 0{,}5$, тогда как при долях вложения $P = 1$ и $P \leq 0{,}1$ все атаки показывают результаты хуже, чем при $P = 0{,}5$. Это свойство СГ-ШПС может быть использовано при разработке секретной СГС.

6.4 Стегоанализ, основанный на методе подсчета нулей гистограммы, применительно к СГ-НЗБ

Данный метод СГА подробно описан в разделе 5.3.

Проверим, работает ли принцип уменьшения количества нулей в гистограмме после вложения методом СГ-НЗБ. Из параграфа 4.2 известно, что если $P = 1$:

$$E\{V_w(2j)\} = E\{V_w(2j+1)\}_{j=0,1,..(L-1)/2} = \frac{1}{2}\big(V(2j) + V(2j+1)\big).$$

Если $V(2j) = 0$, а $V(2j+1) > 0$, то при вложении методом СГ-НЗБ при $P = 1$ получаем:

$$E\{V_w(2j)\} = E\{V_w(2j+1)\}_{j=0,1,..(L-1)/2} = \frac{V(2j+1)}{2}.$$

Следовательно, $V_w(2j) > 0$ и $V_w(2j+1) > 0$. Значит, количество нулей в гистограммеСО, созданного методом СГ-НЗБ меньше, чем в гистограмме ПО.

В таблице 6.6 приведены результаты эксперимента по подсчету количества нулей в гистограмме для 5 различных изображений с долей вложения $P = 1; 0{,}5; 0{,}1; 0$.

Таблица 6.6 – Результаты подсчета количества нулей гистограммы для 5 различных изображений при вложении по методы СГ-НЗБ

№изображения	P	K
1	1	162
	0,5	161
	0,1	164
	0	201
2	1	68
	0,5	68
	0,1	68
	0	162
3	1	148
	0,5	150
	0,1	157
	0	193
4	1	120
	0,5	123
	0,1	130
	0	174
5	1	152
	0,5	153
	0,1	160
	0	200

В таблице 6.7 приведены значения вероятностей ложной тревоги P_{fa} и пропуска P_m для данного метода при различных пороговых значениях K_0.

Таблица 6.7 – Значения P_{fa} и P_m для СГА, основанного на методе подсчета нулей гистограммы при вложении по методы СГ-НЗБ

K_0	P_m			P_{fa}
	$P=1$	$P=0{,}5$	$P=0{,}1$	
200	0,06	0,06	0,06	0,85
170	0,09	0,11	0,11	0,75
160	0,13	0,13	0,15	0,25
140	0,17	0,17	0,20	0,18
120	0,23	0,23	0,26	0,14
100	0,29	0,32	0,33	0,12
70	0,57	0,60	0,71	0,03

Как видно из таблиц 6.6 и 6.7, метод подсчета нулей гистограммы может быть применен для выявления вложений методом СГ-НЗБ. Заметим, что при уменьшении порогового значения K_0 вероятность ложной тревоги P_{fa} уменьшается, а вероятность пропуска P_m увеличивается. Данный СГА позволяет выявить отсутствие или наличие вложения методом СГ-НЗБ с вероятностью более 0,80. Стоит отметить, что рассмотренные в главе 4 методы для обнаружения вложений методом СГ-НЗБ значительно надежнее метода подсчета количества нулей гистограмм.

6.5Стегоанализ, основанный на сравнении соседних значений гистограммы, применительно в СГ-НЗБ

Данный метод СГА подробно описан в разделе 5.4. Он схож с методом СГА, основанном на статистике 1-ого порядка. Следовательно, должен выявлять наличие вложений методом СГ-НЗБ, но не настолько надежно, как атака по (4.10).

В таблице 6.8 приведены результаты экспериментов для данного стегоанализа для пяти различных изображений с долей вложения $P=1;0{,}5;0{,}1;0$.

Таблица 6.8 – Результат СГА, основанного на сравнении соседних значений гистограммы при вложении по методы СГ-НЗБ

№изображения	P	$R^{/}$
1	1	0,4756
	0,5	0,6963
	0,1	0,9288
	0	0,9798
2	1	0,2241
	0,5	0,5421
	0,1	0,9250
	0	0,9999
3	1	0,4233
	0,5	0,5704
	0,1	0,8328
	0	0,8927
4	1	0,3725
	0,5	0,6528
	0,1	0,9095
	0	0,9602
5	1	0,5257
	0,5	0,6462
	0,1	0,9387
	0	0,9998

Вероятности пропуска P_m и ложной тревоги P_{fa} для различных пороговых значений $R_0^{/}$ приведены в таблице 6.9.

Таблица 6.9 – Значения P_{fa} и P_m для СГА, основанного на методе сравнения соседних значений гистограммы при вложении по методы СГ-НЗБ

$R_0^/$	P_m			P_{fa}
	$P = 1$	$P = 0{,}5$	$P = 0{,}1$	
0,9999	0	0	0	0,39
0,97	0	0	0	0,31
0,93	0	0	0,29	0,21
0,90	0	0	0,73	0,18
0,88	0	0	0,74	0,15
0,85	0	0	0,79	0,10
0,80	0	0	0,88	0,03

Как видно из таблиц 6.8 и 6.9, метод сравнения значений соседних отсчетов гистограммы может применяться для выявления вложений методом СГ-НЗБ. Как показано в таблице 6.9, при уменьшении порогового значения λ_0 уменьшается вероятность ложной тревоги P_{fa} и увеличивается вероятность пропуска P_m. Надежность определения наличия или отсутствия вложения СГ-ГЗБ методом сравнения значений соседних отсчетов гистограмм не превышает 0,7. Следовательно, методы, рассмотренные в главе 4 лучше чем данный метод подходят для выявления вложений методом СГ-НЗБ.

6.6 Метод подсчета локальных максимумов

Как было сказано ранее, рассмотренные выше методы СГА оказываются неэффективными при исследовании некачественных изображений, изображений с сильным шумом или зернистостью. Отметим, что хотя обычно шум и зернистость видно невооруженным глазом, встречаются изображения, смотря на которые человек не может распознать эти особенности. Другими словами, не все «шумные» изображения будут выглядеть таковыми без детального

исследования.Отличить случайный шум самого изображения от искажения, привносимого вложением, довольно сложная задача.

Для того, чтобы понять, откуда возникает шум в цифровых изображениях, надо разобраться, как такие изображения создаются.

Принцип фотографирования цифровых фотокамер основан на фотоэффекте, возникающем в решетке кремния (из него состоят фоточувствительные элементы камеры, объединенные в матрицу). Матрица состоит из миллионов фоточувствительных элементов. Действие света на фоточувствительные элементы преобразуется в напряжения. Далее полученный сигнал усиливается и подвергается квантованию. Полученный квантованный сигнал еще внутри фотоаппарата проходит ряд преобразований, таких как цветокоррекция, подавление шума и другие. Далее изображение сохраняется в выбранном заранее формате, например BMP илиJPEG.

Шум может возникнуть на любом этапе создания изображения, как в процессе фотографирования, так и при обработке и пересохранении его уже в компьютере. При этом, по своим свойствам шум может быть разным – постоянным (возникающим во всех изображениях) или случайным (возникающим не во всех изображениях). Одни виды шума поддаются коррекции, другие можно совсем избежать, но некоторые виды не избежать и не откорректировать.

Одно из видов шума – шум от грязной оптики. При этом, чем сильнее загрязнена линза, тем сильнее шум. Такой шум как правило видно невооруженным глазом на полученным снимке. Откорректировать его довольно сложно, но можно избежать совсем, если следить за чистотой оптики у фотоаппарата. Данный шум является случайным.

Еще одной причиной шума, возникающего непосредственно при съемке, может быть квантовая природа света и неравномерное попадание света на фоточувствительный элементы. Данный шум возникает абсолютно во всех изображениях, его нельзя избежать или откорректировать, поэтому он является «фундаментальным ограничением». Пути решения данной проблемы пока не найдены.

Также при съемке может возникнуть засветка как части изображения, так и почти всего изображения. Засветка возникает, если на часть фотоэлементов попадает прямой свет, при этом часть пикселей, соответствующие этим фотоэлементам, оказываются значительно ярче других и влияет на соседние пиксели, при этом самого изображения на данном участке невидно. Из-за влияния засвеченных пикселей на соседние возникают взаимосвязи пикселей. Данного шума можно избежать правильным выбором ракурса фотографирования, при этом надо следить, чтобы прямой свет не попадал в объектив. Но если засветка уже присутствует на снимке, откорректировать ее практически невозможно.

Еще один шум, присутствующий во всех изображениях, возникает из-за квантования уже самого изображения при его обработке в фотоаппарате.В этом случае каждый пиксель немного искажается, и изображения получается неоднородным, даже если фотографировался однородный объект. В процессе квантования также возникает зависимость между соседними пикселями.Эффект данного шума можно уменьшить, совершенствуя методы квантования, но нельзя полностью устранить. Аналогичный шум может возникнуть при обработке и пересохранении изображения уже в компьютере.

В процессе фотографирования могут возникнуть «битые» (испорченные, дефектные) пиксели. Большинство современных фотокамер умеют находить и исправлять битые пиксели. Но при исправлении нарушается существующая взаимосвязь пикселей и возникает новая. Данный шум является случайным, он может и не появиться. Но предугадать его появление невозможно.

Как видно из вышесказанного, различные шумы могут возникнуть на этапе фотографирования и обработки. Часть этих шумов корректируются, часть нет. При этом во время обработки изображения и корректировке шума в фотоаппарате (и в компьютере тоже) возникают локальные взаимозависимости пикселей друг от друга.Бороться с шумами в изображениях очень сложно, при этом в подобных изображениях взаимозависимость пикселей отличаются от взаимозависимости более «чистого» изображения. Поэтому так важно научиться различать шумные изображения от «чистых» и от СО.

Отличительная особенность «шумных» изображений – очень малое количество нулей в гистограмме. При анализе «шумных» изображений с использованием пороговых значений, рассмотренных ранее, вероятность ложной тревоги возрастает до 0,90 для метода, основанного на статистике 1-ого порядка и метода сравнения соседних значений гистограммы, до 0,7 для метода ПВА. Метод подсчета нулей гистограммы для таких изображений непоказателен по определению. Отметим, что в отличие от изображений, рассмотренных в главах 4 и 5, в «шумных» изображениях и изображениях с полной палитрой серого количество нулей в гистограмме, как правило, меньше 60.

В таблице 6.10 приведены результаты стегоанализа «шумных» изображений рассмотренными ранее методами, при чем цифра «1» на первой позиции номера изображения обозначает изображение без визуально различимого шума, цифра «2» – визуально «зашумленное» изображение, знаком «+» отмечен правильный выбор гипотезы, а знаком «-» - неправильный.

Таблица 6.10 –Результаты СГА по (4.10), (4.14), (5.13) и (5.14) на изображенияс полной палитрой оттенков серого

Номер изображения	**Результат СГА по методу χ^2**	**Результат СГА по методу ПВА**	**Результат СГА по методу подсчета нулей гистограммы**	**Результат СГА по методу сравнения соседних значений гистограммы**
110	-	-	-	-
111	-	-	-	-
115	-	-	-	-
116	-	-	-	-
117	-	-	-	-
118	-	-	-	-
119	-	+	-	+
122	-	+	-	-

Продолжение таблицы 6.10

Номер изображения	**Результат СГА по методу χ^2**	**Результат СГА по методу ПВА**	**Результат СГА по методу подсчета нулей гистограммы**	**Результат СГА по методу сравнения соседних значений гистограммы**
201	-	+	-	-
205	-	-	-	-
206	+	-	-	-
207	-	-	-	-
208	-	+	-	-
209	-	+	-	-
211	-	-	-	+

Отметим, что в связи с ростом качества обработки и хранения цифровых фотографий большинство изображений сейчас имеют полную палитру оттенков цветов (в рассматриваемых нами изображениях – полную палитру оттенков серого), что приводит к уменьшению количества нулей гистограммы. Такие полноцветные изображения при СГА ведут себя как «зашумленные», следовательно, рассмотренные ранее атаки для них ненадежны.

При проведении экспериментальных исследований было замечено, что различные шумы увеличивают количество максимумов гистограммы, тогда как изображения с полной палитрой цветов имеют более «гладкую» гистограмму. На основе этого наблюдения и дальнейших экспериментальных исследований был разработан метод локальных максимумов, который подходит как для СГ-НЗБ, так и для СГ-ШПС. В основе метода – определение локальных максимумов на отдельно взятом участке гистограммы.

Алгоритм метода локальных максимумов состоит в следующем:

- на гистограмме исследуемого изображения находится отсчет с максимальным значением;

- выделяется исследуемая область по 25 отсчетов с каждой стороны от отсчета с максимальным значением. Это основная область гистограммы изображения, она состоит из номеров отсчетов, соответствующих основным (наиболее часто встречающимся) оттенкам изображения. Если отсчет с максимальным значением находится менее чем в 25 отсчетах от края гистограммы, то основная область сдвигается так, чтобы один из ее краев совпадал с краем гистограммы;
- в основной области гистограммы определяем все максимумы – локальные максимумы гистограммы. Отсчет j считается локальным максимумом гистограммы, если $V(j) > V(j+1)$ и $V(j) > V(j-1)$. Обозначим количество локальных максимумов гистограммы на основном участке как M ;
- в зависимости от количества локальных максимумов M на основной области гистограммы делаем выводы о наличии или отсутствии вложения:
  - если $M \le 14$ – в исследуемом изображении вложение отсутствует;
  - если $M \ge 19$ – исследуемое изображение содержит скрытую информацию;
  - если $15 \le M \le 18$ – нельзя сделать однозначные выводы о наличии или отсутствии вложения. Требуются дополнительно исследовать изображение рассмотренными ранее методами СГА.

Результаты анализа методом локальных максимумов изображений с полной палитрой серого при использовании метода вложения СГ-НЗБ представлены в таблице 6.11 (успешное обнаружение по данному критерию отмечено знаком «+»).

Таблица 6.11. Результаты СГА изображений с полной палитрой оттенков серого методом подсчета локальных максимумов гистограммы для СГ-НЗБ

Номер изображения	P		
	0	**1**	**0,5**
110	+	-	-
111	+	-	-
115	-	+	+
116	+	+	+
117	-	+	+
118	+	+	+
119	+	+	-
122	+	+	+
201	+	+	-
205	+	+	+
206	+	+	+
207	+	+	+
208	+	+	+
209	+	+	+
211	+	-	+

Данный метод дает надежные результаты при доли вложения $P > 0{,}5$, вероятность верного определения наличия или отсутствия вложения составляет 0,8. Метод позволяет выявлять вложения, сделанные как методом С-НЗБ, так и методом СГ-ШПС. Стоит отметить, что данный метод не рекомендуется рассматривать как самостоятельный метод, поскольку он достаточно надежен только при $P > 0{,}5$. При этом для изображений с неполной палитрой серого рассмотренные ранее методы не менее эффективны при меньших долях вложения. Однако, данный метод может использоваться совместно с уже рассмотренными в главах 4 и 5 методами. В параграфе 6.7 будет рассмотрен комбинированный метод стегоанализа, объединяющий в себе рассмотренные ранее методы.

6.7 Комбинированный метод стегоанализа

Для наиболее эффективного тестирования изображений был разработан комбинированный метод СГА, объединяющий в себе рассмотренные ранее методы СГА. Блок-схема алгоритма данного метода приведена на рисунках 6.4 и 6.5.

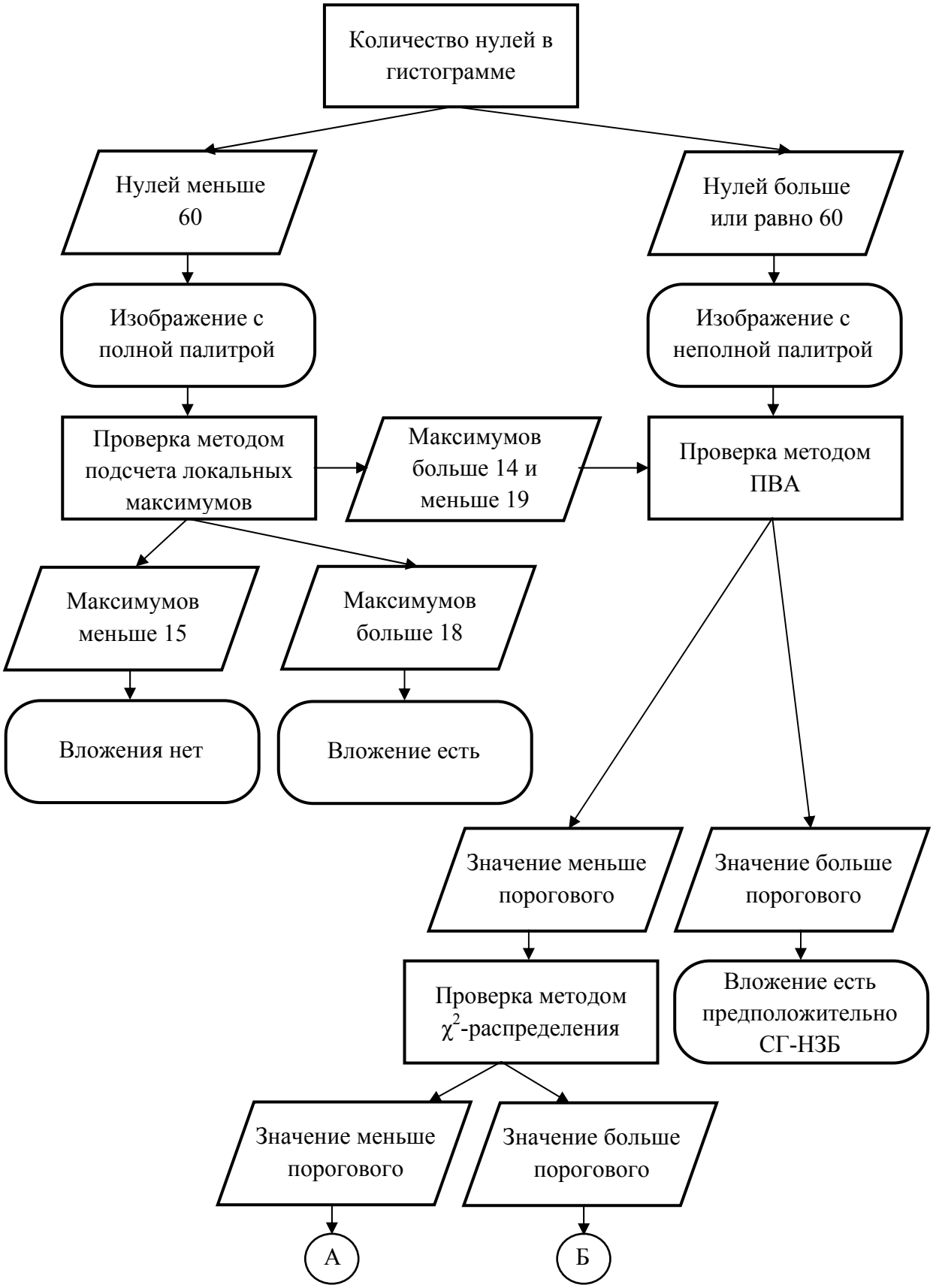


Рисунок 6.4 – Блок-схема комбинированного метода СГА (начало)

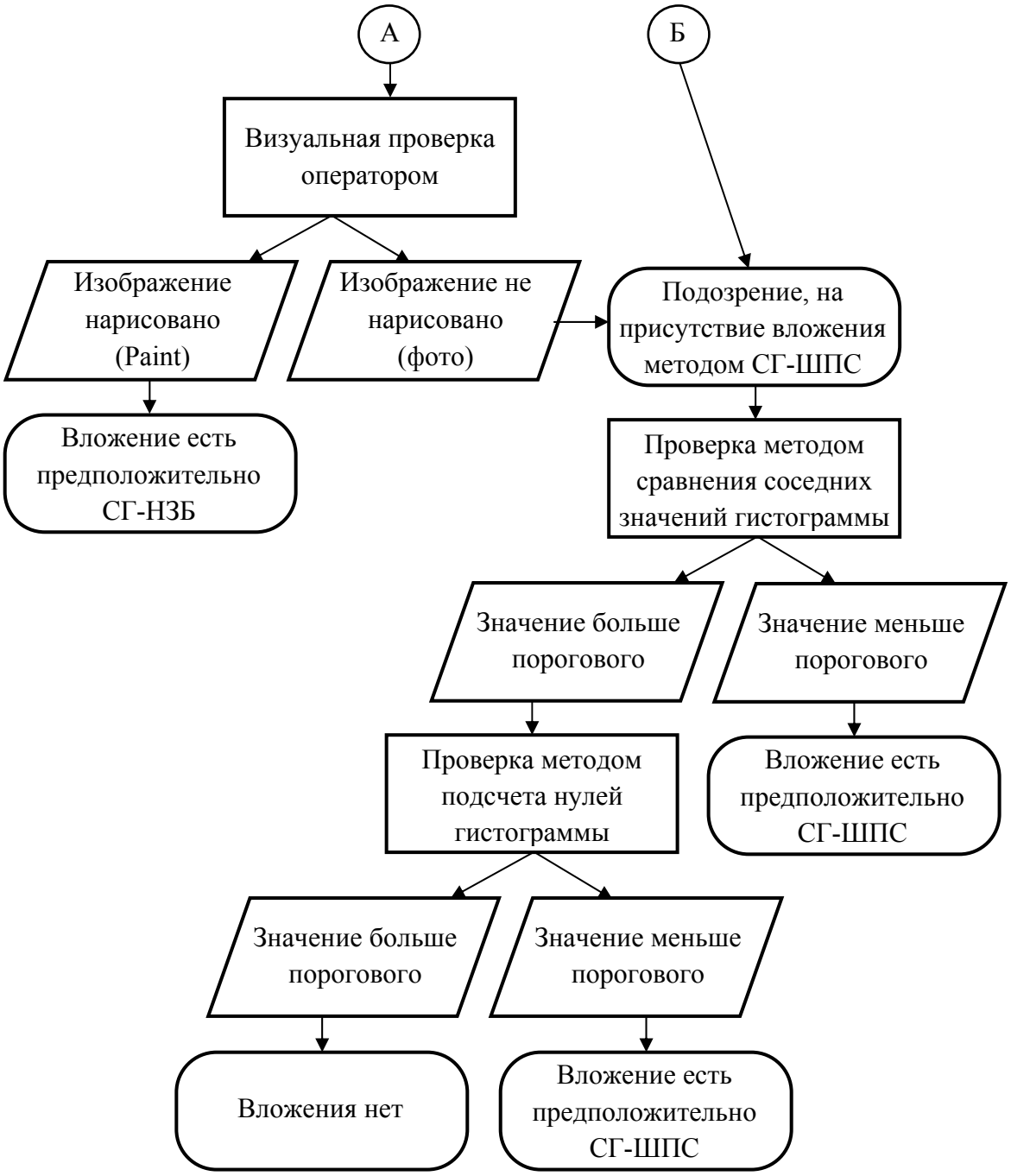


Рисунок 6.5 – Блок-схема комбинированного метода СГА (окончание)

Как видно из рисунков 6.4 и 6.5, комбинированный метод имеет следующий алгоритм:

Вначале производим подсчет количества нулей в гистограмме для выявления характеристик исследуемого изображения. (Стоит отметить, что в этом случае не используетсяСГА, основанныйна подсчете нулей в гистограмме, тогда можно определяется, является ли исследуемое изображение полноцветным («зашумленным») или нет.) Рекомендуемый порог – 60.

Если нулей меньше 60, считаем, что исследуемое изображение либо обладает полной палитрой серого, либо зашумлено. Чтобы определить наличие или отсутствие вложения применяем к исследуемому изображению метод подсчета локальных максимумов и в зависимости о результата делаем выводы:

- если количество локальных максимумов меньше 15 – в исследуемом изображении вложение отсутствует;
- если количество локальных максимумов больше 18 – исследуемое изображение содержит скрытую информацию;
- если количество локальных максимумов больше 14 и меньше 19 – нельзя сделать однозначные выводы о наличии или отсутствии вложения. Проверяем изображение методами СГА для СГ-НЗБ и СГ-ШПС, начиная с метода ПВА.

Если нулей 60 и больше, считаем, что перед нами изображение с неполной палитрой серого. Исследуем изображение методами СГА для СГ-НЗБ и СГ-ШПС.

Начинаем исследовать изображение с метода ПВА. Если полученный результат атаки больше порогового значения $\widetilde{P}_0$, считаем, что изображение содержит вложение СГ-НЗБ. Стоит отметить, что если полученное в результате атаки значение меньше 0,5, то оно является оценкой доли вложенное скрытой информации. Если результат атаки меньше порогового значения, переходим к следующему методуСГА.

Проверяем изображения методом, основанным на подсчете χ^2. Если результат атаки меньше порогового значения, требуется визуальная проверка изображения человеком-оператором.

Человек-оператор должен оценить изображения – является ли оно нарисованным в компьютерной программе, например в программе Paint, или это изображение создано иным способом, например, фотография. Если изображение нарисовано, считаем, что изображение содержит вложение СГ-НЗБ. Если изображение не нарисовано, подозреваем, что изображение содержит вложение

СГ-ШПС, продолжаем исследовать изображение, начиная с метода сравнения соседних значений гистограммы.

Если результат СГА методом χ^2 больше порогового значения, предполагаем, что в изображении содержится вложение методом СГ-ШПС. Продолжаем проверку изображения методом сравнения соседних значений гистограммы.

Если результат СГА методом сравнения соседних значений гистограммы меньше порогового значения, считаем, что в изображение содержит вложение СГ-ШПС. Если результат больше порогового значения – продолжаем проверку методом подсчета нулей гистограммы.

Если результат атаки методом подсчета нулей гистограммы меньше порогового значений, считаем, что изображение содержит вложение СГ-ШПС. Если результат атаки больше порогового значения, считаем, что в исследуемом изображении вложения нет.

Заметим, что в данный алгоритм не позволяет точно определить используемый метод вложения, основной его задачей является выявить только наличие или отсутствие вложения.

Результаты анализа 5изображений комплексным СГА приведены для метода вложения СГ-НЗБ в таблице 6.12 и для метода вложения СГ-ШПС в таблице 6.13. Знаком «+» отмечен правильный выбор гипотезы, а знаком «-» – неправильный.

Таблица 6.12Экспериментальные результаты комплексного СГАдля СГ-НЗБ

Номер изображения	*P*					
	1	0,5	0,1	0,05	0,01	0,001
10	+	+	+	+	-	-
12	+	+	+	+	+	+
20	+	+	+	+	+	+
21	+	+	+	+	+	+
23	+	+	+	+	+	+

Таблица 6.13Экспериментальные результаты комплексного СГА для СГ-ШПС

Номер изображения	P					
	1	0,5	0,1	0,05	0,01	0,001
10	+	+	+	+	+	-
12	+	+	+	+	+	+
20	+	+	+	+	+	+
21	+	+	+	+	+	+
23	+	+	+	+	+	+

Из таблиц 6.10 и 6.11 видно, что комплексный метод эффективен при анализе различных по типу создания и обработки изображений для выявление вложений СГ-НЗБ и СГ-ШПС. Как показал эксперимент, проведенный на 1000 изображений, комбинирование традиционных и предложенных методов СГА позволяют добиться малых значений вероятности ложной тревоги $P_{fa} = 0{,}13$ и вероятности пропуска $P_m = 0{,}12$.

Отметим, что данный метод, как и все предыдущие, имеет свои достоинства и недостатки.

К основным достоинствам относятся:

- возможность анализировать несколько видов вложения;
- возможность легко изменить алгоритм метода (добавить или убрать «блоки», поменять их местами) для расширения области применения;
- уменьшение вероятности ложной тревоги P_{fa} при анализе изображений с полной палитрой серого;
- возможность поточного анализа изображений с помощью компьютеров, поскольку большая часть алгоритма реализуется компьютерной программой и не требует присутствие оператора.

К основным недостаткам относятся:

- необходимость присутствия оператора в спорных случаях;
- невозможность точно определить метод применяемого вложения;

- отсутствие надежных результатов при анализе сильно зашумленных изображений с полной палитрой серого, особенно при малой доле вложения P .

6.8 Выводы

Как показали результаты экспериментов, целевые методы СГА дают не слишком надежные результаты, если применять их для исследования СО, при создании которых применялся не тот метод вложения, на который направлен целевой метод. Другими словами, метод СГА для СГ-ШПС при атаке на СО, созданный методом СГ-НЗБ, дают менее надежные результаты, чем при использовании СГА для СГ-НЗБ. При использовании методов СГА для СГ-НЗБ для атаки на СО, созданный методом СГ-ШПС, наблюдаются аналогичные результаты.

Однако, даже в случае использования целевых методов для СО с их «родным» методом вложения, результаты анализа могут быть ложными. Особенно неэффективны методы целевого СГА при атаках на изображения с полной палитрой или на изображения низкого качества (так называемые «зашумленные» изображения), которые по своим статистическим свойствам близки к изображениям с полной палитрой. При анализе таких изображений вероятность ложной тревоги P_{fa} резко возрастает.

Для более надежного и эффективного анализа был разработан метод локальных максимумов, позволяющий определять, есть ли в данном изображении с полной палитрой вложение или нет.

Для того, чтобы определить, обладает ли исследуемое изображение полной палитрой или нет, был разработан критерий, пороговым значением которого является количество нулей в гистограмме, равное 60.

Объединив критерий определения изображений с полной палитрой, метод подсчета локальных максимумов и наиболее эффективные из рассмотренных ранее целевых методов был разработан метод комплексного СГА(рисунки 6.4 и

6.5), который позволил повысить эффективность атак на любые изображения формата BMP до 0,8.

Стоит отметить, что и метод комплексного СГА имеет свои слабые стороны и не дает абсолютнойточности решения. Например, комплексный СГА малоэффективен, если в изображение с полной палитрой методом СГ-ШПС вкладывается доля скрываемой информации менее 0,1. Используя в качестве ПО изображения с полной палитрой, в качестве метода вложения – СГ-ШПС и вкладывая достаточно малые доли скрываемой информации (что приведет к снижению скорости передачи данных) можно значительно повысить устойчивость СГС к СГА.

Заключение

В современном мире возрастает интерес к методам СГ и СГА. Это связано в том числе и с тем, что СГС часто стали использоваться в криминальных и террористических целях [6,7].Обзор современных методов СГС [5] и СГА показал, что существующие методы СГА, особенно для СГ-ШПС, не достаточно эффективны. Поэтому разработка новых и совершенствование старых методов СГА является актуальной задачей.

В диссертации предложены методы СГА для СО, созданных с помощью метода СГ-ШПС, а также исследованы уже существующие методы СГА для СО, созданных методом СГ-НЗБ, и проведена экспериментальная проверка возможность их применения.

Основными результатами диссертации являются:

- предложен и экспериментально исследован критерий оценки стойкости СГС, основанный на расчете расстояния Бхаттачариа [22];
- экспериментально исследованыэффективности методов СГА, таких как визуальный анализ, СГА, основанный на статистике 1-ого порядка, и СГА, основанный на статистике 2-ого порядка, для СГ-НЗБ [37] срекомендациями по выбору пороговых значений для методов СГА, основанных на статистике 1-ого и 2-ого порядка;
- предложены и экспериментально исследованы методы СГА, основанный на подсчете нулей гистограммы и основанный на сравнении соседних значений гистограммы, даны рекомендации по выбору порогового значения для методов СГА, основанного на подсчете нулей гистограммы и основанного на сравнении соседних значений гистограммы;
- предложен и экспериментально исследован комбинированный метод СГА;
- реализация исследований, проведенных в данной диссертации, в научно-исследовательской работе «Ярус-СГ».

Результаты диссертации были опубликованы в следующих статьях:

- «Обнаружение видеостегосистем при вложении секретной информации в наименьшие значащие биты» [4];
- «Исследование возможностей выявления скрытых сообщений в информационных ресурсах сети интернет» [5];
- «Построение стегосистемы, инвариантной к статистике покрывающего сообщения» [11];
- « On the Use of Bhattacharyya Distance as a Measure of the Detectability of Steganographic Systems » [21];
- «Исследование эффективности методов обнаружения стегосистем, использующих широкополосное вложение» [28];
- «Исследование эффективности методов обнаружения стегосистем, использующих вложение в наименее значащие биты» [37].

Результаты диссертации были использованы в научно-исследовательской работы «Ярус-СГ» [5] и в лабораторной работе и курсовом проекте в курсе «Основы стеганографии» на кафедре Защищенные системы связи Санкт-Петербургского государственного университета телекоммуникаций им. проф. М.А. Бонч-Бруевича, что подтверждается соответствующими актами, приведенными в Приложении В и Приложении Г. Многолетние использование результатов диссертации в курсе «Основы стеганографии» показало актуальность и эффективность результатов исследований. Текст основного файла программы с алгоритмами методов СГА, описанных в диссертации, приведен в Приложении Б.

Отметим, что решения задачи СГА представляет собой весьма сложную научную и техническую задачу. В данной работе проанализированы уже имеющиеся методы СГА и предложены методы, в том числе комбинированный метод, позволяющий повысить эффективность СГА. Они могут найти практическое применение в государственных структурах для обнаружения скрытых каналов передачи данных между террористами и другими криминальными группировками, а также в бизнесе для обнаружения промышленного шпионажа.

Список используемых сокращений

ПО – покрывающий объект;

СО – стеганографический объект;

СГ – стеганография;

СГС – стеганографическая система;

СГА – стеганографический анализ

ССГА – «слепой» стеганографический анализ;

НЗБ – наименее значащий бит;

СГ-НЗБ – метод стеганографии, использующий вложение в наименее значащие биты;

ШПС – широкополосный сигнал;

СГ-ШПС – метод стеганографии, использующий вложение в широкополосный сигнал;

МОВ – метод опорных векторов;

ПВА – парно-выборочный анализ.

Литература

1 Das, S. Steganography and Steganalysis: Different Approaches [Электронный ресурс] / S. Das, S. Das, B. Bandyopadhyay, S. Sanyal // International Journal of Computers, Information Technology and Engineering. – 2008. – Vol. 2. – Режим доступа: http://arxiv.org/ftp/arxiv/papers/1111/1111.3758.pdf (Датаобращения 09.02.214).

2 Westfeld,A. AttacksonSteganographySystems. BreakingtheSteganographicUtilitiesEzStego, Jsteg, SteganosandS-Toolsandsome Lessons Learne / A. Westfeld, A. Pfitzmann // Lecture Notes in Computer Science. – 2000. – Vol. 1765. – Pp. 61-76.

3 Dumitrescu, S. Detection of LSB Steganography via Sample Pair Analysis / S. Dumitrescu, X. Wu, Z. Wang, F.A.P. Petitcolas// Lecture Notes in Computer Science. – 2002. – Vol. 2578. – Pp.355-372.

4 Коржик, В.И. Обнаружение видеостегосистем при вложении секретной информации в наименьшие значащие биты / В.И. Коржик, Е.Ю. Герлинг, А.П. Дубов // Санкт-Петербургский государственный университет телекоммуникаций им. проф. М.А. Бонч-Бруевича. Юбилейная научно-техническая конференция профессорско-преподавательского состава, научных сотрудников и аспирантов, 57-я. 24-28 января 2005 года: материалы. – СПб. :СПбГУТ. – 2005. – С. 148-149.

5 Коржик, В.И. Исследование возможностей выявления скрытых сообщений в информационных ресурсах сети интернет: отчет о НИР «Ярус-СГ» / В.И. Коржик, Р.В. Чесноков, Е.Ю. Герлинг – СПб. :СПбГУТ. – 2010. – 181с.

6 Conway,M. CodeWars: Steganography, SignalsIntelligence, andTerrorism. Knowledge, TechnologyandPolicy / M. Conway // Special issue entitled Technology and Terrorism. – 2003. – Vol. 16, №. 2. – Pp. 45-62.

7 Jesse, D.D. Tactical Means, Strategic Ends: Al Qaeda’s Use of Denial and Deception / D.D. Jesse // Terrorism and Political Violence. – 2006. – Vol. 18. – Pp. 367–388.

8 Betancourt, S.R. Steganography: A New Age of Terrorism [Электронный ресурс] / S.R. Betancourt // SANS Institute. – 2004. – 10 с. – Режим доступа: http://www.giac.org/paper/gsec/3494/steganography-age-terrorism/102620 (Дата обращения 01.02.2014).

9 Грибунин,В.Г. Цифроваястеганография / В.Г. Грибунин, И.Н. Оков, И.В. Туринцев. – М. :СОЛОН-ПРЕСС, 2009. – 265 с.

10 Bandyopadhyay, S.A Tutorial Review on Steganography / S. Bandyopadhyay, D.Bhattacharyya // UFL & JIITU. – IC-2008. – Pp. 105-114.

11 Коржик, В.И. Построение стегосистемы, инвариантной к статистике покрывающего сообщения / В. И. Коржик, Е. Ю. Герлинг // Санкт-Петербургский государственный университет телекоммуникаций им. проф. М.А. Бонч-Бруевича. Научно-техническая конференция профессорско-преподавательского состава, научных сотрудников и аспирантов, 59-я. 22-26 января 2007 года: материалы. – СПб. :СПбГУТ. – 2007. – С. 184-185.

12 Рябко, Б.Я. Основы современной криптографии и стеганографии / Б.Я. Рябко, А.Н. Фионов. – М : Горячая линия–Телеком, 2010. –232 с.

13 Fridrich,J. SteganographyinDigitalMediaPrinciples, Algorithms, andApplications / J. Fridrich. – CambridgeUnivP, 2010. – 462 pp.

14 Luo, W.Edge Adaptive Image Steganography Based on LSB Matching Revisited/ W. Luo, F. Huang, J. Huang // Transactions on Information Forensics and Security. – 2010. – Vol. 5. – Pp. 201-214.

15 Ker, A. Steganalysis of LSB Matching in Grayscale Images / A. Ker // Signal Processing Letters. – 2005. – Vol. 12. – Pp. 441-444.

16 Коржик, В.И. Построение идеально стойкой лингвистической стегосистемы с редактируемым текстом / В.И. Коржик, А.А. Залетов // Санкт-Петербургский государственный университет телекоммуникаций им. проф. М.А. Бонч-Бруевича. Научно-техническая конференция профессорско-преподавательского состава, научных сотрудников и аспирантов, 60-я. 21-25 января 2008 года: материалы. – СПб. :СПбГУТ. – 2008. – С. 168.

17 Kullback, S. On information and sufficiency / S. Kullback, R.A. Leibler // The Annals of Mathematical Statistics. – 1951. – Vol.22. № 1. – Pp. 79-86.

18 Cover, T.M. Elements ofInformation Theory / T.M. Cover, J.A. Thomas. – Chichester. : John Wiley& Sons, Inc, 1991. – 542 pp.

19 Poor, H.V. An Introduction to Signal Detection and Estimation /H.V. Poor – Heidelberg. :Springer, 1994. – 398 pp.

20 Wang, Y. Steganalysis of block-structured stegotext / Y. Wang, P. Moulin // Security, Steganography, and Watermarking of Multimedia Contents. – 2004. – Vol. 5306. – Pp. 477–488.

21 Korzhik, V. On the Use of Bhattacharyya Distance as a Measure of the Detectability of Steganographic Systems / V. Korzhik, H. Imai, J. Shikata, G. Morales-Luna, E. Gerling // Transactions on Data Hiding and Multimedia Security III, Lecture Notes in Computer Science. – 2008. – Vol. 4920. – Pp. 23–32.

22 Kailath, T. The divergence and Bhattacharyya distance measures in signal selection / T. Kailath // IEEE Transactions on Communication Technology. – 1967. – Vol. 15. – Pp. 52–60.

23 Van Trees, H.L. Detection, Estimation, and Modulation Theory, 2nd Edition, Part I / H.L. Van Trees, K.L. Bell. – Chichester. :John Wiley & Sons, Inc, 2013. – 1176 pp.

24 Horn, R.A. Matrix Analysis, 2nd Edition / R.A. Horn, C.R. Johnson – Cambridge. : Cambridge University Press, 2013. – 662 pp.

25 Коржик, В.И. Помехоустойчивое кодирование дискретных сообщений в каналах со случайной структурой / В.И. Коржик, Л.М. Финк. – М : Связь. – 1975. – 272 с.

26 Hall, M. Combinatorial Theory / M. Hall – Chichester. : John Wiley & Sons, Inc, 1986. – 464 pp.

27 Korzhik, V. Stegosystems based on noisy channels / V. Korzhik, M.H. Lee, G. Morales-Luna //Proc. IX Spanish Meeting on Cryptology and Information Security. – 2006. – Pp. 379-387.

28 Герлинг Е.Ю. Исследование эффективности методов обнаружения стегосистем, использующих широкополосное вложение / Е.Ю. Герлинг // Телекоммуникации. – 2014. – № 1. – С. 6-12.

29 Pevny, T.Towards Multi-class Blind Steganalyzer for JPEG Images / T. Pevny, J. Fridrich // Lecture Notes in Computer Science. – 2005. – Vol. 3710. – Pp. 39-53.

30 Воронцов, К.В. Лекции по методу опорных векторов [Электронный ресурс] / К.В. Воронцов – 2007. – 18 с. – Режим доступа: http://www.ccas.ru/voron/download/SVM.pdf (Дата обращения 26.01.2014).

31 Farid, H. Detection Hidden Message Using Higher-Order Statistics and Support Vector Machines / H. Farid, L. Siwei // Lecture Notes in Computer Science. – 2003. – Vol. 2578. – Pp. 350-354.

32 Fridrich, J. Feature-Based Steganalysis for JPEG Image and its Implication for Future Design of Steganographic Schemes / J. Fridrich // Lecture Notes in Computer Science. – 2005. – Vol. 3200. – Pp. 67-81.

33 Salle, P. Model-Based Steganography / P. Salle // Lecture Notes in Computer Science. – 2004. – Vol. 2939. – Pp. 154-167.

34 Коржик, В.И.. Разработка расширенного комплекса функционалов для обнаружения вложений с использованием «слепого» стегоанализа на SVM / В. И. Коржик, Д. В. Цветков // Международная научно-техническая и научно-методическая конференция. Актуальные проблемы инфотелекоммуникаций в науке и образовании.№64. 20-24 февраля 2012 года : материалы. – СПб. : СПбГУТ. – 2012. – С. 233-234.

35 Lee, K. Category Attack for LSB Steganalysis of JPEG / K. Lee, A. Westfeld, S. Lee // Lecture Notes in Computer Science. – 2006. – Vol. 4283. – Pp. 35-48.

36 Lee, Y.-K. An Advanced Least-Significant-Bit Embedding Scheme for Steganographic Encoding / Y.-K. Lee, G. Bell, S.-Y. Huang, R.-Z. Wang, S.-J. Shyu // Lecture Notes in Computer Science. – 2009. – Vol. 5414. – Pp. 349-360.

37 Герлинг Е.Ю. Исследование эффективности методов обнаружения стегосистем, использующих вложение в наименее значащие биты / Е.Ю. Герлинг // Информационные системы и технологии. – 2011. – № 4. – С. 137-144.

38 Ван дер Варден, Б.Л. Математическая статистика / Б.Л. Ван дер Варден. – М. :Москва, 1960. – 435 с.

39 Maes M. Twin Peaks: The Histogram Attack to Fixed Depth Image Watermarking / M. Maes // LNCS. – 1998. – Vol. 1525. – Pp. 290-305

Приложение А

Выводуравнениедля расчета СГА методом ПВА

В работе [3] было приведено доказательство метода ПВА, описанного в главе 4. Однако, поскольку это доказательство является недостаточно полным, в данном приложении приводится более подробное доказательство, выполненное автором диссертации.

Введем некоторые специальные множества пар образцов, называемых следами множеств.

Представим цифровой сигнал изображения в качестве образцов s_0, s_1,…, s_{L-1}, где индекс представляет собой значение яркости, соответствующей данному образцу, а L – количество возможных яркостей в изображении. В нашем случае $L = 256$. Пары образцов имеют вид (s_i, s_j), где $1 \le i, j \le L$. Пусть W представляет собой набор пар образцов, вытащенных из цифрового сигнала, вида (u, v), где u и v – величины двух образцов.

Для анализа эффекта вложения в наименьшие значащие биты введем подмножества, входящие в множество W.

Обозначим D_n как подмножество W, которое состоит из пар образцов вида $(u, u+n)$ или $(u, u+n)$, то есть пара представляет собой две величины, отличающиеся ровно на n, где n – целое число $1 \le n \le 2^t - 1$, где t – количество бит на пиксель изображения. Например, пары образцов D_1 имеют вид $(u, u+1)$ или $(u+1, u)$, тогда водна из пар может представлять собой, например, две двоичные последовательности, приведенные ниже.

u =1001110110,

$u+1$=1001110111.

Для каждого целого числа m, где $0 \le m \le 2^{t-1} - 1$, обозначим C_m как подмножество W, состоящее из пар образцов, чьи величины отличаются на

величину m в первых $t-1$ битах, то есть биты сдвигаются на одну позицию справа, и затем оценивается разница. Пара образцов C_1 может иметь вид двоичных последовательностей:

11001110,

11001100.

А может быть представлена и другими двоичными последовательностями:

11001110,

11001101.

Для данного подмножества величины наименьших значащих бит не имеют значение.

Видно, что подмножество D_{2m} является частью множества C_m. Для примера возьмем $m=1$, тогда подмножество D_2 состоит из пар образцов вида $(u,u+2)$или $(u+2,u)$, а подмножество C_1 состоит из пар образцов вида $(v,v+2)$, $(v+2,v)$, $(v,v+3)$ или $(v+3,v)$, а так же включает в себя часть пар образцов вида $(v,v+1)$ или $(v+1,v)$.Из примера видно, что D_{2m} является частью C_m. При аналогичном рассмотрении подмножества D_{2m+1} видно, что оно находится между множествами C_m и C_{m+1}.

Само множество D_{2m+1} может быть разделено еще на два подмножества X_{2m+1} и Y_{2m+1}, где $X_{2m+1}=D_{2m+1}\cap C_{m+1}$ и $Y_{2m+1}=D_{2m+1}\cap C_m$, для $0\le m\le 2^{t-1}-2$, и $X_{2^t-1}=0$, $Y_{2^t-1}=D_{2^t-1}$. Подмножество X_{2m+1} состоит из пар образцов вида $(2k-2m-1,2k)$ или $(2k,2k-2m-1)$, а подмножество Y_{2m+1} состоит из пар образцов вида $(2k-2m,2k+1)$ или $(2k+1,2k-2m)$.

Если образцы пар множества W распределены равномерно во временной области, то при целом числе m, таком что $0\le m\le 2^{t-1}-2$ получаем $\mathrm{E}\{|X_{2m+1}|\}=\mathrm{E}\{|Y_{2m+1}|\}$. Это одно из основных наблюдений в технологии данного СГА.

Наименьшие значащие биты пар образцов могут принимать четыре состояния, называемые шаблонами $\pi \in \{00,\ 01,\ 10,\ 11\}$. Данные значения отражают возможные комбинации наименьших значащих бит пар образцов. Например, у нас есть пара образцов в виде двоичных последовательностей 01100011 и 01100010, тогда π имеет значение 10.

Для каждого целого числа m, где $1 \le m \le 2^{t-1} - 1$ подмножество C_m можно разделить на три подмножества X_{2m-1}, D_{2m} и Y_{2m+1}. Возьмем произвольную пару образцов (u, v) подмножества X_{2m-1}, тогда $(u, v) = (2k - 2m + 1, 2k)$ или $(u, v) = (2k, 2k - 2m + 1)$. После изменения пар образцов (u, v) по шаблону 10 получим $(u', v') = (2k - 2m, 2k)$ или $(u', v') = (2k + 1, 2k - 2m + 1)$. Если будем изменять по шаблону 01, получим $(u', v') = (2k - 2m + 1, 2k + 1)$ или $(u', v') = (2k, 2k - 2m)$.

Эти наблюдения показывают пригодность множеств X_{2m} и Y_{2m} для стегоанализа. Обозначим X_{2m} как подмножество W, состоящее из всех пар вида $(2k - 2m, 2k)$ или $(2k + 1, 2k - 2m + 1)$, а Y_{2m} как подмножество W, состоящее из всех пар вида $(2k - 2m + 1, 2k + 1)$ или $(2k, 2k - 2m)$. Возьмем $m = 2$. Тогда пары образцы подмножества D_4 имеют вид $(u, u + 4)$ или $(u + 4, u)$, подмножества X_{2m} имеют вид $(2k - 4, 2k)$ или $(2k + 1, 2k - 3)$, а подмножества Y_{2m} имеют вид $(2k - 3, 2k + 1)$ или $(2k, 2k - 4)$. Видно, что подмножество D_{2m} разделяется на два подмножества X_{2m} и Y_{2m}.

Множество C_m при $1 \le m \le 2^{t-1} - 1$ может быть разделено на четыре подмножества X_{2m-1}, X_{2m}, Y_{2m} и Y_{2m+1}, называемые следами множества C_m. Операции, производимые над наименьшими значащими битами, могут быть смоделированы с помощью машин конечных состояний, показанных на рисунок А.1.

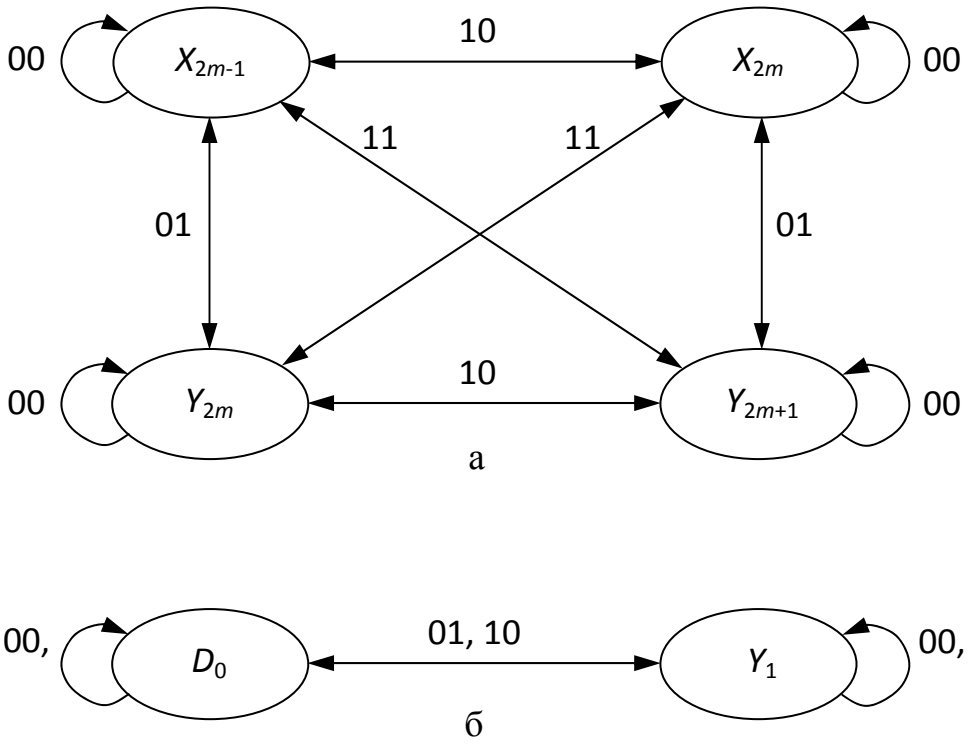


Рисунок А.1 – а – граф машины конечных состояний для множества C_m при m>0,
б – граф машины конечных состояний для множества C_0

На рисунке А.1а показана машина конечных состояний для множества C_m при $1 \le m \le 2^{t-1} - 1$, которая показывает, как перемещаются пары образцов между подмножествами X_{2m-1}, X_{2m}, Y_{2m} и Y_{2m+1}, при изменении наименьших значащих бит по различным шаблонам. На рисунке А.1б показана машина конечных состояний для подмножества C_0, которая показывает, как перемещаются пары образцов между подмножествами D_0 и Y_1, при изменении наименьших значащих бит по различным шаблонам.

Каждая стрелка, нарисованная из множества A к множеству B, означает, что какая-то пара образцов из множества A становится парой образцов из множества B, изменяясь, соответственно шаблону, указанному около стрелки. Машина конечных состояний, показанная на рисунке А.1а, основана на четырех подмножествах C_m. Рассмотрим переход пар образцов из подмножества X_{2m-1} в подмножество X_{2m} при использовании образца 10. Подмножества X_{2m-1} состоит из пар образцов вида $(2k - 2m + 1, 2k)$ или $(2k, 2k - 2m + 1)$, а подмножество X_{2m}

состоит из пар образцов вида $(2k-2m,2k)$ или $(2k+1,2k-2m+1)$. Рассмотрим более подробно взаимодействие подмножество X_{2m-1} с шаблоном 10:

Подмножество X_{2m-1}	Шаблон 10	Подмножество X_{2m}
$2k-2m+1$	$\oplus 1$	$2k-2m$
$2k$	$\oplus 0$	$2k$
	или	
$2k$	$\oplus 1$	$2k+1$
$2k-2m+1$	$\oplus 0$	$2k-2m+1$

Видно, что из подмножества X_{2m-1} получается подмножество X_{2m} при использовании шаблона 10.

Аналогичным образом можно рассмотреть другие переходы из одного множества в другое.

Машина конечных состояний, показанная на рисунке А.1а не применима для множества C_0. Нужно создать отдельную машину конечных состояний с свойствами множества C_0. Множество C_0 можно разделить на два подмножества D_0 и Y_1. Переходы множества C_0 описаны с помощью машины конечных состояний, изображенной на рисунке А.1б.

С помощью машин конечных состояний, изображенных на рисунке А.1, и вероятностей изменения по шаблонам в каждом множестве определяются статистические критерии главных следов множеств до и после вложения в наименьшие значащие биты.

Более того, если наименьшие значащие биты покрывающего сообщения распределены случайным образом во временной области, то вероятность появления пикселя определенной яркости после вложения является функцией от длины секретного сообщения или, что то же самое, от вероятности вложения.

Для каждого шаблона изменений $\pi \in \{00,\ 10,\ 01,\ 11\}$ и любого подмножества $A \subseteq W$ обозначим $P(\pi, A)$ как вероятность того, что пары

образцов множества A изменятся по шаблону π в результате вложения в наименьшие значащие биты. Пусть P – это вероятность вложения секретного сообщения. Тогда доля образцов, измененных после вложения в наименьшие значащие биты равна $P/2$, где P выражено в долях. (На практике это не всегда так, отсюда и возникает погрешность результатов атаки. Эта погрешность особенно велика при больших вероятностях вложения, но при малых атака дает точные результаты.)

Принимая, что биты сообщения, вложенного в наименьшие значащие биты файла, случайно распределены во временной области, получаем вероятности изменений по шаблону:

$$\begin{cases} P'(00,P)=\left(1-\dfrac{P}{2}\right)^2, \\ P'(01,P)=P'(10,P)=\dfrac{P}{2}\left(1-\dfrac{P}{2}\right), \\ P'(11,P)=\left(\dfrac{P}{2}\right)^2. \end{cases} \qquad \text{(A.1)}$$

где P и P' выражены в долях.

Если A и B подмножества W, такие, что $A \subseteq B$. Множество A независимо по отношению к B, если для всех шаблонов изменений $\pi \in \{00, 10, 01, 11\}$ выполняется равенство $P'(\pi, A) = P'(\pi, B)$.

Если все следы подмножеств C_m независимы, то и само множество C_m независимо.

Введем обозначения множества A – если множество получено из покрывающего сообщения, и A' – если множество получено из сигнала, измененного вложением в наименьшие значащие биты.

Аналогично будем отличать пары образцов до вложения (u, v) и после вложения (u', v').

Выведем формулу для множества X_{2m-1}, воспользовавшись машиной конечных состояний для множества C_m, изображенной на рисунке А.1а и взаимосвязь вероятностей изменение по шаблонам, показанную в (А.1).

Множество $X'_{2m-1} \cup X'_{2m}$ состоит из пар образцов множества $X_{2m-1} \cup X_{2m}$, измененных по шаблонам 00 или 10, и из пар образцов множества $Y_{2m} \cup Y_{2m+1}$, измененных по шаблонам 01 или 11. Вероятность того, что произвольная пара образцов множества $X_{2m-1} \cup X_{2m}$ изменится по шаблонам 00 или 10 равна

$$\left(1-\frac{P}{2}\right)^2 + \frac{P}{2}\left(1-\frac{P}{2}\right) = 1 - P + \frac{P^2}{4} + \frac{P}{2} - \frac{P^2}{4} = 1 - \frac{P}{2}.$$

Вероятность того, что произвольная пара образцов множества $Y_{2m} \cup Y_{2m+1}$ изменится по шаблонам 01 или 11 равна

$$\left(\frac{P}{2}\right)^2 + \frac{P}{2}\left(1-\frac{P}{2}\right) = \frac{P^2}{4} + \frac{P}{2} - \frac{P^2}{4} = \frac{P}{2}.$$

Вышесказанное позволяет выразить множество $X'_{2m-1} \cup X'_{2m}$, которое численно равно $\left|X'_{2m-1}\right| + \left|X'_{2m}\right|$ через следующую формулу:

$$\left|X'_{2m-1}\right| + \left|X'_{2m}\right| = \left(\left|X_{2m-1}\right| + \left|X_{2m}\right|\right)\left(1-\frac{P}{2}\right) + \left(\left|Y_{2m}\right| + \left|Y_{2m+1}\right|\right)\frac{P}{2}. \qquad \text{(А.2)}$$

Множество $Y'_{2m} \cup Y'_{2m+1}$ состоит из пар образцов множества $Y_{2m} \cup Y_{2m+1}$, измененных по шаблонам 00 или 10, и из пар образцов множества $X_{2m-1} \cup X_{2m}$, измененных по шаблонам 01 или 11. Вероятность того, что произвольная пара образцов множества $Y_{2m} \cup Y_{2m+1}$ изменится по шаблонам 00 или 10 равна

$$\left(1-\frac{P}{2}\right)^2 + \frac{P}{2}\left(1-\frac{P}{2}\right) = 1 - P + \frac{P^2}{4} + \frac{P}{2} - \frac{P^2}{4} = 1 - \frac{P}{2}.$$

Вероятность того, что произвольная пара образцов множества $X_{2m-1} \cup X_{2m}$ изменится по шаблонам 01 или 11 равна

$$\left(\frac{P}{2}\right)^2+\frac{P}{2}\left(1-\frac{P}{2}\right)=\frac{P^2}{4}+\frac{P}{2}-\frac{P^2}{4}=\frac{P}{2}.$$

Аналогично формуле (А.2) получаем:

$$\left|Y'_{2m}\right|+\left|Y'_{2m+1}\right|=\left(\left|Y_{2m}\right|+\left|Y_{2m+1}\right|\right)\left(1-\frac{P}{2}\right)+\left(\left|X_{2m-1}\right|+\left|X_{2m}\right|\right)\frac{P}{2}. \qquad \text{(А.3)}$$

Вычтем (А.3) из (А.2):

$$\left|X'_{2m-1}\right|+\left|X'_{2m}\right|-\left|Y'_{2m}\right|-\left|Y'_{2m+1}\right|=\left(\left(\left|X_{2m-1}\right|+\left|X_{2m}\right|\right)\left(1-\frac{P}{2}\right)+\left(\left|Y_{2m}\right|+\left|Y_{2m+1}\right|\right)\frac{P}{2}\right)-$$

$$-\left(\left(\left|Y_{2m}\right|+\left|Y_{2m+1}\right|\right)\left(1-\frac{P}{2}\right)+\left(\left|X_{2m-1}\right|+\left|X_{2m}\right|\right)\frac{P}{2}\right),$$

$$\left|X'_{2m-1}\right|+\left|X'_{2m}\right|-\left|Y'_{2m}\right|-\left|Y'_{2m+1}\right|=\left(\left|X_{2m-1}\right|+\left|X_{2m}\right|\right)\left(1-\frac{P}{2}\right)+\left(\left|Y_{2m}\right|+\left|Y_{2m+1}\right|\right)\frac{P}{2}-$$

$$-\left(\left|Y_{2m}\right|+\left|Y_{2m+1}\right|\right)\left(1-\frac{P}{2}\right)-\left(\left|X_{2m-1}\right|+\left|X_{2m}\right|\right)\frac{P}{2},$$

$$\left|X'_{2m-1}\right|+\left|X'_{2m}\right|-\left|Y'_{2m}\right|-\left|Y'_{2m+1}\right|=\left|X_{2m-1}\right|+\left|X_{2m}\right|-\frac{P}{2}\left|X_{2m-1}\right|-\frac{P}{2}\left|X_{2m}\right|+\frac{P}{2}\left|Y_{2m}\right|+$$

$$+\frac{P}{2}\left|Y_{2m+1}\right|-\left|Y_{2m}\right|-\left|Y_{2m+1}\right|+\frac{P}{2}\left|Y_{2m}\right|+\frac{P}{2}\left|Y_{2m+1}\right|-\frac{P}{2}\left|X_{2m-1}\right|+\frac{P}{2}\left|X_{2m}\right|,$$

$$\left|X'_{2m-1}\right|+\left|X'_{2m}\right|-\left|Y'_{2m}\right|-\left|Y'_{2m+1}\right|=(1-P)\left(\left|X_{2m-1}\right|+\left|X_{2m}\right|-\left|Y_{2m}\right|-\left|Y_{2m+1}\right|\right). \qquad \text{(А.4)}$$

Множество $X'_{2m-1}\cup Y'_{2m}$ состоит из пар образцов множества $X_{2m-1}\cup Y_{2m}$, измененных по шаблонам 00 или 01, и из пар образцов множества $X_{2m}\cup Y_{2m+1}$, измененных по шаблонам 10 или 11. Вероятность того, что произвольная пара образцов множества $X_{2m-1}\cup Y_{2m}$ изменится по шаблонам 00 или 01 равна

$$\left(1-\frac{P}{2}\right)^2+\frac{P}{2}\left(1-\frac{P}{2}\right)=1-P+\frac{P^2}{4}+\frac{P}{2}-\frac{P^2}{4}=1-\frac{P}{2}.$$

Вероятность того, что произвольная пара образцов множества $X_{2m}\cup Y_{2m+1}$ изменится по шаблонам 10 или 11 равна

$$\left(\frac{P}{2}\right)^2+\frac{P}{2}\left(1-\frac{P}{2}\right)=\frac{P^2}{4}+\frac{P}{2}-\frac{P^2}{4}=\frac{P}{2}.$$

Получаем:

$$\left|X'_{2m-1}\right|+\left|Y'_{2m}\right|=\left(\left|X_{2m-1}\right|+\left|Y_{2m}\right|\right)\left(1-\frac{P}{2}\right)+\left(\left|X_{2m}\right|+\left|Y_{2m+1}\right|\right)\frac{P}{2}. \qquad \text{(A.5)}$$

Множество $Y'_{2m+1}\cup X'_{2m}$ состоит из пар образцов множества $Y_{2m+1}\cup X_{2m}$, измененных по шаблонам 00 или 01, и из пар образцов множества $Y_{2m}\cup X_{2m-1}$, измененных по шаблонам 10 или 11. Вероятность того, что произвольная пара образцов множества $Y_{2m+1}\cup X_{2m}$ изменится по шаблонам 00 или 01 равна

$$\left(1-\frac{P}{2}\right)^2+\frac{P}{2}\left(1-\frac{P}{2}\right)=1-P+\frac{P^2}{4}+\frac{P}{2}-\frac{P^2}{4}=1-\frac{P}{2}.$$

Вероятность того, что произвольная пара образцов множества $Y_{2m}\cup X_{2m-1}$ изменится по шаблонам 10 или 11 равна

$$\left(\frac{P}{2}\right)^2+\frac{P}{2}\left(1-\frac{P}{2}\right)=\frac{P^2}{4}+\frac{P}{2}-\frac{P^2}{4}=\frac{P}{2}.$$

Получаем:

$$\left|Y'_{2m+1}\right|+\left|X'_{2m}\right|=\left(\left|Y_{2m+1}\right|+\left|X_{2m}\right|\right)\left(1-\frac{P}{2}\right)+\left(\left|X_{2m-1}\right|+\left|Y_{2m}\right|\right)\frac{P}{2}. \qquad \text{(A.6)}$$

Вычтем (A.6) из (A.5):

$$\left|X'_{2m-1}\right|+\left|Y'_{2m}\right|-\left|Y'_{2m+1}\right|-\left|X'_{2m}\right|=\left(\left(\left|X_{2m-1}\right|+\left|Y_{2m}\right|\right)\left(1-\frac{P}{2}\right)+\left(\left|X_{2m}\right|+\left|Y_{2m+1}\right|\right)\frac{P}{2}\right)-$$

$$-\left(\left(\left|Y_{2m+1}\right|+\left|X_{2m}\right|\right)\left(1-\frac{P}{2}\right)+\left(\left|X_{2m-1}\right|+\left|Y_{2m}\right|\right)\frac{P}{2}\right),$$

$$\left|X'_{2m-1}\right|+\left|Y'_{2m}\right|-\left|Y'_{2m+1}\right|-\left|X'_{2m}\right|=\left(\left|X_{2m-1}\right|+\left|Y_{2m}\right|\right)\left(1-\frac{P}{2}\right)+\left(\left|X_{2m}\right|+\left|Y_{2m+1}\right|\right)\frac{P}{2}-$$

$$-\left(\left|Y_{2m+1}\right|+\left|X_{2m}\right|\right)\left(1-\frac{P}{2}\right)-\left(\left|X_{2m-1}\right|+\left|Y_{2m}\right|\right)\frac{P}{2},$$

$$\left|X'_{2m-1}\right|+\left|Y'_{2m}\right|-\left|Y'_{2m+1}\right|-\left|X'_{2m}\right|=\left|X_{2m-1}\right|+\left|Y_{2m}\right|-\frac{P}{2}\left|X_{2m-1}\right|-\frac{P}{2}\left|Y_{2m}\right|+\frac{P}{2}\left|X_{2m}\right|+$$

$$+\frac{P}{2}\left|Y_{2m+1}\right|-\left|Y_{2m+1}\right|-\left|X_{2m}\right|+\frac{P}{2}\left|Y_{2m+1}\right|+\frac{P}{2}\left|X_{2m}\right|-\frac{P}{2}\left|X_{2m-1}\right|+\frac{P}{2}\left|Y_{2m}\right|,$$

$$\left|X'_{2m-1}\right|+\left|Y'_{2m}\right|-\left|Y'_{2m+1}\right|-\left|X'_{2m}\right|=(1-P)\left(\left|X_{2m-1}\right|+\left|Y_{2m}\right|-\left|Y_{2m+1}\right|-\left|X_{2m}\right|\right). \quad \text{(A.7)}$$

Сложив (A.4) и (A.7), получим:

$$\left|X'_{2m-1}\right|+\left|X'_{2m}\right|-\left|Y'_{2m}\right|-\left|Y'_{2m+1}\right|+\left|X'_{2m-1}\right|+\left|Y'_{2m}\right|-\left|Y'_{2m+1}\right|-\left|X'_{2m}\right|=$$

$$=(1-P)\left(\left|X_{2m-1}\right|+\left|X_{2m}\right|-\left|Y_{2m}\right|-\left|Y_{2m+1}\right|\right)+(1-P)\left(\left|X_{2m-1}\right|+\left|Y_{2m}\right|-\left|Y_{2m+1}\right|-\left|X_{2m}\right|\right),$$

$$2\left|X'_{2m-1}\right|-2\left|Y'_{2m+1}\right|=(1-P)\left(\left|X_{2m-1}\right|+\left|X_{2m}\right|-\left|Y_{2m}\right|-\left|Y_{2m+1}\right|+\left|X_{2m-1}\right|+\left|Y_{2m}\right|-\right.$$

$$\left.-\left|Y_{2m+1}\right|-\left|X_{2m}\right|\right),$$

$$2\left|X'_{2m-1}\right|-2\left|Y'_{2m+1}\right|=(1-P)\left(2\left|X_{2m-1}\right|-2\left|Y_{2m+1}\right|\right),$$

$$\left|X'_{2m-1}\right|-\left|Y'_{2m+1}\right|=(1-P)\left(\left|X_{2m-1}\right|-\left|Y_{2m+1}\right|\right). \quad \text{(A.8)}$$

Множество $X'_{2m-1}\cup Y'_{2m+1}$ состоит из пар образцов множества $X_{2m-1}\cup Y_{2m+1}$, измененных по шаблонам 00 или 11, и из пар образцов множества D_{2m} (помним, что $D_{2m}=X_{2m}\cup Y_{2m}$), измененных по шаблонам 01 или 10. Вероятность того, что произвольная пара образцов множества $X_{2m-1}\cup Y_{2m+1}$ изменится по шаблонам 00 или 11 равна

$$\left(1-\frac{P}{2}\right)^2+\left(\frac{P}{2}\right)^2.$$

Вероятность того, что произвольная пара образцов множества D_{2m} изменится по шаблонам 01 или 10 равна

$$P\left(1-\frac{P}{2}\right).$$

Тогда получаем:

$$\left|X'_{2m-1}\right|+\left|Y'_{2m+1}\right|=\left(\left(1-\frac{P}{2}\right)^2+\left(\frac{P}{2}\right)^2\right)\left(\left|X_{2m-1}\right|+\left|Y_{2m+1}\right|\right)+P\left(1-\frac{P}{2}\right)\left|D_{2m}\right|. \quad \text{(A.9)}$$

Как уже было ранее сказано

$$\left|C_m\right|=\left|X_{2m}\right|+\left|Y_{2m}\right|+\left|X_{2m-1}\right|+\left|Y_{2m+1}\right|,$$

$$\left|C_m\right|=\left|D_{2m}\right|+\left|X_{2m-1}\right|+\left|Y_{2m+1}\right|. \quad \text{(A.10)}$$

Следовательно

$$\left|D_{2m}\right|=\left|C_m\right|-\left|X_{2m-1}\right|-\left|Y_{2m+1}\right|. \quad \text{(A.11)}$$

Подставим выражение (A.10) в (A.9):

$$\left|X'_{2m-1}\right|+\left|Y'_{2m+1}\right|=\left(\left(1-\frac{P}{2}\right)^2+\left(\frac{P}{2}\right)^2\right)\left(\left|X_{2m-1}\right|+\left|Y_{2m+1}\right|\right)+P\left(1-\frac{P}{2}\right)\left(\left|C_m\right|-\right.$$

$$\left.-\left|X_{2m-1}\right|-\left|Y_{2m+1}\right|\right),$$

$$\left|X'_{2m-1}\right|+\left|Y'_{2m+1}\right|=\left(1-P+\frac{P^2}{4}+\frac{P^2}{4}\right)\left(\left|X_{2m-1}\right|+\left|Y_{2m+1}\right|\right)+\left(P-\frac{P^2}{2}\right)\left(\left|C_m\right|-\right.$$

$$\left.-\left|X_{2m-1}\right|-\left|Y_{2m+1}\right|\right),$$

$$\left|X'_{2m-1}\right|+\left|Y'_{2m+1}\right|=\left|X_{2m-1}\right|+\left|Y_{2m+1}\right|-P\left|X_{2m-1}\right|-P\left|Y_{2m+1}\right|+\frac{P^2}{2}\left|X_{2m-1}\right|+\frac{P^2}{2}\left|Y_{2m+1}\right|-$$

$$-P\left|X_{2m-1}\right|-P\left|Y_{2m+1}\right|+\frac{P^2}{2}\left|X_{2m-1}\right|+\frac{P^2}{2}\left|Y_{2m+1}\right|+P\left(1-\frac{P}{2}\right)\left|C_m\right|,$$

$$\left|X'_{2m-1}\right|+\left|Y'_{2m+1}\right|=(1-p)^2\left(\left|X_{2m-1}\right|+\left|Y_{2m+1}\right|\right)+P\left(1-\frac{P}{2}\right)\left|C_m\right|. \quad \text{(A.12)}$$

Умножим обе части (A.8) на $(1-P)$:

$$(1-P)\left(\left|X'_{2m-1}\right|-\left|Y'_{2m+1}\right|\right)=(1-P)^2\left(\left|X_{2m-1}\right|-\left|Y_{2m+1}\right|\right). \quad \text{(A.13)}$$

Складывая (A.12) и (A.13), получим:

$$\left|X'_{2m-1}\right|+\left|Y'_{2m+1}\right|+(1-P)\left(\left|X'_{2m-1}\right|-\left|Y'_{2m+1}\right|\right)=(1-P)^2\left(\left|X_{2m-1}\right|+\left|Y_{2m+1}\right|\right)+$$

$$+P\left(1-\frac{P}{2}\right)\left|C_m\right|+(1-P)^2\left(\left|X_{2m-1}\right|-\left|Y_{2m+1}\right|\right),$$

$$\left|X'_{2m-1}\right|+\left|Y'_{2m+1}\right|+\left|X'_{2m-1}\right|-\left|Y'_{2m+1}\right|-P\left|X'_{2m-1}\right|+P\left|Y'_{2m+1}\right|=(1-P)^2\left|X_{2m-1}\right|+$$

$$+(1-P)^2\left|Y_{2m+1}\right|+P\left(1-\frac{P}{2}\right)\left|C_m\right|+(1-P)^2\left|X_{2m-1}\right|-(1-P)^2\left|Y_{2m+1}\right|,$$

$$(2-P)\left|X'_{2m-1}\right|+P\left|Y'_{2m+1}\right|=2(1-P)^2\left|X_{2m-1}\right|+P\left(1-\frac{P}{2}\right)\left|C_m\right|. \qquad \text{(A.14)}$$

Поскольку множество C_m независимо, из графа машины конечных состояний для C_m и (А.10) видно, что

$$\left|C_m\right|=\left|D'_{2m}\right|+\left|X'_{2m-1}\right|+\left|Y'_{2m+1}\right|.$$

Отсюда

$$\left|Y'_{2m+1}\right|=\left|C_m\right|-\left|D'_{2m}\right|-\left|X'_{2m-1}\right|. \qquad \text{(A.15)}$$

Подставляя (А.15) в (А.14) и получаем:

$$(2-P)\left|X'_{2m-1}\right|+P\left(\left|C_m\right|-\left|D'_{2m}\right|-\left|X'_{2m-1}\right|\right)=2(1-P)^2\left|X_{2m-1}\right|+P\left(1-\frac{P}{2}\right)\left|C_m\right|,$$

$$2(1-P)^2\left|X_{2m-1}\right|=(2-P)\left|X'_{2m-1}\right|+P\left(\left|C_m\right|-\left|D'_{2m}\right|-\left|X'_{2m-1}\right|\right)-P\left(1-\frac{P}{2}\right)\left|C_m\right|,$$

$$2(1-P)^2\left|X_{2m-1}\right|=2\left|X'_{2m-1}\right|-P\left|X'_{2m-1}\right|+P\left|C_m\right|-P\left|D'_{2m}\right|-P\left|X'_{2m-1}\right|-$$

$$-P\left|C_m\right|+\frac{P^2}{2}\left|C_m\right|,$$

$$2(1-P)^2\left|X_{2m-1}\right|=2\left|X'_{2m-1}\right|-P\left|X'_{2m-1}\right|-P\left|D'_{2m}\right|-P\left|X'_{2m-1}\right|+\frac{P^2}{2}\left|C_m\right|,$$

$$2(1-P)^2\left|X_{2m-1}\right|=2\frac{P^2}{4}\left|C_m\right|-2\frac{P}{2}\left(\left|D'_{2m}\right|+2\left|X'_{2m-1}\right|\right)+2\left|X'_{2m-1}\right|,$$

$$(1-P)^2\left|X_{2m-1}\right|=\frac{P^2}{4}\left|C_m\right|-\frac{P}{2}\left(\left|D'_{2m}\right|+2\left|X'_{2m-1}\right|\right)+\left|X'_{2m-1}\right|. \qquad \text{(A.16)}$$

В (А.16) $1 \le m \le 2^{t-1} - 1$. Данное равенство является одним из основных принципов, на котором построена статистическая атака ПВА.

Теперь выведем формулу для множества Y_1, воспользовавшись графом машины конечных состояний для множества C_0, изображенной на рисунке А.1б и взаимосвязь вероятностей изменений по шаблонам, показанную в (А.1).

Множество $D_0^{/}$ состоит из пар образцов множества Y_1, измененных по шаблонам 01 или 10, и из пар образцов множества D_0, измененных по шаблонам 00 или 11. Вероятность того, что произвольная пара образцов множества Y_1 изменится по шаблонам 01 или 10 равна

$$P\left(1 - \frac{P}{2}\right).$$

Вероятность того, что произвольная пара образцов множества D_0 изменится по шаблонам 00 или 11 равна

$$\left(\frac{P}{2}\right)^2 + \left(1 - \frac{P}{2}\right)^2 = \frac{P^2}{4} + 1 - P + \frac{P^2}{4} = \frac{P^2}{2} - P + 1.$$

Тогда получаем:

$$\left|D_0^{/}\right| = P\left(1 - \frac{P}{2}\right)\left|Y_1\right| + \left(\frac{P^2}{2} - P + 1\right)\left|D_0\right|. \qquad \text{(А.17)}$$

Множество $Y_1^{/}$ состоит из пар образцов множества D_0, измененных по шаблонам 01 или 10, и из пар образцов множества Y_1, измененных по шаблонам 00 или 11. Вероятность того, что произвольная пара образцов множества D_0 изменится по шаблонам 01 или 10 будет равна

$$P\left(1 - \frac{P}{2}\right).$$

Вероятность того, что произвольная пара образцов множества Y_1 изменится по шаблонам 00 или 11 будет равна

$$\left(\frac{P}{2}\right)^2+\left(1-\frac{P}{2}\right)^2=\frac{P^2}{4}+1-P+\frac{P^2}{4}=\frac{P^2}{2}-P+1.$$

Далее получаем:

$$\left|Y_1^{/}\right|=P\left(1-\frac{P}{2}\right)\left|D_0\right|+\left(\frac{P^2}{2}-P+1\right)\left|Y_1\right|. \qquad \text{(A.18)}$$

Вычитая (A.18) из (A.17), получаем

$$\left|D_0^{/}\right|-\left|Y_1^{/}\right|=P\left(1-\frac{P}{2}\right)\left|Y_1\right|+\left(\frac{P^2}{2}-P+1\right)\left|D_0\right|-$$

$$-\left(P\left(1-\frac{P}{2}\right)\left|D_0\right|+\left(\frac{P^2}{2}-P+1\right)\left|Y_1\right|\right),$$

$$\left|D_0^{/}\right|-\left|Y_1^{/}\right|=P\left|Y_1\right|-\frac{P^2}{2}\left|Y_1\right|+\frac{P^2}{2}\left|D_0\right|-P\left|D_0\right|+\left|D_0\right|-P\left|D_0\right|+\frac{P^2}{2}\left|D_0\right|-\frac{P^2}{2}\left|Y_1\right|+$$

$$+P\left|Y_1\right|-\left|Y_1\right|,$$

$$\left|D_0^{/}\right|-\left|Y_1^{/}\right|=2P\left|Y_1\right|-P^2\left|Y_1\right|+P^2\left|D_0\right|-2P\left|D_0\right|+\left|D_0\right|-\left|Y_1\right|,$$

$$\left|D_0^{/}\right|-\left|Y_1^{/}\right|=\left|D_0\right|(1-P)^2-(1-P)^2\left|Y_1\right|. \qquad \text{(A.19)}$$

Было доказано, что

$$\left|C_0\right|=\left|D_0\right|+\left|Y_1\right|. \qquad \text{(A.20)}$$

Отсюда, следует, что

$$\left|D_0\right|=\left|C_0\right|-\left|Y_1\right|. \qquad \text{(A.21)}$$

Подставляя (A.21) в (A.19), получим

$$\left|D_0^{/}\right|-\left|Y_1^{/}\right|=\left(\left|C_0\right|-\left|Y_1\right|\right)(1-P)^2-(1-P)^2\left|Y_1\right|,$$

$$\left|D_0^{/}\right|-\left|Y_1^{/}\right|=(1-P)^2\left|C_0\right|-(1-P)^2\left|Y_1\right|-(1-P)^2\left|Y_1\right|,$$

$$\left|D_0^{/}\right|-\left|Y_1^{/}\right|=(1-P)^2\left|C_0\right|-2(1-P)^2\left|Y_1\right|,$$

$$2(1-P)^2\left|Y_1\right|=(1-P)^2\left|C_0\right|-\left|D_0^{/}\right|+\left|Y_1^{/}\right|. \qquad \text{(A.22)}$$

Поскольку множество C_0 независимо, из графа машины конечных состояний для C_0 и (А.20) видно, что

$$|C_0| = |D_0'| + |Y_1'|. \qquad \text{(А.23)}$$

Подставляя (А.23) в (А.22), получаем

$$2(1-P)^2|Y_1| = (1-P)^2\left(|D_0'| + |Y_1'|\right) - |D_0'| + |Y_1'|,$$

$$2(1-P)^2|Y_1| = |D_0'| + |Y_1'| - 2P|D_0'| - 2P|Y_1'| + P^2|D_0'| + P^2|Y_1'| - |D_0'| + |Y_1'|,$$

$$2(1-P)^2|Y_1| = P^2\left(|D_0'| + |Y_1'|\right) - 2P\left(|D_0'| + |Y_1'|\right) + 2|Y_1'|.$$

С учетом формулы (А.23) получаем

$$2(1-P)^2|Y_1| = P^2|C_0| - 2P\left(|D_0'| + |Y_1'|\right) + 2|Y_1'|,$$

$$2(1-P)^2|Y_1| = 2\frac{P^2}{2}|C_0| - 2\frac{P}{2}\left(2|D_0'| + 2|Y_1'|\right) + 2|Y_1'|,$$

$$(1-P)^2|Y_1| = \frac{P^2}{2}|C_0| - \frac{P}{2}\left(2|D_0'| + 2|Y_1'|\right) + |Y_1'|. \qquad \text{(А.24)}$$

Равенство (А.24) является вторым основным принципом, на котором построена статистическая атака с учетом корреляции пикселей.

Уже было показано, что $\mathrm{E}\{|X_{2m+1}|\} = \mathrm{E}\{|Y_{2m+1}|\}$. Следовательно $\mathrm{E}\{|X_1|\} = \mathrm{E}\{|Y_1|\}$.

Подставляя в (А.16) $m = 1$, получим

$$(1-P)^2|X_1| = \frac{P^2}{4}|C_1| - \frac{P}{2}\left(|D_2'| + 2|X_1'|\right) + |X_1'|. \qquad \text{(А.25)}$$

Видно, что левые части (А.24) и (А.25) равны, а, следовательно, равны и правые части:

$$\frac{P^2}{2}|C_0| - \frac{P}{2}\left(2|D_0'| + 2|Y_1'|\right) + |Y_1'| = \frac{P^2}{4}|C_1| - \frac{P}{2}\left(|D_2'| + 2|X_1'|\right) + |X_1'|,$$

$$\frac{P^2}{4}\left(2|C_0| - |C_1|\right) - \frac{P}{2}\left(2|D_0'| + 2|Y_1'| - |D_2'| - 2|X_1'|\right) + |Y_1'| - |X_1'| = 0. \qquad \text{(А.26)}$$

Формула (А.26) является уравнением, где вероятность вложения P выступает как неизвестное. Множества, используемые в этой формуле, легко находятся при наличии изображения.

Множество $|C_0|$ состоит из пар образцов, которые, если откинуть наименьшие значащие биты, одинаковых во всех остальных позициях.

Множество $|C_1|$ состоит из пар образцов, которые, отличаются предпоследним битом.

Данные множества не учитывают наименьшие значащие биты.

Множество $|D_0|$ состоит из пар образцов, одинаковых во всех битах.

Множество $|D_2|$ состоит из пар образцов вида $(u, u+2)$ или $(u+2, u)$, то есть яркости пикселей отличаются на 2 единицы.

Множество $|Y_1|$ состоит из пар образцов вида $(2k, 2k+1)$ или $(2k+1, 2k)$, то есть из яркости пикселей отличается в наименьших значащих битах.

Множество $|X_1|$ состоит из пар образцов вида $(2k, 2k-1)$ или $(2k-1, 2k)$, то есть из яркости пикселей отличается в двух последних битах.

Из двух корней уравнения (А.26) выбираем наименьший:

$$P = \frac{\frac{2|D_0'| + 2|Y_1'| - |D_2'| - 2|X_1'|}{2} - \sqrt{\left(-\frac{2|D_0'| + 2|Y_1'| - |D_2'| - 2|X_1'|}{2}\right)^2 - \frac{2|C_0| - |C_1|}{2}}}{\frac{2|C_0| - |C_1|}{4}} - \frac{\sqrt{-4\frac{2|C_0| - |C_1|}{4}\left(|Y_1'| - |X_1'|\right)}}{\frac{2|C_0| - |C_1|}{4}}. \quad \text{(А.27)}$$

Величина P, полученная при расчете по формуле (А.27) и будет являться результатом статистической атаки методом ПВА.

Приложение Б

Текст программы с алгоритмами исследуемых методов стегоанализа

```
#include <stdio.h>
#include <stdlib.h>
#include <conio.h>
#include "stdafx.h"
#include "test.h"
#include "testDlg.h"
#include "KOD.h"
#include "HI.h"
#include "KOD1.h"
#include "ORR.h"
#include "table.h"
#include "math.h"
#include "stat.h"
#include "VLOG.h"
#include "SRIF.h"
#include "first.h"
#include "second.h"
#include "nuli.h"
#include "sos.h"
#ifdef _DEBUG
#define new DEBUG_NEW
#undef THIS_FILE
static char THIS_FILE[] = __FILE__;
#endif
#define COLOR_BCKGR     RGB(224, 0, 0)
class CAboutDlg : public CDialog
{
public:
     CAboutDlg();
     enum { IDD = IDD_ABOUTBOX };
        protected:
     virtual void DoDataExchange(CDataExchange* pDX);
     DECLARE_MESSAGE_MAP()
};
CAboutDlg::CAboutDlg() : CDialog(CAboutDlg::IDD)
{}
void CAboutDlg::DoDataExchange(CDataExchange* pDX)
{
```

```
    CDialog::DoDataExchange(pDX);
}
BEGIN_MESSAGE_MAP(CAboutDlg, CDialog)
END_MESSAGE_MAP()
CTestDlg::CTestDlg(CWnd* pParent /*=NULL*/)
    : CDialog(CTestDlg::IDD, pParent)
{
    m_strText = _T("");
    m_Strksy1 = _T("");
    m_Strksy2 = _T("");
    m_x1 = _T("");
    m_x10 = _T("");
    m_x2 = _T("");
    m_x3 = _T("");
    m_x4 = _T("");
    m_x5 = _T("");
    m_x6 = _T("");
    m_x7 = _T("");
    m_x8 = _T("");
    m_x9 = _T("");
    m_name1 = _T("");
    m_name2 = _T("");
    m_name3 = _T("");
    m_name4 = _T("");
    m_na1 = _T("");
    m_na2 = _T("");
    m_na3 = _T("");
    m_na4 = _T("");
    m_hi2 = _T("");
    m_name21 = _T("");
    m_name22 = _T("");
    m_na21 = _T("");
    m_na22 = _T("");
    m_name23 = _T("");
    m_name24 = _T("");
    m_na23 = _T("");
    m_na24 = _T("");
    m_hIcon = AfxGetApp()->LoadIcon(IDR_MAINFRAME);
}
void CTestDlg::DoDataExchange(CDataExchange* pDX)
{
    CDialog::DoDataExchange(pDX);
    DDX_Control(pDX, IDC_name24, m_na24st);
    DDX_Control(pDX, IDC_name23, m_na23st);
```

```
DDX_Control(pDX, IDC_fname24, m_name24st);
DDX_Control(pDX, IDC_fname23, m_name23st);
DDX_Control(pDX, IDC_name22, m_na22st);
DDX_Control(pDX, IDC_name21, m_na21st);
DDX_Control(pDX, IDC_fname22, m_name22st);
DDX_Control(pDX, IDC_fname21, m_name21st);
DDX_Control(pDX, IDC_HI2, m_hi2st);
DDX_Control(pDX, IDC_name4, m_na4st);
DDX_Control(pDX, IDC_name3, m_na3st);
DDX_Control(pDX, IDC_name2, m_na2st);
DDX_Control(pDX, IDC_name1, m_na1st);
DDX_Control(pDX, IDC_fname4, m_name4st);
DDX_Control(pDX, IDC_fname3, m_name3st);
DDX_Control(pDX, IDC_fname2, m_name2st);
DDX_Control(pDX, IDC_fname1, m_name1st);
DDX_Control(pDX, IDC_X9, m_x9static);
DDX_Control(pDX, IDC_X8, m_x8static);
DDX_Control(pDX, IDC_X7, m_x7static);
DDX_Control(pDX, IDC_X6, m_x6static);
DDX_Control(pDX, IDC_X5, m_x5static);
DDX_Control(pDX, IDC_X4, m_x4static);
DDX_Control(pDX, IDC_X3, m_x3static);
DDX_Control(pDX, IDC_X2, m_x2static);
DDX_Control(pDX, IDC_X10, m_x10static);
DDX_Control(pDX, IDC_X1, m_x1static);
DDX_Control(pDX, IDC_KSY2, m_wndksy2);
DDX_Control(pDX, IDC_KSY1, m_wndksy1);
DDX_Control(pDX, IDC_TEXT, m_wndText);
DDX_Control(pDX, IDC_BUTTON1, m_filename);
DDX_Text(pDX, IDC_TEXT, m_strText);
DDV_MaxChars(pDX, m_strText, 200);
DDX_Text(pDX, IDC_KSY1, m_Strksy1);
DDX_Text(pDX, IDC_KSY2, m_Strksy2);
DDX_Text(pDX, IDC_X1, m_x1);
DDX_Text(pDX, IDC_X10, m_x10);
DDX_Text(pDX, IDC_X2, m_x2);
DDX_Text(pDX, IDC_X3, m_x3);
DDX_Text(pDX, IDC_X4, m_x4);
DDX_Text(pDX, IDC_X5, m_x5);
DDX_Text(pDX, IDC_X6, m_x6);
DDX_Text(pDX, IDC_X7, m_x7);
DDX_Text(pDX, IDC_X8, m_x8);
DDX_Text(pDX, IDC_X9, m_x9);
DDX_Text(pDX, IDC_fname1, m_name1);
```

```
    DDX_Text(pDX, IDC_fname2, m_name2);
    DDX_Text(pDX, IDC_fname3, m_name3);
    DDX_Text(pDX, IDC_fname4, m_name4);
    DDX_Text(pDX, IDC_name1, m_na1);
    DDX_Text(pDX, IDC_name2, m_na2);
    DDX_Text(pDX, IDC_name3, m_na3);
    DDX_Text(pDX, IDC_name4, m_na4);
    DDX_Text(pDX, IDC_HI2, m_hi2);
    DDV_MaxChars(pDX, m_hi2, 1000000);
    DDX_Text(pDX, IDC_fname21, m_name21);
    DDX_Text(pDX, IDC_fname22, m_name22);
    DDX_Text(pDX, IDC_name21, m_na21);
    DDX_Text(pDX, IDC_name22, m_na22);
    DDX_Text(pDX, IDC_fname23, m_name23);
    DDX_Text(pDX, IDC_fname24, m_name24);
    DDX_Text(pDX, IDC_name23, m_na23);
    DDX_Text(pDX, IDC_name24, m_na24);
}
BEGIN_MESSAGE_MAP(CTestDlg, CDialog)
    ON_WM_SYSCOMMAND()
    ON_WM_PAINT()
    ON_WM_QUERYDRAGICON()
    ON_BN_CLICKED(IDC_BUTTON2, OnButton2)
    ON_BN_CLICKED(IDC_BUTTON1, OnButton1)
    ON_BN_CLICKED(IDC_BUTTON3, OnButton3)
    ON_BN_CLICKED(IDC_BUTTON4, OnButton4)
    ON_BN_CLICKED(IDC_BUTTON5, OnButton5)
    ON_BN_CLICKED(IDC_BUTTON6, OnButton6)
    ON_BN_CLICKED(IDC_BUTTON7, OnButton7)
    ON_BN_CLICKED(IDC_BUTTON8, OnButton8)
    ON_BN_CLICKED(IDC_BUTTON9, OnButton9)
    ON_BN_CLICKED(IDC_BUTTON10, OnButton10)
    ON_BN_CLICKED(IDC_BUTTON11, OnButton11)
    ON_BN_CLICKED(IDC_BUTTON12, OnButton12)
    END_MESSAGE_MAP()
BOOL CTestDlg::OnInitDialog()
{
    CDialog::OnInitDialog();
    ASSERT((IDM_ABOUTBOX & 0xFFF0) == IDM_ABOUTBOX);
    ASSERT(IDM_ABOUTBOX < 0xF000);
    CMenu* pSysMenu = GetSystemMenu(FALSE);
    if (pSysMenu != NULL)
    {
       CString strAboutMenu;
```

```
        strAboutMenu.LoadString(IDS_ABOUTBOX);
        if (!strAboutMenu.IsEmpty())
        {
                pSysMenu->AppendMenu(MF_SEPARATOR);
                pSysMenu->AppendMenu(MF_STRING, IDM_ABOUTBOX,
strAboutMenu);
        }
    }
    SetIcon(m_hIcon, TRUE);                 // Set big icon
    SetIcon(m_hIcon, FALSE);        // Set small icon
    return TRUE;  // return TRUE  unless you set the focus to a control
}
void CTestDlg::OnSysCommand(UINT nID, LPARAM lParam)
{
    if ((nID & 0xFFF0) == IDM_ABOUTBOX)
    {
       CAboutDlg dlgAbout;
       dlgAbout.DoModal();
    }
    else
    {
       CDialog::OnSysCommand(nID, lParam);
    }
}
HANDLE BitMap;
CString Str;
CString Str00;
char  strnew[200],strnewat[200], strnewat1[200],str0[200], str00[200], strnewatst[200],
strnewatst1[200];
char strcl[200],strcl1[200];
HDC hdc, hdc1;
BITMAPINFO BMI;
BITMAPINFOHEADER BH;
int fl1=0,fl2=0,fl3=0, hi2z[1],filex=0,fl4=0, flhi0=0,fl10;
float Ebrid, Ebrid1 ;
char strname[22][200], strhi[22][200], strpro[22][200];
CDC *DC;
HANDLE BitMap1;
BITMAPINFO BMI1;
BITMAPINFOHEADER BH1;
TABLE tab;
int flagfi,flagfi1=0;
    HDC hdc2,hdc3,hdc4;
    HANDLE BitMap2,BitMap3, BitMap4;
```

```
    BITMAPINFO BMI2,BMI3,BMI4;
    BITMAPINFOHEADER BH2,BH3,BH4;
void CTestDlg::OnPaint()
{
    if (IsIconic())
    {
       CPaintDC dc(this); // device context for painting
       SendMessage(WM_ICONERASEBKGND, (WPARAM) dc.GetSafeHdc(), 0);
       int cxIcon = GetSystemMetrics(SM_CXICON);
       int cyIcon = GetSystemMetrics(SM_CYICON);
       CRect rect;
       GetClientRect(&rect);
       int x = (rect.Width() - cxIcon + 1) / 2;
       int y = (rect.Height() - cyIcon + 1) / 2;
       dc.DrawIcon(x, y, m_hIcon);
    }
    else
    {
       CDialog::OnPaint();
       if(fl1==1)
    {
    BitMap=LoadImage(GetModuleHandle(NULL),Str,IMAGE_BITMAP,0,0,LR_LO
ADFROMFILE);
    DC=GetDC();
    hdc=::CreateCompatibleDC(DC->m_hDC);
    ::SelectObject(hdc,BitMap);
    BitBlt(DC->m_hDC,150,50,300,200,hdc,0,0,SRCCOPY);
    BH.biSize=sizeof(BH);
    BH.biWidth=300;
    BH.biHeight=200;
    BH.biPlanes=1;
    BH.biBitCount=8;
    BH.biCompression=BI_RGB;
    BH.biSizeImage=0;
    BMI.bmiHeader=BH;
    ::DeleteDC(hdc);
    DeleteObject(BitMap);
    ReleaseDC(DC);
    }
    if(fl2==1)
    {
    BitMap1=LoadImage(GetModuleHandle(NULL),strnew,IMAGE_BITMAP,0,0,LR
_LOADFROMFILE);
    DC=GetDC();
```

```
    hdc1=::CreateCompatibleDC(DC->m_hDC);
    ::SelectObject(hdc1,BitMap1);
    BitBlt(DC->m_hDC,150,315,300,200,hdc1,0,0,SRCCOPY);
    BH1.biSize=sizeof(BH1);
    BH1.biWidth=300;
    BH1.biHeight=200;
    BH1.biPlanes=1;
    BH1.biBitCount=8;
    BH1.biCompression=BI_RGB;
    BH1.biSizeImage=0;
    BMI1.bmiHeader=BH1;
    ::DeleteDC(hdc1);
    DeleteObject(BitMap1);
    ReleaseDC(DC);
    }
    if(fl3==1)
    {
    BitMap2=LoadImage(GetModuleHandle(NULL),strnewat,IMAGE_BITMAP,0,0,L
R_LOADFROMFILE);
    DC=GetDC();
    hdc2=::CreateCompatibleDC(DC->m_hDC);
    ::SelectObject(hdc2,BitMap2);
    BitBlt(DC->m_hDC,470,50,300,200,hdc2,0,0,SRCCOPY);
    BH2.biSize=sizeof(BH2);
    BH2.biWidth=300;
    BH2.biHeight=200;
    BH2.biPlanes=1;
    BH2.biBitCount=8;
    BH2.biCompression=BI_RGB;
    BH2.biSizeImage=0;
    BMI2.bmiHeader=BH2;
    ::DeleteDC(hdc2);
    DeleteObject(BitMap2);
    ReleaseDC(DC);
    }
    if(fl4==1)
    {
    BitMap3=LoadImage(GetModuleHandle(NULL),strnewat1,IMAGE_BITMAP,0,0,
LR_LOADFROMFILE);
    DC=GetDC();
    hdc3=::CreateCompatibleDC(DC->m_hDC);
    ::SelectObject(hdc3,BitMap3);
    BitBlt(DC->m_hDC,470,315,300,200,hdc3,0,0,SRCCOPY);
    BH3.biSize=sizeof(BH3);
```

```
    BH3.biWidth=300;
    BH3.biHeight=200;
    BH3.biPlanes=1;
    BH3.biBitCount=8;
    BH3.biCompression=BI_RGB;
    BH3.biSizeImage=0;
    BMI3.bmiHeader=BH3;
    ::DeleteDC(hdc3);
    DeleteObject(BitMap3);
    ReleaseDC(DC);
    }
    }
}
HCURSOR CTestDlg::OnQueryDragIcon()
{
    return (HCURSOR) m_hIcon;
}
void CTestDlg::OnButton1() //открытиефайла
{
    ZeroMemory(strnew, sizeof(strnew));
    ZeroMemory(strnewat, sizeof(strnewat));
    ZeroMemory(  strnewat1, sizeof(  strnewat1));
    ZeroMemory(str0, sizeof(str0));
    ZeroMemory(  strnewatst, sizeof( strnewatst));
    ZeroMemory(  strnewatst1, sizeof(         strnewatst1));
    ZeroMemory(  strcl, sizeof( strcl));
    m_wndText.ShowWindow(SW_HIDE);
    fl1=0;
    fl2=0;
    fl3=0;
    m_wndksy1.ShowWindow(SW_HIDE);
    m_wndksy2.ShowWindow(SW_HIDE);
    m_na1st.ShowWindow(SW_HIDE);
    m_name1st.ShowWindow(SW_HIDE);
    m_na2st.ShowWindow(SW_HIDE);
    m_name2st.ShowWindow(SW_HIDE);
    m_na3st.ShowWindow(SW_HIDE);
    m_name3st.ShowWindow(SW_HIDE);
    m_na4st.ShowWindow(SW_HIDE);
    m_name4st.ShowWindow(SW_HIDE);
    m_na21st.ShowWindow(SW_HIDE);
    m_name21st.ShowWindow(SW_HIDE);
    m_na22st.ShowWindow(SW_HIDE);
    m_name22st.ShowWindow(SW_HIDE);
```

```
    m_na23st.ShowWindow(SW_HIDE);
    m_name23st.ShowWindow(SW_HIDE);
    m_na24st.ShowWindow(SW_HIDE);
    m_name24st.ShowWindow(SW_HIDE);
    Invalidate();
    HANDLE  TEXT1, TEXT2, TEXT3, TEXT4, PROV;
    CFileDialog FD(1);
    int bResult, b,i,z,flz=0;
    char  str0txt[100],prov1[100];
    ZeroMemory(str0,sizeof(str0));
    ZeroMemory(prov1,sizeof(prov1));
    bResult;
    if(!FD.DoModal())
       return;
    Str=FD.GetPathName();
    lstrcpy(str0,Str);
    z=0;
    while(str0[z]!='.')
    {z++;}
    lstrcpyn(prov1, str0,(z+1));
    lstrcat(prov1, ".txt");
    PROV=::CreateFile(prov1,GENERIC_READ,0,NULL,OPEN_ALWAYS,FILE_A
TTRIBUTE_NORMAL,NULL);
    if(GetLastError()==0)
    {
              m_na1st.ShowWindow(SW_SHOW);
    m_na1 = "";
    flz=1;
    }
    else
    {
    m_na1st.ShowWindow(SW_SHOW);
    m_na1 = "";
    }
    UpdateData(FALSE);
       m_name1st.ShowWindow(SW_SHOW);
    m_name1 = "";
    UpdateData(FALSE);
    CloseHandle(PROV);
    DeleteObject(PROV);
    if(flz==1)
    {
    DeleteFile(prov1);
    }
```

```
    fl1=1;
    Invalidate();
}
void CTestDlg::OnButton2() //Вложитьсообщение
{
    if(strlen(str0)==0)
       MessageBox("Вначале надо открыть файл с
изображением.","Внимание!",MB_OK);
    else
    {
       char stego[60000];
    KOD kod;
    VLOG vlog;
    vlog.DoModal();
    if(strlen(vlog.m_vlogtext)!=0)
       {
    char strold[100],  strnewtxt[100], int itxt[5],strnewtxt14[100];
    int text1[1], oct,poz[60000],p,z;
    int i,a, nBytesToRead, bResult;
    unsigned long nBytesRead, i1, i2, i3, i4,i5;
    int text2[1], oct1;
    HANDLE TEXT1, TEXT2, TEXT3, TEXT4,TEXT14;
    TEXT1=::CreateFile(str0,GENERIC_READ,0,NULL,OPEN_ALWAYS,FILE_AT
TRIBUTE_NORMAL,NULL);
    srand(time(0));
    ZeroMemory(strold,sizeof(strold));
    ZeroMemory(strnewtxt,sizeof(strnewtxt));
    ZeroMemory(strnewtxt14,sizeof(strnewtxt14));
    a=1;
    while(str0[a]!='.')
    {a++;}
    for(i=0;i<60000;i++)
    {
    poz[i]=0;
    }
    p=0;
    lstrcpyn(strnew, str0,(a+1));
    lstrcat(strnew, "msg");
    if((vlog.m_vlogpro[0]=='0')&&(vlog.m_vlogpro[1]=='.')&&(vlog.m_vlogpro[2]=='
5'))
    {
       lstrcat(strnew, "50%");
       for(i=0;i<60000;i++)
       {
```

```
        if((rand()%2)==1)
        {
        poz[p]=i;
        p++;
        }
        }
    }
    else
    {
    if((vlog.m_vlogpro[0]=='0')&&(vlog.m_vlogpro[1]=='.')&&(vlog.m_vlogpro[2]=='
1'))
        {
        lstrcat(strnew, "10%");
        for(i=0;i<60000;i++)
        {
        if((rand()%10)==1)
        {
        poz[p]=i;
        p++;
        }
        }
        }
        else
        {
    if((vlog.m_vlogpro[0]=='0')&&(vlog.m_vlogpro[1]=='.')&&(vlog.m_vlogpro[2]=='
0')&&(vlog.m_vlogpro[3]=='1'))
                {
                lstrcat(strnew, "1%");
                for(i=0;i<60000;i++)
                {
                if((rand()%100)==1)
                {
                poz[p]=i;
                p++;
                }
                }
                }
                else
                {
    if((vlog.m_vlogpro[0]=='0')&&(vlog.m_vlogpro[1]=='.')&&(vlog.m_vlogpro[2]=='
0')&&(vlog.m_vlogpro[3]=='0')&&(vlog.m_vlogpro[4]=='1'))
                        {
                        lstrcat(strnew, "0,1%");
                        for(i=0;i<60000;i++)
```

```
                    {
                    if((rand()%1000)==1)
                    {
                    poz[p]=i;
                    p++;
                    }
                    }
                    }
                    else
                    {
    if((vlog.m_vlogpro[0]=='0')&&(vlog.m_vlogpro[1]=='.')&&(vlog.m_vlogpro[2]=='
0')&&(vlog.m_vlogpro[3]=='5'))
                    {
                    lstrcat(strnew, "5%");
                    for(i=0;i<60000;i++)
                    {
                    if((rand()%20)==1)
                    {
                    poz[p]=i;
                    p++;
                    }
                    }
                    }
                    else
                    {
                    if((vlog.m_vlogpro[0]=='2')&&(vlog.m_vlogpro[1]=='%'))
                    {
                    lstrcat(strnew, "2%");
                    for(i=0;i<60000;i++)
                    {
                    if((rand()%50)==1)
                    {
                    poz[p]=i;
                    p++;
                    }
                    }
                    }
                    else
                    {
    if((vlog.m_vlogpro[0]=='9')&&(vlog.m_vlogpro[1]=='.')&&(vlog.m_vlogpro[2]=='
0')&&(vlog.m_vlogpro[3]=='5')&&(vlog.m_vlogpro[4]=='%'))
                    {
                    lstrcat(strnew, "0,05%");
                    for(i=0;i<60000;i++)
```

```
                    {
                    if((rand()%2000)==1)
                    {
                    poz[p]=i;
                    p++;
                    }
                    }
                    }
                    else
                    {
    if((vlog.m_vlogpro[0]=='9')&&(vlog.m_vlogpro[1]=='.')&&(vlog.m_vlogpro[2]=='
0')&&(vlog.m_vlogpro[3]=='1')&&(vlog.m_vlogpro[4]=='%'))
                    {
                            lstrcat(strnew, "0,01%");
                    for(i=0;i<60000;i++)
                    {
                    if((rand()%10000)==1)
                    {
                    poz[p]=i;
                    p++;
                    }
                    }
                    }
                    else
                    {
                    if((vlog.m_vlogpro[0]=='3')&&(vlog.m_vlogpro[1]=='
')&&(vlog.m_vlogpro[2]=='д')&&(vlog.m_vlogpro[3]=='в')&&(vlog.m_vlogpro[4]=='o
'))
                    {
                    lstrcat(strnew, "1bs");
                    p=rand()%60000;
                    poz[p]=i;
                    p=1;
                    }
                    else
                    {
    if((vlog.m_vlogpro[0]=='0')&&(vlog.m_vlogpro[1]=='.')&&(vlog.m_vlogpro[2]=='
0')&&(vlog.m_vlogpro[3]=='0')&&(vlog.m_vlogpro[4]=='5'))
                    {
                    lstrcat(strnew, "0,5%");
                    for(i=0;i<60000;i++)
                    {
                    if((rand()%200)==1)
                    {
```

```
                    poz[p]=i;
                    p++;
                    }
                    }
                    }
                    else
                    if(vlog.m_vlogpro[0]=='1')
                    {
                    lstrcat(strnew, "100%");
                    for(i=0;i<60000;i++)
                    {
                    poz[p]=i;
                    p++;
                    }
                    }
                    }
                    }
                    }
                    }
                    }
                    }
                    }
        }
    }
    if((vlog.m_vlogglub[0]=='L')&&(vlog.m_vlogglub[1]=='S')&&(vlog.m_vlogglub[2
]=='B'))
    {
    lstrcat(strnew, "-LSB");
    }
if(vlog.m_vlogglub[0]=='1')
{
lstrcat(strnew, "-1");
}
if(vlog.m_vlogglub[0]=='2')
{
lstrcat(strnew, "-2");
}
if(vlog.m_vlogglub[0]=='3')
{
lstrcat(strnew, "-3");
}
lstrcpy(strnewtxt,strnew);
    lstrcpy(strnewtxt14,strnewtxt);
    lstrcat(strnew, ".bmp");
```

```
    lstrcat(strnewtxt, ".txt");
    lstrcat(strnewtxt14, "1.txt");
    TEXT2=::CreateFile(strnew,GENERIC_WRITE,0,NULL,OPEN_ALWAYS,FILE_
ATTRIBUTE_NORMAL,NULL);
    TEXT3=::CreateFile(strnewtxt,GENERIC_WRITE,0,NULL,OPEN_ALWAYS,FIL
E_ATTRIBUTE_NORMAL,NULL);
    if(kod.m_check==1)
    {
    TEXT14=::CreateFile(strnewtxt14,GENERIC_WRITE,0,NULL,OPEN_ALWAYS,
FILE_ATTRIBUTE_NORMAL,NULL);
    }
       char ch[10],EX[100];
       int p8=0;
       ZeroMemory(ch,sizeof(ch));
       ZeroMemory(EX,sizeof(EX));
       lstrcpy(EX,"Всего вложено символов (букв) - ");
       p8=p/8;
    sprintf(ch, "%d", p8);
    lstrcat(EX,ch);
    z=0;
char stego1[60001],textvid[60001];
int len,t14,j,lenvid,vid;
p=0;
len=0;
lenvid=0;
vid=0;
ZeroMemory(stego1,sizeof(stego1));
ZeroMemory(stego,sizeof(stego));
ZeroMemory(ch,sizeof(ch));
ZeroMemory(textvid,sizeof(textvid));
    lstrcpy(textvid,vlog.m_vlogtext);
    lenvid=strlen(textvid);
    for(i=0;i<lenvid; i++)
    {
    vid=(int)textvid[i];//получениекодаобратно textvid[0]=(char)vid[0]
    if(vid<0)
    {
       vid=256+vid;
    }
    sprintf(ch, "%d", vid);
    if(vid>=128)
    {
    lstrcat(stego1,"1");
    vid=vid-128;
```

```
}
else
{
lstrcat(stego1,"0");
}
if(vid>=64)
{
lstrcat(stego1,"1");
vid=vid-64;
}
else
{
lstrcat(stego1,"0");
}
if(vid>=32)
{
lstrcat(stego1,"1");
vid=vid-32;
}
else
{
lstrcat(stego1,"0");
}
if(vid>=16)
{
lstrcat(stego1,"1");
vid=vid-16;
}
else
{
lstrcat(stego1,"0");
}
if(vid>=8)
{
lstrcat(stego1,"1");
vid=vid-8;
}
else
{
lstrcat(stego1,"0");
}
if(vid>=4)
{
lstrcat(stego1,"1");
```

```
vid=vid-4;
}
else
{
lstrcat(stego1,"0");
}
if(vid>=2)
{
lstrcat(stego1,"1");
vid=vid-2;
}
else
{
lstrcat(stego1,"0");
}
if(vid>=1)
{
lstrcat(stego1,"1");
vid=vid-1;
}
else
{
lstrcat(stego1,"0");
}
}
len=strlen(stego1);
ZeroMemory(ch,sizeof(ch));
for(i=0;i<60000;i++)
{
stego[i]=stego1[(i-((len)*p))];
if(i==((len*(p+1))-1))
{
p++;
}
}
for(i=0;i<1078;i++)
   {
   ReadFile(TEXT1,text1,1,&i1,NULL);
   WriteFile(TEXT2,text1,1,&i2,NULL);
   }
for(i=0;i<60000;i++)
   {
   ReadFile(TEXT1,text1,1,&i1,NULL);
   if(poz[z]==i)
```

```
        {
                if(z!=0)
                {
                WriteFile(TEXT3,"*",1,&i3,NULL);
                }
                sprintf(itxt, "%d", i);
                WriteFile(TEXT3,itxt,strlen(itxt),&i3,NULL);
    if((vlog.m_vlogglub[0]=='L')&&(vlog.m_vlogglub[1]=='S')&&(vlog.m_vlogglub[2
]=='B'))
                {
                oct=text1[0]%2;
                if(stego[z]=='0')
                {
                if(oct!=0)    //тогда у нас число заканчивается на 1(нечетное), меняем
все на 0
                {
                text1[0]=text1[0]^01;//получаем на конце все 0!!! УРА!!!
                }
                }
                else
                {
                if(oct==0)            //тогда у нас число заканчивается на 0(четное),
меняем все на 1
                {
                text1[0]=text1[0]^01;//получаем на конце все 1!!! УРА!!!
                }
                }
                }
                if(vlog.m_vlogglub[0]=='1')
                {
                if(stego[z]=='0')
                {
                if(text1[0]!=0)
                text1[0]=text1[0]-1;//отнимаем 1
                }
                else
                {
                if(text1[0]!=255)
                text1[0]=text1[0]+1;//прибавляем 1
                }
                }
                if(vlog.m_vlogglub[0]=='2')
                {
                if(stego[z]=='0')
```

```
            {
            if(text1[0]>1)
            text1[0]=text1[0]-2;//отнимаем 1
            }
            else
            {
            if(text1[0]<254)
            text1[0]=text1[0]+2;//прибавляем 1
            }
            }
            if(vlog.m_vlogglub[0]=='3')
            {
            if(stego[z]=='0')
            {
            if(text1[0]>2)
            text1[0]=text1[0]-3;//отнимаем 1
            }
            else
            {
            if(text1[0]<253)
            text1[0]=text1[0]+3;//прибавляем 1
            }
            }
            z++;
    }
    WriteFile(TEXT2,text1,1,&i2,NULL);
    }
WriteFile(TEXT3,"#",1,&i3,NULL);
CloseHandle(TEXT1);
CloseHandle(TEXT2);
CloseHandle(TEXT3);
DeleteObject(TEXT1);
DeleteObject(TEXT2);
DeleteObject(TEXT3);
if(kod.m_check==1)
{
CloseHandle(TEXT14);
DeleteObject(TEXT14);
}
m_wndText.ShowWindow(SW_HIDE);
m_na2st.ShowWindow(SW_SHOW);
m_na2 = "";
UpdateData(FALSE);
m_name2st.ShowWindow(SW_SHOW);
```

```
    m_name2 = "";
    UpdateData(FALSE);
    fl2=1;
    Invalidate();
}
}
}
void CTestDlg::OnButton3() //нулигистограммы
{
    HANDLE TEXT21, TEXT22;
      int i, text21[1], text22[1],Cn[60000], Sk[256], Ns,Cn1[60000], Sk1[256], Ns1;
      unsigned long i21, i22;
      char hh1[20];
      Ns=0;
      Ns1=0;
      text21[0]=0;
      text22[0]=0;
      nuli nulist;
      for(i=0; i<60000; i++)
      {
            Cn[i]=0;
            Cn1[i]=0;
      }
      for(i=0; i<256; i++)
      {
            Sk[i]=0;
            Sk1[i]=0;
      }
    if(strlen(str0)==0)
      MessageBox("Вначале надо открыть файл с изображением и
сообщением.","Внимание!",MB_OK);
    else
    {
      stat statst;
      nulist.DoModal();
    TEXT21=::CreateFile(str0,GENERIC_READ,0,NULL,OPEN_ALWAYS,FILE_A
TTRIBUTE_NORMAL,NULL);
    for(i=0;i<1078;i++)
      {
      ReadFile(TEXT21,text21,1,&i21,NULL);
      }
    for(i=0;i<60000;i++)
      {
      ReadFile(TEXT21,text21,1,&i21,NULL);
```

```
        Cn[i]=text21[0];
        Sk[Cn[i]]++;
        }
    for(i=0; i<256; i++)
    {
    if(Sk[i]==0)
    {
    Ns++;
    }
    }
    m_na21st.ShowWindow(SW_SHOW);
if(Ns<nulist.m_nul)
m_na21="Есть вложение!";
else
m_na21="Нет вложения!";
if(nulist.m_nulch==1)
{
m_name21st.ShowWindow(SW_SHOW);
ZeroMemory(hh1, sizeof(hh1));
sprintf(hh1, "%d", Ns);
m_name21=hh1;
}
    UpdateData(FALSE);
    CloseHandle(TEXT21);
    DeleteObject(TEXT21);
    }
    if(strlen(strnew)!=0)
    {
TEXT22=::CreateFile(strnew,GENERIC_READ,0,NULL,OPEN_ALWAYS,FILE_AT
TRIBUTE_NORMAL,NULL);
    for(i=0;i<1078;i++)
        {
        ReadFile(TEXT22,text22,1,&i22,NULL);
        }
    for(i=0;i<60000;i++)
        {
        ReadFile(TEXT22,text22,1,&i22,NULL);
        Cn1[i]=text22[0];
        Sk1[Cn1[i]]++;
        }
    for(i=0; i<256; i++)
    {
    if(Sk1[i]==0)
    {
```

```
    Ns1++;
    }
    }
    m_na22st.ShowWindow(SW_SHOW);
if(Ns1<nulist.m_nul)
m_na22="Есть вложение!";
else
m_na22="Нет вложения!";
if(nulist.m_nulch==1)
{
m_name22st.ShowWindow(SW_SHOW);
ZeroMemory(hh1, sizeof(hh1));
sprintf(hh1, "%d", Ns1);
m_name22=hh1;
}
    UpdateData(FALSE);
    CloseHandle(TEXT21);
    DeleteObject(TEXT21);
    }
}
void CTestDlg::OnButton6() //Соседниезначения
{
    HANDLE TEXT23, TEXT24;
       int i, text23[1], text24[1], aa, Cn[256], Cn1[256];
       unsigned long i23, i24;
       char hh1[20];
       float S, S1, sum, sum1;
       S=0;
       S1=0;
       sum=0;
       sum1=0;
       aa=0;
       text23[0]=0;
       text24[0]=0;
       for(i=0; i<256; i++)
       {
       Cn[i]=0;
       Cn1[i]=0;
       }
       sos sosst;
    if(strlen(str0)==0)
       MessageBox("Вначале надо открыть файл с изображением и
сообщением.","Внимание!",MB_OK);
    else
```

```
    {
      sosst.DoModal();
    TEXT23=::CreateFile(str0,GENERIC_READ,0,NULL,OPEN_ALWAYS,FILE_A
TTRIBUTE_NORMAL,NULL);
      for(i=0;i<1078;i++)
      {
      ReadFile(TEXT23,text23,1,&i23,NULL);
      }
      for(i=0;i<60000;i++)
      {
      ReadFile(TEXT23,text23,1,&i23,NULL);
      aa=text23[0];
      Cn[aa]++;
      }
      for(i=0;i<255;i++)
    {
      S=S+((Cn[i]-Cn[i+1])*(Cn[i]-Cn[i+1]));
      sum=sum+(Cn[i]*Cn[i]);
    }
    S=S+(Cn[0]*Cn[0])+(Cn[255]*Cn[255]);
    sum=sum+(Cn[255]*Cn[255]);
    S=S/(2*sum);
    m_na23st.ShowWindow(SW_SHOW);
if(S<sosst.m_sosed)
m_na23="Есть вложение!";
else
m_na23="Нет вложения!";
if(sosst.m_sosedch==1)
{
m_name23st.ShowWindow(SW_SHOW);
ZeroMemory(hh1, sizeof(hh1));
sprintf(hh1, "%f", S);
m_name23=hh1;
}
    UpdateData(FALSE);
    CloseHandle(TEXT23);
    DeleteObject(TEXT23);
    }
if(strlen(strnew)!=0)
    {
      aa=0;
    TEXT24=::CreateFile(strnew,GENERIC_READ,0,NULL,OPEN_ALWAYS,FILE_
ATTRIBUTE_NORMAL,NULL);
      for(i=0;i<1078;i++)
```

```
        {
        ReadFile(TEXT24,text24,1,&i24,NULL);
        }
        for(i=0;i<60000;i++)
        {
        ReadFile(TEXT24,text24,1,&i24,NULL);
        aa=text24[0];
        Cn1[aa]++;
        }
        for(i=0;i<255;i++)
    {
        S1=S1+((Cn1[i]-Cn1[i+1])*(Cn1[i]-Cn1[i+1]));
        sum1=sum1+(Cn1[i]*Cn1[i]);
    }
    S1=S1+(Cn1[0]*Cn1[0])+(Cn1[255]*Cn1[255]);
    sum1=sum1+(Cn1[255]*Cn1[255]);
    S1=S1/(2*sum1);
        m_na24st.ShowWindow(SW_SHOW);
if(S1<sosst.m_sosed)
m_na24="Есть вложение!";
else
m_na24="Нет вложения!";
if(sosst.m_sosedch==1)
{
m_name24st.ShowWindow(SW_SHOW);
ZeroMemory(hh1, sizeof(hh1));
sprintf(hh1, "%f", S1);
m_name24=hh1;
}
    UpdateData(FALSE);
    CloseHandle(TEXT23);
    DeleteObject(TEXT23);
}
}
void CTestDlg::OnButton4() //Визуальнаяатака
{
        if(strlen(str0)==0)
        MessageBox("Вначале надо открыть файл с изображением и
сообщением.","Внимание!",MB_OK);
    else
    {
    char stroldat[100];
    int text1[1], oct, flag;
    int i,a, nBytesToRead, bResult;
```

```
    unsigned long nBytesRead, i1, i2, i3, i4;
    int text3[1], oct2;
    HANDLE TEXT4, TEXT5, TEXT6, TEXT7;
    flag=0;
    TEXT4=::CreateFile(str0,GENERIC_READ,0,NULL,OPEN_ALWAYS,FILE_AT
TRIBUTE_NORMAL,NULL);
    a=1;
    fl3=0;
    while(str0[a]!='.')
    {a++;}
    for(i=0; i<1;i++)
    {
    text3[i]=' ';
    }
    lstrcpyn(strnewat, str0,(a+1));
    lstrcat(strnewat, "attack");
    lstrcat(strnewat, ".bmp");
    TEXT5=::CreateFile(strnewat,GENERIC_WRITE,0,NULL,OPEN_ALWAYS,FIL
E_ATTRIBUTE_NORMAL,NULL);
    for(i=0;i<1078;i++)
       {
       ReadFile(TEXT4,text3,1,&i1,NULL);
       WriteFile(TEXT5,text3,1,&i2,NULL);
       }
    for(i=0;i<60000;i++)
       {
       ReadFile(TEXT4,text3,1,&i1,NULL);
       oct2=text3[0]%2;
              if(oct2==0)//заканчиваетьсяна 0
              {
              text3[0]=0x11;//черный
              }
              else
              {
              text3[0]=0xFF;
              }
       WriteFile(TEXT5,text3,1,&i2,NULL);
       }
    CloseHandle(TEXT4);
    CloseHandle(TEXT5);
    DeleteObject(TEXT4);
    DeleteObject(TEXT5);
    int flz1=0,z1=0;
    char prov2[100];
```

```
    HANDLE PROV1;
    fl3=1;
    Invalidate();
ZeroMemory(prov2,sizeof(prov2));
    while(str0[z1]!='.')
    {
    z1++;
    }
    lstrcpyn(prov2, str0,(z1+1));
    lstrcat(prov2, ".txt");
    PROV1=::CreateFile(prov2,GENERIC_READ,0,NULL,OPEN_ALWAYS,FILE_A
TTRIBUTE_NORMAL,NULL);
    if(GetLastError()==0)
    {
            m_na3st.ShowWindow(SW_SHOW);
    m_na3 = "";
    flz1=1;
    }
    else
    {
    m_na3st.ShowWindow(SW_SHOW);
    m_na3 = "";
    }
    UpdateData(FALSE);
    m_name3st.ShowWindow(SW_SHOW);
    m_name3 = "";
    UpdateData(FALSE);
    CloseHandle(PROV1);
    DeleteObject(PROV1);
    if(flz1==1)
    {
    DeleteFile(prov2);
    }
    if(strlen(strnew)!=0)
    {
    TEXT6=::CreateFile(strnew,GENERIC_READ,0,NULL,OPEN_ALWAYS,FILE_
ATTRIBUTE_NORMAL,NULL);
    lstrcpy(stroldat,strnew);
    a=1;
    while(stroldat[a]!='.')
    {
    a++;
    }
    lstrcpyn(strnewat1, stroldat,(a+1));
```

```
    lstrcat(strnewat1, "attack");
    lstrcat(strnewat1, ".bmp");
    TEXT7=::CreateFile(strnewat1,GENERIC_WRITE,0,NULL,OPEN_ALWAYS,FIL
E_ATTRIBUTE_NORMAL,NULL);
    for(i=0;i<1078;i++)
       {
       ReadFile(TEXT6,text3,1,&i1,NULL);
       WriteFile(TEXT7,text3,1,&i2,NULL);
       }
    for(i=0;i<60000;i++)
       {
       ReadFile(TEXT6,text3,1,&i1,NULL);
       oct2=text3[0]%2;
              if(oct2==0)//заканчиваетьсяна 0
              {
              text3[0]=0x11;//черный
              }
              else
              {
              text3[0]=0xFF;
              }
       WriteFile(TEXT7,text3,1,&i2,NULL);
       }
    CloseHandle(TEXT6);
    CloseHandle(TEXT7);
    DeleteObject(TEXT6);
    DeleteObject(TEXT7);
    m_wndText.ShowWindow(SW_HIDE);
    m_na4st.ShowWindow(SW_SHOW);
    m_na4 = "";
    UpdateData(FALSE);
    m_name4st.ShowWindow(SW_SHOW);
    m_name4 = "";
    UpdateData(FALSE);
    fl4=1;
    Invalidate();
    }
    }
}
void CTestDlg::OnButton5() //Статистическаяатака
{
flagfi=0;
flagfi1=0;
TABLE tab1;
```

```
CORR corr2;
corr2.m_corr1=1;
corr2.m_corr=1;
stat statistic;
if(strlen(str0)==0)
MessageBox("Вначале надо открыть файл с изображением и
сообщением.","Внимание!",MB_OK);
else
    {
       first firstst;
       firstst.DoModal();
    if(corr2.m_corr1==1)
    {
    char ch[20], ch1[20];
    int text4[1], flag;
    int i,a, nBytesToRead, bResult;
    unsigned long nBytesRead, i1, i2, i3, i4;
    int text5[1],brid[300],brid1[300]/*этояркость*/;
    HANDLE TEXT8, TEXT9, TEXT10, TEXT11,TEXT12,TEXT13 ;
    flag=0;
    Ebrid=0;
    Ebrid1=0;
    ZeroMemory(ch,sizeof(ch));
    ZeroMemory(ch1,sizeof(ch1));
    for(i=0;i<300;i++)
    {
    brid[i]=0;
    }
    TEXT8=::CreateFile(str0,GENERIC_READ,0,NULL,OPEN_ALWAYS,FILE_AT
TRIBUTE_NORMAL,NULL);
    a=1;
    text4[0]=0;
    while(str0[a]!='.')
    {
    a++;
    }
    for(i=0;i<20;i++)
    {
    ch[i]=' ';
    }
    lstrcpyn(strnewatst, str0,(a+1));
    lstrcat(strnewatst, "attack");
    lstrcat(strnewatst, ".txt");
```

```
    TEXT9=::CreateFile(strnewatst,GENERIC_WRITE,0,NULL,OPEN_ALWAYS,FI
LE_ATTRIBUTE_NORMAL,NULL);
    for(i=0;i<1078;i++)
       {
       ReadFile(TEXT8,text4,1,&i1,NULL);
       }
    for(i=0;i<60000;i++)
       {
       ReadFile(TEXT8,text4,1,&i1,NULL);
       brid[text4[0]]+=1;
       }
       for(i=0;i<128;i++)
       {
              if((brid[2*i]!=0)||(brid[(2*i)+1]!=0))
              {
              Ebrid+=(((brid[2*i]-brid[(2*i)+1])*(brid[2*i]-
brid[(2*i)+1]))/(brid[2*i]+brid[(2*i)+1]));
              }
       }
    Ebrid=Ebrid/(2*60000);
    m_na1st.ShowWindow(SW_SHOW);
    if(Ebrid<firstst.m_hikv)
m_na1="Есть вложение!";
else
m_na1="Нет вложения!";
if(firstst.m_hikvch==1)
{
m_name1st.ShowWindow(SW_SHOW);
ZeroMemory(ch, sizeof(ch));
sprintf(ch, "%f", Ebrid);
m_name1=ch;
}
       WriteFile(TEXT9,ch,strlen(ch),&i2,NULL);
       if(flagfi==0)
       {
        filex++;
       }
       if(flagfi1==1)
       {
       lstrcpy(strhi[filex-1],ch);
       }
       else
       {
       lstrcpy(strname[filex],str0);
```

```
        lstrcpy(strhi[filex],ch);
       }
        if(filex==10)
       {
       lstrcpy(strname[filex],str0);
       lstrcpy(strhi[filex],ch);
       }
    UpdateData(FALSE);
    CloseHandle(TEXT8);
    CloseHandle(TEXT9);
    DeleteObject(TEXT8);
    DeleteObject(TEXT9);
    HANDLE PROV2;
    int z2=0,flz2=0;
    char prov3[100];
ZeroMemory(prov3,sizeof(prov3));
while(str0[z2]!='.')
    {
    z2++;
    }
    lstrcpyn(prov3, str0,(z2+1));
    lstrcat(prov3, ".txt");
    PROV2=::CreateFile(prov3,GENERIC_READ,0,NULL,OPEN_ALWAYS,FILE_A
TTRIBUTE_NORMAL,NULL);
    UpdateData(FALSE);
    CloseHandle(PROV2);
    DeleteObject(PROV2);
    if(flz2==1)
    {
    DeleteFile(prov3);
    }
    if(strlen(strnew)!=0)
    {
    TEXT10=::CreateFile(strnew,GENERIC_READ,0,NULL,OPEN_ALWAYS,FILE_
ATTRIBUTE_NORMAL,NULL);
    a=1;
    for(i=0;i<300;i++)
    {
    brid1[i]=0;
    }
    while(strnew[a]!='.')
    {
    a++;
    }
```

```
    for(i=0;i<20;i++)
    {
    ch[i]=' ';
    }
    text5[0]=0;
    lstrcpyn(strnewatst1, strnew,(a+1));
    lstrcat(strnewatst1, "attack");
    lstrcat(strnewatst1, ".txt");
    TEXT11=::CreateFile(strnewatst1,GENERIC_WRITE,0,NULL,OPEN_ALWAYS,
FILE_ATTRIBUTE_NORMAL,NULL);
    for(i=0;i<1078;i++)
       {
       ReadFile(TEXT10,text5,1,&i3,NULL);
       }
    for(i=0;i<60000;i++)
       {
       ReadFile(TEXT10,text5,1,&i3,NULL);
       brid1[text5[0]]+=1;
       }
       for(i=0;i<128;i++)
       {
              if((brid1[2*i]!=0)||(brid1[(2*i)+1]!=0))
              {
              Ebrid1+=(((brid1[2*i]-brid1[(2*i)+1])*(brid1[2*i]-
brid1[(2*i)+1]))/(brid1[2*i]+brid1[(2*i)+1]));
              }
       }
       Ebrid1=Ebrid1/(2*60000);
       sprintf(ch, "%f", Ebrid1);
     WriteFile(TEXT11,ch,strlen(ch),&i4,NULL);
     m_na3st.ShowWindow(SW_SHOW);
    if(Ebrid1<firstst.m_hikv)
m_na3="Есть вложение!";
else
m_na3="Нет вложения!";
if(firstst.m_hikvch==1)
{
m_name3st.ShowWindow(SW_SHOW);
ZeroMemory(ch, sizeof(ch));
sprintf(ch, "%f", Ebrid1);
m_name3=ch;
}
       if(flagfi==0)
       { filex++; }
```

```
        lstrcpy(strname[filex],strnew);
         lstrcpy(strhi[filex],ch);
      UpdateData(FALSE);
      CloseHandle(TEXT10);
      CloseHandle(TEXT11);
      DeleteObject(TEXT10);
      DeleteObject(TEXT11);
      UpdateData(FALSE);
      }
      else
      {
        if(strlen(strcl)!=0)
      {
      TEXT10=::CreateFile(strcl,GENERIC_READ,0,NULL,OPEN_ALWAYS,FILE_A
TTRIBUTE_NORMAL,NULL);
      a=1;
      for(i=0;i<300;i++)
      {brid1[i]=0;}
      while(strcl[a]!='.')
      {a++;}
      for(i=0;i<20;i++)
      {ch[i]=' ';}
      text5[0]=0;
      lstrcpyn(strnewatst1, strcl,(a+1));
      lstrcat(strnewatst1, "attack");
      lstrcat(strnewatst1, ".txt");
      TEXT11=::CreateFile(strnewatst1,GENERIC_WRITE,0,NULL,OPEN_ALWAYS,
FILE_ATTRIBUTE_NORMAL,NULL);
      for(i=0;i<1078;i++)
        {
        ReadFile(TEXT10,text5,1,&i3,NULL);
        }
      for(i=0;i<60000;i++)
        {
        ReadFile(TEXT10,text5,1,&i3,NULL);
        brid1[text5[0]]+=1;
        }
        for(i=0;i<128;i++)
        {
                if((brid1[2*i]!=0)||(brid1[(2*i)+1]!=0))
                {
                Ebrid1+=(((brid1[2*i]-brid1[(2*i)+1])*(brid1[2*i]-
brid1[(2*i)+1]))/(brid1[2*i]+brid1[(2*i)+1]));
                }
```

```
        }
        sprintf(ch, "%f", Ebrid1);
     WriteFile(TEXT11,ch,strlen(ch),&i4,NULL);
        if(flagfi==0)
        { filex++; }
         lstrcpy(strname[filex],strcl);
        lstrcpy(strhi[filex],ch);
    UpdateData(FALSE);
    CloseHandle(TEXT10);
    CloseHandle(TEXT11);
    DeleteObject(TEXT10);
    DeleteObject(TEXT11);
    UpdateData(FALSE);
    }
    }
    }
    }
}
void CTestDlg::OnButton7() //очиститьтаблицу
{}
void CTestDlg::OnButton8() //удаление
{
        if(strlen(str0)==0)
        MessageBox("Вначале надо открыть файл с изображением и
сообщением.","Внимание!",MB_OK);
    else
    {
    char stroldat[100];
    int text6[1], oct5,oct6, flag;
    int i,a, nBytesToRead, bResult;
    unsigned long nBytesRead, i1, i2, i3, i4;
    int text5[1], r;
    HANDLE TEXT14, TEXT15, TEXT16, TEXT17;
    flag=0;
    i1=0;
    i2=0;
    i3=0;
    i4=0;
    TEXT14=::CreateFile(str0,GENERIC_READ,0,NULL,OPEN_ALWAYS,FILE_A
TTRIBUTE_NORMAL,NULL);
    a=1;
    fl3=0;
    while(str0[a]!='.')
    {
```

```
    a++;
    }
    ZeroMemory(text5,sizeof(text5));
    lstrcpyn(strcl, str0,(a+1));
    lstrcat(strcl, "clear");
    lstrcat(strcl, ".bmp");
    TEXT15=::CreateFile(strcl,GENERIC_WRITE,0,NULL,OPEN_ALWAYS,FILE_
ATTRIBUTE_NORMAL,NULL);
    for(i=0;i<1078;i++)
       {
       ReadFile(TEXT14,text5,1,&i1,NULL);
       WriteFile(TEXT15,text5,1,&i2,NULL);
       }
    for(i=0;i<60000;i++)
       {
       ReadFile(TEXT14,text5,1,&i1,NULL);
       r=rand()%2;
       oct5=text5[0]%2;
              if(r==0)
              {
              if(oct5!=0)   //тогда у нас число заканчивается на 1(нечетное), меняем
все на 0
              {
              text5[0]=text5[0]^01;//получаем на конце все 0!!! УРА!!!
              }
              }
              else
              {
              if(oct5==0)          //тогда у нас число заканчивается на 0(четное),
меняем все на 1
              {
              text5[0]=text5[0]^01;//получаем на конце все 1!!! УРА!!!
              }
              }
       WriteFile(TEXT15,text5,1,&i2,NULL);
       }
    CloseHandle(TEXT14);
    CloseHandle(TEXT15);
    DeleteObject(TEXT14);
    DeleteObject(TEXT15);
    m_wndText.ShowWindow(SW_HIDE);
    m_wndText.ShowWindow(SW_HIDE);
    m_na2st.ShowWindow(SW_SHOW);
    m_na2 = "";
```

```
    UpdateData(FALSE);
    m_name2st.ShowWindow(SW_SHOW);
    m_name2 = "";
    UpdateData(FALSE);
    fl4=1;
    Invalidate();
    }
}
voidCTestDlg::OnButton9() //Пороговоезначениехиквадрат
{
    hi2z[0]=0;
    CHI hi2;
    hi2.DoModal();
    char histr[100],hiz[50];
    sscanf(hi2.m_newhi2,"%d",hi2z);
    ZeroMemory(histr,sizeof(histr));
    lstrcat(histr, "Пороговоезначениехиквадрат - ");
    lstrcat(histr, hi2.m_newhi2);
    tab.m_hi0st.ShowWindow(SW_SHOW);
    tab.m_hi0=histr;
    UpdateData(FALSE);
}
void CTestDlg::OnButton10()
{
    char ch[20], ch1[20];
    ZeroMemory(ch,sizeof(ch));
    ZeroMemory(ch1,sizeof(ch1));
    sprintf(ch, "%f", Ebrid);
    if(filex>0)
    {
        if(filex>=1)
        {
       tab.m_tabname1st.ShowWindow(SW_SHOW);
       tab.m_tabhi1st.ShowWindow(SW_SHOW);
       tab.m_tabpro1st.ShowWindow(SW_SHOW);
        tab.m_tabname1=strname[1];
        tab.m_tabhi1=strhi[1];
       tab.m_tabpro1=strpro[1];
        }
        if(filex>=2)
        {
        tab.m_tabname2st.ShowWindow(SW_SHOW);
       tab.m_tabhi2st.ShowWindow(SW_SHOW);
       tab.m_tabpro2st.ShowWindow(SW_SHOW);
```

```
 tab.m_tabname2=strname[2];
 tab.m_tabhi2=strhi[2];
 tab.m_tabpro2=strpro[2];
 }
 if(filex>=3)
 {
 tab.m_tabname3st.ShowWindow(SW_SHOW);
tab.m_tabhi3st.ShowWindow(SW_SHOW);
tab.m_tabpro3st.ShowWindow(SW_SHOW);
 tab.m_tabname3=strname[3];
 tab.m_tabhi3=strhi[3];
 tab.m_tabpro3=strpro[3];
 }
 if(filex>=4)
 {
 tab.m_tabname4st.ShowWindow(SW_SHOW);
tab.m_tabhi4st.ShowWindow(SW_SHOW);
tab.m_tabpro4st.ShowWindow(SW_SHOW);
 tab.m_tabname4=strname[4];
 tab.m_tabhi4=strhi[4];
  tab.m_tabpro4=strpro[4];
 }
 if(filex>=5)
 {
 tab.m_tabname5st.ShowWindow(SW_SHOW);
tab.m_tabhi5st.ShowWindow(SW_SHOW);
tab.m_tabpro5st.ShowWindow(SW_SHOW);
 tab.m_tabname5=strname[5];
 tab.m_tabhi5=strhi[5];
 tab.m_tabpro5=strpro[5];
 }
 if(filex>=6)
 {
tab.m_tabname6st.ShowWindow(SW_SHOW);
tab.m_tabhi6st.ShowWindow(SW_SHOW);
tab.m_tabpro6st.ShowWindow(SW_SHOW);
 tab.m_tabname6=strname[6];
 tab.m_tabhi6=strhi[6];
 tab.m_tabpro6=strpro[6];
 }
 if(filex>=7)
 {
 tab.m_tabname7st.ShowWindow(SW_SHOW);
tab.m_tabhi7st.ShowWindow(SW_SHOW);
```

```
        tab.m_tabpro7st.ShowWindow(SW_SHOW);
         tab.m_tabname7=strname[7];
         tab.m_tabhi7=strhi[7];
         tab.m_tabpro7=strpro[7];
         }
         if(filex>=8)
         {
        tab.m_tabname8st.ShowWindow(SW_SHOW);
        tab.m_tabhi8st.ShowWindow(SW_SHOW);
        tab.m_tabpro8st.ShowWindow(SW_SHOW);
         tab.m_tabname8=strname[8];
         tab.m_tabhi8=strhi[8];
         tab.m_tabpro8=strpro[8];
         }
         if(filex>=9)
         {
         tab.m_tabname9st.ShowWindow(SW_SHOW);
        tab.m_tabhi9st.ShowWindow(SW_SHOW);
        tab.m_tabpro9st.ShowWindow(SW_SHOW);
        tab.m_tabname9=strname[9];
         tab.m_tabhi9=strhi[9];
         tab.m_tabpro9=strpro[9];
         }
         if(filex>=10)
         {
        tab.m_tabname10st.ShowWindow(SW_SHOW);
        tab.m_tabhi10st.ShowWindow(SW_SHOW);
        tab.m_tabpro10st.ShowWindow(SW_SHOW);
         tab.m_tabname10=strname[10];
         tab.m_tabhi10=strhi[10];
         tab.m_tabpro10=strpro[10];
         }
    }
else
{
char hi0[10];
     lstrcat(hi0, " ");
     tab.m_tabname1=hi0;
     tab.m_tabhi1=hi0;
     tab.m_tabpro1=hi0;
     tab.m_tabname2=hi0;
     tab.m_tabhi2=hi0;
     tab.m_tabpro2=hi0;
     tab.m_tabname3=hi0;
```

```
tab.m_tabhi3=hi0;
tab.m_tabpro3=hi0;
tab.m_tabname4=hi0;
tab.m_tabhi4=hi0;
tab.m_tabpro4=hi0;
tab.m_tabname5=hi0;
tab.m_tabhi5=hi0;
tab.m_tabpro5=hi0;
tab.m_tabname6=hi0;
tab.m_tabhi6=hi0;
tab.m_tabpro6=hi0;
tab.m_tabname7=hi0;
tab.m_tabhi7=hi0;
tab.m_tabpro7=hi0;
tab.m_tabname8=hi0;
tab.m_tabhi8=hi0;
tab.m_tabpro8=hi0;
tab.m_tabname9=hi0;
tab.m_tabhi9=hi0;
tab.m_tabpro9=hi0;
tab.m_tabname10=hi0;
tab.m_tabhi10=hi0;
tab.m_tabpro10=hi0;
tab.m_tabhi1st.ShowWindow(SW_HIDE);
tab.m_tabhi2st.ShowWindow(SW_HIDE);
tab.m_tabhi3st.ShowWindow(SW_HIDE);
tab.m_tabhi4st.ShowWindow(SW_HIDE);
tab.m_tabhi5st.ShowWindow(SW_HIDE);
tab.m_tabhi6st.ShowWindow(SW_HIDE);
tab.m_tabhi7st.ShowWindow(SW_HIDE);
tab.m_tabhi8st.ShowWindow(SW_HIDE);
tab.m_tabhi9st.ShowWindow(SW_HIDE);
tab.m_tabhi10st.ShowWindow(SW_HIDE);
tab.m_tabname1st.ShowWindow(SW_HIDE);
tab.m_tabname2st.ShowWindow(SW_HIDE);
tab.m_tabname3st.ShowWindow(SW_HIDE);
tab.m_tabname4st.ShowWindow(SW_HIDE);
tab.m_tabname5st.ShowWindow(SW_HIDE);
tab.m_tabname6st.ShowWindow(SW_HIDE);
tab.m_tabname7st.ShowWindow(SW_HIDE);
tab.m_tabname8st.ShowWindow(SW_HIDE);
tab.m_tabname9st.ShowWindow(SW_HIDE);
tab.m_tabname10st.ShowWindow(SW_HIDE);
tab.m_tabpro1st.ShowWindow(SW_HIDE);
```

```
    tab.m_tabpro2st.ShowWindow(SW_HIDE);
    tab.m_tabpro3st.ShowWindow(SW_HIDE);
    tab.m_tabpro4st.ShowWindow(SW_HIDE);
    tab.m_tabpro5st.ShowWindow(SW_HIDE);
    tab.m_tabpro6st.ShowWindow(SW_HIDE);
    tab.m_tabpro7st.ShowWindow(SW_HIDE);
    tab.m_tabpro8st.ShowWindow(SW_HIDE);
    tab.m_tabpro9st.ShowWindow(SW_HIDE);
    tab.m_tabpro10st.ShowWindow(SW_HIDE);
}
    tab.DoModal();
}
void CTestDlg::OnButton11()//атакастатистикой 2-гопорядкаПВА
{
flagfi=0;
flagfi1=0;
TABLE tab1;
CORR corr2;
corr2.m_corr1=1;
corr2.m_corr=1;
stat statistic;
    if(strlen(str0)==0)
    MessageBox("Вначале надо открыть файл с изображением и
сообщением.","Внимание!",MB_OK);//**************
    else
    {
       second secondst;
       secondst.DoModal();
       if(corr2.m_corr==1)
    {
    HANDLE TEXT18, TEXT19;
    int text8[1],text9[1],aa,bb,c0,i,c1,d0,d2,y1,x1;
    float D,sD;
    float p1,p2,A,B,C;
    unsigned long i5,i6;
    char hh1[20],hh2[20];
    text8[0]=0;
    text9[0]=0;
    aa=0;
    bb=0;
    c0=0;
    c1=0;
    d0=0;
    d2=0;
```

```
    y1=0;
    x1=0;
    A=0;
    B=0;
    C=0;
    D=0;
    sD=0;
    p1=0;
    p2=0;
    ZeroMemory(hh1,sizeof(hh1));
    TEXT18=::CreateFile(str0,GENERIC_READ,0,NULL,OPEN_ALWAYS,FILE_A
TTRIBUTE_NORMAL,NULL);
    for(i=0;i<1078;i++)
       {ReadFile(TEXT18,text8,1,&i5,NULL);}
    for(i=0;i<60000;i++)
       {
       bb=aa;
       ReadFile(TEXT18,text8,1,&i5,NULL);
       aa=text8[0];
       if(i>0)
       //c0-совпадают все биты кроме НЗБ
                        if(aa!=bb)
                        if(((aa/2)*2)==((bb/2)*2))
                        {c0++;}
       if((((aa/2)*2)-((bb/2)*2)==2)||(((bb/2)*2)-((aa/2)*2)==2))
       {c1++;}
       if(aa==bb)
       {d0++;}
       if(((aa-bb)==2)||((bb-aa)==2))
       d2++;
       if(((aa/2)*2)!=aa)
       if(((aa/2)*2)==bb)
       y1++;
       if(((bb/2)*2)!=bb)
       if(((bb/2)*2)==aa)
       y1++;
       if(((bb/2)*2)!=bb)
       if(((bb/2)*2)==(aa-2))
       x1++;
       if(((aa/2)*2)!=aa)
       if(((aa/2)*2)==(bb-2))
       x1++;
    }
       A=((2*c0)-c1)/4;
```

```
        B=((2*d0)-d2+(2*y1)-(2*x1))/2;
        C=(y1-x1);
        D=(B*B)-(4*A*C);
        if(D>=0)
        {
          sD=sqrt(D);
        }
        p1=(B-sD)/(2*A);
        p1=p1;
        if(p1>100)
        p1=100;
        if(p1<0)
        p1=0;
        filex++;
     sprintf(hh1, "%f", p1);
     lstrcpy(strname[filex],str0);
     lstrcpy(strpro[filex],hh1);
     flagfi=1;
        m_na2st.ShowWindow(SW_SHOW);
if(p1>secondst.m_pva)
m_na2="Есть вложение!";
else
m_na2="Нет вложения!";
if(secondst.m_pvach==1)
{
m_name2st.ShowWindow(SW_SHOW);
ZeroMemory(hh1, sizeof(hh1));
sprintf(hh1, "%f", p1);
m_name2=hh1;
}
     UpdateData(FALSE);
     CloseHandle(TEXT18);
     DeleteObject(TEXT18);
     if((strlen(strnew)!=0)||(strlen(strcl)!=0))
     {
     aa=0;
     bb=0;
     c0=0;
     c1=0;
     d0=0;
     d2=0;
     y1=0;
     x1=0;
     A=0;
```

```
    B=0;
    C=0;
    D=0;
    sD=0;
    p2=0;
    if(strlen(strnew)!=0)
    TEXT19=::CreateFile(strnew,GENERIC_READ,0,NULL,OPEN_ALWAYS,FILE_
ATTRIBUTE_NORMAL,NULL);
    if(strlen(strcl)!=0)
    TEXT19=::CreateFile(strcl,GENERIC_READ,0,NULL,OPEN_ALWAYS,FILE_A
TTRIBUTE_NORMAL,NULL);
    for(i=0;i<1078;i++)
       {ReadFile(TEXT19,text9,1,&i6,NULL);}
    for(i=0;i<60000;i++)
       {
       bb=aa;
       ReadFile(TEXT19,text9,1,&i6,NULL);
       aa=text9[0];
       if(i>0)
       if(aa!=bb)
       if(((aa/2)*2)==((bb/2)*2))
       c0++;
       if((((aa/2)*2)-((bb/2)*2)==2)||(((bb/2)*2)-((aa/2)*2)==2))
       c1++;
       if(aa==bb)
       d0++;
       if(((aa-bb)==2)||((bb-aa)==2))
       d2++;
       if(((aa/2)*2)!=aa)
       if(((aa/2)*2)==bb)
       y1++;
       if(((bb/2)*2)!=bb)
       if(((bb/2)*2)==aa)
       y1++;
       if(((bb/2)*2)!=bb)
       if(((bb/2)*2)==(aa-2))
       x1++;
       if(((aa/2)*2)!=aa)
       if(((aa/2)*2)==(bb-2))
       x1++;
    }
       A=((2*c0)-c1)/4;
       B=((2*d0)-d2+(2*y1)-(2*x1))/2;
       C=(y1-x1);
```

```
        D=(B*B)-(4*A*C);
        if(D>=0)
        { sD=sqrt(D);}
        p2=(B-sD)/(2*A);
        ZeroMemory(hh2,sizeof(hh2));
        p2=p2;
        if(p2>100)
                p2=100;
        if(p2<0)
                p2=0;
        filex++;
     sprintf(hh2, "%f", p2);
     if(strlen(strnew)!=0)
     lstrcpy(strname[filex],strnew);
     if(strlen(strcl)!=0)
     lstrcpy(strname[filex],strcl);
     lstrcpy(strpro[filex],hh2);
     flagfi1=1;
     m_na4st.ShowWindow(SW_SHOW);
     if(p2>secondst.m_pva)
m_na4="Есть вложение!";
else
m_na4="Нет вложения!";
if(secondst.m_pvach==1)
{
m_name4st.ShowWindow(SW_SHOW);
ZeroMemory(hh2, sizeof(hh2));
sprintf(hh2, "%f", p2);
m_name4=hh2;
}
UpdateData(FALSE);
     CloseHandle(TEXT19);
     DeleteObject(TEXT19);
}
}
}
}
voidCTestDlg::OnButton12() //соседниезначенийгистограммы
{}
void CTestDlg::OnOK()
{exit(NULL);}
```

209 Приложение В

digiton®
research division

ООО "Дигитон", 197110, Санкт-Петербург, Крестовский пр., д.15, 12-I
тел.: (812)601-68-17, факс: (812)601-68-12
www.digiton-rd.com, e-mail: info@digiton.ru

Акт
Об использовании результатов диссертационной работы
аспирантки Государственного университета телекоммуникаций
им. проф. М.А. Бонч-Бруевича Гелинг Екатерины Юрьевны
на тему

«Исследование целевых методов обнаружения стегосистем»

Мы, нижеподписавшиеся, Грудинин Владимир Алексеевич и Касимов Денис Васильевич составили настоящий акт в том, что разработанные аспиранткой Герлинг Е.Ю. методы целевого обнаружения стегосистем с вложением в наименьшие значащие биты и с использованием для вложения широкополосных сигналов, а также оценки их эффективности, выполненные на достаточно большом объеме статистики неподвижных изображений, были использованы в научно-исследовательской работе "Ярус-СГ", выполнявшейся по заказу ООО «Дигитон».

24.05.2010

Генеральный директор ООО «Дигитон» В.А. Грудинин

Инженер-аналитик ООО «Дигитон» Д.В. Касимов

Общество с ограниченной ответственностью «Дигитон» * Санкт-Петербург *

210
Приложение Г

**Министерство
связи и массовых коммуникаций
Российской Федерации**

ФЕДЕРАЛЬНОЕ АГЕНТСТВО СВЯЗИ

федеральное государственное
образовательное бюджетное
учреждение высшего
профессионального образования

**"САНКТ-ПЕТЕРБУРГСКИЙ
ГОСУДАРСТВЕННЫЙ
УНИВЕРСИТЕТ ТЕЛЕКОММУНИКАЦИЙ
им. проф. М.А. Бонч-Бруевича"
(СПбГУТ)**

Юридический адрес: набережная реки
Мойки, д. 61,Санкт-Петербург, 191186

Почтовый адрес: пр. Большевиков, д. 22,
корп. 1, Санкт-Петербург, 193232
Тел.(812) 3263156, Факс: (812) 3263159
E-mail: rector@sut.ru
ИНН 7808004760 КПП 784001001
ОГРН 1027809197635 ОКПО 01179934

_____№_____

на №_____ от_____

УТВЕРЖДАЮ
Первый проректор – проректор по учебной работе
д-р техн. наук, профессор Машков Г. М.

«11» 12 2014 года

Акт
об использовании научных исследований
Герлинг Екатерины Юрьевны
на тему:
«Исследование и разработка методов обнаружения
стеговложений в неподвижных изображениях»

Комиссия в составе:

председатель – Просихин В.П., д-р техн. наук, профессор, заведующий кафедрой «Защищенные системы связи»

члены комиссии – Коржика В. И., д-р техн. наук, профессор, профессор кафедры, «Защищенные системы связи»

– Красов А. В., канд. техн. наук, доцент, профессор кафедры, «Защищенные системы связи»

удостоверяет, что результаты диссертационной работы Герлинг Екатерины Юрьевны используются в лабораторной работе «Методы обнаружения стегосистем НЗБ и ШПС» в курсе «Основы стеганографии» на кафедре «Защищенные системы связи». Также в лабораторной работе используется программа, написанная Герлинг Е.Ю. в которой наглядно демонстрируются методы стеговложения и стегоанализа, рассмотренные в диссертации.

Лабораторная работа наглядно показывает студентам методы вложения в наименее значащие биты и вложения в широкополосные сигналы и методы стегоанализа,

основанные на визуальной атаки, статистике 1-ого и 2-ого порядка, подсчете нулей гистограммы и сравнении соседних значений гистограммы, а также позволяет сравнить эффективность используемых методов стегоанализа.

Использование результатов диссертационной работы Герлинг Е.Ю. в учебном процессе повышает уровень знаний в области защиты информации.

Председатель комиссии: Просихин В.П.

11.02.2014

Члены комиссии: Коржик В.И.

Красов А.В.