

**МИНИСТЕРСТВО ОБРАЗОВАНИЯ И НАУКИ
РОССИЙСКОЙ ФЕДЕРАЦИИ
ФЕДЕРАЛЬНОЕ АГЕНТСТВО ПО ОБРАЗОВАНИЮ
ЮЖНЫЙ ФЕДЕРАЛЬНЫЙ УНИВЕРСИТЕТ**

На правах рукописи



104.200.8 18209

Алиев Александр Тофикович

**РАЗРАБОТКА МОДЕЛЕЙ, МЕТОДОВ И АЛГОРИТМОВ
ПЕРСПЕКТИВНЫХ СРЕДСТВ ЗАЩИТЫ ИНФОРМАЦИИ
В СИСТЕМАХ ЭЛЕКТРОННОГО ДОКУМЕНТООБОРОТА
НА БАЗЕ СОВРЕМЕННЫХ ТЕХНОЛОГИЙ СКРЫТОЙ СВЯЗИ**

Специальность:

05.13.19 – Методы и системы защиты информации, информационная безопасность

Диссертация на соискание ученой степени
кандидата технических наук

Научный руководитель:
д.т.н., профессор Аграновский А.В.

Ростов-на-Дону

2008

СОДЕРЖАНИЕ

	Стр.
ВВЕДЕНИЕ.....	6
ГЛАВА 1. ИССЛЕДОВАНИЕ И АНАЛИЗ СУЩЕСТВУЮЩИХ РЕШЕНИЙ В ОБЛАСТИ ЗАЩИТЫ ЭЛЕКТРОННЫХ ДОКУМЕНТОВ.....	16
1.1 Электронный документооборот	16
1.1.1 Понятие электронного документа	16
1.1.2 Системы электронного документооборота.....	20
1.2 Анализ программно реализуемых механизмов защиты информации циркулирующей в СЭД	22
1.2.1 Методы и средства обеспечения безопасности информации	22
1.2.2 Криптографические методы и механизмы защиты информации.....	24
1.2.3 Стеганографические методы защиты информации.....	29
1.2.4 Технологии цифровых водяных знаков	33
1.2.5 Классификации атак на системы связи	35
1.3 Анализ современных направлений в стеганографии	39
1.3.1 Стеганографические методы в текстовых файлах.....	39
1.3.2 Стеганографические методы в мультимедиа	42
1.3.3 Стеганографические методы реального времени	45
1.3.4 Классификация атак на системы скрытой передачи электронных документов в зависимости от используемых уязвимостей.....	47
1.4 Современные модели стеганографических систем	54
1.4.1 Стеганографическая система как система связи.....	54
1.4.2 Математическая модель стеганографической системы	57
1.4.3 Модели «криптография + стеганография»	61
1.5 Выводы по главе	63
ГЛАВА 2. АНАЛИЗ И ОЦЕНКА ЗАЩИЩЕННОСТИ ИНФОРМАЦИИ В СИСТЕМАХ СКРЫТОГО ЭЛЕКТРОННОГО ДОКУМЕНТООБОРОТА	66
2.1 Оценка возможности и разработка способов противодействия методам современного стеганоанализа	66
2.1.1 Визуальный стеганоанализ	66
2.1.2 Поиск сигнатур.....	68
2.1.3 Статистический стеганоанализ.....	69
2.1.4 Корреляционный анализ.....	72
2.1.5 Методы универсального стеганоанализа.....	74
2.2 Теоретическая стойкость стеганографических систем	76
2.2.1 Совершенная стеганографическая система	76
2.2.2 Совершенная стеганографическая система на битовых строках содержащих длинные серии одинаковых битов	78
2.3 Методика оценки практической стойкости.....	89
2.3.1 Критический коэффициент сокрытия	89
2.3.2 Теоретико-множественный подход к оценке эффективности многомодульных систем стеганографического анализа	93
2.3.3 Практическая стойкость и предельный коэффициент сокрытия	97
2.4 Выводы по главе	102
ГЛАВА 3. РАЗРАБОТКА МОДЕЛЕЙ И МЕТОДОВ СИСТЕМ СКРЫТОГО ЭЛЕКТРОННОГО ДОКУМЕНТООБОРОТА	103
3.1 Криптостеганографические системы связи: базовые принципы, модель и определение	103
3.1.1 Критерии стойкости систем скрытой передачи ЭД.....	103
3.1.2 Гибридная – криптостеганографическая система.....	105

3.2 Криптографическая часть.....	111
3.2.1 Требования к криптографическим алгоритмам	111
3.2.2 Гибкость архитектуры	112
3.3 Стеганографическая часть.....	114
3.3.1 Требования к стеганографическим методам и алгоритмам	114
3.3.2 Стеганографические методы на базе пространственно-частотных фильтров усредняющих масок.....	119
3.4 Особенности реализации алгоритмов согласования	125
3.4.1 Требования к алгоритмам согласования, базовый алгоритм	125
3.4.2 Метод сдвига битовых последовательностей	127
3.5 Мультиплексирование канала.....	133
3.6 Выводы по главе	135
ГЛАВА 4. ПРИМЕНЕНИЕ СКРЫТОЙ ПЕРЕДАЧИ ИНФОРМАЦИИ В РАСПРЕДЕЛЕННЫХ СИСТЕМАХ ЭЛЕКТРОННОГО ДОКУМЕНТООБОРОТА.....	137
4.1 Архитектура многопользовательских распределенных систем скрытого электронного документооборота.....	137
4.1.1 Организация процесса скрытой передачи ЭД.....	139
4.1.2 Общая схема крипостеганографической системы скрытой передачи ЭД на стороне отправителя	141
4.1.3 Общая схема крипостеганографической системы скрытой передачи ЭД на стороне адресата.....	146
4.2 Скрытая маркировка электронных документов в СЭД.....	149
4.2.1 Многопользовательская система скрытой уникальной маркировки электронных документов	149
4.2.2 Модель системы скрытой маркировки ЭД на базе клиент-серверной архитектуры.....	151
4.2.3 Методы встраивания данных ЦВЗ в системе маркировки ЭД	155
4.2.4 Формирование данных ЦВЗ	158
4.2.5 Проверка ЦВЗ в электронных документах	162
4.3 Выводы по главе	166
ЗАКЛЮЧЕНИЕ	167
СПИСОК ЛИТЕРАТУРЫ	170
ПРИЛОЖЕНИЯ.....	188
Приложение А. Термины и определения в теории стеганографии.....	189
Приложение Б. Результаты экспериментального применения МСБП совместно с методом замены младших значащих битов.....	194
Приложение В. Таблицы кодов символов с одинаковым начертанием	205
Приложение Г. Способ борьбы с инъективными ошибками в каналах передачи данных малой пропускной способности	208

ПРИНЯТЫЕ ОБОЗНАЧЕНИЯ И СОКРАЩЕНИЯ

С	– множество (пустых) контейнеров
Е	– криптографические алгоритмы
Н	– множество скрывающих преобразований
К	– множество ключей
$K_{\text{доп}}$	– допустимый коэффициент сокрытия
$K_{\text{крит}}$	– критический коэффициент сокрытия
$K_{\text{мкс}}$	– минимальный предельный коэффициент сокрытия
$K_{\text{скс}}$	– средний коэффициент сокрытия
Л	– множество тестов, методов стеганоанализа
LSB	– метод замены младших значащих битов
М	– множество сообщений
МТ	– алгоритмы согласования
Р	– множество правил извлечения сообщений
RTP	– Real-time Transport Protocol
С	– стеганографические алгоритмы
Q	– подмножество заполненных контейнеров
VoIP	– IP-телефония, Voice over IP
UID	– уникальный идентификатор
АШ	– ассиметричный шифр
ГСК	– генератор сеансового ключа
ЗИ	– защита информации
ЛВС	– локальная вычислительная сеть
МСБП	– метод сдвига битовых последовательностей
НСД	– несанкционированный доступ
ПС	– программное средство
СУБД	– система управления базой данных
СЭД	– система электронного документооборота
СШ	– симметричный шифр

ЦВЗ	– цифровой водяной знак
ЭВМ	– электронно-вычислительная машина
ЭВТ	– электронно-вычислительная техника
ЭД	– электронный документ
ЭЦП	– электронная цифровая подпись

ВВЕДЕНИЕ

В последние годы наметился переход от традиционной формы представления документов к электронным документам. В России не так давно были приняты государственный стандарт электронной цифровой подписи ГОСТ Р 34.10-2001 [132] и Федеральный закон РФ «Об электронной цифровой подписи» [155], что является верным свидетельством серьезных шагов в данном направлении. С принятием в скором времени Федерального закона РФ «Об электронном документе» [147] электронные документы (ЭД) обретут юридическую силу и смогут заменить традиционные документы.

Переход к электронному документообороту несет целый ряд преимуществ. Прежде всего, введение ЭД позволит существенно сократить сроки разработки и прохождения новых документов в структуре предприятия, упростить работу по формированию и пересылке пакетов документов между предприятиями. Использование систем электронного документооборота послужит фундаментом для формирования единого информационного пространства предприятия. Введение электронных архивов документов позволит значительно сократить бумажный архив любого предприятия и обеспечить возможность быстрого поиска и предоставления электронных копий документов. Предполагается также ощутимая экономическая выгода.

Вместе с тем вопросы, связанные с противодействием разглашению, перехвату и передаче третьей стороне электронных документов в настоящий момент все еще остаются нерешенными. В тоже время, все большее количество коммерческих организаций сталкивается с жесткой конкурентной борьбой. Современные условия требуют обязательного документирования по сути всех стратегических и тактических решений принимаемых руководством в целях обеспечения более эффективной работы на рынке. Получение или перехват подобных документов или их копий конкурентами, может повлечь за собой серьезные финансовые потери и (или) подорвать

имидж организации в глазах потенциальных клиентов. Растущая информатизация современного общества и переход к электронным формам хранения и представления информации несут за собой новые потенциальные угрозы информационной безопасности коммерческих организаций.

Существующие механизмы обеспечения защиты информации не в состоянии решить ряд специфических задач характерных для электронного документооборота. В частности использование открытых каналов связи для предоставления, передачи и распространения ЭД чревато возможным перехватом документов третьей стороной. В отличие от бумажного документа, передаваемого в единичном экземпляре, копии ЭД создаваемые при его передаче по каналам связи могут долгое время храниться на почтовых ящиках пользователей и серверах провайдеров. В результате, если доступ к обычным документам возможен только физически непосредственно в процессе их передачи, то для получения доступа к электронному документу в современных условиях злоумышленнику предоставляется большое разнообразие возможностей. Доступ к множеству копий ЭД может быть осуществлен удаленно без непосредственного физического доступа к материальным носителям и растянуть во времени на период хранения электронных копий. При этом ни отправитель, ни получатель электронных документов могут и не догадываться о наличии хранимых копий и факте получения доступа к ним и перехвата исходных электронных документов в процессе передачи третьей стороной.

В условиях невозможности обеспечения абсолютного контроля каналов связи и недопущения несанкционированного доступа (НСД) к информации со стороны третьих лиц, защита информации может быть основана на применении средств криптографии и стеганографии. При этом ни одно из указанных направлений на текущем уровне развития не в состоянии самостоятельно решить все задачи связанные с защитой информации в электронном документообороте. Очевидно, что решение некоторых задач

возможно только при совместном согласованном применении методов криптографии и стеганографии.

Вопросами стеганографической защиты информации в нашей стране в разные годы занимались: Аграновский А.В., Архипов О.П., Барсуков В.С., Быков С. Ф., Варновский Н.П., Голубев Е.А., Грибунин В.Г., Ковалев Р.М., Кустов В.Н., Логачёв О.А., Макаревич О.Б., Оков И.Н., Романцов А.П., Рублев Д.П., Сидоров М.А., Зыков З.П., Федоров В.М, Федчук А.А, и другие.

Возможности криптографии известны, она уже прошла определенные этапы становления, и существующие методы могут быть легко адаптированы к задачам электронного документооборота. В тоже время стеганография в цифровых системах связи направление сравнительно молодое, а известные методы практически не приспособлены к задачам связанным с защитой электронных документов. Кроме того они обладают такими серьезными недостатками, как невысокая надежность и стойкость. Также отсутствуют методы оценки степени защищенности информации. Вместе с тем применение методов стеганографии к электронному документообороту позволит решить целый ряд значимых задач. Среди таких задач можно отметить обеспечение скрытого хранения и передачи электронных документов, сокрытие факта электронного взаимодействия между отправителем и адресатом, предоставление проектов документов для предварительного ознакомления и на этапе согласования, одновременное предоставление копий электронных документов большому количеству удаленных пользователей, выявление каналов утечки ЭД и внесение скрытой электронной цифровой подписи.

Таким образом, представляется актуальной и своевременной задача исследования возможностей и разработки соответствующих методов и моделей систем защиты электронных документов с применением в комплексе современных, перспективных методов скрытой передачи информации и криптографии, а также проработки соответствующей теоретической базы. Следует отметить, что вопросы обеспечения

безопасности электронных документов, при их создании, обработке, хранении и передаче, учитывая современное состояние и тенденции развития средств вычислительной техники и соответствующей законодательно-правовой базы в области электронного документооборота, являются на настоящий момент актуальными как никогда ранее.

Объектом настоящего исследования является информационный обмен в системах электронного документооборота.

Предмет исследования – методы и модели систем защиты ЭД при их передаче по открытым каналам связи, определения каналов утечки и вторичных источников информации в многопользовательских системах.

Цель диссертационной работы – повышение эффективности защиты информации в многопользовательских распределенных системах электронного документооборота на базе современных технологий скрытой связи.

Для достижения поставленной цели в диссертационной работе определены следующие задачи:

1. Анализ уязвимостей и разработка классификации атак на системы скрытой передачи электронных документов.
2. Оценка эффективности современных стеганографических методов и определение границ их применимости, разработка метода и определение критериев оценки практической стойкости стеганографических методов защиты информации.
3. Исследование возможности построения теоретически стойких стеганографических методов и систем.
4. Разработка моделей, принципов и проектных решений на базе методов криптографии и стеганографии для создания перспективных средств защиты электронных документов, обладающих высокой теоретической и практической стойкостью, обоснование эффективности предложенных решений.

5. Разработка новых стеганографических методов, ориентированных на использование в разрабатываемых системах защиты информации, которые бы отвечали необходимым требованиям и обладали высоким уровнем стойкости.
6. Разработка методов, алгоритмов, модели и архитектуры системы скрытой маркировки и проверки маркировки электронных документов в системах электронного документооборота и базах данных.

Методы исследования. Для решения задач использованы методы теории информации и связи, теории вероятностей и математической статистики, методы вычислительной математики, теории принятия решения, теории информационной безопасности и распределенных систем.

Научная новизна исследования заключается в совершенствовании теоретических положений, разработке оригинальных методов и моделей систем технической защиты электронных документов на базе современных положений криптографии и стеганографии.

1. Представлена новая классификация атак на системы скрытого электронного документооборота, основанная на уязвимостях существующего программного обеспечения.
2. Разработан метод и определены критерии оценки практической стойкости стеганографических систем связи.
3. На базе теории конечных автоматов и теории информации доказано существование и показана возможность построения теоретически совершенных стеганографических систем, использующих в качестве контейнеров не отвечающие критериям случайности битовые строки.
4. Введено новое понятие криптостеганографической системы связи, как трехкомпонентной системы включающей криптографические методы и алгоритмы, стеганографические методы и алгоритмы, а также алгоритмы согласования. Определены условия совмещения компонентов в рамках

единой системы, выработаны основные принципы и разработана базовая модель.

5. Предложен новый метод записи информации в битовые строки конечной длины на базе теоретических совершенных стеганографических систем и показано его применение в качестве согласующего алгоритма в крипостеганографических системах скрытой передачи электронных документов.

6. Разработаны новые стеганографические методы, отвечающие уточненным с учетом современных условий функционирования требованиям к стеганографическим методам и алгоритмам защиты информации.

7. Предложена новая обобщенная архитектура многопользовательской распределенной системы скрытого электронного документооборота с детальной проработкой соответствующих протоколов взаимодействия, принципов и схематических решений.

8. Впервые разработаны модель системы, методы и алгоритмы скрытой маркировки и проверки маркировки электронных документов, позволяющие отслеживать перемещение электронных документов, а также локализовать и выявлять каналы утечки информации.

Практическая ценность исследования заключается в том, что его результаты могут быть использованы при проектировании и разработке новых, а также совершенствовании уже существующих систем электронного документооборота, систем управления базами данных и знаний, систем защищенной передачи информации по открытым каналам связи и различных систем защиты с использованием технологии цифровых водяных знаков.

На защиту выносятся следующие основные результаты работы:

1. Классификация атак на системы скрытого электронного документооборота.
2. Критерии оценки практической стойкости стеганографических методов защиты информации.

3. Теоретически совершенная стеганографическая система связи, использующая в качестве контейнеров не отвечающие критериям случайности и содержащие длинные серии одинаковых битов потенциально бесконечные битовые строки.
4. Определение, базовые принципы построения, обобщенная модель и проработанное схематическое решение криптостеганографической системы связи.
5. Новое семейство стеганографических методов защиты информации на базе пространственно-частотных фильтров.
6. Методы, алгоритмы и архитектура системы скрытой маркировки и проверки маркировки электронных документов.

Достоверность результатов обусловлена корректной постановкой задач, использованием современного научного аппарата, применением математически обоснованных методов, использованием известных положений фундаментальных наук, а также сходимостью полученных теоретических результатов с данными экспериментов.

Использование результатов исследования.

Основные результаты исследований использованы:

- при выполнении НИР «Картина-А» в ГНИИ ПТЗИ ФСТЭК России;
- при выполнении НИР «Полтава-ТС» в МТУСИ;
- при выполнении ОКР «Сверчок», НИР «Маркер», СЧ НИР «Медовуха-1-СВ» в ФГНУ НИИ «Спецвузавтоматика».

Апробация работы.

Результаты работы докладывались и обсуждались на следующих конференциях и симпозиумах: I и IV Международная научно-практическая конференция «Теория, методы проектирования, программно-техническая платформа корпоративных информационных систем», Новочеркасск, 2003, 2006 гг.; VI-X Международная научно-практическая конференция «Информационная безопасность», Таганрог, 2004, 2005, 2006, 2007, 2008 гг.; V Международная научно-практическая конференция «Методы и алгоритмы

прикладной математики в технике, медицине и экономике», Новочеркасск, 2005 г.; VI Всероссийский симпозиум по прикладной и промышленной математике, Дагомыс, 2005 г.; Всероссийская научно-практическая конференция «Охрана, безопасность и связь – 2005», Воронеж, 2005 г.; XXXIV Международная конференция «Информационные технологии в науке, социологии, экономике и бизнесе», IT+SE'07, Ялта-Гурзуф, 2007 г.; VIII Всероссийский симпозиум по прикладной и промышленной математике, Адлер, 2007 г. Основные результаты работы были представлены на пленарных заседаниях IX (2007 г.) и X (2008 г.) Международной научно-практической конференции «Информационная безопасность», Таганрог, Россия.

Публикации. По теме диссертации опубликовано 27 печатных работ, в том числе: 1 монография, 8 статей в рецензируемых научных изданиях, из которых 4 – в журналах, рекомендованных ВАК для публикации основных результатов диссертационной работы, получены 2 патента РФ на изобретения, 1 авторское свидетельство об официальной регистрации программ для ЭВМ; 3 статьи и 12 тезисов докладов представлено в материалах международных и всероссийских конференций.

Структура и объем работы. Работа состоит из введения, четырех глав, заключения, списка литературы из 159 позиций, а также четырех приложений. Объем основной части – 169 страниц, 5 таблиц, 27 рисунков.

Первая глава посвящена рассмотрению механизмов обеспечения безопасности электронных документов и защиты информации в системах электронного документооборота. Особо рассматриваются такие направления в технологиях защиты информации как криптография, стеганография и цифровые водяные знаки. Ввиду высокой проработанности теоретической и практической базы методов криптографической защиты информации, особое внимание уделено смежному направлению – стеганографии. Рассмотрены существующие подходы к построению систем скрытого электронного документооборота, используемые на настоящее время методы стеганографии

и стеганоанализа, введена классификация атак на системы скрытого электронного документооборота, определены наиболее перспективные пути совершенствования данных систем.

Во второй главе рассматриваются вопросы теоретической и практической стойкости стеганографических систем связи. Проведен анализ существующих методов стеганоанализа на предмет границ применимости и возможности эффективного противодействия. Введены критерии сравнения и предложены методы оценки практической стойкости стеганографических методов используемых в системах скрытого электронного документооборота. Также введены понятия стеганографической системы совершенно стойкой практически к выбранному методу стеганоанализа и Δ_L -стойкой к атакам пассивного противника по множеству методов анализа L .

На основе теории совершенных систем стеганографических систем; в главе приводится доказательство существования и показана возможность построения совершенных стеганографических систем на потенциально бесконечных битовых строках, не отвечающих требованиям случайности и содержащих длинные серии нулей и единиц.

В третьей главе исследуется возможность построения гибридных систем скрытого электронного документооборота на основе совместного использования криптографических и стеганографических методов защиты информации. Вводится понятие криптостеганографической системы связи. Определены требования и ограничения к используемым алгоритмам. На основе теоретической базы приведенной во второй главе показано, что стойкость криптостеганографических систем к раскрытию факта передачи информации определяется стойкостью используемых криптографических алгоритмов. Также показана высокая гибкость криптостеганографических систем и возможность построения систем, как с симметричными, так и с открытыми ключами. На основе теоретически совершенного стеганографического алгоритма для битовых строк, содержащих длинные серии нулей и единиц, разработан новый метод сдвига битовых

последовательностей. Метод может быть использован при построении различных реальных систем с целью обеспечения согласования криптографической и стеганографической частей.

Для отработки и проверки предложенных решений разработаны новые стеганографические методы на базе пространственно-частотных фильтров обеспечивающие скрытие информации в графических изображениях и аудиозаписях. Разработанные методы отвечают всем заданным требованиям и ограничениям.

В четвертой главе рассматриваются вопросы практического применения крипостеганографических систем связи для решения таких задач электронного документооборота, как скрытая передача электронных документов и контроль за распространением копий электронных документов. На основе обобщения полученного опыта построения экспериментальных крипостеганографических систем предложена базовая архитектура, протоколы, описаны принципы построения и представлены детальные схематические решения. Большое внимание уделено вопросу применения крипостеганографических систем для целей защиты авторского права. Представлены методы, алгоритмы и система скрытой уникальной маркировки и проверки маркировки электронных документов в многопользовательских распределенных СЭД.

В заключении приведены основные результаты работы.

В приложениях представлены дополнительные материалы и некоторые результаты экспериментов. В частности приведены используемые в данной работе обобщенные по ряду источников и уточненные термины и определения из теории стеганографии. В виде отдельного приложения вынесены результаты экспериментального применения метода сдвига битовых последовательностей совместно со стеганографическим методом замены младших значащих битов в графических изображениях. Описаны особенности использования скрытых каналов передачи информации с малой пропускной способностью.

ГЛАВА 1. ИССЛЕДОВАНИЕ И АНАЛИЗ СУЩЕСТВУЮЩИХ РЕШЕНИЙ В ОБЛАСТИ ЗАЩИТЫ ЭЛЕКТРОННЫХ ДОКУМЕНТОВ

Данная глава посвящена подробному рассмотрению предметной области исследования. Определены понятия электронного документа и электронного документооборота, проведен анализ существующего программного обеспечения и известных программно реализуемых механизмов защиты информации. Выделены такие направления как криптография и технологии скрытой связи (стеганография и цифровые водяные знаки). Проведен анализ существующих решений в области технологий скрытой связи, исследованы существующие модели, методы и алгоритмы. Введена классификация атак на системы скрытой передачи электронных документов. Определены пути совершенствования систем скрытой передачи документов. В качестве наиболее перспективного направления предложено совместное использование криптографии и технологий скрытой связи.

1.1 Электронный документооборот

1.1.1 Понятие электронного документа

Несмотря на активное использование электронных документов множеством коммерческих и государственных предприятий, до настоящего момента закон об электронном документе так и не принят. И потому нет точного и однозначного определения, что же представляет собой электронный документ. В федеральном законе «Об обязательном экземпляре документов» [154] введено следующее понятие документа:

Документ – материальный носитель с зафиксированной на нем информацией в виде текста, звукозаписи (фонограммы), изображения или их сочетания, предназначенный для передачи во времени и пространстве в целях общественного использования и хранения.

Согласно ГОСТ Р 51141-98 «Делопроизводство и архивное дело. Термины и определения» [137] приняты следующие определения:

Документ; документированная информация – зафиксированная на материальном носителе информация с реквизитами, позволяющими ее идентифицировать;

Изобразительный документ – документ, содержащий информацию, выраженную посредством изображения какого-либо объекта;

Графический документ – изобразительный документ, в котором изображение объекта получено посредством линий, штрихов, светотени;

Аудиовизуальный документ – документ, содержащий изобразительную и звуковую информацию;

Кинодокумент – изобразительный или аудиовизуальный документ, созданный кинематографическим способом;

Фотодокумент – изобразительный документ, созданный фотографическим способом;

Фонодокумент – документ, содержащий звуковую информацию, зафиксированную любой системой звукозаписи;

Текстовый документ – документ, содержащий речевую информацию, зафиксированную любым типом письма или любой системой звукозаписи;

Письменный документ – текстовый документ, информация которого зафиксирована любым типом письма;

Машинописный документ – письменный документ, при создании которого знаки письма наносятся техническими средствами;

Документ на машинном носителе – документ, созданный с использованием носителей и способов записи, обеспечивающих обработку его информации электронно-вычислительной машиной.

Таким образом, из анализа существующей нормативно-правовой базы можно заключить, что понятие электронного документа должно быть довольно широким, учитывая современный уровень развития вычислительной техники. Так под понятие электронного документа должны подпадать не только текстовые документы, представленные в электронном виде, но и графические, фотографические, аудиовизуальные и видео

документы. Следует отметить, что все указанные виды документов с использованием электронно-вычислительной техники (ЭВТ) изначально могут создаваться в электронном виде или же могут быть приведены к электронному виду (закодированы, оцифрованы). Хранение таких документов в ЭВМ реализовано в виде файлов в файловых хранилищах, на магнитных, оптических дисках, или же в виде записей в базах данных. Таким образом, для ЭВМ документ представляет собой упорядоченный определенным образом набор битов (состояний элементов ЭВТ), содержащий идентифицирующую информацию (заголовок, информацию о типе и формате хранения данных) и непосредственно информацию самого электронного документа (данные).

Понятие электронного документа представлено в федеральном законе «Об электронной цифровой подписи» [155]:

Электронный документ – документ, в котором информация представлена в электронно-цифровой форме.

Учитывая данное определение, электронный документ может быть представлен в виде строки бит ограниченной длины, содержащей информацию (данные документа) в виде текста, звукозаписи (фонограммы), изображения или их сочетания и информацию о реквизитах, позволяющих идентифицировать информацию документа. Таким образом, при хранении и передаче электронных документов, на уровне отвечающего за хранение и передачу документов программного обеспечения, электронный документ может быть представлен в виде строки бит ограниченной длины. Такое представление электронных документов позволяет оперировать с ними как с любой другой информацией представленной в электронно-цифровой форме. Для целей сжатия, проверки целостности и защиты электронных документов при их хранении и передаче могут быть использованы стандартные алгоритмы, реализующие указанные процедуры.

Несколько отличающееся определение электронного документа представлено в проекте федерального закона «Об электронном документе» [147], согласно которому:

Электронный документ – форма подготовки, отправления, получения или хранения информации с помощью электронных технических средств, зафиксированная на магнитном диске, магнитной ленте, лазерном диске и ином электронном материальном носителе.

При этом отдельно указывается, что информация, зафиксированная на электронном материальном носителе, признается электронным документом, если она:

- создается, обрабатывается, хранится и передается с помощью электронных технических средств;
- подписана с соблюдением требований, предусмотренных действующим законодательством;
- может быть представлена в форме, пригодной для восприятия человеком, не обладающим специальными техническими навыками;
- если при его составлении, хранении, передаче использован предусмотренный государственными или международными стандартами либо соглашением сторон способ, позволяющий достоверно идентифицировать составителя электронного документа.

Представленное определение, а также первый и третий пункты уточнения не противоречат представлению электронного документа на уровне средств обработки в виде строки бит конечной длины. Информация должна быть представлена в форме пригодной для восприятия человеком лишь на этапе визуализации. Второй и четвертый пункты в неявной форме говорят о необходимости использования механизмов электронно-цифровой подписи. Однако электронно-цифровая подпись является необходимой лишь для придания юридической силы электронному документу, что не всегда

является обязательным. Вместе с тем, электронно-цифровая подпись, как и прочая информация о документе относится к его реквизитам и также может быть представлена строкой бит конечной длины. В результате, электронный документ, снабженный ЭЦП, также как и электронный документ без ЭЦП может быть представлен в виде строки бит ограниченной длины.

1.1.2 Системы электронного документооборота

Замена обычных документов на их электронные аналоги сама по себе не дает существенных преимуществ. Существенно повысить эффективность работы с документами позволяет использование на предприятиях автоматизированных систем электронного документооборота (СЭД). Само понятие документооборот использовалось еще с обычными бумажными документами и согласно ГОСТ Р 51141-98 [137] означает движение документов в организации с момента их создания или получения до завершения исполнения или отправления. Согласно проекту федерального закона «Об электронном документе» [147], для электронных документов введено следующее понятие:

Электронный документооборот – система составления, использования, хранения и обмена электронными документами с использованием электронных средств массовых коммуникаций.

В целях организации электронного документооборота используются различные СЭД. В большинстве систем документы хранятся в специальных хранилищах или с использованием обычных файловых систем. При этом механизмы разграничения доступа не всегда реализованы грамотно, и порой позволяют получить доступ к электронным документам в обход СЭД. Отсутствие же в ряде систем встроенных средств шифрования позволяет легко получить доступ к документам посредством стандартного файлового браузера. В таблице 1 по обзору [146] представлены наиболее популярные в России системы электронного документооборота, отмечены используемые в них механизмы криптографической защиты информации.

Таблица 1 – Системы электронного документооборота

СЭД	Способ представления ЭД	Шифрование и ЭЦП	СЭД	Способ представления ЭД	Шифрование и ЭЦП
CORPORATE BUSINESS	Файл любого расширения	+ PGP, CryptoC, КриптоПро, CryptoCOM	СУПеР	Файл любого расширения	+
DocsVision	Файл любого расширения	+ КриптоПро	PayDox	Файл любого расширения	-
LanDocs	Файл любого расширения	+ КриптоПро, Верба, LanCrypto	ИНТАЛЕВ-Документооборот	Файл любого расширения	-
БОСС-Референт	Файл любого расширения	+ Верба, Криптон, КриптоПро, Домен-К	OPTIMA-WORKFLOW	Файл любого расширения	+ КриптоПро
DIRECTUM	Файл и поля на учетной карточке	+ КриптоПро, LanCrypto	МОТИВ	Файл любого расширения	+ КриптоПро
ДЕЛО	Файл любого расширения	+ КриптоПро	Effect Office	Файл + текстовые поля на карточке + дополнительные атрибуты	-
NAUDOC	Файл любого расширения	-	jDocflow	Файл любого расширения	-
ЕВФРАТ-Документооборот	Файл + текстовые поля на карточке + дополнительные атрибуты	+	Avacco	Файл любого расширения	-

Как видно из представленной таблицы, далеко не все системы электронного документооборота в полной мере предоставляют необходимые механизмы защиты данных. Следует заметить, что большая часть систем позволяет использовать механизмы электронной цифровой подписи, но при этом не позволяет обеспечить защиту данных посредством алгоритмов шифрования. Наиболее популярным средством криптографической защиты данных является КриптоПро (<http://www.cryptopro.ru/>).

Следуя определению электронного документооборота, приведенному выше, системы электронного документооборота должны обеспечивать обмен электронными документами с использованием электронных средств

массовых коммуникаций. Использование электронных средств массовых коммуникаций подразумевает передачу электронных документов посредством открытых каналов связи. Вместе с тем, как было отмечено выше, во многих системах механизмы криптографической защиты информации не используются. Вследствие чего, информация оказывается легко доступной для третьей стороны.

1.2 Анализ программно реализуемых механизмов защиты информации циркулирующей в СЭД

1.2.1 Методы и средства обеспечения безопасности информации

Под *безопасностью информации* понимается состояние устойчивости информации к случайным или преднамеренным воздействиям, исключающее недопустимые риски ее уничтожения, искажения и раскрытия, которые приводят к материальному ущербу владельца или пользователя информации. Обеспечение безопасности информации неразрывно связано со следующими основными принципами:

1. *Конфиденциальность* – секретные данные должны быть доступны только авторизованным пользователям;
2. *Целостность* – неавторизованные пользователи не должны обладать возможностью изменения и модификации данных;
3. *Доступность* – защищаемая информация всегда должна быть доступна авторизованным пользователям.

Под авторизованным пользователем здесь понимается владелец или же законный пользователь информации. Любое потенциальное действие, которое направлено на нарушение конфиденциальности, целостности или доступности информации называется *угрозой*. Любая реализованная угроза называется *атакой*. Основные понятия безопасности информационных технологий регламентированы в основополагающих документах [130, 131, 133, 134, 135, 136, 148]. В ГОСТ Р 50922-2006 [136] определено четыре вида методов и средств защиты информации:

- правовая защита информации;
- техническая защита информации;
- криптографическая защита информации;
- физическая защита информации.

В большей части литературных источников представлена несколько иная классификация, согласно которой среди основных направлений в методах и средствах обеспечения безопасности информации выделяют:

- технические методы и средства защиты информации;
- программно-аппаратные методы и средства защиты;
- криптографические и стеганографические методы и средства;
- организационные методы защиты информации.

Несмотря на достаточно серьезные различия в представленных классификациях общим для них является то, что в обоих вариантах криптографические методы защиты информации представлены обособленно. Учитывая заданную предметную область данного исследования, связанную с защитой ЭД при их передаче по открытым каналам связи, в качестве основного направления исследований в диссертационной работе выбраны криптографические и стеганографические методы и средства защиты информации. Выбор данного направления обусловлен также возможностью программной реализации методов и средств защиты электронных документов. Возможность программной реализации методов и средств защиты следует считать приоритетной, ввиду того, что работа с ЭД неразрывно связана с их обработкой и хранением с использованием ЭВМ, вычислительных комплексов и сетей передачи данных [98]. Кроме того, использование открытых сетей передачи данных, автоматических делает невозможным применение технических и организационных методов защиты информации. В тоже время криптографические и стеганографические методы защиты могут найти отражение в программно-аппаратных средствах защиты.

Особое место в криптографических и стеганографических методах защиты информации занимает направление, связанное с защитой авторского

права, доказательством целостности и аутентичности сообщений. Сравнительно молодое направление цифровых водяных знаков, фактически сформировалось на стыке методов криптографии и стеганографии. Вопросы защиты авторского права на электронные произведения и электронные документы в частности сейчас являются весьма актуальными. Учитывая это, в настоящей работе данному направлению уделено отдельное внимание.

1.2.2 Криптографические методы и механизмы защиты информации

Криптография изучает методы преобразования информации, обеспечивающие конфиденциальность, контроль целостности и аутентичность информации. В современной криптографии можно выделить четыре относительно обособленных раздела [140, 157, 159]: симметричные крипtosистемы, крипtosистемы с открытым ключом, системы контроля целостности и электронно-цифровой подписи, управление ключами.

В основе любой симметричной криптографической системы, как и крипtosистемы с открытым ключом лежит шифр. Под *шифром* понимается совокупность обратимых преобразований множества возможных открытых данных на множество зашифрованных данных, осуществляемых по определенным правилам с применением ключей [130]. Применение шифров позволяет преобразовать обычное текстовое сообщение в совершенно непонятный и с виду бессмысленный набор символов. Такое преобразование открытых данных называется процедурой *зашифрования данных*, а обратное преобразование (восстановление исходного сообщения) – *расшифровкой*. Зашифрованные данные становятся непонятными и не поддающимися прочтению, для каждого кто не знает секрета шифрования, который бы позволил восстановить исходные данные. Этим секретом может быть как сам алгоритм шифрования, так и некоторая другая информация, управляющая процессами зашифрования и расшифрования.

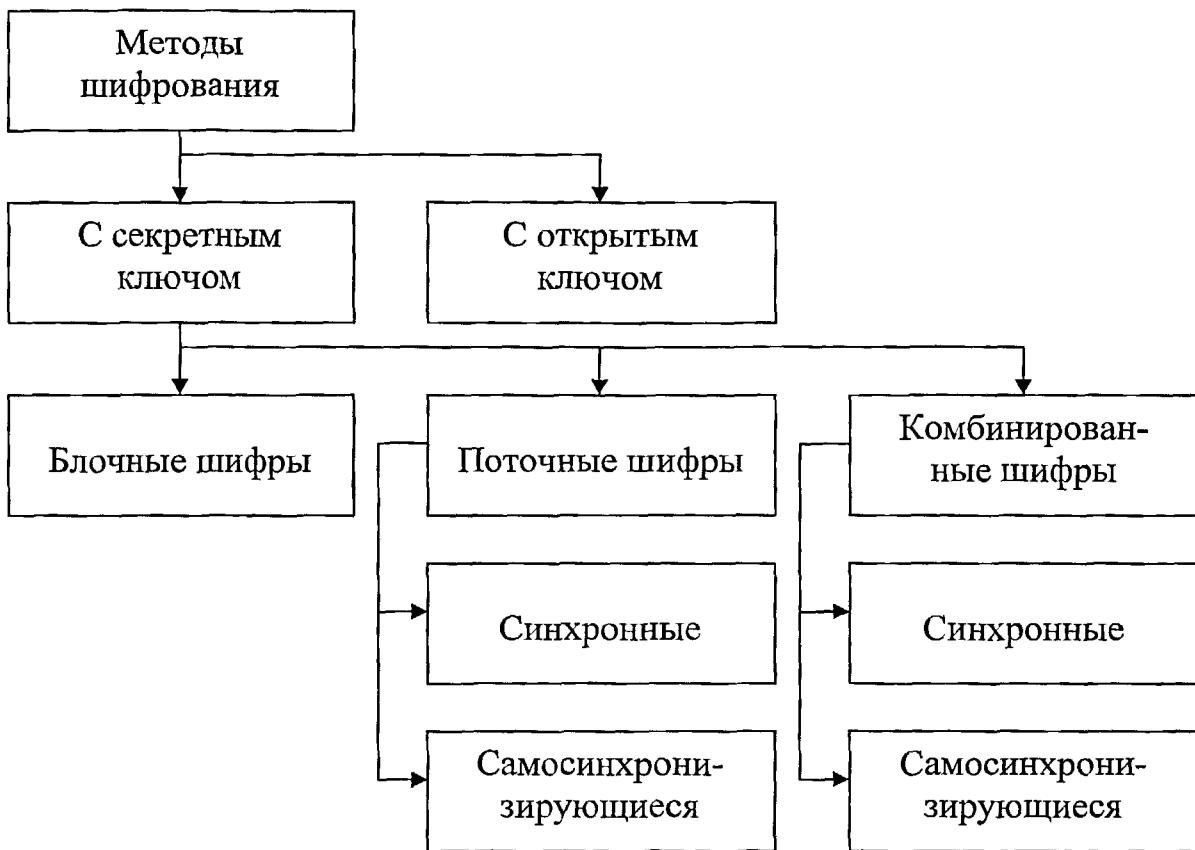


Рисунок 1 – Классификация методов шифрования информации

Особенностью симметричных криптографических методов и систем является то, что один и тот же секрет (или секретная информация) используется как на стороне отправителя сообщения, так и на стороне получателя. В современных крипtosистемах в качестве секрета выступает *криптографический ключ* – конкретное секретное состояние некоторых параметров алгоритма криптографического преобразования данных, обеспечивающее выбор одного преобразования из совокупности всевозможных для данного алгоритма преобразований. Непременным требованием к современным симметричным крипtosистемам является требование открытости алгоритмов и протоколов. Единственным элементом, который допускается хранить в секрете, является ключ шифрования. На сегодняшний день среди множества алгоритмов симметричного шифрования присутствуют алгоритмы, ставшие государственными стандартами. Среди них российский ГОСТ 28147-89 [130], стандарты шифрования США – AES, DES

и международный европейский стандарт – IDEA. То обстоятельство, что эти алгоритмы стали открытыми стандартами шифрования, говорит о высокой надежности алгоритмов и их востребованности как в государственной, так и в коммерческой сфере. Следует особо отметить, что данные алгоритмы широко применяются в различных системах защищенного электронного документооборота. Недостатком симметричных алгоритмов является необходимость использования одного и того же ключа на стороне отправителя и на стороне получателя сообщений. Сложность заключается в необходимости обеспечения секретности ключа шифрования при его передаче адресату, что требует использования дополнительных защищенных каналов связи.

В 1976 году Уитфилдом Дифи и Мартином Хеллманом была опубликована работа «Новые направления в криптографии» [24] в которой было представлено уникальное решение задачи шифрования данных, в котором впервые было введено понятие систем с открытым ключом. По сути, предложенное решение стало революционным, так как позволило использовать на стороне отправителя открытый ключ шифрования, о секретности которого можно не беспокоиться. Даже если открытый ключ станет известен злоумышленнику, восстановить по нему ключ расшифрования, и, следовательно, прочесть сообщение он не сможет. Прочесть зашифрованное сообщение сможет только непосредственный его получатель, используя вторую секретную половину ключа. Такое решение значительно упрощает порядок обмена зашифрованной информацией, так как не требует наличия предварительной договоренности о секретных ключах шифрования. Сегодня криптографические алгоритмы с открытым ключом широко используются в коммерческой деятельности. Стандартом де-факто в этой области стал алгоритм RSA, предложенный еще в 1977 году Роном Ривестом, Ади Шамиром и Леонардо Адлеманом [73].

Существенным для текущего уровня развития техники недостатком алгоритмов шифрования с открытым ключом является их крайне низкая скорость работы, что не позволяет им вытеснить алгоритмы симметричной

криптографии. Учитывая плюсы ассиметричной криптографии, популярной является схема, когда симметричные алгоритмы используются для шифрования данных, а ассиметричные для обмена сеансовыми ключами. Применяемые в такой схеме ассиметричные алгоритмы позволяют фактически организовать необходимый для симметричных алгоритмов дополнительный защищенный канал обмена ключами шифрования.

Отдельное направление в теории криптографии связано с методами контроля целостности и имитозащиты (аутентичности) данных. Под целостностью данных понимается свойство, при выполнении которого сохраняются неизменность данных и существование в том виде, в котором они были созданы. В целях обеспечения возможности контроля целостности данных в условиях возможных воздействий на них при хранении и передаче по каналам связи используются хэш-функции – односторонние сжимающие функции [157, 159]. Это математические или иные функции, которые принимают на вход строку переменной длины и преобразуют ее в строку фиксированной длины, элементы которой зависят от всех элементов входной строки. Хэш-функции делятся на безключевые, и хэш-функции с секретным ключом. Примерами безключевых хэш-функций являются функции MD2, MD4, MD5, SHA-1, SHA-2, HEVAL, RIPEMD-160, ГОСТ Р 34.11-94 [133]. В качестве примера хэш-функции с секретным ключом можно привести ГОСТ 28147-89 [130]. В стандарте, помимо уже стандартных режимов шифрования (простой замены, гаммирования, гаммирования с обратной связью), описан режим выработки имитовставки. Этот режим позволяет параллельно с процессом зашифрования/расшифрования данных вычислить имитовставку – отрезок информации фиксированной длины полученной по определенному правилу из открытых данных и ключа, и добавленный к зашифрованным данным для обеспечения имитозащиты.

Логическим продолжением алгоритмов имитозащиты и ассиметричных алгоритмов шифрования стали алгоритмы электронной цифровой подписи

(ЭЦП). Алгоритмы электронной цифровой подписи для электронных документов являются по своей сути, прямым аналогом обычной рукописной подписи, поставленной под бумажным документом. Аналогичны и требования предъявляемые к ЭЦП. Главным из них является возможность проверки достоверности подписи любым из участников электронного обмена, в то время как подписать электронный документ может только законный владелец ЭЦП. Данное требование выполняется благодаря использованию асимметричных криптографических алгоритмов. Типичная схема в этом случае предполагает использование безключевой хэш-функции для генерации свертки (хэш-образа документа) и дальнейшее ее шифрование на секретном ключе подписывающей стороны. Для проверки подписи используется открытый ключ. Учитывая свойство асимметричных алгоритмов, восстановить секретный ключ по открытому ключу практически невозможно. В результате любой желающий, зная открытый ключ проверки подписи, может ее проверить, но не имеет возможности ее подделать.

Электронная цифровая подпись является основой современного электронного документооборота. В России действует государственный стандарт на электронную цифровую подпись ГОСТ Р 34.10-2001 [132].

Практика использования криптографических систем показала, что для обеспечения их эффективного использования необходима детальная проработка решения ряда дополнительных задач связанных с наиболее значимой компонентой – секретными ключами шифрования. Задачи генерации, распространения, хранения, уничтожения относят к отдельному разделу в криптографии, который получил название управление ключами. Управление ключами – это совокупность технологий и процедур, посредством которых устанавливаются и поддерживаются ключевые отношения между участниками криптографического протокола. Вопросы грамотного управления ключами столь серьезны, что им посвящен отдельный международный стандарт ISO/IEC 11770. Жизненный цикл криптографических ключей и вопросы управления ключами подробно

рассмотрены в [140, 157]. Следует отметить, что построение любой защищенной системы электронного документооборота, в основе которой лежат криптографические алгоритмы требует соответствующей проработки вопросов генерации, хранения, передачи и уничтожения ключей.

1.2.3 Стеганографические методы защиты информации

Используемые в различных системах скрытой передачи электронных документов стеганографические методы относятся к сравнительно молодым направлениям технической стеганографии – цифровой и компьютерной стеганографии. Первые признаки начала формирования данных направлений относят к концу XX века. Так, например, единая терминология была принята только в 1996 году на первом симпозиуме, посвященном вопросам скрытой передачи информации "Information Hiding – First information Workshop" [70].

Согласно общепринятыму определению, *стеганография* – направление в технологиях сокрытия информации, рассматривающее вопросы скрытого взаимодействия, обмена и передачи информации с использованием открытых каналов связи. Целью стеганографии является организация защищенного обмена сообщениями между двумя сторонами, при котором от третьей стороны, осуществляющей контроль над используемым каналом связи, скрывается не только содержимое сообщений, но и факт их передачи. То есть методы стеганографии позволяют обеспечить конфиденциальность информации, но в более широком смысле, по сравнению с криптографией. Вместе с тем, обеспечивая конфиденциальность важной информации, стеганография не предполагает противодействия нацеленным на перехват открыто передаваемых сообщений попыткам третьей стороны. На первый взгляд имеется явное противоречие, которое следует прояснить. Важный момент заключается в том, что применение методов стеганографии позволяет на основе имеющегося открытого канала связи организовать в некотором роде второй параллельный ему канал, скрытый от третьей стороны. Соответственно сообщения, передаваемые по открытому каналу, остаются доступными третьей стороне, в то время как существование параллельного

канала тщательно скрывается. Скрытый канал передачи сообщений формируется на основе имеющегося открытого канала путем незначительных и незаметных изменений сообщений, передаваемых по открытому каналу.

Проблема организации скрытого или потайного канала в условиях строго контроля осуществляемого третьей стороной была достаточно подробно рассмотрена Густавом Симмонсом в 1983 году в работе [77], получившей название «Проблема заключенных». Им был приведен пример демонстрирующий суть проблемы и предложена классическая модель стеганографического канала связи. Данная модель, с учетом с учетом терминологии приведенной в статье [70], а также по примерам, приведенным в работах [70, 44, 21] представлена на рисунке 2.

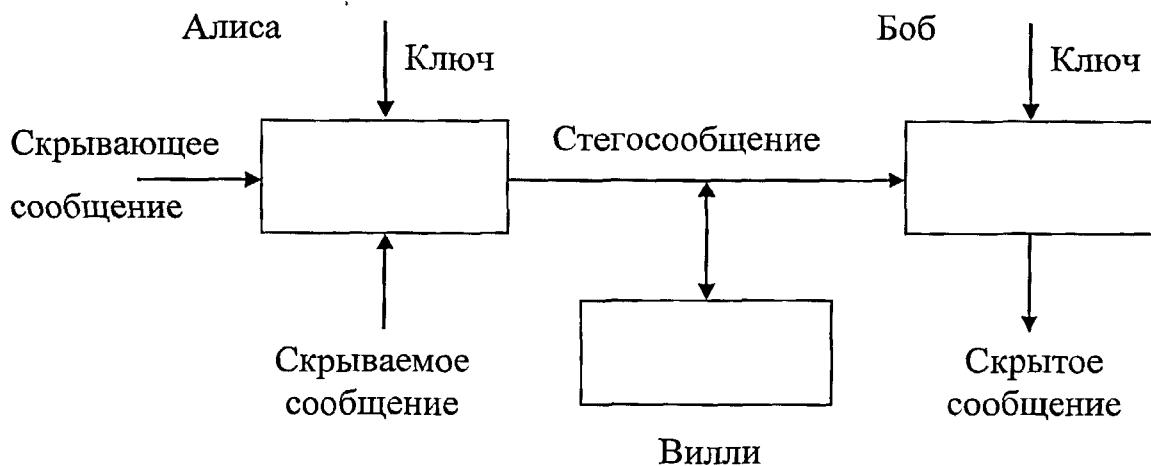


Рисунок 2 – Модель стеганографического канала в проблеме заключенных

Симмонсом приведен следующий пример. Двоих обвиняемых в преступлении заключенных Алису (Alice) и Боба (Bob), сажают в разные тюремные камеры. С того момента, как они оказались запертыми в своих камерах, единственной возможностью для общения у них остается только возможность передачи посланий через одного из охранников. С помощью этих сообщений Алиса и Боб должны договориться о плане совместного побега. Сложность заключается в том, что посредник является агентом надзирателя Вилли (Willy), который внимательно просматривает все передаваемые послания и не позволяет обмениваться зашифрованными или

подозрительными сообщениями. Надзиратель, зная о том, что заключенные обмениваются посланиями, не противодействует этому в надежде контролировать их планы. Он может пытаться обмануть и заставить раскрыться хотя бы одного из них путем подмены посланий или путем различных манипуляций с передаваемыми сообщениями. Как следствие, перед Алисой и Бобом возникает сложная задача, которая заключается в том, чтобы придумать способ обмена секретными сообщениями через открытые послания. При этом надзиратель не должен ничего заподозрить и суметь исказить смысл передаваемых сообщений. Данный пример демонстрирует основные идеи, и принципы организации стеганографического канала связи. Так в модели электронного документооборота роль посредника выполняют сети передачи данных, например открытая сеть Интернет, отправитель и адресат (в примере заключенные Алиса и Боб) могут представлять собой филиалы предприятия, а злоумышленника (Вилли) – конкуренты. В отличие от криптографии в случае применения для защиты электронных документов стеганографических методов, для конкурентной разведки будет недоступна информация не только о самом передаваемом документе, но и о том передавался ли этот или какой либо другой документ вообще.

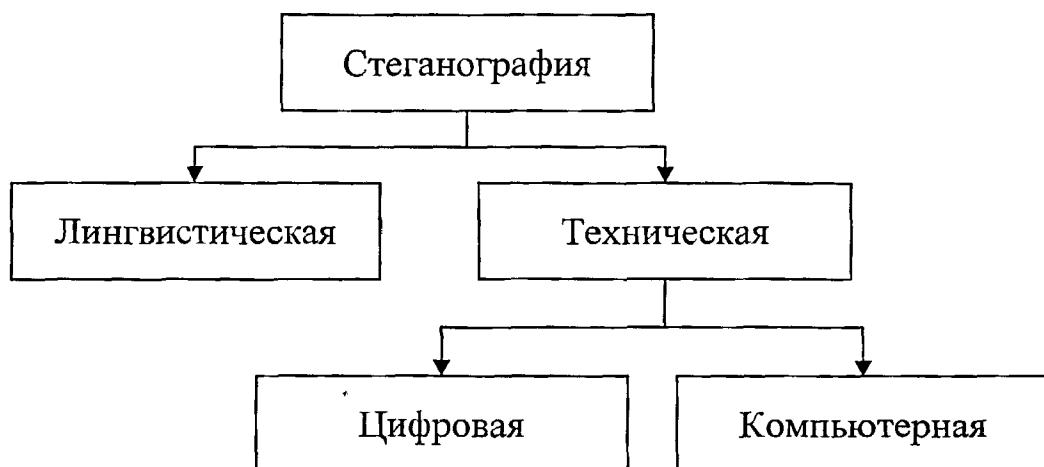


Рисунок 3 – Основные направления современной стеганографии

Современная стеганография практически не ограничена в выборе носителей и методов для передачи скрытой информации. Развитие средств

вычислительной техники и быстрый рост объемов электронного документооборота в конце XX века, равно как все более широкое использование вычислительных сетей в начале XXI века, дало толчок появлению и развитию соответствующих направлений в стеганографических системах скрытой связи. Во второй половине XX века стали выделяться два новых направления связанных с обработкой данных на ЭВМ и представлением информации в цифровой форме – компьютерная и цифровая стеганография. Отличия между этими двумя направлениями и современная терминология подробно представлены в приложении А, а современные методы цифровой и компьютерной стеганографии, в зависимости от используемых типов контейнеров приведены в пп. 1.3.2-1.3.3.

Говоря об электронных документах нельзя не упомянуть методы лингвистической стеганографии. Их важной особенностью является то, что в качестве открытых маскирующих сообщений используются обычные текстовые сообщения, а механизм сокрытия информации основан на использовании неочевидных правил записи секретного сообщения. Методы лингвистической стеганографии сыграли большую роль в истории, но и сегодня, с переходом к электронным формам представления текстовых документов они не утратили своей актуальности. Большинство методов основано на использовании достаточно простых правил сокрытия одного текстового сообщения в другом. Более того, зачастую скрывающее сообщение формируется непосредственно на основе данных скрываемого сообщения. Выявить в скрывающих сообщениях какие-либо специфичные особенности даже в современных условиях с применением ЭВМ во многих случаях является сложной задачей. В тоже время отправителю и получателю сообщений порой достаточно просто договориться о правилах извлечения скрытых сообщений, даже не оговаривая порядка формирования маскирующих сообщений. Подробно наиболее известные и популярные методы лингвистической стеганографии рассмотрены в п. 1.3.1.

1.2.4 Технологии цифровых водяных знаков

Направление, связанное с защитой авторского права прочно занимает лидирующие позиции, что не в последнюю очередь обусловлено достаточно серьезным коммерческим интересом. Начало данному направлению было положено принятием 6 сентября 1952 года в Женеве «Всемирной конвенции об авторском праве». Особенность авторского права согласно принятой конвенции заключается в том, что права возникают и охраняются с момента создания произведения в силу самого факта его создания. По принятым международным соглашениям выполнение каких-либо специальных процедур регистрации авторских прав не требуется. Согласно законодательству РФ [139] в области защиты авторского права, авторские права вступают в силу с момента создания произведения. Вместе с тем, сами авторы для подтверждения своих прав часто прибегают к процедуре депонирования и регистрации своих произведений в авторско-правовых организациях. Причиной тому служит то, что в действительности имеет место правоприменительная практика не по дате создания, а по дате публикации произведения. В результате, если в период предварительной подготовки (до публикации автором и правообладателем) электронного произведения соответствующие ему данные оказались в руках третьей стороны, то она может заявить свои претензии на авторство, путем их публикации от своего имени. Не менее остро стоят вопросы защиты интеллектуальной собственности и в сфере промышленности. Далеко не все используемые в производстве технологии могут быть своевременно запатентованы и на них получены охранные документы. Защита же этой информации необходима.

С развитием ЭВТ все больше объектов авторского права создается и хранится на электронных вычислительных машинах и передается с использованием цифровых сетей связи. Одним из возможных механизмов технических средств защиты авторского права, согласно [139, ст. 48] является сопровождение электронного произведения идентифицирующей

информацией. При этом в качестве идентифицирующей информации может выступать не только текстовое описание, но и специальные коды. К этим кодам относятся *цифровые водяные знаки* (ЦВЗ) – специальные цифровые метки, вносимые в данные электронного произведения [145]. Цифровые водяные знаки являются наиболее известной и популярной на настоящий момент технологией в области защиты авторского права на графические и музыкальные произведения, представленные в электронном виде. Основными их преимуществами являются глубокая интеграция в данные электронного произведения и скрытность от рядового пользователя. В ряде случаев только авторы и законные правообладатели произведений, могут представить необходимые доказательства.

Широкий интерес к данному направлению обусловлен особенностями защиты авторского права на электронные произведения, и аналогией с обычными водяными знаками. Как и в случае водяных знаков на бумажных документах, технология цифровых водяных знаков предполагает встраивание неразличимых для человека, но в тоже время легко выявляемых в определенных условиях, специальных цифровых меток. В случае использования ЦВЗ для защиты цифровых изображений, достаточно часто применяется подход, основанный на суммировании информации защищаемого изображения с информацией практически прозрачного идентифицирующего изображения – логотипа. Наличие невидимого логотипа может быть установлено перемножением данных проверяемого изображения с данными идентифицирующего изображения.

Высокие требования к скрытности ЦВЗ и объемам записываемой информации обычно не предъявляются, так как для доказательства авторства не требуется больших объемов идентифицирующей информации. Идентифицирующие данные ЦВЗ зачастую представляет собой небольшие метки объем информации, в которых, исчисляется единицами или десятками байтов. Основным же требованием, предъявляемым к системам цифровых водяных знаков, является требование обеспечения высокой стойкости

записанных данных к уничтожению и модификации вследствие целенаправленного искажения носителя [19, 41, 51, 53, 69, 95].

Более подробно применение ЦВЗ для защиты авторского права на электронные произведения рассмотрены в монографии [105]. Здесь же следует отметить, что вопросы использования ЦВЗ для защиты авторского права и отслеживания каналов перемещения электронных произведений представляющих собой текстовую информацию на настоящее время остаются практически непроработанными.

1.2.5 Классификации атак на системы связи

Оценка надежности криптографических систем связи предполагает наличие противника обладающего возможностью контролировать используемый канал передачи данных, или, по крайней мере, перехватывать передаваемые по нему зашифрованные сообщения. В тоже время, криптографическая система может считаться надежно только в том случае, если противник при всех имеющихся у него возможностях не сможет подменить или хотя бы прочесть ни одного из перехваченных им сообщений. Соответственно раскрытие криптографической системы предполагает раскрытие сути передаваемых сообщений, восстановление открытого текста по шифротексту, или в более сложном варианте определение используемого секретного ключа. *Атакой на криптографическую систему* принято называть действия, предпринимаемые противником с целью раскрытия сути передаваемых сообщений или восстановления секретного ключа. Общепринятая классификация классических атак на криптографические системы приведена в [141, 157, 159].

В теории стеганографии противник также наблюдает передачу сообщений по контролируемому им каналу связи. Несмотря на то, что все передаваемые сообщения являются открытыми, он не может обнаружить и прочесть скрытых в них секретных посланий. Противник может перехватывать, анализировать, а возможно и подменять передаваемые открытые сообщения. Под *раскрытием (взломом)* стеганографической

системы принято понимать нахождение такой ее конструктивной либо иной уязвимости, которая позволяет определить факт сокрытия информации в контейнере, и возможность доказать данное утверждение третьей стороне с высокой степенью достоверности [126]. В широком смысле под *уязвимостью* понимается слабое место в системе, с использованием которого может быть осуществлена атака. Другими словами уязвимость системы это некоторая ее неудачная особенность, незамеченная ошибка или какой-либо недостаток, которые делают возможным возникновение угрозы. Соответственно любая атака на систему может быть осуществлена только при наличии тех или иных уязвимостей. *Атакой на стеганографическую систему* называют действия противника с целью определить наличие скрытого сообщения, оказать влияние на скрытое сообщение, установить его суть, восстановить алгоритм скрывающего преобразования или используемый секретный ключ.

При проведении атаки на стеганографическую систему противник может преследовать различные цели, как-то обнаружить скрытый канал, перехватить контейнер с сообщением, извлечь или подменить само сообщение, попытаться оказать воздействие на систему в целом. Так же противник может обладать различной информацией о системе, используемых контейнерах или передаваемых сообщениях. Все это определят цели и возможности противника по проведению успешной атаки. По характеру воздействия на стеганографическую систему возможные со стороны противника, согласно [42], атаки можно разделить на четыре группы:

1. Установление факта скрытой передачи информации;
2. Извлечение скрытого сообщения;
3. Искажение сути или подмена скрытого сообщения;
4. Уничтожение скрытого сообщения.

Наиболее распространенная классификация атак на стеганографические системы [13, 23, 42], принятая по аналогии с классификацией атак на криптографические системы, представлена ниже.

Атака на основе известного заполненного контейнера – аналог атаки по шифротексту в криптографии. Противник располагает только одним или несколькими заполненными контейнерами. Обладая только модифицированными контейнерами, противник может провести атаку с целью обнаружения факта скрытой передачи сообщений, восстановить стеганографический алгоритм, а так же попытаться извлечь передаваемые сообщения и определить ключ стеганографического преобразования.

Атака на основе известного скрытого сообщения – противник располагает парами «скрытое/открытое» сообщение. Такая атака характерна для систем цифровых водяных знаков, в случае с которыми противнику изначально может быть известно скрытое сообщение, торговая марка или логотип. Задача противника состоит в подборе ключа стеганографической системы, с тем, чтобы в последующем иметь возможность самому встраивать аналогичные сообщения или модифицировать уже имеющиеся.

Атака на основе выборочного скрытого сообщения – аналогична предыдущей атаке, с той лишь разницей, что противник сам способен навязывать стеганографической системе сообщения, предназначенные для последующей передачи, но при этом не может влиять на выбор контейнера.

Атака на основе выбранного сообщения – под сообщением в данном случае подразумевается и скрываемое сообщение, и контейнер. Противник в данном случае может оказывать влияние на стеганографическую систему, как со стороны отправителя, так и со стороны получателя. Данная атака применяется для случая «черного ящика», когда используемая стеганографическая система, принципы ее работы или используемый ключ неизвестны. Атака зачастую проводится с целью выявления характерных для данной стеганографической системы искажений вносимых в передаваемые контейнеры, а также разбора алгоритмов работы и используемых ключей, в случае если последние держаться в секрете.

Помимо указанных классов атак, имеющих прямые аналоги в классификации атак на криптографические системы, в стеганографии противнику доступны еще три специфичных класса:

Атака на основе известного оригинального контейнера – наиболее простая и в тоже время серьезная из атак, которые могут быть предприняты противником. В данном случае противник обладает парами «пустой/заполненный» контейнер. Естественно, что решение задачи обнаружения скрытого канала в данном случае является тривиальным.

Атака на основе известной математической модели контейнера. По заданному множеству пустых (заполненных) контейнеров противник имеет возможность построить некоторую усредненную математическую модель выбранного множества контейнеров. В том случае, если перехваченный им впоследствии контейнер не будет соответствовать имеющейся модели, он может быть легко выделен и отмечен как подозрительный, возможно содержащий скрытое сообщение. Достоверность результатов данной атаки напрямую зависит от качества используемых математических моделей и полноты выборки исходного эталонного множества.

Противодействие передаче информации – данный класс атак нацелен на уничтожение скрытой информации, путем снижения качества канала (отношения сигнал/шум) и его пропускной способности. Возможно также применение специальных алгоритмов сжатия с потерей качества и искажения малозначащей составляющей информационного потока, которая могла бы быть использована для целей скрытой передачи информации.

Системы цифровых водяных знаков во многом схожи с системами скрытой передачи информации на базе стеганографических методов и алгоритмов. По этой причине атаки на системы ЦВЗ во многом аналогичны атакам на стеганографические системы. Основная разница заключается лишь в приоритете атак. Среди атак на стеганографические системы наиболее серьезной считается выявление факта скрытой передачи информации. Для систем ЦВЗ аналогичной по значимости атакой является *уничтожение*

данных ЦВЗ с минимальным искажением контейнера. В классификации атак на системы ЦВЗ [105, 138] выделяют также *атаки против декодера ЦВЗ, атаки против протокола и атаки против встроенного сообщения.* Подробно все атаки против систем ЦВЗ рассмотрены в [105].

1.3 Анализ современных направлений в стеганографии

1.3.1 Стеганографические методы в текстовых файлах

Появление вычислительной техники позволило значительно упростить обработку и хранение текстовой информации. Малые требования к вычислительным ресурсам и объемам памяти, предназначенной для оперативного и длительного хранения информации способствовали все более широкому использованию ЭВМ в обработке, хранении и оперативном предоставлении необходимой текстовой информации. Текстовые файлы, являются наиболее простыми файлами ЭВМ, так как помимо самих данных не содержат никаких дополнительных специальных полей и заголовков. Форматирование текста при отображении или выводе на печать осуществляется благодаря использованию специальных (неалфавитных) символов перевода каретки, переноса, табуляции и символа конца файла.

Как было показано в начале данной главы, текст может быть эффективно использован для целей стеганографии. Помимо приведенного метода нулевого шифра, в лингвистической стеганографии выделяют следующие наиболее популярные направления: использование особенностей символов, кодирование смещением строк, кодирование смещением слов, синтаксические и семантические методы.

Использование особенностей символов: В рукописном тексте написание отдельных символов может заметно варьироваться, помимо явных различий в начертании символов, может отличаться высота букв, их ширина, высота средней линии, угол наклона и т.д. Все это может эффективно использоваться для передачи скрытых посланий. Основная сложность методов основанных на использовании особенностей символов заключается

лишь в формировании правил, как отличить букву открытого текста от аналогичной буквы скрытого сообщения. В простейшем случае, возле отдельных букв могут встречаться «случайные» точки или едва заметные подчеркивания. Так как символы текста в электронном виде идентичны, для целей цифровой стеганографии данный подход малоприменим. Но отдельные публикации все же встречаются [6, 62, 33, 76, 88].

Кодирование смещением строк: В основу методов положено изменение интервала между строками сообщения. Каждая строка маскирующего текста сдвигается немного вверх или вниз относительно своего исходного положения (базовой линии), соответственно смещением строки вверх можно закодировать, например, единицу, а вниз ноль очередного двоичного символа скрываемого сообщения [9, 58, 55]. Так же может использоваться и сам межстрочный интервал. Метод достаточно часто применяется для целей скрытой маркировки твердых копий электронных документов при печати на сетевых принтерах [61].

Кодирование смещением слов и символов: Методы, заключающиеся в изменении горизонтального интервала между отдельными словами или символами [17, 37, 55, 96]. Кодирование с использованием данного подхода наиболее эффективно при выборе в качестве маскирующего сообщения больших текстов с выравниванием по ширине, так как в данном случае расстояние между словами может меняться в достаточно широких пределах. В ряде случаев применяется кодирование не только длиной пробельных символов, но и их числом [4]. Так два пробела в интервале между предложениями могут кодировать очередной двоичный символ скрытого сообщения со значением равным единице, а один со значением нуля. Аналогично могут быть использованы пробельные символы в конце строки.

К недостаткам представленных методов следует отнести высокую вероятность разрушения скрытого сообщения при повторном наборе текста или использовании более сложных текстовых редакторов способных осуществлять ряд автоматических операций над текстом. Такие операции как

форматирование, замена символов табуляции пробелами, удаление лишних пробелов в конце строк и т.д., приведут к порче или же полному уничтожению скрытого сообщения. Значительно большей стойкостью к подобным искажениям обладают методы, оперирующие непосредственно самим текстом, отдельными его предложениями и словами [5].

Синтаксические методы: Методы, использующие особенности пунктуации, аббревиатуры и сокращения [63, 91]. Хотя правила пунктуации достаточно строго оговорены правилами используемого языка, существуют случаи, когда эти правила оказываются неоднозначными или же отклонение от них не ведет к существенному искажению смысла скрывающего текста. В качестве примера можно привести два следующих предложения:

«Используют красный, синий, зеленый цвета».

«Используют красный, синий и зеленый цвета».

Одна из запятых в исходном предложении была заменена союзом «и». В результате мы получили два варианта одного предложения. Первому варианту можно поставить в соответствие «0», а второму «1» двоичного символа передаваемого сообщения. Совершенно аналогичным образом могут использоваться сокращения и аббревиатуры. К синтаксическим методам относят также методы, основанные на изменении стиля и структуры предложения без заметного искажения исходной смысловой нагрузки [66].

Семантические методы, пожалуй, наиболее интересное направление в лингвистической стеганографии. Оно отличается высокой эффективностью, обусловленной применением различных методов манипулирования не второстепенными элементами и незначительными особенностями текстов, а непосредственно самими предложениями и словами. Ряд методов относящихся к данному направлению основан на использовании синонимов [7, 8, 10, 91, 97]. Практически в любом достаточно длинном предложении встречаются слова, которые без потери смысла могут быть заменены синонимами. Если для некоторого слова существует набор более чем из одного синонима, то возможно формирование специальных таблиц замен. В

таких таблицах каждому синониму может быть поставлено в соответствие некоторое кодовое слово, состоящее более чем из одного двоичного символа. Однако необходимо отметить, что в ряде случаев использование методов осложнено определенными нюансами и оттенками ключевых слов в предложениях, что несколько ограничивает их применение.

1.3.2 Стеганографические методы в мультимедиа

На сегодня большая часть мировых вычислительных ресурсов отведена под обработку и хранение мультимедиа информации. Технология мультимедиа предполагает возможность предоставления пользователю информации не только в текстовом виде, но также и в виде графических или фото изображений, анимаций, аудио и видео потоков. В последние годы визуальное представление информации становится все более популярным. Уже сейчас, на смену традиционным книгам и учебникам приходят аудиокниги, интерактивные электронные учебные пособия и курсы.

Большое количество различных стандартов и способов представления мультимедиа информации, её явная избыточность и неоднозначность, могут быть использованы для целей скрытой передачи информации. На этом фоне не вызывает удивления тот факт, что подавляющее большинство существующих алгоритмом и программных средств в области цифровой и компьютерной стеганографии направлено именно на работу с данными мультимедиа. Среди наиболее популярных направлений можно выделить:

а) *Стеганографию в звуке*. В подавляющем большинстве случаев упор делается на использование шумовой составляющей, которая присутствует практически в любой естественной аудиозаписи. Человеческий слух неспособен различить слабый фоновый шум, и поэтому он может быть эффективно использован для целей скрытой передачи информации. Решение заключается в замене или дополнении существующего шума передаваемыми данными. К наиболее известным стеганографическим методам данного направления можно отнести: замену младших битов [11, 22, 26, 94], фазовое кодирование [49, 60, 127, 151], использование эхо сигнала [32, 38, 47, 68,

138], частотное кодирование [11, 31, 45, 46]. Весьма интересным решением [28, 67, 83] является использование для целей скрытой передачи информации формата цифрового кодирования со сжатием аудиоданных алгоритмом MP3 (MPEG-1/2/2.5 Layer 3).

б) *Стеганография в изображениях.* Данное направление является на настоящий момент безусловным фаворитом, по количеству публикаций, разработанных алгоритмов и выпущенных программных продуктов. Популярность направления обусловлена тем, что графические картинки и фотоизображения являются, чуть ли не наиболее распространенным форматом представления информации в глобальной сети Интернет. Среди всех органов чувств человека, наиболее информативным является зрение. Материал представленный набором изображений привлекает к себе больше зрительного внимания, воспринимается и запоминается проще даже самого удачного текстового описания. Графические изображения, фотографии и графики, дополняющие текстовую информацию, делают текстовые описания более наглядными и доступными. Следует отметить и простоту получения изображений с цифровых фотоаппаратов, WEB-камер и сканеров в виде уже готовых и прошедших необходимую предварительную обработку файлов.

Популярность цифровых изображений в качестве контейнеров для стеганографических систем связи помимо их широкой распространенности обусловлена довольно большой избыточностью и неоднозначностью данных, а также особенностями зрительного восприятия, позволяющими вносить весьма значительные изменения в исходные данные при отсутствии визуально заметных искажений. Для целей стеганографии с использованием цифровых изображений разработано большое количество различных методов и алгоритмов основанных на использовании особенностей форматов хранения графической информации, избыточности самих данных и применяемых алгоритмов сжатия информации [78, 138, 156].

в) *Стеганография в видео.* Данное направление развито не так сильно как стеганография в изображениях и звуке, но вместе с тем является

наиболее наглядным примером перспективных технологий скрытой передачи информации с использованием мультимедиа контейнеров. На настоящий момент, большинство методов скрывающих информацию в видеоданных ориентировано на встраивание цифровых водяных знаков [40, 34, 35, 36, 52, 54, 81, 82, 85, 87, 93], так как именно это направление в наибольшей степени подкреплено коммерческим интересом. Существуют и методы направленные именно на скрытую передачу информации [14, 64, 65, 80, 90]. В качестве примера можно привести методы организации стеганографического канала связи на базе цифровой видеоконференцсвязи [39, 79]. С развитием сетей связи и увеличением пропускной способности используемых каналов передачи данных данное направление может стать наиболее перспективным.

г) *Стеганография в текстовых документах.* Текстовые документы, несмотря на кажущуюся строгость оформления и невозможность изменения отдельных битов, в действительности являются одними из наиболее интересных контейнеров в стеганографии. Используемые сегодня для создания и редактирования текстовых документов специализированные текстовые процессоры предоставляют пользователям весьма широкие возможности по предпечатной подготовке текстовых документов. Это и отдельное форматирование символов, абзацев, страниц, разделов документа, использование различных шрифтов, всевозможных списков, вставка и отображение при предварительном просмотре непосредственно в самом документе таблиц, рисунков, различных диаграмм и графиков. Естественно, что такое многообразие скрывает под собой весьма сложную структуру файлов современных текстовых документов. Разбиение документа на блоки, нетривиальные связи между различными его фрагментами, разнообразие в рамках одного файла типов данных (текст, таблицы, графики, рисунки, анимации и др.) множество различных полей и специальных заголовков, наличие больших объемов неотображаемой служебной информации, предоставляют весьма широкие возможности для целей стеганографии. В качестве примера свободно доступного программного продукта можно

привести Merge Streams 1.0 (<http://www.ntkernel.com/w&p.php?id=23>), позволяющего встраивать документы MS Excel в документы MS Word.

Ввиду закрытости наиболее популярных форматов файлов текстовых документов и их высокой сложности, использование указанных возможностей в стеганографии, в настоящее время, ограничено. Но, вместе с тем, весьма актуальны лингвистические методы, оперирующие разметкой документов и самими текстовыми данными.

1.3.3 Стеганографические методы реального времени

Современный уровень развития вычислительной техники и телекоммуникаций позволяет уже сейчас говорить не только о перспективах, но и о реально существующей возможности создания систем скрытой передачи информации работающих в масштабе реального времени. Появление таких сервисов как передача речи в телефонном режиме по цифровым каналам связи, развитие технологий VoIP (Voice over IP) и видеоконференцсвязи на цифровых каналах, предоставляют принципиально новые потенциальные возможности для систем скрытой связи. Даже сейчас вычислительных ресурсов персонального компьютера среднего уровня достаточно для осуществления операций перекодирования аналогового аудио сигнала в цифровую форму и обратно практически в режиме реального времени и с минимальными временными задержками. Использование более производительных ЭВМ позволит обеспечить своеобразный резерв вычислительных ресурсов, который и может быть предоставлен для работы стеганографических алгоритмов.

Коренной особенностью стеганографических алгоритмов реального времени является непрерывность контейнеров. В текущий момент времени алгоритму известна лишь уже ранее обработанная часть контейнера и некоторый текущий его фрагмент в рамках заданного временного окна. Размер временного окна определяется допустимой величиной временной задержки для выбранного канала связи. При работе с потоковым

контейнером, алгоритму заранее неизвестно ни поведение контейнера, ни его состояние в последующий момент времени, ни общая длительность.

Несмотря на высокую перспективность стеганографии реального времени, лишь сравнительно небольшое количество разработчиков уделяют ей соответствующее внимание. На сегодня, предложены только отдельные алгоритмы и некоторые подходы к построению стеганографических систем реального времени. В частности рассматриваются системы на основе временной задержки передачи или порядка следования пакетов по каналам TCP/IP и использования особенностей отдельных полей заголовков сетевых пакетов [2, 48, 86]. Так же встречаются наработки по использованию каналов VoIP [84] и видеоконференцсвязи [39, 79] на основе протоколов семейства RTP [74, 3]. Работы фактически носят лабораторный характер и их результаты неприменимы на практике в условиях реальных каналов, так как авторами не затрагивается весь комплекс сопутствующих проблем.

Проектирование стеганографических систем и алгоритмов реального времени требует решения целого ряда задач связанных со следующими непременно возникающими моментами:

- ограниченность во времени и вычислительных ресурсах;
- необходимость организации пакетной передачи данных;
- сложность синхронизации;
- невозможность заранее спрогнозировать емкость контейнера;
- возможное выпадение в канале связи фрагментов контейнера;
- появление не только аддитивных, но и инъективных ошибок;
- необходимость использования протоколов гарантированной доставки;
- канал связи может быть односторонним;
- необходимость проверки целостности и сложность подтверждения доставки полностью принятого сообщения.

Указанные моменты требуют построения полноценного узкополосного скрытого канала связи включающего: аутентификацию, пакетную передачу

данных, контроль целостности, коды исправляющие ошибки, алгоритмы синхронизации и удержания несущей, протоколы двусторонней передачи данных, алгоритмы разбавления данных и закрытия канала в случае многопользовательских систем.

1.3.4 Классификация атак на системы скрытой передачи электронных документов в зависимости от используемых уязвимостей

На настоящий момент на рынке представлено более сотни различных программ и подключаемых модулей, реализующих стеганографическое скрытие электронных документов в различных цифровых контейнерах и на электронных носителях. Все представленные продукты обладают различными уязвимостями и могут быть подвержены различным атакам [104, 105, 115]. Классификация атак на стеганографические системы в зависимости от используемых уязвимостей представлена на рисунке 4.

Успех в проведении атаки нацеленной на установление факта скрытой передачи информации обусловлен наличием уязвимостей связанных с внесением алгоритмами скрытия неестественных искажений, нарушением формата хранения данных, разрушением существующих корреляционных связей и изменением статистических показателей. Именно эти уязвимости используются всеми современными методами стеганоанализа. Можно ввести следующую классификацию методов специального и универсального стеганоанализа согласно используемым уязвимостям:

1. *Методы визуального анализа* – используют уязвимость, связанную с нарушением свойства неразличимости и основаны на выявлении визуально заметных искажений контейнеров;
2. *Поиск сигнатур* – методы основаны на уязвимости, связанной с внесением специальных маркеров, характерных искажений, определенных и специфичных для конкретных стеганографических алгоритмов сигнатур;

3. *Методы анализа структуры* (соответствия формату) – основаны на уязвимости, заключающейся в неестественном изменении формата файлов, организации хранения и представления данных;
4. *Статистического анализа* – методы, включающие частотный анализ и выявление отклонений статистических показателей, используемые уязвимости связаны с нарушением тех или иных статистических свойств контейнеров;
5. *Методы корреляционного анализа* – методы основаны на неявных корреляционных связях в данных контейнера, нарушаемых в результате применения стеганографического преобразования;
6. *Поиск и анализ аномалий* – универсальные методы стеганоанализа, основанные на автоматическом поиске зачастую неизвестных изначально уязвимостей.

Атаки, нацеленные на извлечение сообщений, могут быть основаны только на полном знании используемой абонентами стеганографической системы связи. Без этой информации, атакам данного класса всегда предшествуют атаки, нацеленные на выявление стеганографических вложений. Целью стеганоанализа в данном случае является выделение заполненных контейнеров и определение возможных методов и алгоритмов сокрытия, а также установление программных средств используемых для целей скрытого электронного документооборота наиболее вероятных в данном конкретном случае. Атаки, нацеленные на извлечение скрытых сообщений, всегда используют только уязвимости отдельно взятого программного обеспечения. Эти уязвимости могут быть связаны, например, с использованием каких-то стандартных безключевых алгоритмов сокрытия, нестойких криптографических элементов и заданием слабых ключей сокрытия и шифрования.



Рисунок 4 – Классификация атак на стеганографические системы в зависимости от используемых уязвимостей

Модификация скрытых сообщений подразумевает такое изменение содержимого передаваемого сообщения, при котором принимающая сторона способна его извлечь. При модификации сообщений в системах скрытого электронного документооборота принимающая сторона должна быть убеждена в целостности принятого электронного документа. Следствием данного ограничения является то, что атаки данного класса могут быть основаны только на полном знании алгоритмов сокрытия и извлечения сообщений используемой стеганографической системы. Дополнительным

ограничением атак данного класса, является то, что при проведении атаки злоумышленник должен иметь возможность выступить в качестве посредника, т.е. должен быть способен вклиниваться в открытый канал связи между отправителем и получателем сообщений.

Простейшей атакой является замена передаваемого сообщения пустым. В большинстве случаев, атака может быть успешно проведена даже без знания используемых в системе стеганографического и криптографического ключей. Зная алгоритм встраивания сообщений, противник может повторно использовать перехваченный им контейнер для записи пустого сообщения с произвольными ключами. В данном случае уязвимость системы связана с тем, что используемые процедуры предварительной проверки наличия сообщения и определения его длины унифицированы и не зависят от сеансовых секретных ключей. В результате, при получении модифицированного контейнера, принимающая сторона решает, что в данном случае передачи сообщения не происходило. Здесь интересным моментом является то, что фактически скрытое отправителем в контейнере сообщение в процессе его модификации злоумышленником удалено не было, и осталось в принятом получателем контейнере, изменению же был подвергнут только заголовок скрытого сообщения.

Атака, нацеленная на нарушение целостности передаваемого сообщения, также может быть проведена со стороны активного противника без знания используемых в системе секретных ключей. Уязвимость связана с повторным использованием одного пространства сокрытия. Эффективное противодействие атакам данного типа в системах скрытого электронного документооборота является практически неразрешимой, так как противнику известны алгоритмы встраивания сообщений, и операция встраивания сообщений может быть проведена многократно, что может привести к полному разрушению данных исходного скрытого сообщения. Данная атака схожа с атаками противодействия передачи, с той лишь разницей, что для её

проведения используются используемые в стеганографической системе алгоритмы встраивания сообщений.

Наиболее сложной и серьезной атакой на системы скрытого электронного документооборота является подмена передаваемых сообщений [100]. Для осуществления данной атаки противнику необходима наиболее полная информация о системе связи. Возможны два типа данной атаки. В первом случае атака проводится путем повтора перехваченного ранее сообщения, во втором предполагается полная подмена скрытого сообщения специальным сформированным сообщением злоумышленника. Для осуществления успешной атаки первого типа противник должен располагать алгоритмами сокрытия и извлечения скрытых сообщений, при этом ключ шифрования ему может быть неизвестен. Проведение атаки включает два этапа. На первом этапе противник осуществляет сбор передаваемых контейнеров и извлекает из них зашифрованные сообщения. На втором этапе противник выбирает одно из извлеченных им ранее сообщений и подменяет им сообщение в текущем перехваченном им контейнере. Возможность атаки обусловлена наличием уязвимости связанной с несогласованностью криптографической и стеганографической частей системы. Атака второго типа предполагает знание атакующим не только деталей реализации стеганографической системы, но так же и секретных ключей. Зная секретные ключи и алгоритмы сокрытия, противник может сам выступать в роли отправителя сообщений, читать и подменять любые из передаваемых документов. Атаки второго типа могут являться следствием целого ряда уязвимостей как стеганографической, так и криптографической частей системы. Кроме того, необходимая для успешного проведения атаки информация может быть тем или иным образом получена от авторизованных пользователей системы.

Ввиду того, что в стеганографических системах связи информация передается скрытно, и модуляция несущего сигнала осуществляется фактически на уровне естественных шумов, стеганографические методы

наиболее подвержены атаке по противодействию передачи информации. Причиной нарушения работы системы могут быть естественные искажения и шум в канале связи. Создание систем скрытой связи на открытых каналах связи со значительными помехами, фоновым шумом, пространственными и временными искажениями сигнала требует учета всех факторов, которые могут оказывать негативное влияние как на отдельно взятое скрытое сообщение, так и на систему в целом. Игнорирование подобных факторов, в конечном итоге может привести к угрозе нарушения доступности информации.

Помимо искажений несущего сигнала вызванных естественными факторами, негативное воздействие на канал связи может быть также оказано и со стороны активного противника. Целью противника в данном случае является оказание противодействия скрытому информационному обмену. Достижение указанной цели, в зависимости от особенностей открытого и стеганографического каналов может быть обеспечено различными методами. В цифровых каналах связи, где скрытая информация передается на уровне ошибок квантования, эффективной атакой является простое добавление шума в младшие биты оцифрованного сигнала. Противоположный подход, наоборот, основан на использовании дополнительных фильтров в канале между абонентами, обеспечивающих сглаживание сигнала и тем самым фактически удаляющих шумовую составляющую сигнала вместе с данными скрытого сообщения. К фиксированным контейнерам могут также применяться атаки по преобразованию формата хранения данных и дополнительное сжатие. Так если в качестве контейнеров в канале связи используются графические изображения, атака может быть проведена путем повторного сохранения изображения в том же или другом формате, дополнительно возможно реализующем сжатие информации с потерями. Так же к цифровым изображениям применимы атаки на основе линейных искажений, как-то изменение масштаба, линейных размеров, яркости изображения, усечение границ, незначительные сдвиги и вращения.

Указанные атаки по противодействию скрытой передаче эффективны против стеганографических алгоритмов предназначенных для передачи больших объемов информации. В этом случае они приводят к полному или частичному разрушению скрытой информации. В том и другом случае скрытая информация, не может быть однозначно восстановлена адресуемой стороной. Вместе с тем, во втором случае, если оказываемое воздействие на открытый канал связи в некоторой степени известно, то возможно обеспечить противодействие данной атаке. Для этой цели в системе скрытой связи необходимо обеспечить пакетную передачу данных с использованием кодов контроля целостности и исправления ошибок. Также необходимо использование специальных протоколов гарантированной доставки и повторения пакетов. В случае одностороннего открытого канала связи, повтор пакетов необходимо реализовать в рамках одной сессии.

Дальнейшее увеличение стойкости стеганографической системы к атакам типа противодействие передачи, возможно за счет увеличения амплитуды и мощности скрываемого сигнала. Так, сейчас уже существует достаточно большое множество различных методов встраивания цифровых водяных знаков, обеспечивающих достаточную стойкость скрытой информации к описанным выше атакам. Однако данный подход ограничен двумя факторами: значительное снижение информационной емкости и увеличение вносимых искажений. Кроме того, существует еще один класс атак – внесение нелинейных искажений. В этом случае сигнал подвергается сложным нелинейным искажениям, порой даже ощутимым на слух и заметным визуально. Если подобные искажения к тому же носят случайный характер, выработать эффективные методы борьбы с ними, в условиях сохранения пусть и минимальной, но определенной информационной емкости, не представляется возможным.

Таким образом, скрытие информации на уровне шумовой составляющей методами, используемыми в известных стеганографических системах, без применения дополнительных кодов исправляющих ошибки,

является наиболее уязвимым к атакам по противодействию передачи. Наиболее рациональным решением по увеличению стойкости к атакам со стороны активного противника является использование протоколов пакетной передачи данных, с применением кодов контроля целостности и исправления ошибок. Так же более высокая стойкость может быть обеспечена за счет использования для целей скрытой передачи информации более значимой части информационного сигнала контейнера.

1.4 Современные модели стеганографических систем

1.4.1 Стеганографическая система как система связи

В работе [89] стеганографическая система представлена как система связи с дополнительной информацией. Соответствующая модель системы приведена на рисунке 5. Изначально предполагается, что и на стороне отправителя, и на стороне получателя известны используемые алгоритмы скрывающего преобразования и извлечения информации, а так же общий секретный ключ. Помимо этого, отправителю и получателю может быть известна и некоторая другая дополнительная информация об используемом канале связи. Доступность информации для каждой из сторон в приведенной модели определяется положениями переключателей А и Б.

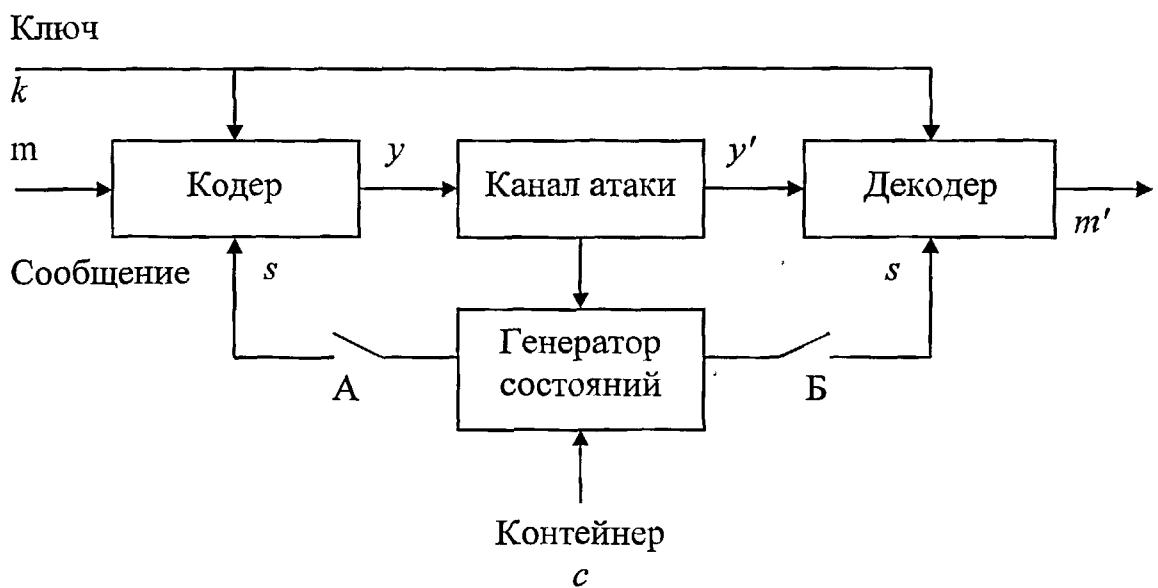


Рисунок 5 – Модель стеганографической системы как системы связи

Два переключателя введенные в модель позволяют определить четыре класса стеганографических систем в зависимости от доступности дополнительной информации о канале связи:

Класс I: Стеганографические системы, в которых дополнительная информация о канале связи не доступна (переключатели А и Б разомкнуты). В таких системах предполагается, что сокрытие информации, а также ее извлечение из заполненных (сгенерированных) контейнеров осуществляется без учета свойств используемых контейнеров и вносимых в ходе передачи по каналу связи искажений. Декодер в данном случае может оперировать только данными полученного сообщения (контейнера). Разумеется, системы, построенные по данной схеме, не способны оптимально использовать канал связи, а также противостоять искажениям и ошибкам, возникающим в процессе передачи информации.

Класс II: Информация доступна только кодеру, переключатель А замкнут, Б разомкнут. Системы данного класса предполагают использование более сложных алгоритмов, способных адаптироваться к используемым контейнерам. Теоретически, пропускная способность подобных систем близка к пропускной способности систем, в которых на стороне получателя априори известен используемый контейнер [18, 20]. К недостаткам систем данного класса следует отнести довольно высокую сложность построения эффективных алгоритмов и невозможность учета искажений вносимых каналом на стороне получателя, для обеспечения возможности адаптивного извлечения информации в зависимости от состояния канала связи. Вместе с тем, на стороне отправителя может учитываться самый худший случай искажений в канале связи, что позволит адаптировать алгоритмы, отвечающие за скрывающее преобразование, и с достаточной степенью надежности передать скрытое сообщение. Обратной стороной такой подхода является неизбежное снижение пропускной способности скрытого канала.

Класс III: Дополнительная информация доступна только декодеру, переключатель А разомкнут, Б замкнут. В данном случае декодеру может быть известен исходный контейнер и характер искажений вносимых каналом. Для эффективного извлечения скрытого сообщения на принимающей стороне могут использоваться адаптивные алгоритмы способные осуществлять сложную одно или многопроходную обработку. Наибольшее распространение такие схемы получили в области цифровых водяных знаков, где при проведении анализа изначально известно исходное и скрытое сообщение, а задача заключается лишь в установлении и возможности предоставить доказательства факта наличия заданного секретного сообщения в данном контейнере.

Класс IV: Стеганографические системы, в которых информация о канале и характере искажений известна как на стороне отправителя, так и на стороне получателя (переключатели А и Б замкнуты). Класс наиболее сложных систем, которые в тоже время обеспечивают значительный выигрыш по эффективности использования канала. Кроме того, возможность адаптировать алгоритмы встраивания и извлечения скрытых сообщений к текущим параметрам канала позволяет достаточно надежно противостоять определенным разрушающим воздействиям в канале передачи данных, направленным на уничтожение скрытой информации. Особо следует отметить, что разрушающее воздействие может быть оказано не преднамеренно или специально с целью противодействия скрытым каналам, а быть следствием применения стандартных алгоритмов перекодирования и сжатия сообщений в процессе передачи. Искажения в некоторых случаях могут носить нелинейный характер, что является критичным для систем трех других классов. В целом, стеганографические системы данного класса являются весьма перспективными ввиду обеспечения наиболее оптимального использования канала передачи информации при одновременном учете возможных искажений.

1.4.2 Математическая модель стеганографической системы

В истории стеганографии можно встретить множество примеров простых стеганографических систем, не предполагающих использования какой либо дополнительной «секретной» информации в процессах сокрытия и последующего извлечения сообщений, кроме самих передаваемых сообщений [142, 125, 25, 72]. Подобные системы относят к системам *чистой стеганографии*, так как секретность в данном случае определяется лишь секретностью методов сокрытия информации и их стойкостью к случайному обнаружению или целенаправленному анализу. Процесс сокрытия информации в данном случае представляет собой простое отображение $h: C \times M \rightarrow C$, где M – множество образованное секретными сообщениями, C – множество контейнеров. Соответственно процесс извлечения сообщений представляет собой обратное отображение $r: C \rightarrow M$. Для подобных систем обязательным является выполнение условия $|C| \geq |M|$. На практике весьма часто для каждого из передаваемых сообщений в рамках одного стеганографического канала используются различные методы сокрытия и извлечения информации. Соответственно в общем случае вместо пары отображений h и r , оперируют конечными множествами H и R .

Чистой (или бесключевой) стеганографической системой называется четверка множеств $\Sigma_S = \langle C, M, H, R \rangle$, состоящая из множества всех возможных контейнеров C , множества сообщений M , такого что выполняется условие $|C| \geq |M|$, множества скрывающих преобразований $H: C \times M \rightarrow C$ и множества правил извлечения сообщений $R: C \rightarrow M$, в котором для каждого $h \in H$, существует $r \in R$, такое что для всех $m \in M, c \in C$ выполняется равенство: $r(h(c, m)) = m$.

В описанном случае классических чистых стеганографических систем предполагается выполнение условия: заполненный контейнер неотличим и также принадлежит множеству возможных исходных контейнеров. Так как множество заполненных контейнеров не всегда является полностью

вложенным во множество пустых контейнеров, то более верным было бы разделить исходное множество контейнеров C на два подмножества: подмножество пустых контейнеров, обозначим его также как и исходное через C , и подмножество заполненных контейнеров – Q . Имеющаяся модель стеганографической системы в данном случае может быть представлена пятеркой элементов $\Sigma_S = \langle C, M, Q, H, R \rangle$.

В восьмидесятых годах XIX века голландским криптографом Огюстом Керкгоффсом был сформулирован ряд из шести требований к системам шифрования. Несмотря на то, что некоторые из них уже устарели, второе требование, известное как «принцип Керкгоффса», многие считают основным требованием применимым ко всем самым современным системам секретной связи. Согласно принципу Керкгоффса [141]: *компрометация системы не должна причинять неудобств пользователем*. Суть правила заключается в том, что противнику, при проведении им атаки на какую-либо систему секретной связи, могут быть известны все детали ее реализации. К примеру, рассмотренная выше чистая стеганографическая система не может считаться надежной, так как раскрытие алгоритмов сокрытия и извлечения информации приводит к полной компрометации системы с раскрытием канала связи и всех передаваемых сообщений. Защиту передаваемой информации в случае раскрытия канала связи может обеспечить использование дополнительной информации – секретного ключа, остающегося в секрете от третьей стороны и отвечающего за возможность перестройки алгоритмов сокрытия и извлечения информации.

В общем случае любая стеганографическая система может быть представлена в виде совокупности объектов $\Sigma_S = \langle C, M, K, Q, H, R \rangle$, где C – множество пустых контейнеров, Q – множество модифицированных контейнеров, M – множество скрываемых сообщений. Множества H , R и K определяют саму стеганографическую часть конечной системы. Множество $H = \{h_k, k \in K\}$ представляет собой множество скрывающих преобразований, обеспечивающих сокрытие сообщений $m \in M$ в контейнерах $c \in C$ в

соответствии с элементами множества ключей $K = \{0, 1, \dots, n-1\}$. Соответственно извлечение сообщений из контейнеров, принадлежащих множеству Q , осуществляется согласно правилам множества $R = \{r_k, k \in K\}$, которые для обеспечения однозначности процедуры извлечения информации поставлены в строгое соответствие элементам множества H по ключам $k \in K$.

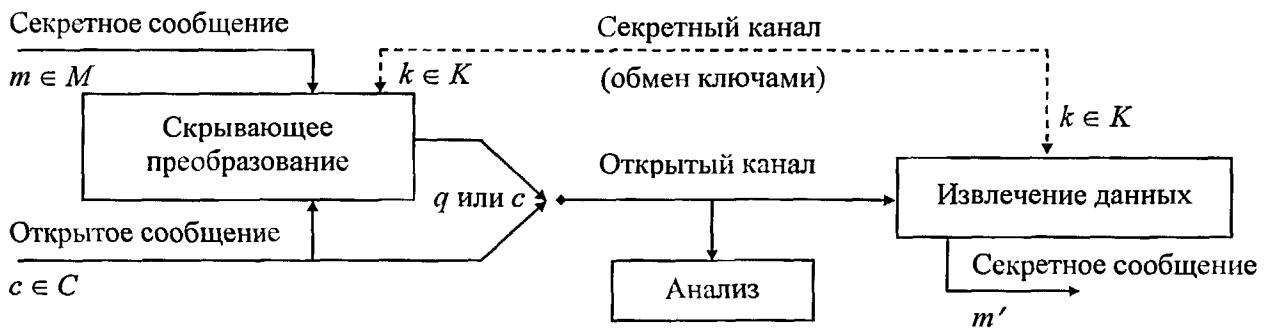


Рисунок 6 – Модель стеганографической системы с секретным ключом в присутствии пассивного противника

Стеганографической системой с секретным ключом называется совокупность $\Sigma_S = \langle C, M, K, Q, H_K, R_K \rangle$, где C и Q представляют собой множества пустых и модифицированных контейнеров, M – множество сообщений, такое что выполняется условие $|C| \geq |Q| \geq |M|$, K – множество ключей, в случае конечного множества $K = \{0, 1, \dots, n-1\}$, $H_K: C \times M \times K \rightarrow C$ и $R_K: Q \times K \rightarrow M$ множества преобразований, таких что для всех $m \in M, c \in C, k \in K$ выполняется равенство: $R_K(H_K(c, m, k)) = m$.

Для того чтобы передать секретное сообщение $m \in M$ по имеющемуся открытому каналу связи, отправитель выбирает некоторое открытое сообщение $c \in C$ и с помощью скрывающего преобразования $h_k \in H_K$ встраивает в него данные передаваемого секретного сообщения m . Сформированное в результате преобразования исходного открытого сообщения, сообщение $q \in Q | q = c \oplus m$, передается по открытому каналу

адресату. В зависимости от ситуации, а, также учитывая то, что открытый канал контролируется противником, который стремиться обнаружить секретное сообщение, по открытому каналу передаются либо обычные сообщения, либо сообщения с внедренными в них данными. В общем случае предполагается, что между отправителем и получателем помимо открытого канала, контролируемого противником, существует некоторый дополнительный секретный защищенный канал передачи информации, посредствам которого отправитель и получатель могут договориться об используемом методе скрытия информации, протоколе передачи стегосообщений и обменяться ключами. Наличие дополнительного канала не противоречит принципам скрытой связи, так как может представлять собой просто встречу двух сторон, задолго до того как появиться необходимость в организации скрытого канала. В частности в «проблеме заключенных» Алиса и Боб могли договориться о способе передачи секретных посланий до того как оказались в тюрьме под контролем строго охранника.

Отметим также, что в модели канала предложенной Кэчином [15] с целью обеспечения некоторой степени неопределенности (случайности) в выборе контейнера и его параметрах, как элемент модели на стороне отправителя дополнительно введен генератор случайных чисел. Однако в данном случае мы уходим в сторону более частных случаев реализации стеганографических систем. Поэтому генератор случайных чисел не включен в обобщенную модель стеганографической системы, представленную выше.

В реальных системах, сложность перестройки алгоритмов внедрения сообщений в зависимости от ключа, привела к тому, что в большинстве существующих решений ключи встраивания не используются. В тех же случаях, когда в программном продукте предлагается использование дополнительных ключей, в действительности оказывается, что эти ключи не являются ключами встраивания, а представляют собой ключи используемых в неявном виде шифров. Зачастую ими оказываются тривиальные шифры перестановки и разбавления. В обоих случаях стеганографическая система

представима в виде $\Sigma_S = \langle C, M, Q, h, r \rangle$, где h и r – независящие от ключа, бесключевые правила внедрения и извлечения информации.

1.4.3 Модели «криптография + стеганография»

Разработка стеганографических систем с секретным ключом на практике является весьма сложной задачей, так как требует построения сложных зависящих от ключа и перестраиваемых по нему скрывающих преобразований. Даже просто задача поиска возможных свободных ресурсов, которые могли бы быть использованы для сокрытия информации в существующих контейнерах, зачастую является нетривиальной. Ввиду этого, разработчиками программных средств, больше предпочтения отдается стеганографическим системам, построенным на основе явного или неявного использования элементов криптографии. В данном случае, о криптографии говориться как о дополнительном рубеже защиты, хотя упускается тот факт, что для стеганографических систем связи основной угрозой является уже само обнаружение факта скрытой передачи информации, а не ее извлечение и декодирование. Гибридные системы, построенные по принципу, когда криптографическая часть может рассматриваться отдельно от стеганографии, будем называть системами типа «криптография + стеганография». Модели подобных систем, с явным и неявным использованием криптографии представлены на рисунке 7.

В некоторых случаях встречаются модели, представляющие собой смесь двух представленных на рисунке 7 моделей. В них передаваемое сообщение сначала подвергается процедуре зашифрования на заданном пользователем ключе шифрования. Далее от этого же ключа или некоторой функции от сообщения вычисляется второй ключ, который используется неявным алгоритмом шифрования при встраивании сообщения в контейнер. Чаще всего, в качестве своего рода неявного шифра используются наиболее простые типы шифров, такие как псевдослучайная перестановка и разбавление. Источником случайности для них зачастую является стандартный машинный генератор псевдослучайных чисел.

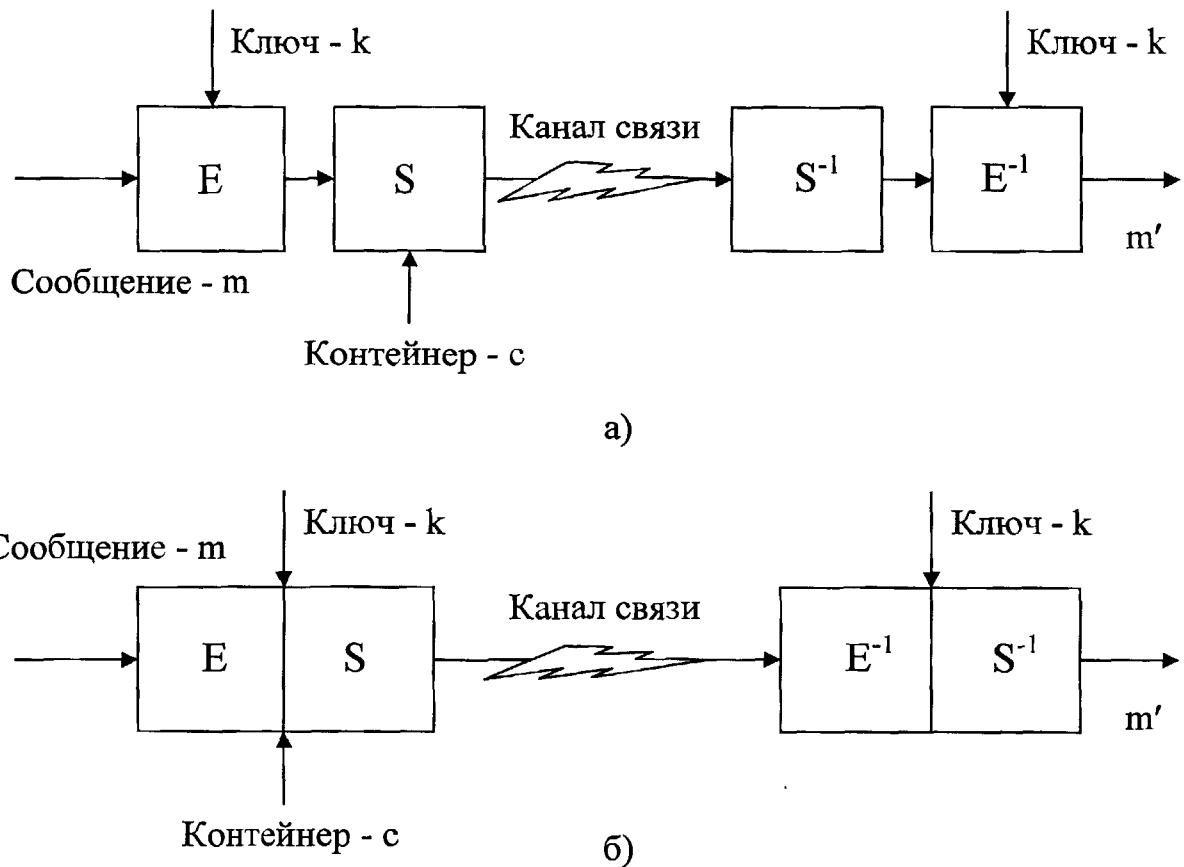


Рисунок 7 – Модели систем типа «криптография + стеганография»

К основным недостаткам, общим для всех описанных в данном пункте моделей стеганографических систем, следует отнести следующие:

- фактически используются безключевые алгоритмы сокрытия;
- стойкость системы к раскрытию факта скрытой передачи сообщений в значительной мере зависит от особенностей контейнеров и скрываемой информации;
- секретность системы в большей части зависит от секретности применяемых алгоритмов сокрытия;
- модель не учитывает особенностей контейнеров и скрываемой информации;
- сокрытие информации увеличивает энтропию скрывающего сообщения;

- криптографические алгоритмы только увеличивают стойкость к раскрытию, модификации и подмене содержания передаваемых секретных сообщений, а не стойкость системы к стеганоанализу.

Таким образом, применение криптографии здесь хоть и оправдано, но оказывается влияние только на степень защиты передаваемой информации от просмотра и модификации, а не от обнаружения факта её передачи. Стеганография в данном случае находится на второй роли, скрывая передаваемую информацию лишь от глаз простого обывателя. О стойкости подобных систем к сколь либо серьезным атакам со стороны даже только пассивного противника, говорить, к сожалению, не приходиться.

1.5 Выводы по главе

1. По результатам анализа существующей нормативно правовой базы обобщено понятие электронного документа, которое будет использовано далее в данной работе. Определено, что любой электронный документ на программном уровне представим в виде конечной битовой строки, к которой могут быть применены стандартные алгоритмы компрессии, помехоустойчивого кодирования и шифрования. Существующие системы электронного документооборота рассмотрены на предмет реализованных в них механизмов защиты.

2. Проведен анализ основных направлений в технических средствах защиты информации, которые используются или могут быть использованы в целях защиты информации в системах электронного документооборота. Выделено три базовых направления: криптографические методы защиты информации, стеганографические методы защиты информации и технологии цифровых водяных знаков. Последние два направления относятся к технологиям скрытой связи и выделены особо как наиболее молодые и малоисследованные направления. Отмечено, что данные направления обладают большими перспективами по применению в системах электронного документооборота.

3. Проведен общий анализ стеганографических методов и алгоритмов, используемых на настоящий момент в программных средствах скрытой передачи электронных документов. Исследованы специфичные уязвимости программных средств и систем скрытой передачи электронных документов. Введена общая классификация атак на системы скрытой передачи электронных документов в зависимости от используемых уязвимостей.

4. Исследованы современные модели стеганографических систем, представленные в публикациях, и построенные по результатам анализа существующего программного обеспечения.

5. По результатам проведенных исследований известных методов, моделей систем и программных средств, которые могут быть использованы для целей скрытой передачи информации, определены следующие общие недостатки существующих решений:

- отсутствие общих единых подходов и проработанных базовых решений к построению стеганографических систем;
- низкая стойкость к различным методам стеганоанализа;
- низкая стойкость к разрушающим воздействиям;
- низкая помехозащищенность;
- сильная зависимость степени скрытности от особенностей контейнера;
- отсутствие методов оценки уровня надежности и доказательства стойкости к атакам пассивного противника;
- слабость применяемых алгоритмов преобразования сообщений;
- сложность перестройки стеганографических алгоритмов в зависимости от используемого ключа сокрытия;
- сложность построения надежных систем с симметричными и открытыми ключами;
- общая надежность систем сильно зависит от объемов скрываемой информации.

Кроме того, коммерческое использование программных средств ЗИ накладывает определенные дополнительные ограничения. В первую очередь ограничения касаются возможностей широкого распространения ПС, что в свою очередь приводит к требованию обеспечения высокой стойкости даже в случае открытости исходного кода. Таким образом, первостепенными становятся вопросы обеспечения теоретической и практической стойкости. Применение методов стеганографии в коммерческих продуктах защиты ЭД требует проработки соответствующих вопросов.

6. Наиболее перспективным видится направление построения гибридных систем скрытой передачи электронных документов на основе плотного взаимодействия или даже синтеза методов криптографии и стеганографии, когда с одной стороны применяются зарекомендовавшие себя криптографические алгоритмы, а с другой отвечающие определенным требованиям и ограничениям новые стеганографические методы. При этом учитывая малую проработанность ряда теоретических и практических вопросов касающихся стеганографических методов защиты информации и эффективного противодействия методам стеганоанализа, предлагается уделить им первостепенное значение.

ГЛАВА 2. АНАЛИЗ И ОЦЕНКА ЗАЩИЩЕННОСТИ ИНФОРМАЦИИ В СИСТЕМАХ СКРЫТОГО ЭЛЕКТРОННОГО ДОКУМЕНТООБОРОТА

В данной главе рассматриваются вопросы, связанные с анализом защищенности информации и оценкой уязвимости систем электронного документооборота. Основной целью главы является анализ возможностей противодействия современным методам стеганоанализа и разработка методов практической оценки стойкости стеганографических систем связи. Для достижения указанной цели проведены исследования существующих методов стеганоанализа на предмет определения пределов чувствительности, границ применимости и возможности противодействия. Приводится понятие теоретической стойкости стеганографических систем связи к атакам пассивного противника. Исследуется возможность построения теоретически стойких стеганографических систем. Формулируется понятие, и вводятся критерии оценки практической стойкости для систем скрытой передачи электронных документов.

2.1 Оценка возможности и разработка способов противодействия методам современного стеганоанализа

2.1.1 Визуальный стеганоанализ

Визуальный стеганоанализ является самым простым, но в тоже время в ряде случаев весьма эффективным способом анализа графических и аудио файлов. Уязвимость обусловлена в первую очередь тем, что многие методы сокрытия информации не учитывают определенные особенности контейнеров и не контролируют полученные после преобразования результаты. В результате внесенные в контейнер искажения могут оказаться заметными для человека при простом просмотре или воспроизведении [12, 29]. Однако в большинстве случае, если стеганографический алгоритм реализован грамотно, обнаружить факт внедрения дополнительной информации, таким способом не удастся.

Более эффективным методом анализа растровых изображений является метод визуального анализа битовых срезов [43, 92]. Идея метода заключается в сравнении исходного изображения с изображениями, построенными по его битовым срезам, то есть с изображениями его отдельных битовых слоев. Пример, иллюстрирующий визуальный анализ битовых срезов, представлен на рисунке 8.



Рисунок 8 – Пример визуального анализа битовых срезов.

Для встраивания данных в исходное изображение был использован метод замы младших значащих битов [4, 1, 26, 143]. Метод применяется в таких известных программных продуктах как Steganos, S-Tools, Steghide, Contraband Hell Edition, Wb Stego, Encrypt Pic, StegoDos, Wnstorm, Invisible Secrets Pro и многих других. Слева показано исходное изображение, на среднем рисунке его битовый срез без дополнительной информации и на правом рисунке – срез изображения полученного после сокрытия информации методом замены младших значащих битов.

Существенным недостатком метода визуального анализа является невозможность обработки большого количества контейнеров. Так как метод требует непосредственного участия оператора, он не может быть автоматизирован. Ему также свойственны ограничения по обнаруживаемым стеганографическим методам [103, 106]. Так если метод достаточно точно учитывает структуру контейнера, равномерно с минимальными искажениями

распределяет информацию по всему контейнеру и не вносит характерных, визуально заметных искажений он не может быть обнаружен.

2.1.2 Поиск сигнатур

Сравнительно большой процент стеганографического программного обеспечения может быть выявлен путем поиска и анализа определенных характерных особенностей у уже обработанных контейнеров – сигнатур. Сколь странным бы это не казалось, но многие стеганографические методы вносят нехарактерные, порой даже заметные визуально, искажения и неестественную упорядоченность данных, или же добавляют специальные маркеры необходимые для последующего извлечения информации. Методы поиска сигнатур основаны на выявлении таких признаков, артефактов и маркеров, которые нехарактерны для общей массы контейнеров [43].

Методы, основанные на поиске сигнатур, достаточно легко автоматизировать и они могут быть эффективно использованы при обработке большого количества контейнеров без непосредственного участия человека. Метод легко применим к форматным стеганографическим алгоритмам, а также к алгоритмам, осуществляющим сокрытие информации в графических файлах на основе палитры цветов. В качестве примера программного обеспечения реализующего сигнатурный поиск можно привести коммерческий сканер StegAlyzerSS (<http://www.sarc-wv.com/stegalyzers.aspx>) от компании Backbone Security и бесплатную программную утилиту StegSpy (<http://www.spy-hunter.com/stegspydownload.htm>).

Естественным недостатком сигнатурного подхода является невозможность выявления новых алгоритмов, информация о которых и соответствующие сигнатуры еще не были добавлены в используемую базу знаний программного обеспечения. Однако факт того, что некоторый стеганографический алгоритм и его программная реализация на некоторый момент времени не обнаруживаются сканерами сигнатур, еще не говорит о том, что с обновлением базы знаний метод останется не обнаруживаемым.

Возможность эффективного выявления стеганографических вложений методом поиска сигнатур обусловлена нарушением стеганографическим программным обеспечением естественности контейнеров. Любые особенности контейнеров, возникающие в результате применения к ним стеганографических преобразований, являются очевидной будущей сигнатурой. Именно по этой причине, все форматные стеганографические методы, уже по определению являются теоретически обнаруживаемыми методом поиска сигнатур. С другой стороны, если применяемые к заданному контейнеру стеганографические преобразования не нарушают его естественной структуры, не оставляют артефактов и не вносят неестественных искажений и, т.е. нехарактерных даже для сравнительно небольшой части изначально пустых контейнеров особенностей, ничто не может выступать в качестве сигнатуры.

Таким образом, необходимым и достаточным условием для того, что бы стеганографический метод обладал свойством необнаруживаемости по отношению к методу поиска сигнатур, является условие сохранения естественности контейнеров.

2.1.3 Статистический стеганоанализ

Методы статистического стеганоанализа являются типичными представителями методов современного стеганоанализа, которые не требуют непосредственного участия оператора при проведении анализа перехваченных контейнеров. Статистические методы нетребовательны к вычислительным ресурсам, могут быть легко автоматизированы и скорректированы под конкретные частные задачи, что делает их применение весьма популярным в современных системах поиска скрытых вложений.

Все статистические методы в той или иной степени основаны на математической модели типового пустого контейнера. Так, если в качестве контейнеров используются графические изображения, то эталонная модель строится по множеству изображений путем проведения анализа их частотных и статистических характеристик. Последующее применение аналогичных

тестов к заполненным контейнерам позволяет выявить критерии сравнения. В качестве таких критериев выбираются параметры, результаты по которым сильно отличаются от средних показателей пустых контейнеров. Далее уже проверенные критерии используются в автоматизированной системе для оценки вероятности наличия скрытого сообщения в очередном контейнере.

Одним из наиболее простых статистических тестов является анализ длин серий. Запись дополнительной информации в битовую строку (например, младшие значащие биты изображения) путем её полной или частичной замены, преобразования её фрагментов и вставки новых битов, приводит к изменению распределения частот встречаемости по длинам битовых серий одинаковых битов. Встраивание информации указанными методами ведет к разрушению длинных серий одинаковых битов и соответствующему росту числа коротких серий. Этот факт лежит в основе теста длин серий. На схожем принципе работает метод анализа числа переходов в тестируемой битовой строке [103, 138]. Аналогом является также достаточно простой тест, основанный на отношении числа переходов к общему числу битов. Так экспериментально установлено, что для графических изображений, не подвергавшихся стеганографическим преобразованиям, указанное отношение находится в пределах 0,34 – 0,39. Изменение отношения в сторону увеличения является явным признаком применения стеганографических методов.

Применительно к графическим изображениям, наиболее интересным статистическим тестом является тест на основе критерия Пирсона. Впервые он был предложен в работе [92]. В основе теста лежит тот факт, что в результате сокрытия в изображении методом замены младших значащих битов дополнительной информации происходит усреднение частот встречаемости соседних цветов. В свою очередь это отражается на общей гистограмме изображения. Пусть n_i это число пикселей цвета i в рассматриваемом изображении. Тогда, для выявления метода замены младших битов ожидаемое распределение можно определить как среднее

значение для каждой пары $y_i^* = (n_{2i} + n_{2i+1})/2$, при наблюдаемом распределении $y_i = n_{2i}$. Величина χ^2 позволяет оценить разницу между двумя распределениями $\chi^2_{k-1} = \sum_{i=1}^k \frac{(y_i - y_i^*)^2}{y_i^*}$ по $k - 1$ числу степеней свободы.

Сравнение двух распределений y_i и y_i^* осуществляют согласно следующей формуле:

$$p = 1 - \frac{1}{2^{\frac{k-1}{2}} \Gamma\left(\frac{k-1}{2}\right)} \int_0^{\chi^2_{k-1}} e^{-\frac{x}{2}} \cdot x^{\frac{k-1}{2}-1} dx,$$

где Γ есть гамма-функция Эйлера. Соответственно, чем больше значение p , тем выше вероятность встраивания скрытой информации.

Следует отметить, что основными преимуществами всех статистических методов стеганоанализа является простота и высокая скорость обработки контейнеров. Недостатком, скорее являющимся естественной особенностью тестов, является вероятностное решение. Т.е. результатом работы тестов является некоторое значение вероятности наличия скрытого сообщения в анализируемом контейнере, а не однозначный ответ о его наличии или отсутствии. Следствием чего, являются присущие данным методам анализа вероятности ошибок первого и второго рода. Методы статистического анализа являются узкоспециализированными и фактически каждый из них направлен на выявление определенного алгоритма сокрытия. Результат анализа порой весьма сильно зависит даже от конкретной реализации алгоритма и ее особенностей. Что отмечается и самими авторами методов [71].

Методы противодействия статистическому анализу, могут быть достаточно просты, учитывая, что большинство методов статистического анализа усредняют значения анализируемых параметров и как следствие не чувствительны к искажениям контейнеров, носящим локальный характер. Для обеспечения эффективного противодействия статистическим методам стеганоанализа достаточно обеспечить сохранение среднего значения

конкретного анализируемого параметра на уровне значения того же параметра для исходного пустого контейнера. Так, например, если анализируемым параметром является частота встречаемости длинных последовательностей одинаковых битов и их длина, то в случайном потоке скрываемых данных достаточно периодически осуществлять вставку битовых последовательностей определенной длины.

В действительности наблюдается большое разнообразие методов статистического анализа и соответственно используемых ими параметров. Ввиду этого наиболее эффективным будет подход, направленный на сохранение основных статистических свойств контейнеров, при встраивании дополнительной информации. Наиболее стойкими следует считать методы сокрытия, обеспечивающие статистическую неразличимость скрываемой битовой строки и замещаемых в процессе записи данных контейнера [107].

2.1.4 Корреляционный анализ

Описанные выше методы стеганоанализа используют относительно небольшую часть информации контейнера и не учитывают существующие второстепенные связи между отдельными элементами контейнера. Методы стеганоанализа учитывающие не только статистические отклонения отдельных показателей, но также и существующие неявные связи между отдельными элементами контейнера принято называть методами корреляционного анализа. Одним из первых методов корреляционного анализа был метод анализа двойственных статистик RS-анализ [27, 59, 127]. Метод применим к монохроматическим фотoreалистичным изображениям при выявлении методов, осуществляющих сокрытие информации в пространственной области. В основе метода лежат неявные связи между цветами соседних точек изображения, которые в естественных изображениях дают достаточно большую корреляцию. Практическое применение метода RS-анализа показало наибольшую эффективность при выявлении вложений, скрытых методом замены младших значащих битов. При этом лучший результат достигается в случае сокрытия информации с использованием

случайного выбора точек по всему полю изображения. Очевидным является возможность применения метода при выявлении методов сокрытия информации в изображениях формата JPEG, а также аудиофайлах.

Метод RS-анализа не является единственным представителем класса методов корреляционного анализа. Близким аналогом метода RS-анализа является метод стеганоанализа на основе скрытых марковских моделей, предложенный в работе [153]. Метод анализа, применим к различным типам контейнеров, данные в которых могут быть описаны как выход некоторого марковского источника. На схожих принципах основаны методы LR-куб анализа [56, 57] и метод обнаружения скрытых сообщений на основе статистических моделей высших порядков [152].

Несмотря на высокую эффективность описанных методов, все они в большей степени ориентированы на определенные методы сокрытия информации и, по сути, также являются методами специального стеганоанализа. Выбор иного пространства сокрытия и использование методов более точно учитывающих сложную структуру контейнеров позволяют противостоять наиболее современным методам корреляционного анализа. Кроме того, как и статистические методы, описанные методы корреляционного стеганоанализа являются вероятностными. Сокрытие небольших объемов информации порядка 0,5% от возможного объема скрываемых данных с равномерным распределением по всему пространству сокрытия если и будет обнаружено, то с крайне малой степенью достоверности. На практике вероятность ложного срабатывания сильно зависит от природы контейнера. В ряде случаев она может достигать нескольких процентов.

Эффективное противодействие методам корреляционного анализа может быть основано на сохранении существующих скрытых и неявных связей и общей естественной структуры контейнеров. Альтернативным подходом можно так же считать снижение процента скрываемой информации за счет использования для передачи одного сообщения

множества контейнеров или же одного контейнера большей емкости. В обоих случаях объем скрываемой в отдельно взятом контейнере информации должен исчисляться долями процента, что обеспечит уверенное прохождение тестов со стандартным порогом принятия решения. Попытка увеличить чувствительность методов путем снижения порога принятия решения в целях выявления скрытых указанным образом сообщений приведет к значительному росту числа ложных срабатываний и как следствие к малой достоверности результатов анализа.

2.1.5 Методы универсального стеганоанализа

Рассмотренные выше методы относят к специализированным методам стеганоанализа, т.е. они являются ориентированными на конкретные скрывающие преобразования, и как следствие не могут быть эффективно использованы при выявлении новых методов сокрытия. К методам универсального стеганоанализа относят целую группу методов, особенностью которых, прежде всего, является их универсальность. Методы универсального стеганоанализа ориентированы на поиск аномалий и неестественности контейнеров, причиной которых могут быть примененные к контейнерам стеганографические преобразования.

В качестве теоретического примера метода универсального стеганоанализа приведем оценку энтропии источника данных контейнера. Энтропия позволяет оценить информативность (количество информации на символ) источника данных. Следуя Шеннону [158], будем полагать, что информация равна 1, в случае если произошло одно из двух, равновероятных, независимых событий. Энтропия H случайной величины X с распределением вероятностей P_X над алфавитом A определяется как

$$H\{X\} = - \sum_{X \in A^n} p_X(X) \log p_X(X).$$

Определим, чему равна энтропия стего S полученного в результате внесения дополнительной информации E в контейнер C :

$$H(S) = H(C) + H(E).$$

То есть при внесении дополнительной информации в контейнер его энтропия увеличивается. В случае если объем внедряемой информации достаточно большой, энтропия заполненного контейнера будет стремиться к единице. Получив оценки энтропии для пустых и заполненных контейнеров легко определить порог принятия решения, так как разница для большинства существующих стеганографических методов оказывается весьма большой. Очевидной причиной увеличения энтропии контейнера при записи дополнительной информации является фактическое дополнение имеющейся изначально в контейнере информации новой.

На практике применение нашли методы универсального анализа основанные на различных методиках и алгоритмах оценки качества и естественности контейнеров. Так к изображениям применяются стандартные методы оценки качества изображений, оцениваются параметры шумов и отношение сигнал/шум. Более сложные методы основаны на наиболее современных достижениях в цифровой обработке сигналов и теории принятия решений. Так в работе [150] для оценки качества изображений и определения степени их естественности используются методы вейвлет-декомпозиции.

К недостаткам методов универсального стеганоанализа следует отнести тот факт, что даже для самых современных из них характерны весьма значительные величины ошибок первого и второго рода. В ряде случаев, по экспериментам самих авторов методов, вероятность ложного срабатывания нередко достигает 10-20 %. Что в условиях необходимости проверки большого числа контейнеров является недопустимым. Для последующего анализа системой будет отбираться чуть ли не каждый пятый контейнер. Тем не менее, методы универсального анализа интересны, прежде всего, тем, что в отличие от методов сигнатурного поиска и статистического стеганоанализа, позволяют обнаруживать новые или еще неизвестные системе методы сокрытия информации.

В случае если авторам методов универсального стеганоанализа удастся значительно снизить вероятность ложного срабатывания, противодействие этим методам может быть основано только на полном сохранении естественности контейнеров. Даже незначительные отклонения могут оказаться критическими в будущем. На настоящий же момент времени, для противодействия методам универсального стеганоанализа достаточно обеспечить естественность контейнеров на уровне необходимом для обеспечения стойкости стеганографического метода к статистическим методам стеганоанализа.

2.2 Теоретическая стойкость стеганографических систем

Современная теория криптографии во многом основана на понятии совершенно стойкой системы секретной связи. Совершенная стойкость предполагает под собой математическое доказательство отсутствия даже теоретической возможности восстановления подвергнутого обратимому криптографическому преобразованию сообщения без знания секретного ключа расшифрования. Другими словами, противник, перехвативший зашифрованное сообщение, обладая безграничными ресурсами и вычислительными возможностями, без знания секретного ключа будет не в состоянии восстановить перехваченное сообщение. В данном разделе исследуется возможность построения подобных стеганографических систем связи и предлагается теоретическая база для их практического построения.

2.2.1 Совершенная стеганографическая система

Определение совершенной стеганографической системы может быть дано на основе теоретико-информационного подхода предложенного Клодом Шенноном для криптографических систем защиты информации [158]. Количество информации содержащейся в случайной последовательности $Y \in B^n$ относительно $X \in A^n$ определяется как:

$$I\{X, Y\} = H\{X\} - H\{X|Y\}$$

$$H\{X\} = - \sum_{X \in A^n} p_X(X) \log p_X(X), H\{X|Y\} = - \sum_{\substack{X \in A^n \\ Y \in B^n}} p_{X,Y}(X,Y) \log p_{X|Y}(X|Y).$$

Согласно Шеннону, необходимым и достаточным условием совершенной стойкости является $P_M(E) = P(E)$ для всех M и E , где M – множество открытых текстов, E – множество зашифрованных сообщений. Т.е. предполагается, что $P_M(E)$ не зависит от M .

Аналогично приведенному определению, можно сформулировать необходимое условие совершенной стойкости и для стеганографических систем связи [104]. Так как вероятность появления на выходе системы того или иного контейнера должна быть $P_C(c) = 1/|C|$ и не должна зависеть от передаваемого сообщения, то для совершенной стеганографической системы должно выполняться условие: $I\{M,C\} = I\{C,M\} = 0$.

В работах [15, 16] дано определение совершенной стеганографической системы на основе относительной энтропии двух распределений P_{Q0} и P_{Q1} , определяемой как:

$$D(P_{Q0}\|P_{Q1}) = \sum_q P_{Q0}(q) \log \frac{P_{Q0}(q)}{P_{Q1}(q)}, \text{ с } 0 \log \frac{0}{0} = 0 \text{ и } p \log \frac{p}{0} = \infty \text{ если } p > 0.$$

Относительная энтропия двух распределений принимает только положительные значения и равна нулю в том и только в том случае, когда распределения совпадают. Пусть P_C и P_Q распределения над множеством пустых и заполненных контейнеров соответственно, тогда определение совершенно стойкой стеганографической системы может быть сформулировано следующим образом:

Определение. Стеганографическая система называется *совершенно стойкой* (к атакам пассивного противника) или *совершенной*, если выполняется условие $D(P_C\|P_Q) = 0$; стеганографическая система называется ε -*стойкой* (к атакам пассивного противника) если $D(P_C\|P_Q) \leq \varepsilon$.

В случае совершенной стеганографической системы, пассивный противник, обладающий сколь угодно большими вычислительными ресурсами, неспособен различить два распределения и не может получить

никакой информации о встроенном сообщении, равно как и о его наличии. Рассмотрим приведенное выше условие более подробно. Допустим, все элементы C и Q принадлежат одному множеству, обозначим его через X . Тогда выполнение условия $D(P_C \parallel P_Q) = 0$ возможно, например, в том случае, если все исходные контейнеры и контейнеры на выходе стеганографической системы равновероятны и не зависят от скрываемого сообщения. В качестве примера подобной стеганографической системы можно рассматривать систему, использующую в качестве контейнеров и скрываемых сообщений случайные последовательности. Запись информации в контейнер в данном случае предполагает просто замену одной последовательности на другую. В работе [15, 16] приведен пример подобной системы, схожий с примером совершенной криптографической системы и основанный на использовании свойств равномерной гаммы. Пусть контейнер C представляет собой равномерно распределенную n -битную последовательность, у отправителя и получателя имеется некоторый общий ключ K также представленный n -битной последовательностью с равномерным распределением. В качестве скрывающего преобразования выбрана операция сложения по модулю два XOR двух n -битных строк M и K . В этом случае скрываемое сообщение $S = M \oplus K$, также представляет собой n -битную строку с равномерным распределением, которая может быть использована в качестве контейнера результата. Таким образом, совершенная стеганографическая система теоретически существует, так как выполняется условие $D(P_C \parallel P_Q) = 0$.

Вместе с тем на практике, в чистом виде такие системы неприменимы, так как сам факт передачи сообщений с равномерным распределением уже является серьезным демаскирующим признаком.

2.2.2 Совершенная стеганографическая система на битовых строках содержащих длинные серии одинаковых битов

Совершенная стеганографическая система, предполагающая использование в качестве контейнеров случайных битовых строк с

равномерным распределением не может быть применена на практике в своем исходном виде. Причиной тому является практическое отсутствие случайных битовых последовательностей передаваемых в чистом виде по каналам связи, которые могли бы быть использованы в качестве контейнеров. Кроме того, передача случайной последовательности по открытому каналу связи, по сути, является явным демаскирующим признаком, говорящим об использовании специальных алгоритмов перекодирования информации (как например, криптографии). Также редко случайные битовые последовательности встречаются и в других естественных контейнерах.

Определение. Серией одинаковых битов называется подстрока длины l , при $1 \leq l \leq n$, всех биты которой имеют равное значение (только ноли или единица), битовой строки длины n .

В качестве примера, обращаясь к уже описанному ранее методу замены младших значащих битов можно отметить, что битовые строки нижних битовых плоскостей не отвечают равномерному случайному распределению, так как содержат длинные серии одинаковых битов. Именно благодаря длинным сериям одинаковых битов метод легко обнаруживается средствами статистического анализа. В целях же противодействия методам статистического стеганоанализа в процессе записи информации необходимо обеспечить как можно более точное сохранение статистических свойств заменяемых битовых строк.

В данной работе предлагается решение, основанное не на прямой замене битовых строк контейнеров строками скрываемых данных, а на их модуляции псевдослучайной последовательностью от скрываемых данных. В этом случае исходные битовые строки выступают в качестве своего рода несущей, которая и определяет все свойства конечной последовательности. Подход, основанный на модуляции изначально присутствующих в сложных контейнерах битовых строк на псевдослучайных скрываемых данных, позволяет сохранить исходную энтропию и большинство корреляционных

связей с другими данным контейнера, так как затрагивает только наиболее случайную составляющую модулируемых последовательностей.

Определение. Границей перехода серий одинаковых битов в битовой строке будем называть положение элемента битовой строки, при движении вправо по строке, следующего за конечной серией одинаковых битов, в случае если его значение отлично от значения битов серии.

В целях сохранения длинных битовых последовательностей, модуляция исходной битовой строки контейнера битовой строкой скрываемых данных может быть осуществлена за счет изменения только тех битов исходной строки, которые расположены вблизи границ переходов серий одинаковых битов. Так, например, в модулируемой последовательности 0000011111... изменению могут быть подвергнуты только пятый или шестой биты, и только один из них. Соответственно могут быть получены два варианта измененной последовательности:

0000111111... – изменение пятого бита;

0000001111... – изменение шестого бита.

Как можно заметить, такие изменения в битовой последовательности не приводят к разрушению длинных серий одинаковых битов и появлению новых переходов (сохраняется общее количество информации в строке), как в случае изменения других битов последовательности:

0101011011... – изменение второго, четвертого и восьмого битов;

0000101111... – изменение пятого и шестого битов.

Лемма. Существует конечный автомат Мили, способный распознавать границы переходов серий одинаковых битов в битовых строках конечной длины.

Доказательство. Рассмотрим автомат Мили, заданный пятеркой множеств $A = (X, S, Y, h, f)$, где $X = \{0, 1\}$ – входной алфавит, S – множество состояний, $Y = \{a, b, \varepsilon\}$ – множество выходных сигналов, $h : S \times X \rightarrow S$ – функции переходов, $f : S \times X \rightarrow Y$ – функции выходов. Таблица автомата:

Таблица 2 – Состояния автомата

	S_{t+1}		y_t	
	0	1	0	1
s_0	s_1	s_3	ε	ε
s_1	s_2	s_4	ε	a
s_2	s_1	s_3	ε	b
s_3	s_2	s_4	a	ε
s_4	s_1	s_3	b	ε

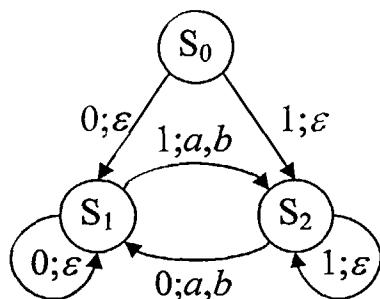
Как видно из представленной таблицы, все переходы по элементам одной серии одинаковых битов являются замкнутыми. Если серия состоит из битов со значением ноль, образуется цикл: $s_1 \rightarrow s_2 \rightarrow s_1$. Для серии битов со значением единица – цикл: $s_3 \rightarrow s_4 \rightarrow s_3$. Состояние s_0 не входит ни в один из указанных циклов и является начальным состоянием автомата. Таким образом, состояния автомата можно разделить на три множества:

$S_0: \{ s_0 \}$ – начальные состояния;

$S_1: \{ s_1, s_2 \}$ – состояния постоянства по нулю;

$S_2: \{ s_3, s_4 \}$ – состояния постоянства по единице;

Множества S_1 и S_2 являются циклическими по нулю и единице соответственно. Любой переход между элементами внутри одного множества сопровождается выводом сигнала " ε " – пустого символа. Множества состояний S_0, S_1, S_2 и соответствующие переходы между множествами можно представить в виде ориентированного графа:



Анализ возможных переходов по множеству элементов S_0 согласно представленной таблице автомата Мили, позволяет заключить, что при любом входном значении, осуществляется переход в одно из состояний множеств S_1 и S_2 , при этом образование циклов внутри S_0 и возврат к элементам множества S_0 невозможны. Любой переход по состояниям множества S_0 сопровождается выводом пустого символа " ε ". Таким образом, множество S_0 является множеством начальных состояний, и переходы по состояниям данного множества не приводят к выдаче выходных сигналов, обозначающих границы переходов серий одинаковых битов.

Рассмотрим возможные переходы между состояниями множеств S_1 и S_2 . Между указанными множествами возможны только следующие переходы:

- 1: $s_1 \rightarrow s_4;$
- 2: $s_2 \rightarrow s_3;$
- 3: $s_3 \rightarrow s_2;$
- 4: $s_4 \rightarrow s_1.$

При этом переходы 1 и 2 возможны тогда и только тогда, когда во входной последовательности автомата с текущим состоянием из S_1 встречен символ "1"; переходы 3 и 4 возможны тогда и только тогда, когда во входной последовательности встречен символ "0" и автомат находится в одном из состояний множества S_2 . Все указанные переходы 1 – 4 сопровождаются выводом одного из сигналов "a" или "b", при этом выдача пустого символа " ε " - невозможна. Переходы 1 и 2 переводят автомат в состояние из S_2 , т.е. на обработку серии битов из единиц, а переходы 3 и 4 в состояние из S_1 (обработка серии из нулей).

Еще раз отметим, что множество состояний автомата S_1 образует цикл по нулю и позволяет замкнуто (без перехода к состояниям множеств S_0 и S_2) обработать битовую серию из нулей с выдачей на выход автомата, для каждого бита серии, пустого символа " ε ". Нахождение автомата в одном из состояний множества S_1 (S_2) однозначно говорит о том, что предыдущим символом входной строки был символ "0" (символ "1"). Таким образом,

переходы 1 и 2 соответствуют встрече во входной последовательности (битовой строке поданной на вход автомата) границы перехода от серии битов из нулей к серии из единиц и сопровождают встречу данной границы сигналом "a" или "b". Аналогично случаи 3 и 4 сопровождают встречу границы перехода от серии битов из единиц к серии битов из нулей. Во всех прочих случаях на выходе автомата присутствует пустой символ " ε ".

Таким образом, автомат Мили с заданной таблицей переходов и выходов, позволяет распознавать границы переходов серий одинаковых битов.

□

Пример. Рассмотрим битовую строку:

$$r = 1110101010000111100001101\dots$$

Распишем последовательность состояний и выходов автомата:

$$\begin{array}{ccccccccccccccccccccc} r & = & 1 & 1 & 1 & 0 & 1 & 0 & 1 & 0 & 1 & 0 & 0 & 0 & 0 & 1 & 1 & 1 & 1 & 0 & 0 & 0 & \dots \\ s & & s_0 & s_3 & s_4 & s_3 & s_2 & s_3 & s_2 & s_3 & s_2 & s_1 & s_2 & s_1 & s_4 & s_3 & s_4 & s_3 & s_2 & s_1 & \dots \\ y & & \varepsilon & \varepsilon & \varepsilon & a & b & a & b & a & b & a & \varepsilon & \varepsilon & \varepsilon & a & \varepsilon & \varepsilon & \varepsilon & a & \varepsilon & \dots \end{array}$$

В представленном фрагменте битовой строки встречаются серии из 1, 3 и 4 одинаковых битов. Как можно заметить, символы "a" и "b" на выходе автомата в действительности соответствуют положениям границ переходов серий одинаковых битов.

Если из полученной на выходе последовательности удалить пустой символ ε , то полученная в результате строка $y^t = abababaaa\dots$ будет нести информацию обо всех переходах в исходной битовой строке.

Замечание 1. Если распределение битов, расположенных на границах серий одинаковых битов, носит случайный характер и отвечает равномерному распределению, то на выходе представленного в доказательстве леммы автомата Мили последовательность $y^t \in \{a, b\}^*$ будет также случайна и будет иметь тот же характер распределения.

Доказательство. В представленном выше примере видно, что в выходной последовательности автомата помимо пустого символа " ε " встречаются также символы " a " и " b ", причем не какой-то один из них, а оба. В основе доказательства замечания лежит причина данного явления. Определим, с чем связано появление символов " a " и " b " в выходной строке автомата Мили, представленного в доказательстве леммы.

Из доказательства леммы следует, что символы " a " и " b " в выходной строке автомата появляются только в случае переходов между состояниями множеств S_1 и S_2 и их позиции в выходной строке в точности совпадают с положениями границ переходов во входной битовой строке. Переходы внутри множеств состояний S_1 и S_2 образуют циклы с периодом повторения $T = 1$. Такие циклы внутри указанных множеств позволяют отслеживать состояние четности для текущего шага работы автомата при обработке фрагмента входной строки представленного серией одинаковых битов. При этом с учетом начального состояния s_0 , текущие состояния автомата s_1 и s_3 соответствуют четным позициям текущих входных символов автомата в обрабатываемой битовой строке, а состояния s_2 и s_4 нечетным позициям. Можно заметить, что переходы между состояниями 1-4 (в доказательстве леммы) переводящие автомат от обработки одной серии одинаковых битов к другой так же сохраняют информацию о четности позиции текущего входного символа автомата относительно начала строки. В результате, символ " a " в выходной строке автомата соответствует обнаружению границы перехода серий одинаковых битов на четной позиции во входной битовой строке, а символ " b " обнаружению границы на нечетной позиции. Таким образом, выходная строка автомата несет в себе информацию не только об обнаруженных границах переходов серий одинаковых битов, но и о том, является ли позиция бита во входной строке соответствующего данной границе четной или нечетности относительно начала строки.

Покажем, что положение границ переходов серий одинаковых битов (как следствие, является ли позиция бита соответствующего границе четной

или нечетной относительно начала строки) определяется битами, расположенными на этих границах. В данном случае предполагается, что битовая строка находится в состоянии формирования как результат работы некоторого источника. При этом само исходное положение границ серий одинаковых битов на момент рассмотрения неопределенно.

Рассмотрим границы двух возможных типов:

- 1: "...01...";
- 2: "...10...".

Согласно определению границы перехода серий одинаковых битов, границ других типов быть не может.

Для первого типа границы в зависимости от бита, следующего за границей перехода серий одинаковых битов, возможны два варианта "...010..." и "...011...". Пусть бит расположенный на границе является случайным, т.е. с равной вероятностью принимает значение «0» и «1». Если этот бит принимает значение «1» то граница перехода серий одинаковых битов остается на исходной позиции, относительно рассматриваемого фрагмента битовой строки. Если бит принимает значение «0» в первом случае граница переходов серий одинаковых битов будет потеряна. Данный случай мы исключаем из рассмотрения, так как в нем исчезает сама граница перехода серий одинаковых битов. Во втором же случае, граница смещается на одну позицию вправо от исходной.

Для границы второго типа все аналогично. Возможны варианты "...100..." и "...101...". Если значение граничного бита «0» положение границы сохраняется. При значении граничного бита «1», в первом случае граница смещается вправо на одну позицию, а во втором граница перехода серий одинаковых битов исчезает.

Возможное смещение границы влево не рассматривается, так как эти случаи являются проявлением случаев рассмотренных выше, когда граница остается на исходной позиции.

Таким образом, является ли положение границы переходов серий одинаковых битов четным или нечетным относительно начала битовой строки в полной мере определяются битами, расположенными на границе переходов серий одинаковых битов. Так как биты расположенные на границах переходов серий одинаковых битов случайны, то появление в выходной строке автомата Мили, представленного в доказательстве леммы, символов " a " и " b " будет иметь случайный характер.

□

Замечание 2. Если распределение битов расположенных на границах монотонных последовательностей не является случайным, то возможно последовательное применение автоматов, каждый из которых работает с выходом предыдущего, и так до получения на выходе последовательности с требуемым распределением.

Рассмотрим вопрос, является ли возможным построение реальных стойких стеганографических методов на основе модуляции битовых строк контейнера скрываемой псевдослучайной последовательностью данных. Для этого еще раз обратимся к понятию совершенно стойкой стеганографической системы, представленной в предыдущем пункте.

Теорема. Совершенная стеганографическая система, использующая в качестве контейнеров потенциально бесконечные битовые строки, содержащие длинные серии нулей и единиц, существует, если граничные элементы всех серий одинаковых битов в исходной строке случайны и их распределение является равномерным.

Доказательство. Из предыдущей леммы и первого замечания к лемме следует, что существует алгоритм способный распознавать границы переходов битовых последовательностей и строить однозначное инъективное отображение множества битовых строк $\{0, 1\}^*$ содержащих длинные серии

нулей и единиц T^* (границные элементы всех последовательностей одинаковых битов случайны) во множество случайных последовательностей $\{a, b\}^*$. Другими словами, существует отображение $\gamma: T^* \rightarrow R^*$, где правое множество представляет собой множество случайных последовательностей $R^* = \{a, b\}^*$, или над двоичным алфавитом $R^* = \{0, 1\}^*$. Данное отображение позволяет сделать переход от рассматриваемой стеганографической системы, к классической совершенной стеганографической системе, использующей в качестве контейнеров и скрываемых сообщений случайные двоичные последовательности. Представленный выше автомат Мили извлекает случайную составляющую исходных битовых строк, содержащих длинные последовательности одинаковых битов, которая, так же как и в случае классической совершенной стеганографической системы, может быть заменена другой случайной последовательностью – сообщением.

Таким образом, с учетом первого замечания к предыдущей лемме, для доказательства представленного утверждения необходимо и достаточно доказать существование операций прямого и обратного преобразований, позволяющих осуществлять запись и извлечение информации по границным элементам битовых последовательностей, поведение которых согласно условию случайно. Для доказательства существования любого конечного алгоритма, достаточно доказать существование реализующей данный алгоритм машины Тьюринга.

Рассмотрим машину Тьюринга заданную пятеркой $M = (Q, \Sigma, \delta, q_0, F)$, где Q – конечное множество состояний, q_0 – начальное состояние $q_0 \in Q$, F – множество заключительных состояний (выходов), $\Sigma = \{0, 1, x, a, b, c, d\}$, $\delta: Q \times \Sigma \rightarrow Q$ – функция переходов.

Машина Тьюринга, реализующая прямое преобразование может быть задана набором правил, представленным в таблице 3.

Таблица 3 – Прямое преобразование

Набор правил			
$q_0^* \rightarrow q_0 R$	$q_5 0 \rightarrow q_6 d L$	$q_9 x \rightarrow q_9 x R$	$q_{12} 1 \rightarrow q_{13} 0 H$
$q_0 0 \rightarrow q_0 R$	$q_6 0 \rightarrow q_6 0 L$	$q_9 0 \rightarrow q_9 0 R$	$q_{13} 0 \rightarrow q_2 0 H$
$q_0 1 \rightarrow q_0 R$	$q_6 1 \rightarrow q_6 1 L$	$q_9 1 \rightarrow q_9 1 R$	$q_9 a \rightarrow q_{14} 0 R$
$q_0 x \rightarrow q_1 R$	$q_6 x \rightarrow q_7 x L$	$q_8 a \rightarrow q_5 1 R$	$q_{14} 0 \rightarrow q_{15} 1 H$
$q_1 0 \rightarrow q_2 0 R$	$q_7 e \rightarrow q_7 e L$	$q_9 b \rightarrow q_4 1 R$	$q_{14} 1 \rightarrow q_{15} 1 H$
$q_1 1 \rightarrow q_4 1 R$	$q_7 0 \rightarrow q_8 e R$	$q_8 c \rightarrow q_3 0 R$	$q_{15} 1 \rightarrow q_5 1 H$
$q_2 0 \rightarrow q_3 0 R$	$q_7 1 \rightarrow q_9 e R$	$q_9 d \rightarrow q_2 0 R$	$q_9 c \rightarrow q_{16} 1 R$
$q_3 0 \rightarrow q_2 0 R$	$q_7^* \rightarrow q_{18} H$	$q_8 b \rightarrow q_{10} 0 R$	$q_{16} 0 \rightarrow q_{17} 0 H$
$q_2 1 \rightarrow q_6 a L$	$q_8 e \rightarrow q_8 e R$	$q_{10} 0 \rightarrow q_{11} 1 H$	$q_{16} 1 \rightarrow q_{17} 0 H$
$q_3 1 \rightarrow q_6 b L$	$q_8 x \rightarrow q_8 x R$	$q_{10} 1 \rightarrow q_{11} 1 H$	$q_{17} 0 \rightarrow q_3 0 H$
$q_4 1 \rightarrow q_5 1 R$	$q_8 0 \rightarrow q_8 0 R$	$q_{11} 1 \rightarrow q_4 1 H$	
$q_5 1 \rightarrow q_4 1 R$	$q_8 1 \rightarrow q_8 1 R$	$q_8 d \rightarrow q_{12} 1 R$	
$q_4 0 \rightarrow q_6 c L$	$q_9 e \rightarrow q_8 e R$	$q_{12} 0 \rightarrow q_{13} 0 H$	

Из анализа представленного набора правил, очевидно существование машины Тьюринга, реализующей обратное преобразование.

□

Замечание. Исходная модулируемая случайной последовательностью бесконечная битовая строка, и битовая строка, полученная на выходе описанного выше стеганографического преобразования, будут идентичны по своим статистическим свойствам.

Следствие. Если в исходном битовом потоке граничные элементы серий одинаковых битов случайны и отвечают распределению Гаусса, то возможно формирование статистически неразличимого битового потока, содержащего в себе поток скрытых случайных данных.

Под статистической неразличимостью понимается определение, введенное в работе [30], также [50]: Пусть для каждого фиксированного

слова x $A(x)$ и $B(x)$ – это случайные величины, значения которых являются двоичными словами. Семейства случайных величин $\{A(x)\}$ и $\{B(x)\}$ называются статистически неразличимыми на множестве X , если

$$\sum_{s \in \{0,1\}} |P[A(x) = s] - P[B(x) = s]| < |x|^{-c}$$

для любой константы $c > 0$ и всех достаточно длинных $x \in X$.

При построении реальных стеганографических систем, в качестве критериев при извлечении информации из битовой последовательности могут выступать длины последовательностей состоящих из одинаковых битов и положения границ (позиции элементов в начале и в конце последовательностей одинаковых битов). При этом использование в качестве критерия положений границ последовательностей одинаковых битов следует считать предпочтительным вследствие большей информационной емкости.

2.3 Методика оценки практической стойкости

2.3.1 Критический коэффициент скрытия

Проведенный выше анализ известных методов пассивного стеганоанализа показал, что большая часть методов основана на анализе определенных статистических характеристик и представляет собой различные вероятностные алгоритмы, которые не дают однозначного ответа, содержатся или нет в анализируемом сообщении скрытые данные. Зачастую конечное решение просто остается за оператором. Необходимость же проверки большого числа контейнеров неизбежно ведет к выделению множества подозрительных контейнеров, среди которых в действительности может не оказаться ни одного стегосообщения. Даже если процент невелик, скажем порядка 3-5 %, и при анализе 10-20 контейнеров им можно пренебречь, то при анализе 1000 контейнеров для последующего более детального анализа будут отобраны уже 30-50 контейнеров. В результате такой, даже незначительный, процент влечет за собой весьма существенную нагрузку на операторов, проводящих последующий анализ. Автоматизация

процесса анализа требует задания определенных, достаточно строгих критериев принятия решения, по которым, исходя из вероятностей полученных на выходе алгоритмов стеганоанализа, может быть получен однозначный ответ. Таким образом, необходимо строгое задание порогов обнаружения. Конкретные их значения определяются по принципу допустимости и минимизации вероятностей ошибок первого и второго рода.

Уточним понятие ошибок первого и второго рода. Для этого введем две гипотезы, которые могут быть отнесены к заданному контейнеру. Контейнер может быть признан содержащим или не содержащим некоторое секретное сообщение, следовательно, можно задать две гипотезы:

H_0 – "контейнер не содержит скрытого сообщения";

H_1 – "контейнер содержит скрытое сообщение".

Соответственно под вероятностью ошибки первого рода α будем понимать вероятность обнаружения скрытого сообщения в пустом контейнере, а под вероятностью ошибки второго рода β – пропуск заполненного контейнера, вероятность принятия отрицательного решения по наличию скрытого сообщения в заполненном контейнере.

Ошибки первого и второго рода имеют различное значение в автоматизированной системе анализа. Так, минимизация величины ошибки первого рода α оказывается на уменьшении загруженности системы анализа сообщений за счет снижения числа постобработок. Именно величина ошибки первого рода говорит о возможном количестве в действительности пустых контейнеров, которые будут выделены в качестве подозрительных и предоставлены системой для более детального анализа. В свою очередь, для некоторого отдельного алгоритма наиболее важной характеристикой будет величина ошибки второго рода – β , так как именно она отвечает за эффективность применяемого алгоритма анализа сообщений. Обе величины отражают за конечную надежность системы и достоверность результатов проводимого с её помощью стеганоанализа.

Рассмотрим стеганографическое отображение $\gamma: C \times M \rightarrow Q$ множества пустых контейнеров C и сообщений M на множество стегосообщений Q . Пусть для выбранного отображения существует некоторое различительное правило (метод стеганоанализа), позволяющее построить отображение $\tau: C \rightarrow \Gamma$, где $\Gamma \subseteq \mathbb{R}$. По значению результата данного отображения для некоторого контейнера $w \in C \cup Q$ величины x , требуется принять решение о справедливости одной из гипотез: $H_0: w \in C \setminus Q$ или $H_1: w \in Q$. Для этого необходимо построить разбиение $\Gamma = (\Gamma_0, \Gamma_1)$ пространства всех возможных значений величины x , для чего определить решающее правило:

$$\delta(x) = \begin{cases} 0, & \text{если } x \in \Gamma_0 \\ 1, & \text{если } x \in \Gamma_1 \end{cases}.$$

Очевидным способом разбиения пространства Γ , является введение некоторого порога принятия решения $T \in \mathbb{R}$, разбивающего пространство на два непересекающихся подпространства Γ_0 и Γ_1 . Разбиение пространства Γ заданием определенного порога принятия решения T играет решающую роль, так как именно этот шаг определяет в итоге величину ошибок первого и второго рода. Так как подпространства Γ_0 и Γ_1 связаны между собой, свести к минимуму одновременно ошибки и первого и второго рода невозможно. Расширение Γ_0 ведет к сужению Γ_1 и увеличению вероятности ошибки второго рода β , а расширение Γ_1 соответственно к увеличению α . Таким образом, сведение α и β к нулю на больших множествах C и Q , при применении методов статистического стеганоанализа фактически не представляется возможным. Для получения оптимального решения может быть применен Байесовский подход или подход Немана-Пирсона, позволяющей по заданной величине вероятности ошибки первого рода α минимизировать β – величину ошибки второго рода. Таким образом, для заданного метода стеганоанализа выбор порога принятия решения T будет однозначно определять эффективность метода и всей системы стеганоанализа. На практике в подавляющем большинстве случаев данный

параметр выбирается экспериментальным путем и, как правило, сильно зависит как от характеристик и распределения используемых в конкретном стеганографическом канале контейнеров и скрываемых сообщений.

Рассмотрим связь между количеством скрываемой в контейнере информации и вероятностью обнаружения скрытой информации статистическими методами стеганоанализа. Для этого определим коэффициент сокрытия $k = |m|/|c|$ – отношения объема скрытого сообщения m в контейнере c к общему объему контейнера. Ранее было отмечено, что при проведении статистических тестов объем скрытой в контейнере информации играет достаточно существенную роль. Так методы стеганоанализа эффективно работающие с контейнерами, заполненными на $2/3$ пространства сокрытия, при сокрытии незначительных объемов информации часто дают отрицательный результат. Причина этого явления как раз и кроется в ошибках первого и второго рода, а, следовательно, связана с заданным порогом принятия решения [109, 110].

Пусть для выбранного стеганографического отображения γ существует статистический метод стеганоанализа заданный отображением τ , такой, что на его выходе выдается вычисленное значение вероятности присутствия в анализируемом контейнере $q \in Q$ скрытой информации, обозначим его $P_s(q)$. Тогда, для выбранных отображений γ и τ , по заданной величине порога принятия решения T , можно определить критический коэффициент сокрытия $K_{\text{крит}}$ – такое значение коэффициента сокрытия, при котором для $\forall q \in Q$: $q = c \oplus m \mid c \in C, m \in M, k = |m|/|c| < K_{\text{крит}}$ выполняется неравенство $P_s(q) \leq T$.

Если для заданного метода сокрытия информации и заданного метода анализа коэффициент сокрытия $k < K_{\text{крит}}$ для выбранной пары контейнер, сообщение, то величина ошибки второго рода будет столь значительной, что противник с большой долей вероятности не сможет выявить факт наличия скрытой информации.

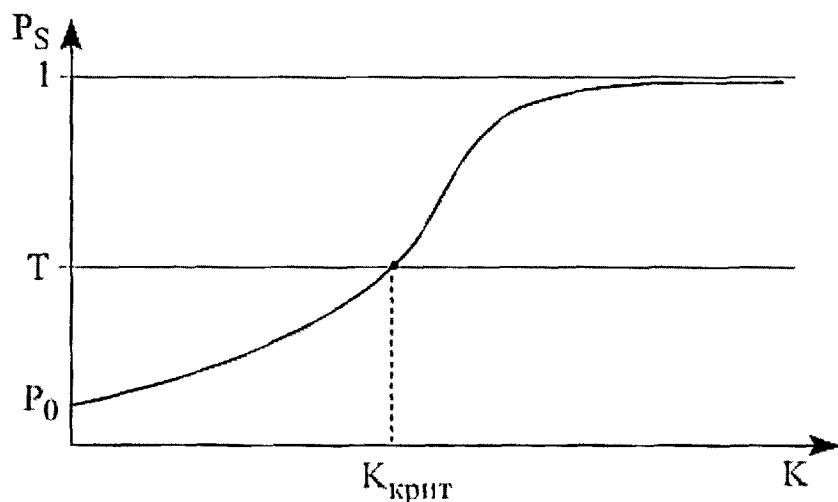


Рисунок 9 – График вероятности обнаружения стегосообщения

На рисунке 9 представлен график отображающий зависимость между коэффициентом сокрытия k , критическим коэффициентом сокрытия $K_{\text{крит}}$ и вероятностью $P_S(q)$ обнаружения скрытого сообщения выбранным методом стеганоанализа. Из графика видно, что с ростом объемов скрываемой информации вероятность обнаружения стеганографической системы растет, и в случае если коэффициент сокрытия превысит величину $K_{\text{крит}}$, системой анализа будут получены значения превышающие порог принятия решения и наличие скрытой информации будет установлено с вероятностью $P_S(q) \geq \alpha$. На графике также видно, что вероятность обнаружения стеганографической системы растет начиная с некоторого значения p_0 , данная величина как раз и соответствует вероятности совершения ошибки первого рода.

2.3.2 Теоретико-множественный подход к оценке эффективности многомодульных систем стеганографического анализа

Как известно, на настоящий момент существует уже достаточно большое число стеганографических методов, которые позволяют скрывать информацию в контейнерах различных типов. Разумеется, без точного знания используемого в данной стеганографической системе метода сокрытия, проведение атаки нацеленной на какой-то определенный стеганографический метод будет малоэффективным. Поэтому в условиях необходимости

обеспечения контроля большого потока различных контейнеров, перед атакующей стороной стоит достаточно сложная задача. С одной стороны необходимо обеспечить высокую достоверность результатов анализа, минимизировав вероятность ошибки первого рода, с другой проверить как можно большее количество возможных стеганографических методов, для того, чтобы минимизировать вероятность пропуска заполненного контейнера. Заметим, что метод сокрытия априори неизвестен атакующей стороне. В этих условиях невозможно ограничиться одним решающим правилом. Для проведения успешной атаки, необходимо для каждого анализируемого типа контейнеров $z \in Z$, например файлов формата JPEG, PCX, WAV, MP3, MPEG и т.п. определить наиболее вероятное подмножество $H_z \subseteq H$ множества методов сокрытия информации. Очевидно, что множество H и его подмножества H_z должны содержать не только существующие методы но, также и все возможные методы, которые уже построены или могут быть построены на их основе или с использованием тех же принципов. Общность принципов позволяет использовать одни методы стеганоанализа против целой группы методов сокрытия. Для каждого элемента, или некоторого множества элементов из H могут быть подобраны соответствующие один или несколько методов анализа (тестов) $l \in L$. В целях обобщения все тесты будем считать вероятностными, в независимости от выдаваемого ими на выходе вероятностного или однозначного решения.

Таким образом, для некоторого $h \in H_z$ можно определить набор тестов $L_{h,z} \subseteq L_z$ соответствующих заданному методу сокрытия информации в контейнерах заданного типа $z \in Z$. Так как все тесты являются вероятностными, то каждый из тестов $l_i \in L_{h,z}$, где $i = 1 \dots |L_{h,z}|$, обладает определенными ошибками первого и второго рода. Отметим, что для любого h подмножество $L_{h,z}$ не пусто, так как оно содержит, по крайней мере, один элемент l , который представляет собой вероятностный тест с вероятностями ошибок $\alpha = \beta = 0.5$.

В [99] предложен теоретико-множественный подход к оценке стойкости стеганографических методов. Пусть $C_{h,z} \subseteq C$ – подмножество контейнеров типа $z \in Z$ к которым может быть применен некоторый стеганографический метод $h \in H_z$, тогда, в общем случае, для заданного метода анализа $l \in L_{h,z}$ вероятность ошибки первого рода может быть определена как отношение:

$$\alpha_{l,h} = \frac{|F_{l,z} \cap \overline{F_{h,z}}|}{|C_{h,z}|},$$

где $F_{h,z} \subset C_{h,z}$ – подмножество действительно заполненных с применением метода $h \in H_z$ контейнеров, $F_{l,z}$ – подмножество контейнеров из $C_{h,z}$, которые выбранный тест l отнес к заполненным.

Аналогичным образом определим вероятность ошибки второго рода для заданного метода анализа, которая состоит в том, что контейнер, содержащий информацию, был признан пустым:

$$\beta_{l,h} = \frac{|F_{h,z} \cap \overline{F_{l,z}} \cap \overline{F_{h,z}}|}{|C_{h,z}|}.$$

Рассмотрим одну из возможных стратегий действий противника стремящегося обнаружить стеганографический канал связи, а также по возможности попытаться определить используемые алгоритмы и установить отправителя и получателя сообщений. Введем ограничение, что ресурсы противника конечны, а контролируемый им открытый канал связи является достаточно широким. В открытом канале связи могут передаваться как сообщения одного типа, так и ограниченного набора различных типов. Будучи поставленным в подобные условия противник будет вынужден использовать систему стеганоанализа состоящую из довольно большого числа отдельных модулей анализа [149], реализующих тот или иной тест, для заданного типа контейнеров и возможного стеганографического алгоритма.

Очевидным преимуществом такого подхода является возможность обнаружения гораздо более широкого множества методов сокрытия

информации. Попытаемся определить, каковы будут вероятности ошибок первого и второго рода многомодульной системы анализа при проведении атаки нацеленной на выявление стеганографических вложений полученных с применением заданного метода сокрытия информации. Опираясь на представленные выше вероятности ошибок первого и второго рода одиночного теста, вероятности ошибок многомодульной системы анализа для заданного метода сокрытия можно определить следующим образом:

$$\alpha_h = \frac{\left| \bigcup_{i=1 \dots |L_{h,z}|} (F_{l_i,z} \cap \overline{F_{h,z}}) \right|}{|C_{h,z}|}, \quad \beta_h = \frac{\left| \bigcap_{i=1 \dots |L_{h,z}|} (F_{h,z} \cap \overline{F_{l_i,z} \cap F_{h,z}}) \right|}{|C_{h,z}|}.$$

Из представленных равенств видно, что увеличение числа различных тестов, позволяет уменьшить вероятность ошибки второго рода, т.е. уменьшить вероятность пропуска заполненного контейнера. Однако с другой стороны из первого равенства следует, что с увеличением количества различных тестов вероятность ложных срабатываний системы может только увеличиться. В свою очередь, увеличение вероятности ошибки первого рода, может заставить противника повысить порог принятия решения, что в свою очередь может привести к увеличению вероятности пропуска действительно заполненных контейнеров. Оптимальное решение может быть достигнуто только путем грамотного задания порогов принятия решения и возможно введением дополнительных коэффициентов надежности, характеризующих каждый из методов анализа. Так при равнозначности большого числа тестов вероятность ошибки первого рода α может устремиться к единице. В случае же задания коэффициентов надежности вероятность ложного срабатывания может быть определена как: $\alpha_z = \sum_{l \in L_z} k_l \cdot \alpha_l$, где k_l – коэффициент надежности метода анализа $l \in L_z$, α_l – вероятность ошибки первого рода для заданного метода анализа по всему множеству допустимых контейнеров. Введение коэффициентов надежности также может позволить несколько снизить вероятность ошибки второго рода, путем применения алгоритмов

голосования к результатам работы отдельных тестов. Однако и в этом случае сохраняются определенные пороги принятия решения, и соответственно свои ошибки первого и второго рода.

В заключение, отметим, что сведение вероятности ошибки второго рода к нулю, простым наращиванием различных тестов в реальных системах невозможно, так как такой подход неизбежно ведет к лавинообразному увеличению числа ложных срабатываний комплексной системы. Минимизация величины ошибки второго рода возможна только путем совершенствования алгоритмов стеганографического анализа и построения более достоверной математической модели естественных контейнеров.

2.3.3 Практическая стойкость и предельный коэффициент сокрытия

Выше были введены понятия коэффициента сокрытия и критического коэффициента сокрытия. Рассмотрим, как эти коэффициенты могут быть связаны с вероятностями ошибок первого и второго рода методов стеганоанализа. Основной целью является определение максимально допустимых объемов скрываемой информации при обеспечении высокой практической стойкости. То есть необходимо решить задачу оценки допустимых объемов скрываемой информации, при которых методы стеганоанализа не позволяют вынести однозначного решения о наличии скрытой информации в некотором контейнере. Еще раз обратимся к совершенно стойким стеганографическим системам. На рисунке 10 представлен график вероятности обнаружения стегосообщения для случая совершенной стеганографической системы.

График отражает тот факт, что для случая совершенных стеганографических систем при записи дополнительной информации в контейнер, объемом не более допустимого, вероятность обнаружения стегосообщения не превышает таковую для случая пустого контейнера. Таким образом, для всех методов стеганоанализа вероятность обнаружения стеганографического вложения не превышает вероятности ложного срабатывания. Отталкиваясь от данного

положения, введем определение теоретической стойкости стеганографической системы [109, 110].

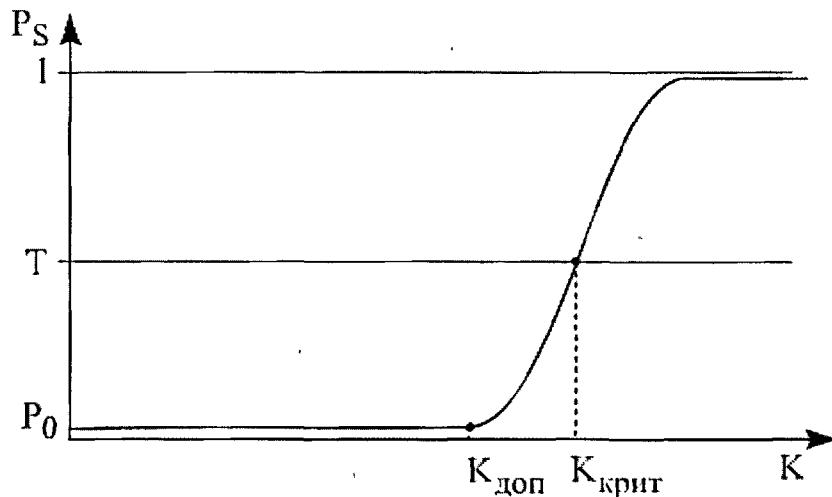


Рисунок 10 – График вероятности обнаружения стегосообщения в случае совершенной стеганографической системы

Определение. Стеганографическую систему основанную на отображении γ будем считать стойкой практически к выбранному методу стеганоанализа, если существует допустимый коэффициент сокрытия $0 < K_{\text{доп}} < K_{\text{крит}}$ – такой что, для $\forall q \in Q: q = c \oplus m \mid c \in C, m \in M, k = |m|/|c| < K_{\text{доп}}$ выполняется $P_S(q) \leq \alpha$, где $P_S(q)$ – апостериорная вероятность наличия скрытого сообщения в контейнере q , α – вероятность ошибки первого рода для выбранного метода стеганоанализа.

Если допустимый коэффициент сокрытия $K_{\text{доп}} > 0$ существует, то для любого контейнера c из множества возможных контейнеров C можно определить максимальный объем скрываемого сообщения:

$$\nu = \max_{m \in M: |m| < |c| K_{\text{доп}}} |m|.$$

Очевидно, что объемы скрываемой информации в значительной степени зависят как от качества алгоритмов сокрытия и особенностей их конкретной реализации, так и от надежности и эффективности систем стеганоанализа. Таким образом, на практике, допустимый коэффициент

сокрытия для некоторого стеганографического метода в действительности зависит от множества факторов. Так, если некоторый метод сокрытия информации является не обнаружимым для одного метода анализа, он может быть обнаружен другим более совершенным или просто более новым методом стеганоанализа. Исходя из этого, более надежная оценка допустимого максимального объема скрываемой информации для данного стеганографического метода может быть получена только по множеству методов анализа. Допустимый коэффициент сокрытия для данного случая обозначим через $K'_{\text{доп}}$. Положим, что противник располагает некоторым конечным множеством методов анализа L , которые могут быть применены против выбранного метода сокрытия информации. Тогда максимальный объем скрываемого сообщения для текущего контейнера, передача которого не приведет к компрометации стеганографического канала связи, может быть определен как:

$$V = \max_{m \in M : |m| < |c| \cdot K'_{\text{доп}} / \min_L (K'_{\text{доп}})} |m|.$$

Если значение V для текущего контейнера $c \in C$ по заданному множеству атак L больше нуля, то можно говорить, о том, что передача сообщения $m \in M$ объемом $v = |m| \leq V$ не предоставит возможность атакующей стороне обнаружить стеганографический канал и однозначно установить факт передачи скрытого сообщения. Разумеется, данное утверждение не может быть отнесено к ошибкам, допущенным в реализации стеганографической системы, которые могут оказаться серьезное влияние на допустимый объем скрываемых сообщений.

Представленный параметр позволяет сравнить два стеганографических метода по необнаруживаемости, надежности и информационной емкости на отдельных выборочных контейнерах. В тоже время, для практической оценки необходим критерий, позволяющий проводить сравнение не по отдельным контейнерам, а по некоторому достаточно широкому множеству различных контейнеров. Для целей практической оценки стойкости стеганографической

системы построенной на основе выбранного стеганографического метода рекомендуется использовать минимальный предельный коэффициент сокрытия, определяемый как:

$$K_{\text{мкс}} = \min_{c \in C} \left(\frac{\max_{m \in M : |m| < |c| \cdot K'_{\text{доп}} |K'_{\text{доп}} = \min_L} |m|}{|c|} \right).$$

Для оценки средней теоретической информационной емкости может быть использован средний коэффициент сокрытия, определяемый, аналогично минимальному предельному коэффициенту сокрытия по тестовому множеству контейнеров:

$$K_{\text{скс}} = \frac{\sum_{c \in C} \left(\frac{\max_{m \in M : |m| < |c| \cdot K'_{\text{доп}} |K'_{\text{доп}} = \min_L} |m|}{|c|} \right)}{|C|}.$$

Значения данных коэффициентов для выбранного метода сокрытия могут быть получены экспериментальным путем на основе тестовой выборки по множеству контейнеров и сообщений. При проведении экспериментов необходимо использовать наиболее полную библиотеку тестов, включающую как методы стеганоанализа, так и методы оценки вносимых искажений. Полученные в результате для некоторой стеганографической системы значения минимального предельного и среднего коэффициентов сокрытия будут однозначно характеризовать как стойкость, так и пропускную способность стегоканала.

Представленные коэффициенты позволяют наиболее точно оценить стойкость стеганографической системы к существующим на текущий момент методам анализа. Точность оценки во многом зависит от адекватности и репрезентативности выбранных тестовых множеств контейнеров, сообщений и методов анализа. Если выбранные тестовые множества в полной мере отражают реальную систему связи, то различные стеганографические методы можно сравнить, используя единый критерий.

Определение. Стеганографическую систему на стеганографическом методе $h \in H$ будем считать Δ стойкой к атакам пассивного противника, если при расширении множества L существует предел $\lim(K_{\text{мкс}}) = \Delta$.

Если для некоторой стеганографической системы значение $\Delta = 0$, то стегосистема не является совершенной и факт ее использования может быть установлен. Т.е. существует практическая возможность выявления скрытых вложений и может быть построен автоматизированный метод анализа, позволяющий обнаружить скрытый канал. Если $\Delta > 0$, то выявить стеганографический канал связи невозможно.

При решении практических задач множество методов анализа всегда конечно и ограничивается тестами, существующими на текущий момент времени. В этих условиях более удобной является оценка стойкости стеганографического метода по некоторому множеству методов анализа.

Определение. Стеганографическую систему на стеганографическом методе $h \in H$ будем считать Δ_L стойкой по множеству методов анализа L , где Δ_L определяется как $\Delta_L = K_{\text{мкс}}^L$, если $\Delta_L > 0$.

Если стеганографическая система является Δ_L стойкой по множеству всех известных методов анализа, то существующими на текущий момент времени средствами анализа выявить скрытый канал связи невозможно. Если для двух стегосистем построенных на основе стеганографических методов h_1 и h_2 соответственно $\Delta_{L,h_1} > \Delta_{L,h_2}$, то первая стегосистема обладает большей стойкостью к анализу и обеспечивает построение скрытого канала с более высокой пропускной способностью.

Предложенные в данном разделе методы оценки по заданному множеству атак L позволяют произвести практическую оценку стойкости выбранной для анализа стеганографической системы, а также сравнить стегосистемы на базе различных методов сокрытия, как по надежности, так и по эффективности процедур сокрытия информации.

ГЛАВА 3. РАЗРАБОТКА МОДЕЛЕЙ И МЕТОДОВ СИСТЕМ СКРЫТОГО ЭЛЕКТРОННОГО ДОКУМЕНТООБОРОТА

Современное быстрое развитие информационных систем и систем электронного документооборота, включая системы мультимедиа, превращают мир информационных технологий в невероятно динамично развивающуюся среду. В этих условиях системы скрытого электронного документооборота должны быть легко адаптируемы к новым типам контейнеров и должны обеспечивать возможность легкой перестройки под все более жесткие требования. Существующие коммерческие решения заточены под конкретные типы контейнеров, и порой претерпевают лишь незначительные изменения в целях повышения стойкости к новым атакам. Существенным ограничением в большинстве случаев является закрытость не только программного кода, но и описания используемых методов.

Данная глава посвящена разработке базовых принципов и построению систем скрытой передачи информации, отличающихся открытостью методов и алгоритмов, высокой гибкостью и простотой перестройки под новые требования и условия функционирования. Предложены новые подходы по проектированию гибридных – криптостеганографических систем скрытой передачи электронных документов. Введены ограничения и заданы требования к составным частям конечной системы, позволяющие в итоге обеспечить гибкость архитектуры и высокий уровень защиты информации. Представлены унифицированный метод, обеспечивающий оптимальное согласование криптографических и стеганографических алгоритмов в рамках единой системы, и новые стеганографические методы.

3.1 Криптостеганографические системы связи: базовые принципы, модель и определение

3.1.1 Критерии стойкости систем скрытой передачи ЭД

Прежде чем перейти к рассмотрению предлагаемого в данной главе подхода к построению систем скрытого электронного документооборота,

ГЛАВА 3. РАЗРАБОТКА МОДЕЛЕЙ И МЕТОДОВ СИСТЕМ СКРЫТОГО ЭЛЕКТРОННОГО ДОКУМЕНТООБОРОТА

Современное быстрое развитие информационных систем и систем электронного документооборота, включая системы мультимедиа, превращают мир информационных технологий в невероятно динамично развивающуюся среду. В этих условиях системы скрытого электронного документооборота должны быть легко адаптируемы к новым типам контейнеров и должны обеспечивать возможность легкой перестройки под все более жесткие требования. Существующие коммерческие решения заточены под конкретные типы контейнеров, и порой претерпевают лишь незначительные изменения в целях повышения стойкости к новым атакам. Существенным ограничением в большинстве случаев является закрытость не только программного кода, но и описания используемых методов.

Данная глава посвящена разработке базовых принципов и построению систем скрытой передачи информации, отличающихся открытостью методов и алгоритмов, высокой гибкостью и простотой перестройки под новые требования и условия функционирования. Предложены новые подходы по проектированию гибридных – криптостеганографических систем скрытой передачи электронных документов. Введены ограничения и заданы требования к составным частям конечной системы, позволяющие в итоге обеспечить гибкость архитектуры и высокий уровень защиты информации. Представлены унифицированный метод, обеспечивающий оптимальное согласование криптографических и стеганографических алгоритмов в рамках единой системы, и новые стеганографические методы.

3.1 Криптостеганографические системы связи: базовые принципы, модель и определение

3.1.1 Критерии стойкости систем скрытой передачи ЭД

Прежде чем перейти к рассмотрению предлагаемого в данной главе подхода к построению систем скрытого электронного документооборота,

определим основные свойства, которыми должна обладать современная стеганографическая система. Далее будем рассматривать систему в условиях противодействия обнаружению и перехвату передаваемых сообщений со стороны пассивного противника обладающего равными или превосходящими вычислительными возможностями. В общем случае, для того чтобы стеганографическая система вида $\Sigma_S = \langle C, M, K, Q, H_K, R_K \rangle$ обладала высокой стойкостью к атакам пассивного противника необходимо выполнение следующих условий:

1. Должна сохраняться функциональность контейнеров. Заполненный контейнер должен обладать свойством естественности, т.е. принадлежать множеству всех возможных контейнеров $Q \subseteq C$.
2. Распределение $P(C \cup Q)$ должно быть равномерным, т.е. должно выполняться условие $\forall c \in C \cup Q : p(c) = 1 / |C \cup Q| = 1 / |C|$. Данное требование говорит о равной вероятности появления в канале связи любого контейнера $c \in C$.
3. Распределение $P(M)$ должно быть равномерным, т.е. должно выполняться условие $\forall m \in M : p(m) = 1 / |M|$. Передача любого из возможных сообщений равновероятна.
4. Множества пустых C и заполненных Q контейнеров, ключей шифрования, а также передаваемых сообщений M должны находиться в отношениях $|C| \geq |Q| > 1$, $|Q| \geq |M| > 1$, $|K| \geq |Q|$.
5. Однозначность ключа $|K(c, m)| = 1$ для любой пары (c, m) , где $c \in C \cup Q$, $m \in M$.
6. Распределение $P(K)$ должно быть равномерным, использование в сеансе связи любого ключа из K должно быть равновероятным.
7. Для любого контейнера $c \in C \cup Q$ вероятность наличия в нем дополнительной информации должна быть равна 1, т.е. любой пустой контейнер тоже содержит в себе информацию.

8. В процессе записи информации в контейнер должны сохраняться все статистические характеристики любого из распределений, полученных в результате вычисления всех возможных функций $f: C \rightarrow X$ для $Q \rightarrow X$, где X – некоторое произвольное множество.
9. Внедрение информации в контейнер должно осуществляться не за счет записи в него дополнительной информации, а за счет изменения уже существующей $\forall m \in M, c \in C: H(q) = H(c), q = h(m, c)$, где H – функция меры количества информации.
10. Для $\forall m \in M$ и $\forall c \in C$ сообщения $m = r(q)$ и $m' = r(c) : q = h(m, c)$ должны быть статистически неразличимы.

Для того чтобы стеганографическая система обладала высоким уровнем стойкости, применяемые в ней стеганографические алгоритмы должны максимально полно отвечать представленным выше требованиям.

3.1.2 Гибридная – криптостеганографическая система

В первой главе данной работы были рассмотрены различные модели стеганографических систем предназначенных для обеспечения скрытого электронного документооборота. В качестве наиболее перспективного направления было выделено направление построения гибридных систем с явным применением элементов криптографии и стеганографии [101, 114, 116, 120, 121, 122, 124]. Среди основных преимуществ подобных систем было отмечено, что явное применение в системах современных криптографических алгоритмов шифрования позволяет значительно повысить уровень защищенности передаваемой информации к раскрытию сути передаваемых сообщений. Вместе с тем обзор существующих решений в области систем скрытой передачи электронных документов, позволяет судить о преимущественном применении в настоящее время неперестраиваемых стеганографических алгоритмов. Что является серьезным ограничением к применению ключей на этапе скрывающего преобразования. Более того, можно сделать вывод о бесперспективности данного подхода в

ближайшее время, так как использование зависимых от ключей скрывающих преобразований требует от авторов предоставления дополнительных доказательств надежности и стойкости. Разумеется, что данная задача нетривиальна. Основной причиной является отсутствие на настоящий момент необходимой для решения задачи теоретической базы, что ставит данное направление в тупик. В сложившейся ситуации альтернативным подходом, позволяющим уже сейчас эффективно решать ряд задач в системах скрытого электронного документооборота, является совмещение криптографии и стеганографии. Но следует оговориться, что действительно положительных результатов можно добиться только путем их согласованного применения.

Рассмотрим стеганографическую систему $\Sigma_S = \langle C, M, Q, h, r \rangle$. В представленном наборе множеств, множество скрывающих преобразований и множество правил извлечения сообщений являются одноэлементными, т.е. в системе используются не перестраиваемые безключевые стеганографические алгоритмы. Секретность систем данного вида в первую очередь определяется секретностью используемых алгоритмов и способов сокрытия информации. Раскрытие сведений о системе, ее компрометация позволяют противнику получить практически полный контроль над используемым скрытым каналом связи. Противник может просматривать все передаваемые сообщения, вносить в них изменения и подменять. Так же он может полностью разрушить сам скрытый канал. В криптографии при оценке надежности той или иной системы зачастую обращаются к правилам Керкгоффса [141]. Напомним, что согласно второму правилу Керкгоффса: *компрометация системы не должна причинять неудобств пользователем*. Суть правила заключается в том, что противнику, при проведении им атаки на какую-либо систему, могут быть известны все детали ее реализации. Единственный неизвестный параметр – секретный ключ преобразования. Таким образом, второе правило Керкгоффса можно перефразировать следующим образом: *надежность системы должна определяться лишь секретностью ключа*. Ввести секретный ключ в случае безключевых стеганографических

алгоритмов позволяет использование в конечных системах согласованной криптографической части.

Использование шифрования исключительно как дополнительного этапа преобразования скрываемой информации с целью обеспечения стойкости к установлению сути передаваемых сообщений в действительности не решает всех необходимых задач. К примеру, в большинстве современных стеганографических систем связи используются алгоритмы встраивания, необнаруживаемость (степень скрытности) и надежность (стойкость) которых сильно зависит от статистических свойств передаваемых данных.

Выполнение части из указанных в предыдущем пункте требований к стеганографическим системам связи может быть обеспечено применением криптографических алгоритмов. Так, например, применение современных шифров позволяет обеспечить выполнение требования равномерности распределения над M . Кроме того, на криптографическую часть комплексной системы может быть переложена вся работа с секретной составляющей – ключом. Перенос ряда задач на криптографическую часть позволяет значительно упростить разработку всей стеганографической системы, так как позволяет использовать в качестве основы наиболее простые скрывающие преобразования, относящиеся к системам вида $\Sigma_S = \langle C, M, Q, h, r \rangle$. Помимо прочего, упрощается также и доказательство стойкости конечной системы. Вся часть наиболее сложных доказательств, связанных с использованием секретных ключей, автоматически переносится в область криптографии. Использование же в системе уже известных, отработанных и даже стандартизованных криптографических алгоритмов позволяет и вовсе оставить значительную часть вопросов без доказательств.

Заметим, что согласно седьмому требованию, противник может извлечь сообщение из любого контейнера. Отметим также требования 8-10. Так, если характеристики сообщений, извлеченных из пустого и заполненного контейнеров, совпадут, то, без знания самого передаваемого сообщения или какой-либо другой дополнительной информации, однозначно установить

факт скрытой передачи сообщения будет невозможен. Однозначность характеристик сообщений может быть обеспечена криптографической частью совместно с алгоритмами согласования, при этом секретный ключ шифрования фактически превращается в секретный ключ всей стеганографической системы.

Определение. Крипстеганографической системой будем называть систему скрытой передачи информации на открытых каналах связи, основанную на совместном применении криптографических алгоритмов, стеганографических методов, а также алгоритмов согласования входных и выходных данных указанных алгоритмов и методов.

Обобщенная модель крипстеганографической системы представлена на рисунке 11.

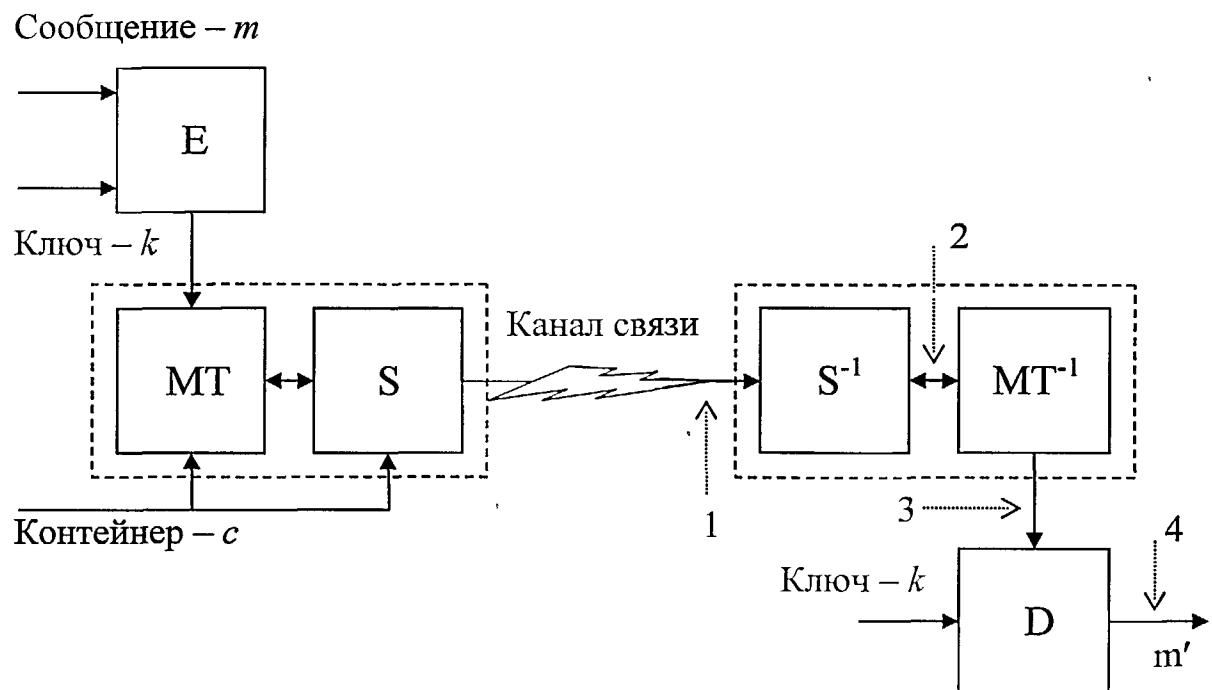


Рисунок 11 – Модель крипстеганографической системы связи

Криптографическая часть (E, D) обеспечивает криптографическое закрытие (предварительное шифрование) передаваемых сообщений. Отвечает за преобразование передаваемых сообщений к псевдослучайному виду с равномерным распределением.

Стеганографическая часть (S) осуществляет непосредственное скрытие и извлечение передаваемых данных, прошедших процедуру предварительного шифрования, в контейнерах из С.

Алгоритмы согласования (MT) обеспечивают согласование криптографической и стеганографической частей системы по входным и выходным данным. Отвечают за прямое приведение и обратное преобразование полученных с выхода криптографической части данных к двоичным последовательностям аналогичным по своим статистическим свойствам двоичным последовательностям, извлекаемым из пустых контейнеров.

Рассмотрим возможности противника по выявлению скрытого канала и восстановлению передаваемых сообщений, в условиях наличия у него определенной информации об используемом канале. Введем следующую градацию степени информированности потенциального противника:

1. Известна информация только об открытом канале связи, информация о скрытом канале не известна;
2. Известно только стеганографическое преобразование;
3. Известно стеганографическое преобразование и используемые в системе алгоритмы согласования;
4. Известны все составляющие системы за исключением секретного ключа криптографической части.

Для каждого из указанных случаев на исходной модели определим точки съема информации с канала (см. рисунок 11). В точке 1 противник контролирует открытый канал связи и ему доступны сообщения вида $c_i \in C$ и $q_j = S(c_j, MT(c_j, E(m, k)))$, где $i \neq j$, $c_i \in C$, $q_j \in Q$, $m \in M$, $k \in K$. Задачей противника является выделить среди перехватываемых открытых сообщений стегосообщения, т.е. различить подмножества образованные рядом перехваченных сообщений c_i и q_j . В данном случае противник ничего не знает о скрытом канале связи и может строить свои выводы только на основе общего анализа перехватываемых открытых сообщений, соответственно

стойкость крипостеганографической системы определяется надежностью стеганографической части S . Если пустые и заполненные контейнеры статистически неразличимы, противник не в состоянии обнаружить скрытый канал и передаваемые по нему сообщения.

Второй случай для большинства современных систем является критическим, ведь противник располагает теми же алгоритмами извлечения информации, что и непосредственные пользователи системы. В этом случае, в точке 2 противник может легко извлекать из всех перехватываемых им контейнеров сообщения $m' = S^{-1}(c) : c \in C \cup Q$ для последующего более детального анализа. Отметим, что в крипостеганографических системах одним из требований является возможность извлечения сообщений стандартными алгоритмами даже из пустых контейнеров, при этом извлекаемая из пустых контейнеров информация должна быть статистически неотличима от информации, извлекаемой из заполненных контейнеров. Учитывая данные требования в случае крипостеганографических систем, полная информация об используемых стеганографических алгоритмах не упрощает стоящих перед противником задач, т.к. для противника сообщения $m'_i = S^{-1}(c_i) : c_i \in C$ и $m'_{j'} = S^{-1}(q_j) : q_j \in Q$ будут эквивалентны $m'_i \approx m'_{j'}$. В результате, даже имея возможность извлекать из контейнеров скрытую в них информацию, противник будет не в состоянии определить, действительно ли извлеченные им данные являются скрытой информацией или же информацией изначально присутствовавшей в контейнере.

Аналогичная ситуация складывается для противника и в третьем случае. Имея полную реализацию стеганографической части и алгоритмов согласования, на выходе (в точке 3) из каждого анализируемого контейнера противником будет получена псевдослучайная последовательность с равномерным распределением $E(m, k)$ или R , где R – случайная составляющая данных контейнера. Что опять-таки не позволит однозначно судить о наличии или же отсутствии скрытой информации.

Таким образом, единственной точкой в системе, снимая информацию с которой можно определить наличие скрытого сообщения является точка 4 на представленной модели. Очевидно, что применение современных шифров не позволит противнику получить какую либо информацию о зашифрованном сообщении без знания секретного ключа. В тоже время, конечный адресат, будучи авторизованным пользователем системы, знающим секретный ключ, может легко определить и восстановить полученное сообщение. Для этого необходимо вычислить $m = D(MT^{-1}(S^{-1}(S(c_j, MT(c_j, E(m, k)))), k))$. Заметим, что для восстановления сообщения, знание исходного контейнера $c_j \in C$ не требуется, а секретный ключ шифрования фактически превращается в ключ всей стеганографической системы.

В общем случае криптостеганографическая система обладает стойкостью эквивалентной стойкости используемых в ней алгоритмов шифрования, но только тогда, когда применяемые стеганографические алгоритмы отвечают описанным выше требованиям и являются совершенными или Δ стойкими к атакам пассивного противника, с $\Delta > 0$.

3.2 Криптографическая часть

3.2.1 Требования к криптографическим алгоритмам

В системах связи согласно предложенному подходу, основным, предъявляемым к криптографической части, требованием является равномерное распределение на множестве зашифрованных сообщений. Криптографические алгоритмы должны обеспечивать однозначное прямое и обратное отображение передаваемых сообщений $m \in M$ на множество случайных строк R над алфавитом A используемым алгоритмами согласования. Должно выполняться равенство $m = d_k(e_k(m))$, где e_k и d_k соответствующие правила зашифрования и расшифрования на ключе $k \in K$.

В общем случае, для любого сообщения $m \in M$, строки $r \in R$ и пары ключей $k_1 \neq k_2: k_1, k_2 \in K$ должно выполняться:

$$1. m = d_{k1}(e_{k1}(m)), m = d_{k2}(e_{k2}(m));$$

2. $m \neq d_{k2}(e_{kl}(m))$, $d_{k2}(e_{kl}(m)) \in M$;
3. $m' = d_{k1}(r)$, $m'' = d_{k2}(r) : m', m'' \in M$.

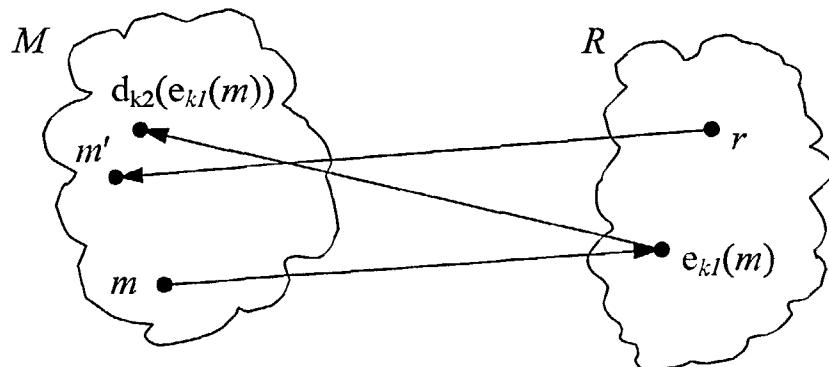


Рисунок 12 – Отображение сообщений на множество случайных строк

Представленные требования говорят о том, что любое сообщение $m \in M$ может быть представлено случайной строкой из R , и любая случайная строка из R может быть отображена в сообщение из M . Данные требования также могут быть поставлены в соответствие требованиям к стеганографическим системам связи по пунктам 5, 6 и 10. Следует отметить, что практически любой современный шифр в полной мере отвечает представленным требованиям и может быть легко использован в предлагаемой модели.

3.2.2 Гибкость архитектуры

Требования, предъявляемые к криптографической части, практически не накладывают ограничений по ее внутренней структуре. Серьезным требованием является лишь равномерность и высокая случайность выходного потока, что дает большую свободу по выбору компонентов и алгоритмов, используемых в криптографической части. В частности криптографическая часть может быть представлена алгоритмами, реализующими современные симметричные блочные или поточные шифры. Кроме того, в криптографической части могут быть реализованы так же и элементы асимметричной криптографии, такие как цифровая подпись и использование открытых ключей.

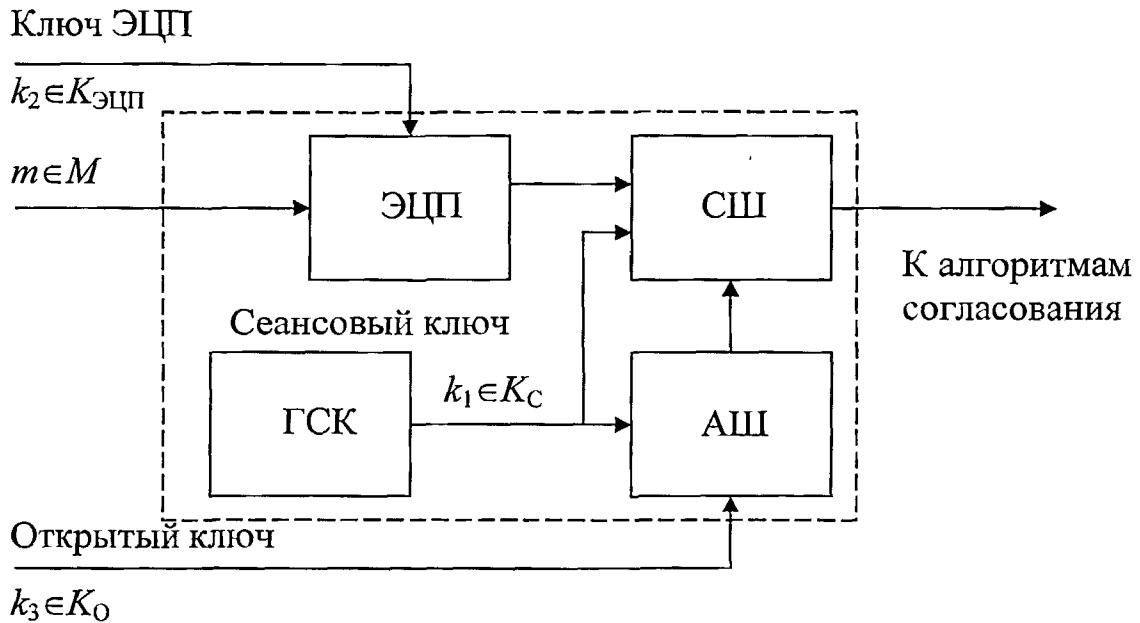


Рисунок 13 – Гибридная криптографическая часть на стороне отправителя

На рисунке 13 представлен один из возможных вариантов гибридной криптографической части, совмещающей алгоритмы симметричной и асимметричной криптографии. В предложенной схеме в дополнение к базовому симметричному шифру (СШ), непосредственно реализующему отображение передаваемых сообщений на множество псевдослучайных последовательностей, добавлены модули генерации сеансового ключа (ГСК), электронно-цифровой подписи (ЭЦП) и асимметричного шифрования (АШ). На вход схемы подается передаваемое сообщение $m \in M$, секретный ключ ЭЦП $k_2 \in K_{\text{ЭЦП}}$ и открытый ключ адресата $k_3 \in K_O$. Сеансовый симметричный ключ $k_1 \in K_C$ формируется модулем ГСК либо на основе данных генератора случайных чисел, либо на основе специализированных алгоритмов генерации ключей, например ANSI X9.17. Подготовка сообщения для передачи разделяется на три этапа. На первом для поступившего на вход сообщения $m \in M$ выполняется процедура вычисления ЭЦП на заданном секретном ключе отправителя $k_2 \in K_{\text{ЭЦП}}$, результат вычисления дописывается в конец сообщения. На втором этапе, модулем ГСК формируется сеансовый ключ

$k_1 \in K_C$, который в дальнейшем используется для шифрования всего сообщения в модуле СШ. Сеансовый ключ $k_1 \in K_C$ проходит процедуру зашифрования в модуле АШ на открытом ключе адресата $k_3 \in K_O$. Результат добавляется в заголовок сообщения. На третьем этапе исходное сообщение $m \in M$, дополненное ЭЦП и заголовком, проходит процедуру зашифрования с использование симметричных криптографических алгоритмов в модуле СШ на сеансовом ключе $k_1 \in K_C$.

Представленная схема в полной мере позволяет реализовать такие услуги безопасности как конфиденциальность, аутентичность, целостность, и принадлежность передаваемых электронных документов. Применение гибридной криптографической части позволяет максимально оптимизировать систему под конкретные условия эксплуатации, обеспечивает общую высокую гибкость канала передачи информации и позволяет отказаться от использования прочих дополнительных надстроек над конечной системой. Все это в свою очередь снижает конечную стоимость и повышает общую надежность реализуемого решения по защите электронных документов.

3.3 Стеганографическая часть

3.3.1 Требования к стеганографическим методам и алгоритмам

Как было отмечено выше, стеганографическая часть может быть представлена стеганографической системой вида $\Sigma_S = \langle C, M, Q, h, r \rangle$, где h и r некоторые фиксированные безключевые алгоритмы сокрытия и извлечения информации (сообщений $m \in M$) в контейнерах из C . Очевидным преимуществом систем данного вида является простота их разработки и конечной реализации, что является следствием отсутствия сложных преобразований и необходимости использования секретных ключей. На практике отсутствие лишних элементов, меньшая неопределенность относительно скрываемых данных и применение безключевых правил внедрения/извлечения информации позволяет применять значительно более простые стеганографические алгоритмы.

Для обеспечения общего высокого уровня стойкости к атакам со стороны пассивного противника в рамках крипстеганографической системы, непосредственно на стеганографическую часть накладываются следующие ограничения:

1. Заполненный контейнер должен сохранять свою функциональность и обладать свойством естественности:

$$\forall q = h(c, m) : c \in C, m \in M \mid q \in C;$$

2. Стеганографический алгоритм r должен извлекать сообщения как из заполненных, так и из пустых контейнеров, при этом вероятность извлечения сообщений не должна зависеть от выбранного контейнера:

$$\forall m = r(c) : c \in C \cup Q \mid m \in M, p(m) = 1/|M|;$$

3. Запись сообщения в контейнер должна осуществляться путем замены исходно содержащегося в нем сообщения, а не путем внесения в контейнер дополнительной информации, $H(q) \approx H(c)$, где H – функция меры количества информации;
4. Процесс записи данных нового сообщения должен учитывать структуру контейнера и существующие корреляционные связи между элементами контейнера;

Дополнительно к общим ограничениям в крипстеганографических системах связи к стеганографической части предъявляются следующие ограничения:

5. Извлекаемые и записываемые сообщения должны представлять собой битовые строки конечной длины с $p_0 > 0, p_1 > 0$;
6. Стеганографический алгоритм должен быть совершенным или Δ стойким к атакам пассивного противника при $\Delta > 0$.

Требования 1-3 обоснованы с точки зрения теории информации (затрагиваются вопросы теоретической стойкости), требования 4-6 связаны с практической стойкостью стеганографических алгоритмов. Допустим, что не выполняется или выполняется, но не в полной мере, первое ограничение, т.е.

при записи информации в контейнер нарушаются его функциональность и (или) естественность. В обоих случаях случае результат не может принадлежать множеству естественных контейнеров, и как следствие может быть легко отфильтрован либо визуально, либо автоматически с использованием достаточно простых правил. Так, например, передача файла, содержимое которого явно не соответствует указанному расширению, фактически ведет к раскрытию канала, так как файл не может быть нормально отображен у перехватившего его противника. В результате нарушается базовый принцип стеганографии – заполненный контейнер не должен вызывать подозрений у перехватившего его противника. Под нарушением естественности понимается сильное искажение контейнера, возможные естественные причины которого отсутствуют. В качестве примера можно привести сильные визуальные искажения и артефакты в графических изображениях неестественные для обычных изображений.

Требования 2 и 3 связаны между собой. Согласно теории информации количество информации содержащейся в случайной последовательности $Y \in B^n$ относительно $X \in A^n$ определяется как: $I\{X, Y\} = H\{X\} - H\{X|Y\}$. Если стеганографическая система является совершенной, то для нее выполняется равенство: $I\{X, Y\} = I\{Y, X\} = 0$. В данном случае количество информации содержащейся в некотором контейнере Y о передаваемом сообщении X , а также о наличии последнего, равны нулю. В случае, когда система не является совершенной, условная энтропия $H\{Y|X\} < H\{Y\}$ и существует ненулевая вероятность обнаружения факта передачи информации вследствие выполнения неравенства: $I\{Y, X\} = H\{Y\} - H\{Y|X\} > 0$. Таким образом, если функций сокрытия осуществляет дополнение, а не замену информации в контейнере, то это ведет к увеличению энтропии $H\{Y\}$, что в свою очередь увеличивает количество информации о сообщении X в контейнере Y . Отсюда очевидна связь между требованиями 2 и 3. Если встраивание дополнительной информации основано на замещении уже изначально имеющейся в контейнере информации скрываемыми данными, то алгоритмы извлечения

информации будут способны извлекать исходную информацию и из пустых контейнеров. Под исходной информацией в данном случае понимается информация изначально присутствующая в пустом контейнере, которая может быть замещена при встраивании скрываемого сообщения.

Еще раз отметим, что в представленной стеганографической системе не используются секретные ключи. Однако для совершенных систем требуется выполнение равенства $I\{X, Y\} = I\{Y, X\} = 0$. Это означает, что в случае успешного перехвата со стороны противника передаваемого открыто сообщения-контейнера, извлечение скрытой составляющей не должно привести к уменьшению неопределенности атакующей стороны о скрытно передаваемом сообщении и собственно факте передачи. Добиться выполнения данного условия возможно лишь за счет равновероятности появления в скрытом канале сообщений из M , причем распределение самих сообщений так же должно быть равномерным. Обеспечить равномерность сообщений, как отмечалось ранее, позволяет этап предварительного шифрования всех передаваемых сообщений.

В тоже время, очевидно, что в условиях применения естественных контейнеров и реальных стеганографических алгоритмов извлекаемая из пустых контейнеров информация не всегда будет псевдослучайной и отвечающей равномерному распределению. В данном случае выполнение требования по п. 2 обеспечиваются алгоритмами согласования.

Пункты 4 – 6 относятся к конечной реализации стеганографических алгоритмов. Для того, что бы прояснить данные пункты рассмотрим непосредственно процедуру сокрытия информации. Как показано на рисунке 11 на первом шаге скрываемое сообщение проходит процедуру предварительного шифрования в криптографической части (блок Е). Далее зашифрованное сообщение поступает на вход алгоритмов согласования. Туда же передается битовая строка, извлеченная блоком S из пустого контейнера выбранного для передачи скрытого сообщения. Алгоритмы согласования осуществляют замену информации в битовой строке пустого контейнера

данными скрываемого сообщения (осуществляют модуляцию битовой строки). После модифицированная битовая строка подается на вход алгоритмов стеганографического преобразования, осуществляющих сокрытие информации (обратное встраивание ранее извлеченной битовой строки) в выбранном контейнере.

Так как скрывающие преобразования и алгоритмы согласования являются детерминированными, это накладывает ограничение по п. 5, извлекаемые и записываемые строки должны быть конечной длины. Данное ограничение в большей части касается фиксированных контейнеров и позволяет заранее определить допустимый объем скрываемых данных. В свою очередь работа с поточными контейнерами может быть реализована путем пакетной передачи с использованием временных буферов и (или) технологии бегущего окна, что позволяет представить поточный контейнер как с последовательный ряд фиксированных контейнеров.

Ограничения же 4 и 6 напрямую связаны с практической стойкостью всей системы. С выхода алгоритмов согласования в случае встраивания дополнительной информации ненулевого объема на вход алгоритмов встраивания подается измененная битовая строка. Отдельные элементы этой строки могут иметь неявные связи с другими элементами контейнера. В последствии это может привести к возможному выявлению скрытых вложений, методами стеганоанализа учитывающими эти неявные связи. Учитывая такую возможность в ряде случаев при встраивании битовой строки обратно в контейнер необходимо обеспечить восстановление нарушенных корреляционных связей. Требование по п. 6 является более жестким, и обеспечивает реальную практическую стойкость всей комплексной системы. Нарушение данного пункта говорит о возможности выявления факта передачи скрытой информации с использованием современных и перспективных методов стеганоанализа.

3.3.2 Стеганографические методы на базе пространственно-частотных фильтров усредняющих масок

Человеческое зрение обладает определенными особенностями. Одно из таких особенностей проявляется в том, что с ростом пространственной частоты, анализируемого изображения, падает контрастная характеристика чувствительности зрительной системы. Вследствие этого даже сравнительно большие искажения яркостной составляющей легко заметные на низкочастотных участках изображения окажутся незаметными на участках с высокой пространственной частотой. Этим феноменом объясняется то, что цифровой шум в изображениях гораздо более заметен на фоне неба, нежели на фоне земли или деревьев. Следует также отметить эффект маскирования заключающийся в том, что если некоторый фрагмент изображения имеет достаточно сложную структуру, элементы которой не могут быть предсказаны зрительной системой, то отдельные, порой даже весьма значительные, искажения на данном фрагменте останутся незаметными. Таким образом, выделение высокочастотных участков изображения, учитывая также их большую непредсказуемость и высокую случайность, для целей стеганографического скрытия информации, очевидно, является перспективным. Ниже представлен пример одномерной функции с участками с преобладанием низкочастотной и высокочастотной составляющих.

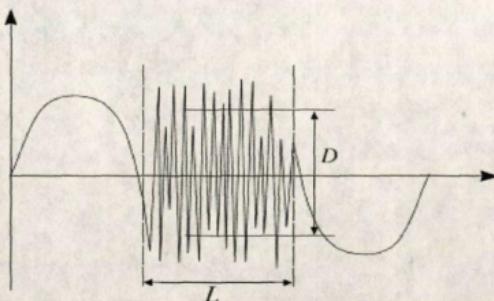


Рисунок 14 – Одномерная функция с низкочастотными участками и участком с преобладанием высокочастотной составляющей

На представленном примере видно, что предсказать поведение функции (путем аппроксимации и интерполяции) на высокочастотном участке L практически невозможно, хотя средний уровень сигнала D сохраняется. Любой из фрагментов функции на участке L может быть увеличен или уменьшен по абсолютному значению амплитуды на значительную величину, в то время как аналогичное изменение фрагмента на низкочастотном участке сигнала может быть легко обнаружено. Данное свойство и используется в стеганографических методах на базе пространственно-частотных фильтров [111, 112].

Растровое графическое изображение может быть представлено двумерной дискретной функцией $f(m, n)$, где m и n – координаты точки изображения (целочисленные индексы в направлении строк и столбцов). Значение, принимаемое функцией $f(m, n)$ в каждой точке координатной плоскости является дискретным значением яркости соответствующей точки монохромного изображения. В качестве такого монохромного изображения далее рассматриваются растровое графическое изображение, представленное в оттенках серого. Для фотоизображений яркость точек будем считать заданной дискретно в диапазоне от 0 до 255, где 0 соответствует яркости точек изображения черного цвета, 255 – белого. Любое цветное растровое фотоизображение также может быть представлено в оттенках серого путем выделения его яркостной составляющей. Наиболее широко в теории обработки фотоизображений для этих целей применяется схема $YC_R C_B$. Это трехкомпонентная схема представления цвета, в которой Y представляет собой яркостную составляющую изображения, C_R и C_B цветоразностные составляющие. Преобразование изображения из наиболее естественного его представления RGB (заданные отдельно интенсивности красной (R), зеленой (G) и синей (B) составляющих) в режиме True Color для представления в цветовой схеме $YC_R C_B$ осуществляют по следующим формулам:

$$Y = 0,299 \cdot R + 0,587 \cdot G + 0,114 \cdot B;$$

$$C_R = 0,5 \cdot R - 0,4187 \cdot G + 0,0813 \cdot B;$$

$$C_B = 0,1687 \cdot R - 0,3313 \cdot G + 0,5 \cdot B.$$

Монохромное изображение, полученное в результате вычисления яркостной составляющей Y , в действительности несет в себе большую часть воспринимаемой человеком информации об анализируемом им изображении и фактически может рассматриваться как представленное в оттенках серого цвета исходное изображение. Цветоразностные составляющие C_R и C_B используются в предлагаемом методе только для коррекции искажений отдельных точек изображения при обратном переходе к представлению RGB. Изменение значения Y при обратном преобразовании может привести к выходу значений интенсивности цветовых составляющих RGB за допустимый предел 0-255. Исключить это можно двумя способами, либо корректировать непосредственно выходные значения RGB для каждой компоненты в отдельности, либо предварительно подвергнуть коррекции составляющие C_R и C_B . Последний вариант лучше, так как снижает вероятность ошибки при последующем извлечении информации.

Выделение для записи информации отдельных точек в областях с высокой пространственной частотой предлагается осуществлять с помощью пространственных фильтров на основе матриц:

$$h_{M1} = \begin{pmatrix} 1 & 1 & 1 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \quad h_{M2} = \begin{pmatrix} 0 & 1 & 1 & 1 \\ 1 & 2 & 2 & 2 \\ 1 & 2 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}.$$

Использование указанных масок пространственных фильтров позволяет определить, сколь сильно отличается текущая точка от своего окружения. Для случая использования h_{M1} :

$$t(m, n) = \frac{1}{4} (f(m-1, n-1) + f(m-1, n) + f(m-1, n+1) + f(m, n-1)) - f(m, n).$$

Если $|t(m, n)| \geq T$, где T – порог принятия решения, значение яркости точки изображения с координатами (m, n) может быть подвергнуто изменению. Превышение порога T в данном случае говорит о том, что текущая точка изображения сильно отличается от своего окружения.

Особенностью представленных фильтров является то, что в процессе определения возможности использования текущей точки изображения для записи информации участвуют только уже обработанные ранее точки. Объясняется это тем, что для представленных фильтров обработка изображения осуществляется построчно сверху вниз с направлением по строкам слева на право, при этом точкам изображения, которые еще не были обработаны, соответствуют коэффициенты 0 в матрице фильтров. Основным результатом такого подхода является то, что процесс выбора точек в измененном изображении для извлечения информации полностью идентичен процессу выбора точек при сокрытии.

Порог T определяется с учетом предполагаемого уровня развития методов стеганоанализа вероятного противника и уровнем вносимых искажений в процессе хранения и передачи изображения по каналам связи. Очевидно, что последующее возможное использование алгоритмов сжатия с потерями требует задания более высокого порога. Вместе с тем, чем выше порог T , тем меньше точек будет выделено для записи информации и соответственно тем меньше будет информационная емкость контейнеров.

Каждая из точек изображения выделенная в процессе встраивания данных с помощью указанных пространственно-частотных фильтров может быть использована для записи от одного до трех информационных битов. Под информационными битами в данном случае понимаются биты сообщения с выхода алгоритмов согласования. Количество записываемых битов определяется исходя из критериев стойкости к возможным искажениям в процессе хранения и передачи, качества исходного изображения, порогов зрительного восприятия, а также степени вносимых искажений допускаемой существующими методами стеганоанализа. Использование одной точки

изображения для встраивания одного информационного бита, при отсутствии последующих искажений (например, в процессе сжатия изображения) достаточно очевидно. Рассмотрим более интересный случай, в котором точки изображения используются для записи сразу двух информационных битов.

Примем порог T равным 5 и введем переменную $b_2(i)$ – десятичное значение вектора из двух записываемых на очередном шаге i информационных битов, тогда процесс записи информации в контейнер можно описать следующим образом:

$$f(m,n) = \begin{cases} f(m,n) - t(m,n) \bmod 5 + b_2(i), & t(m,n) \geq 5 \\ f(m,n) + |t(m,n)| \bmod 5 - b_2(i), & t(m,n) \leq -5 \\ f(m,n), & -5 < t(m,n) < 5 \end{cases}$$

В указанном преобразовании есть два особых случая, которые связаны с диапазоном допустимых значений для функции $f(m,n)$. Данная функция, для обычных графических изображений, может принимать значения в диапазоне от 0 до 255. Для исключения ситуаций связанных с выходом за допустимый диапазон, после внесения изменений в значение яркости очередной точки изображения необходима дополнительная проверка. Так если новое вычисленное значение $f(m,n) \geq 255$ ($f(m,n) \leq 0$), то значение ограничивается $f(m,n) = 255$ ($f(m,n) = 0$), а изменение индекса i не осуществляется. Если же значение $f(m,n)$ после изменения находится в допустимом диапазоне (1;254), запись очередной пары информационных битов считается проведенной успешно и значение индекса i увеличивается на единицу.

Извлечение информационных битов из заполненного контейнера, как отмечалось ранее, осуществляется в точности согласно процессу записи, за исключением последнего шага. На последнем шаге осуществляется вычисления переменной $b_2(i)$:

$$b_2(i) = |t(m,n)| \bmod 5, |t(m,n)| \geq 5 \wedge f(m,n) > 0 \wedge f(m,n) < 255.$$

На рисунке 15 представлен результат применения описанного метода с порогом $T=5$, для случая внесения дополнительной информации на основе фильтра с маской h_{M1} и кодера обеспечивающего запись максимально

допустимых объемов скрываемых данных. На рисунке показано изменение яркостной составляющей для исходного и обработанного изображений. В представленном примере амплитудные искажения отдельных точек достигают десятков процентов, и в тоже время они остаются незаметными вследствие расположения в высокочастотных областях изображения.

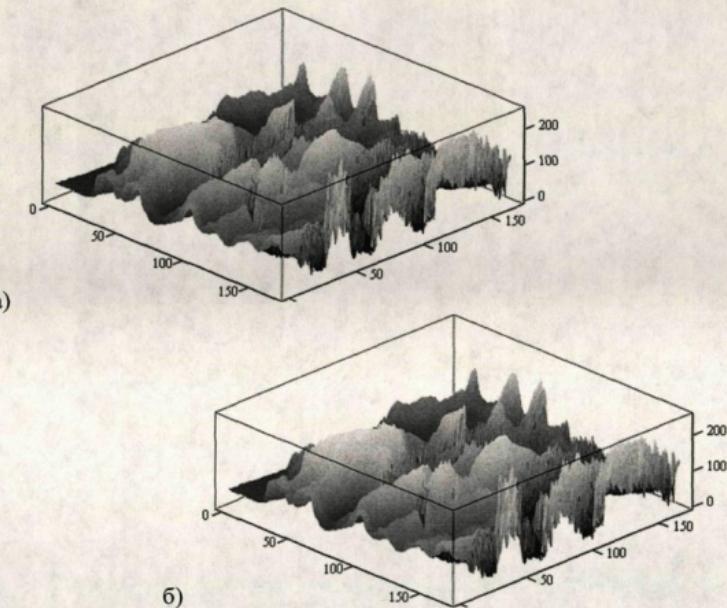


Рисунок 15 – Сравнение яркостных составляющей на участке исходного (а) и обработанного изображений (б)

Проведенные эксперименты показали высокую стойкость предложенного метода сокрытия информации, к различным методам стеганоанализа включая также и методы, основанные на анализе пространственных корреляций, такие как метод учета двойственных статистик на основе анализа RS-диаграмм. Метод обладает высоким

коэффициентом сокрытия, объем скрываемых данных для предложенного метода может составлять 1/5 от объема исходного изображения. Кроме того, на базе данного метода возможно построение метода с использованием кодов исправляющих ошибки, что позволит обеспечить стойкость скрываемых данных при последующем сжатии с потерей качества по алгоритму JPEG.

Отметим, что представленный метод отвечает всем требованиям, предъявляемым к стеганографическим методам, рекомендуемым к использованию в крипстеганографических системах связи.

3.4 Особенности реализации алгоритмов согласования

3.4.1 Требования к алгоритмам согласования, базовый алгоритм

Крипстеганографическая система представляет собой сложный комплекс, общая стойкость которого не определяется только лишь стойкостью используемого скрывающего преобразования. Большую роль для надежности системы играет правильное согласование криптографической и стеганографической частей системы. На выходе криптографической части получить псевдослучайную битовую строку, отвечающую равномерному распределению можно известными и стандартными алгоритмами. В тоже время задача разработки стеганографических алгоритмов способных извлекать аналогичные строки из пустых контейнеров может оказаться весьма сложной. В рамках данной работы было проведено исследование по анализу перспективных подходов к построению стеганографических алгоритмов на множестве контейнеров различных типов и различной природы. В качестве исследуемых мультимедиа контейнеров использовались цифровые графические изображения (форматы без сжатия, со сжатием без потери и с потерей качества) и аудиозаписи (речь в различной обстановке, аудиокомпозиции). Проведенные исследования с применением современных наиболее распространенных стеганографических алгоритмов подтвердили предположение о неравномерном характере распределения извлекаемых из изначально пустых контейнеров битовых строк. Зачастую извлекаемые

строки имели длинные монотонные последовательности нулей и единиц. В подавляющем большинстве случаев случайный характер распределения имели лишь биты расположенные на границах переходов указанных монотонных последовательностей и сами позиции переходов.

Результаты экспериментов говорят в пользу применения в качестве основы теоретической базы для построения алгоритмов согласования, способа записи дополнительной информации в битовую строку предложенного в разделе 2.3.2 данной работы. Напомним, что согласно приведенной в указанном разделе теореме, существует возможность построения совершенных стеганографических систем на битовых строках содержащих длинные серии одинаковых битов. Следует отметить, что в нашем случае выполняются все необходимые условия:

- a) граничные элементы всех серий одинаковых битов в исходной строке случайны и их распределение является равномерным;
- б) скрываемые сообщения представляют собой псевдослучайные битовые строки конечной длины;
- в) распределение нулей и единиц в скрываемом сообщении является равномерным;

Учитывая соблюдение всех необходимых условий, можно считать, что алгоритмы согласования в данном случае фактически выступают в роли совершенных стеганографических алгоритмов. Они осуществляют сокрытие псевдослучайных битовых строк конечной длины (строки с выхода криптографического преобразования) в битовых строках содержащих длинные серии одинаковых битов (строки, извлеченные основным стеганографическим алгоритмом из пустых контейнеров).

Однако применение на практике сложных алгоритмов реализующих предложенную выше машину Тьюринга не всегда оправдано. В ряде случаев может оказаться допустимым использование несовершенных, но в тоже время более простых алгоритмов. При проектировании специфичных

алгоритмов согласования необходимо учитывать, что конечный алгоритм должен отвечать следующим основным требованиям:

1. Должно обеспечиваться точное согласование входных и выходных данных криптографической и стеганографической частей системы;
2. Алгоритм должен осуществлять преобразование битовой строки с равномерным распределением к битовой строке с распределением соответствующим распределению битовых строк извлекаемых из пустых контейнеров;
3. Запись информации в битовую строку должна осуществляться с сохранением свойств, статистических характеристик и специфичных особенностей исходной последовательности;
4. В результирующей битовой строке должны отсутствовать признаки модификации и очевидные отличительные особенности;
5. Должна обеспечиваться возможность однозначного восстановления исходной (скрываемой) битовой строки из строки битов полученной в результате проведенных преобразований.

Помимо базового алгоритма согласования существует также возможность задания табличного кода реализующего встраивание дополнительной информации в битовые строки.

3.4.2 Метод сдвига битовых последовательностей

С целью упрощения программной или аппаратной реализации и сведения к минимуму числа дополнительно контролируемых параметров для записи информации в битовую строку предлагается использовать метод сдвига битовых последовательностей (МСБП) заданный табличным кодом [107, 108]. Применение МСБП позволяет автоматизировать процесс записи информации и свести его к простой замене четырехбитных векторов, сформированных из идущих последовательно битов исходной битовой строки векторами таблицы замен. Процесс записи информации представлен на рисунке 16. В каждый момент времени кодер просматривает последовательность битов контейнера по четверкам, при этом изменению

подвергаются только средние два бита. Биты слева и справа позволяют более точно определить границу переходов и исключить ситуации, когда граница оказывается между рассматриваемыми парами. Смещение окна просмотра кодера на каждом шаге составляет две позиции, что позволяет осуществлять запись информации без пропуска переходов между последовательными сериями одинаковых битов и одиночными единицами и нулями.

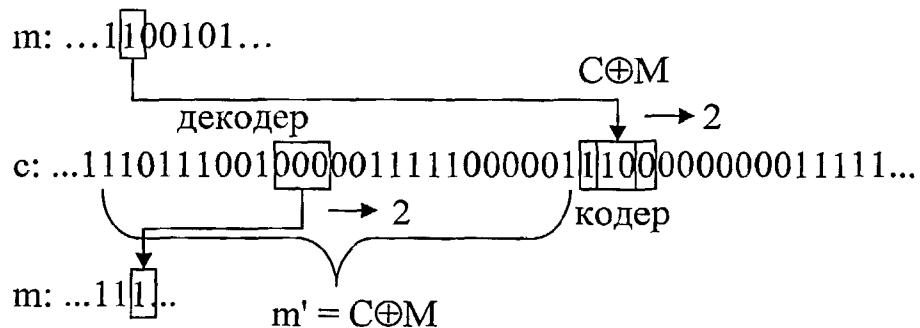


Рисунок 16 – Запись зашифрованного сообщения в битовую строку

Предлагаемый кодер осуществляет замену элементов битовой последовательности в соответствии с записываемыми данными согласно кодовой таблице. Кодовая таблица состоит из двух столбцов содержащих вектора замены соответственно для нуля и единицы записываемой информации. Строки таблицы соответствуют всем возможным четырехбитным векторам. На каждом шаге, окно просмотра кодера выделят четырехбитный отрезок строки (вектор) и в соответствии с кодовой таблицей по текущему информационному символу (биту данных) осуществляет его замену другим вектором, согласно кодовой таблице.

Для получения оптимальной кодовой таблицы была проведена серия экспериментов, в частности были проведены исследования статистических характеристик растровых фотoreалистичных изображений. После детального анализа полученных результатов была разработана наиболее оптимальная кодовая таблица. Использование приведенной ниже (таблица 4) кодовой таблицы позволило построить код обеспечивающий сокрытие максимально

возможных объемов информации (за счет рационального использования всех кодовых слов) при условии минимального искажения характеристик используемых контейнеров. Предлагаемая таблица является наиболее оптимальной в отношении эффективность/вычислительная сложность. Вместе с тем она не исключает возможность использования других кодовых таблиц. В зависимости от свойств контейнеров выбранного типа и специфики решаемых задач в конечной реализации кодовая таблица и ее использование могут отличаться и иметь определенные особенности.

Таблица 4 – Кодовая таблица прямого преобразования

Вход (4-х битный вектор)	Запись нуля	Запись единицы
0000	—	—
0001	0011	—
0010	0010	0100
0011	0011	0111
0100	0010	0100
0101	0011	0101
0110	0010	0100
0111	0011	0111
1000	1000	1100
1001	1011	1101
1010	1010	1100
1011	1011	1101
1100	1000	1100
1101	1011	1101
1110	—	1100
1111	—	—

Обратное преобразование (восстановление скрытой в битовой строке информации) осуществляется с помощью специального упрощенного кода. Предложенный декодер анализирует битовую строку не по четверкам, как

скрывающий код, а по тройкам и извлекает информацию в соответствии со своей кодовой таблицей (таблица 5). Смещение окна просмотра декодера на каждом шаге так же составляет две позиции (см. рисунок 16). Использование в декодере трех битов объясняется отсутствием необходимости в прогнозировании вероятности встречи перехода битовых серий на следующем шаге. При анализе битовой строки декодер не вносит в нее никаких изменений и, следовательно, также отсутствует необходимость в контроле возможных вносимых искажений.

Таблица 5 – Кодовая таблица обратного преобразования

Вход	Результат
000	—
001	0
010	1
011	1
100	0
101	0
110	1
111	—

Из анализа кодовой таблицы кодера можно предположить, возможность ситуации, когда уже внесенные изменения повлекут за собой цепочку неверных изменений последующих бит. Однако это не так. При использовании предложенного кодера подобный эффект не наблюдается вследствие быстрого затухания наследуемой ошибки. Связано это с тем, что должна сложиться ситуация, когда на достаточно протяженном участке некоторая последовательность битов контейнера должна четко соответствовать определенной цепочке битов скрываемой строки. Но, учитывая то, что как минимум срываемая битовая строка представляет собой последовательность битов с равномерным распределением нулей и единиц, на каждом шаге вероятность такой ситуации будет уменьшаться вдвое.

Так же можно заметить, что некоторые замены из таблицы 4 могут привести к изменению числа единиц или нулей. Это изменение заметно лишь в локальном масштабе. Благодаря удачному сочетанию отдельных замен их воздействие на статистику взаимно компенсируется. Экспериментально было установлено, что на реальных битовых стоках извлекаемых из пустых контейнеров (фотореалистичных графических изображений) максимальное отклонение соотношения нулей и единиц в заполненном и исходном контейнере не превосходит 1.45%. Это крайнее отклонение, когда скрываемые данные представляют собой строку, состоящую из одинаковых битов (только из нулей или только из единиц). Чем распределение нулей и единиц в скрываемом сообщении более равномерно, тем меньше возможное отклонение пустого и заполненного контейнеров. При использовании в качестве скрываемой строки возможное отклонение стремиться к нулю.

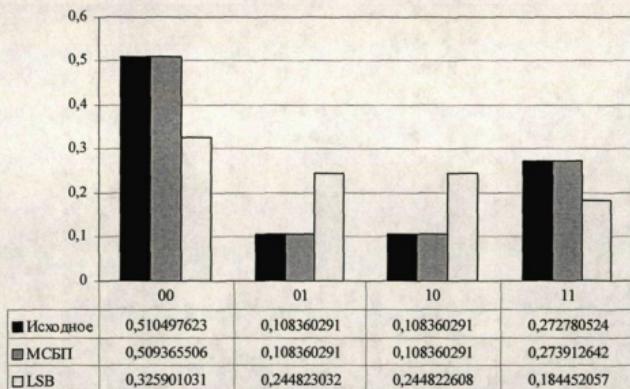


Рисунок 17 – Соотношение числа переходов

На рисунке 17, показаны результаты одного из экспериментов по применению МСБП совместно с наиболее популярным методом скрытия информации – методом замены младших значащих битов (LSB). Для

проведения тестов случайным образом было отобрано несколько графических изображений полученных обычной цифровой фотокамерой. Эти изображения использовалось в качестве контейнера для сокрытия информации. Скрываемые данные представляли собой сжатые (ZIP) заархивированные данные других изображений. Для каждого изображения, по очереди выполнялись две процедуры встраивания данных. В первом случае сокрытие информации осуществлялось непосредственно самим методом замены младших значащих битов, без какой либо модификации. Во втором случае с применением алгоритмов согласования, роль которых выполнял описанный выше код. Во всех случаях заполнение контейнера осуществлялось более чем на 98,5% свободного пространства сокрытия контейнера. Оцениваемым параметром было отношение числа переходов. Результаты всех экспериментов оказались очень близки. Так для приведенного случая, разница между числом переходов от единицы к единице в исходном и заполненном с применением метода сдвига битовых последовательностей контейнере составила 0,0011321. То есть отклонение частотных характеристик заполненного контейнера не превысило 0.23%. Для стандартного метода замены младших значащих битов величина отклонения данного показателя составила 36%. Количество же переходов между последовательностями одинаковых битов как видно из представленного графика для метода сдвига битовых последовательностей сохранилось практически с абсолютной точностью. Данные результаты позволяют говорить о достаточно высокой эффективности предложенного кода.

Результаты экспериментального применения метода сдвига битовых последовательностей совместно со стеганографическим методом замены младших значащих битов представлены в приложении Б. Здесь же отметим, что непосредственное применение МСБП позволило в несколько раз увеличить емкость контейнеров и при этом значительно повысить стойкость к ряду статистических методов анализа.

3.5 Мультиплексирование канала

Помимо прочих особенностей криптостеганографические системы скрытой передачи электронных документов обладают еще одним важным свойством. Они позволяют организовать протокол передачи информации не только в режимах 1:1 и 1:n, но протокол передачи информации в режиме n:n с использованием одного общего открытого канала связи [117]. При этом ни один из законных пользователей системы, даже зная все используемые в ней алгоритмы, будет неспособен определить идет ли в канале связи обмен скрытой информацией между другими пользователями системы. Данная возможность обусловлена тем, что в криптостеганографических системах проверить факт наличия в контейнере скрыто передаваемой информации можно только после процедуры расшифровывания извлеченных данных. При этом если получатель скрыто передаваемого сообщения не знает секретного ключа, последний не сможет не только расшифровать уже извлеченные им открытым стеганографическим алгоритмом данные, но даже проверить являются ли полученные данные зашифрованным сообщением или же просто «мусором». Именно на этом факте и основан механизм мультиплексирования стеганографического канала.

На рисунке 18 приведена модель мультиплексированного канала с множеством отправителей и получателей сообщений. Поясним работу данной модели на примере ее возможного использования в условиях глобальной сети Internet. В этой сети существует множество серверов, сайтов и файловых хранилищ позволяющих создавать общие разделяемые ресурсы для большого числа пользователей. Допустим, что разделяемый ресурс позволяет каждому из пользователей периодически просматривать все хранящиеся в открытом доступе мультимедиа файлы (контейнеры), размещенные на данном ресурсе. В данном случае применение криптостеганографического подхода позволяет каждому пользователю системы в автоматизированном режиме проводить извлечение предназначеннной непосредственно ему информации. Для этого программно

реализуется сбор всех доступных контейнеров, для каждого из контейнеров выполняется процедура извлечения скрытых данных. И далее, если ключ пользователя позволяет расшифровать извлеченные данные, на выходе он получает предназначеное ему сообщение (сообщение для которого он знает секретный ключ расшифрования). Если ключ пользователя не позволяет расшифровать извлеченные данные и эта операция завершается выдачей вместо осмысленного сообщения, нераспознаваемого «мусора», системой принимается решением об отсутствии скрытого сообщения. При таком подходе пользователь видит только предназначенные ему документы, а информация, предназначенная другим пользователям системы, будет для него недоступна. В случае если один из секретных ключей известен сразу нескольким пользователям, то все они имеют возможность ознакомиться с информацией защищенной данным ключом.

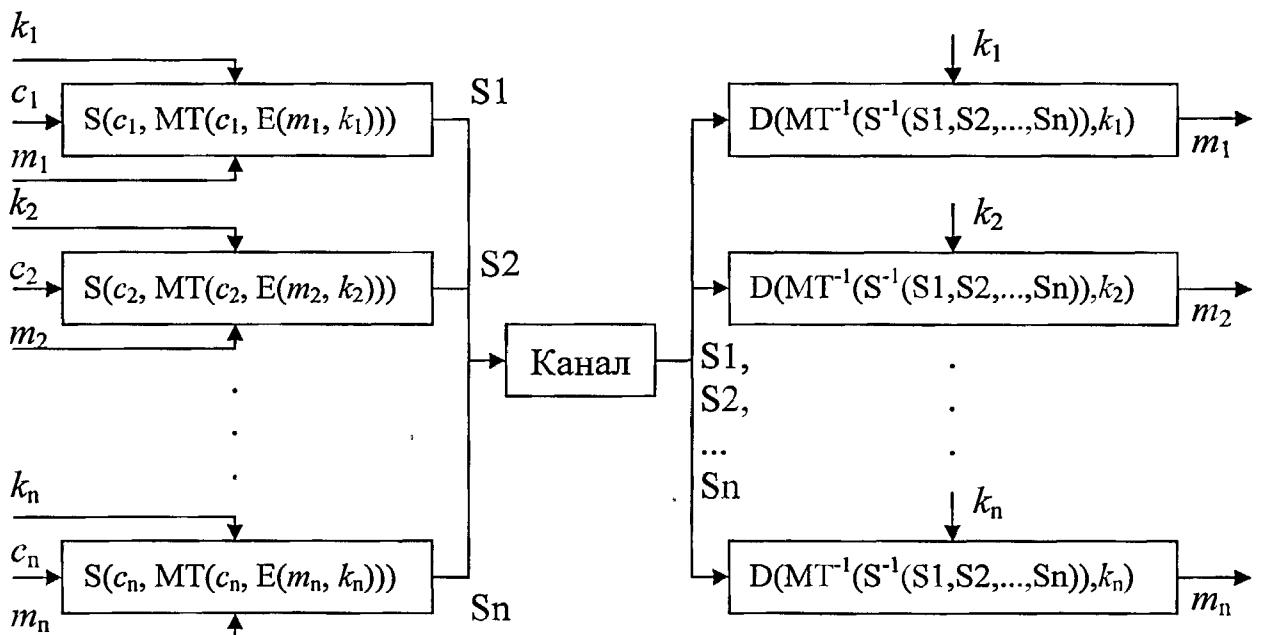


Рисунок 18 – Мультиплексирование канала

Использование общего Интернет ресурса позволяет большому числу пользователей организовывать свои скрытые каналы за счет размещения новых контейнеров. Так, например, в случае использования в качестве контейнеров фотоизображений каждый пользователь может разместить в

сети свою фотогалерею и предоставить всем желающим возможность просматривать и размещать в ней свои фотографии. Часть фотографий в галерее может являться контейнерами, содержащими скрытую информацию. Таким образом, совершенно обычный сетевой ресурс может быть превращен в мультиплексированный канал скрытого электронного документооборота.

3.6 Выводы по главе

1. Определены критерии стойкости систем скрытой передачи электронных документов проектируемых для целей обеспечения скрытого электронного документооборота на базе общих открытых систем и сетей передачи данных, таких как Интернет.

2. Представлены криптостеганографические системы связи, которые могут быть использованы для целей скрытой передачи электронных документов. Введено определение, предложена модель, определены требования ко всем базовым элементам системы. Предложены методы и алгоритмы, отвечающие указанным требованиям. В частности разработаны новые стеганографические методы на базе пространственно-частотных фильтров и метод сдвига битовых последовательностей, позволяющий упростить реализацию алгоритмов, согласования сведя их к алгоритму табличной замены.

3. Очевидным преимуществом представленного решения является возможность использования известных симметричных и ассиметричных крипtosистем. Показано, что криптостеганографические системы обладают высокой гибкостью и просты в перестройке под новые задачи и требования. Так на одном стеганографическом алгоритме может быть построен целый ряд конечных систем с различными характеристиками. Основными преимуществами предложенного подхода по совместному использованию криптографических и стеганографических алгоритмов являются:

- Возможность использования безключевых алгоритмов сокрытия и извлечения информации;

- Стойкость системы к раскрытию факта скрытой передачи информации определяется стойкостью криптографической части;
- Используемые в системе алгоритмы, особенности их взаимодействия и реализации, равно как и сама структура системы могут быть открыты;
- В криптографической части могут быть использованы существующие алгоритмы, без их перестройки и дополнительной оптимизации под «новые» задачи;
- Система с симметричными ключами может быть легко трансформирована в систему с открытыми ключами;
- Значительно упрощается процедура доказательства соответствия определенному уровню стойкости к атакам пассивного противника.

4. Особо отмечена возможность мультиплексирования канала связи,

т.е. возможность использования одного открытого канала связи множеством пользователей системы, при котором они не могут просматривать чужие сообщения и оказывать на них влияние.

ГЛАВА 4. ПРИМЕНЕНИЕ СКРЫТОЙ ПЕРЕДАЧИ ИНФОРМАЦИИ В РАСПРЕДЕЛЕННЫХ СИСТЕМАХ ЭЛЕКТРОННОГО ДОКУМЕНТООБОРОТА

В данной главе рассматриваются вопросы практического применения и особенности реализации криптостеганографических систем связи для решения различных задач электронного документооборота. Выделены два основных направления:

- скрытая передача электронных документов;
- контроль распространения копий электронных документов.

Для каждого из направлений предложены обобщенная архитектура, протоколы, принципы построения и детальные схематические решения, которые могут быть использованы при проектировании реальных систем. Представленные решения позволяют реализовать все плюсы предложенного подхода по построению криптостеганографических систем скрытой передачи информации. В данной главе так же показано, что криптостеганографические системы могут быть использованы не только для целей скрытой передачи информации, но и для целей защиты авторского права.

4.1 Архитектура многопользовательских распределенных систем скрытого электронного документооборота

В современных открытых сетях, существует возможность построения каналов передачи данных различного рода. Так структура и особенности организации хранения и передачи данных в сети Интернет позволяют организовать различные каналы связи, сильно отличающиеся временными задержками, надежностью доставки и скоростью передачи данных. Возможность организации множества скрытых каналов с различными характеристиками позволяет выбрать оптимальное решение для каждой из целого круга задач. Например, на основе использования файлов мультимедиа в сети Интернет могут быть организованы системы скрытой передачи данных для больших объемов информации, с другой стороны сообщения на

форумах могут быть использованы для сигнальных систем. Наибольшее внимание в данной работе уделяется скрытой передаче больших объемов информации, что позволяет использовать предложенные в работе подходы для целей скрытой передачи электронных документов. Применение технологии мультиплексирования канала позволяет создавать скрытые электронные хранилища и системы скрытой передачи документов. В этом случае доступ на передачу и получение документов имеют только авторизованные пользователи. Для всех прочих пользователей сети, скрытое электронное хранилище выступает лишь в роли некоторого открытого сервера, предоставляющего различную открытую информацию и ряд дополнительных услуг. В этом случае, предоставляемая сервером открытая информация, доступна большому числу пользователей системы и гостей, что позволяет использовать ее в качестве контейнеров для скрытого предоставления информации в соответствии с ключами пользователей.

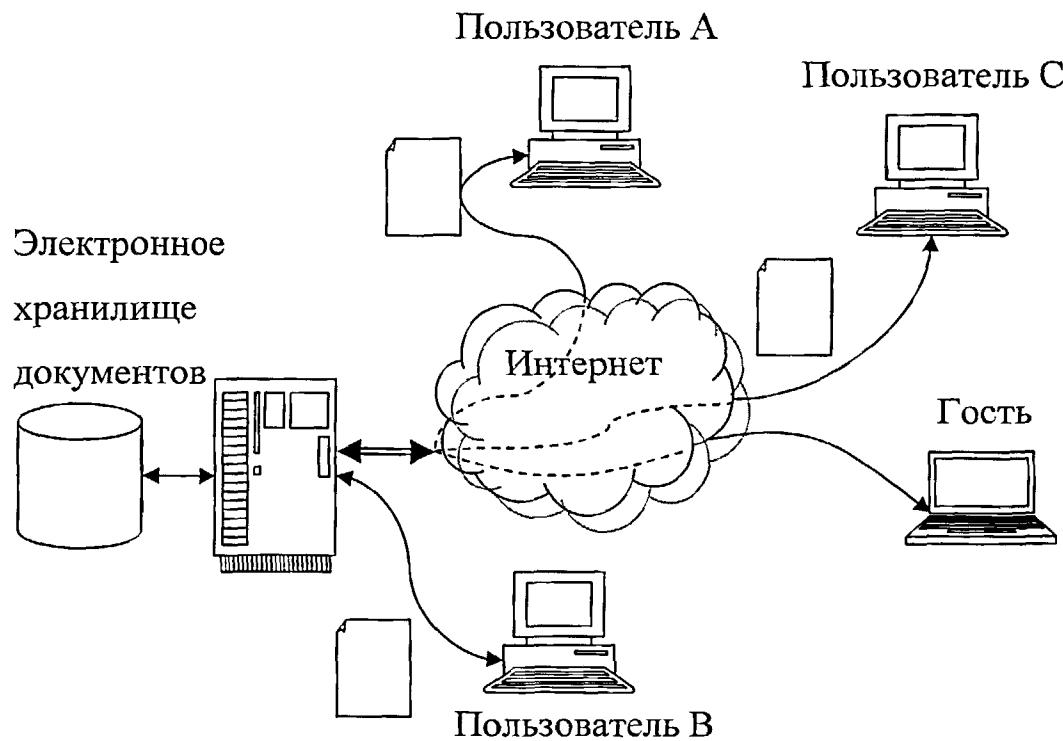


Рисунок 19 – Многопользовательская СЭД с возможностью удаленного доступа посредствам скрытых каналов связи

4.1.1 Организация процесса скрытой передачи ЭД

Любая надежная система передачи информации представляет собой сложный комплекс, общая надежность которого не определяется только использованием в составе системы наиболее современных алгоритмов. Большую роль для надежности системы играет правильное согласование всех компонентов и точное следование всем заданным ограничениям.

Процесс передачи информации в современных системах связи состоит из целого ряда отдельных процедур, таких как предварительная подготовка сообщения, шифрование, помехоустойчивое кодирование, канальное кодирование. В классических стеганографических системах связи роль процедуры канального кодирования выполняет стеганографическое сокрытие данных передаваемого сообщения в некотором контейнере.

Рассмотрим процесс сокрытия информации в цифровом мультимедиа контейнере в криптостеганографической системе связи. В общем случае процесс записи информации в контейнер в криптостеганографической системе содержит следующие основные этапы:

1. Выбор уникального и неиспользовавшегося ранее контейнера $c \in C$;
2. Извлечение из выбранного контейнера c битовой строки $t \in T$;
3. Оценка объема данных, которые можно записать в битовую строку t , с учетом ограничений, накладываемых стеганографическим алгоритмом встраивания сообщений (например, элементов строки, которые не могут быть изменены);
4. Предварительная подготовка сообщения $m \in M$;
5. Шифрование на секретном (открытом) ключе получателя $k \in K$ прошедших предварительную подготовку данных сообщения, для получения битовой строки $y \in Y$ с равномерным распределением;
6. Сведение алгоритмами согласования полученных на этапах 2 и 5 битовых строк $t' = g(t, y)$;

7. Запись путем прямого стеганографического преобразования в исходный контейнер взамен строки t полученной на этапе 6 битовой строки t' ;
8. Уничтожение использованного исходного контейнера;
9. Выдача заполненного контейнера $q \in Q$.

Соответственно процесс извлечения сообщения из ранее заполненного контейнера, согласно предлагаемому протоколу, включает следующие этапы:

1. Извлечение из полученного контейнера $c \in C$ битовой строки $t \in T$;
2. Извлечение из битовой строки t алгоритмами согласования псевдослучайной составляющей $y = g^{-1}(t)$, где $y \in Y$;
3. Расшифровывание битовой строки y на секретном ключе $k \in K$;
4. Восстановление сообщения $m \in M$;
5. Проверка заголовка и контрольной суммы сообщения;
6. Вывод результатов (сообщение без заголовка и контрольной суммы или отрицательное заключение).

При извлечении сообщения из контейнера контрольная сумма сообщения и заголовок используются в качестве некоторого индикатора, что извлечено именно сообщение, а не случайные данные («мусор»). Если восстановленное сообщение содержит верный заголовок и правильную контрольную сумму, процесс извлечения сообщения завершился успехом. В противном случае, если не восстановлен заголовок сообщения или не совпала контрольная сумма, делается вывод о том, что был получен пустой контейнер. Отметим, что контрольная сумма, заголовок сообщения, равно как и сами данные сообщения оказываются доступными только на выходе криптографической части системы. Организация процесса указанным образом обеспечивает выполнение одно из главных требований предъявляемых к стойким стеганографическим системам – пустой контейнер тоже содержит в себе сообщение.

Шаги этапа предварительной подготовки сообщений сильно зависят от особенностей используемого канал и объемов передаваемой информации. В

общем случае этап может содержать следующие шаги: сжатие сообщения; разбиение больших сообщений на блоки; дополнение сообщений и каждого из блоков заголовком и вычисленной контрольной суммой; если в процессе передачи есть вероятность искажения данных, то дополнительно на каждый из передаваемых блоков накладываются коды коррекции ошибок.

4.1.2 Общая схема криптостеганографической системы скрытой передачи ЭД на стороне отправителя

Общая схема предлагаемой модели стегоканала, с выделением всех основных этапов формирования стегосообщений на стороне отправителя [114, 116] представлена на рисунке 20. Согласно предложенной схеме этап предварительной подготовки сообщения может включать следующие шаги:

1. Определение способа представления информации, анализ сообщения, определение типа и начального размера;
2. Приведение сообщений к единому типу, дополнение передаваемого сообщения заголовком и ЭЦП (или же общей контрольной суммой). На выходе сообщение $m \in M$ содержащее заголовок, данные передаваемого сообщения и ЭЦП.
3. Сжатие исходного сообщения $m \in M$ с целью увеличения энтропии и уменьшения объемов передаваемой информации $|m'| \leq |m|$;
4. Приведение размера сжатого сообщения в соответствие емкости текущего выбранного контейнера и дополнение сообщения канальным заголовком. На данном шаге большое сообщение может быть разбито на блоки, а короткое сообщение дополнено случайными данными $m' \rightarrow m'_1, m'_2, \dots, m'_n$, где размер $|m'_i|$ определяется емкостью соответствующего контейнера $c_i \in C$;
5. Формирование заголовка содержащего информацию о длине, номере и прочих параметрах блока: $b_i = \{\text{Заголовок}, m'_i, 0\dots0\}$. Заголовок первого блока m'_1 помимо прочего должен содержать

информацию об общей длине сообщения, параметрах кодирования и общем количестве передаваемых блоков;

6. Вычисление контрольной суммы от данных для каждого из блоков и запись результата в конец блока: $b_i = \{\text{Заголовок}, m'_i, \text{CRC}(m'_i)\}$;
7. Наложение кодов коррекции ошибок: $b'_i = ErrC(b_i)$.

Если открытый канал предполагает гарантированную доставку сообщений, седьмой шаг может быть опущен. Полученные на седьмом, или соответственно на шестом шаге блоки подаются на вход алгоритмов шифрования. В криптографической части могут быть реализованы различные режимы, такие как использование только одного секретного ключа, использование сеансовых ключей, использование открытых ключей. Если в системе используется один общий секретный ключ $k \in K$, то шифрование блока b'_i осуществляется простым симметричным алгоритмом $b''_i = E_k(b'_i)$. Если в системе необходимо реализовать использование открытых или сеансовых ключей, то выполняется процедура генерирования случайного сеансового ключа $k \in K$. Сгенерированный сеансовый ключ, совместно с другой необходимой дополнительной информацией формирует новый общий заголовок сообщения. Данный заголовок проходит процедуру наложения кодов коррекции ошибок, так же как и все блоки сообщения. Результат подвергается процедуре зашифрования на открытом ключе адресата или же общем для отправителя и адресата секретном ключе $k' \in K$. Далее все сформированные ранее блоки данных проходят процедуру зашифрования: $b''_i = E_k(b'_i) | i \neq 1, b''_1 = \{E_{k'}(ErrC(\{k, \text{CRC}(k)\})), E_k(b'_1)\}$. На выходе алгоритмов шифрования блоки данных b''_i , по сути, представляют собой битовые строки $y \in Y$ конечной длины с равномерным распределением.

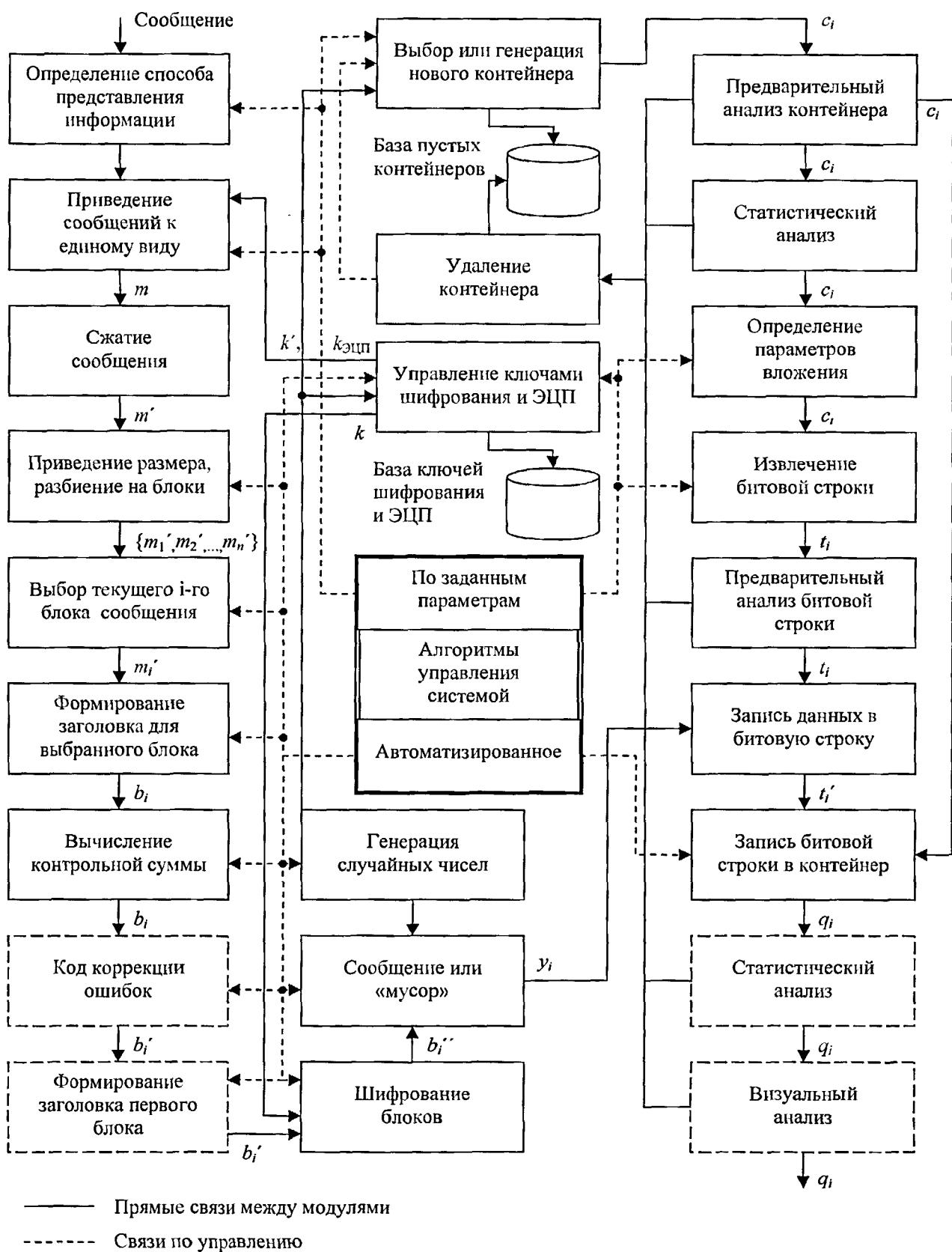


Рисунок 20 – Схема криптостеганографической системы
на стороне отправителя.

Не всегда выбранный случайным образом контейнер является походящим для целей скрытой передачи информации. Так вполне вероятна ситуация, что изначально пустой контейнер одним из вероятностных методов статистического стеганоанализа будет определяться как заполненный. Разумеется, использовать такой контейнер нельзя. По возможности он должен быть автоматически отбракован. Другим фактором, который также может служить причиной отбраковки контейнера является его низкая емкость. Все аналогичные факторы должны быть непременно учтены в конечной системе. Для этого необходимо введение в общую схему системы процедур предварительной проверки, анализа и подготовки контейнеров.

Необходимо также учитывать и то, что в ходе записи сообщений в контейнеры с помощью реальных алгоритмов все же возможно искажение некоторых параметров контейнеров, что может быть выявлено противником при сопоставлении параметров пустых и заполненных контейнеров. Для того чтобы исключить такую возможность необходимо при передаче пустых контейнеров заполнять их каким-либо «мусором», т.е. случайными данными не несущими никакой полезной информации. При этом все статистические характеристики «мусора» должны соответствовать характеристикам передаваемых сообщений. В реальных системах может существовать ненулевая вероятность, что при определенных сочетаниях параметров контейнеров и скрываемых сообщений стегосистема может быть обнаружена одним из известных методов стеганоанализа. Не допустить такой ситуации позволяет автоматический и визуальный контроль стегосообщений непосредственно перед их отправкой адресату. Не менее важным является разовость контейнеров. В случае повторения контейнера противник простым сопоставлением (побитовым сравнением) двух перехваченных им контейнеров может легко обнаружить скрытый канал.

Таким образом, учитывая приведенные выше замечания, использование контейнеров в стеганографической системе так же должно быть хорошо продумано и автоматизировано, как и работа с самими

передаваемыми сообщениями. В представленной на рисунке 20 схеме все отмеченные выше моменты, связанные с выбором и обработкой контейнеров уже учтены. На первом шаге осуществляется случайный выбор контейнера из базы имеющихся и отобранных ранее, но еще не использованных контейнеров. На этом шаге может быть осуществлена и автоматическая генерация нового контейнера. Выбранный контейнер проходит процедуру предварительного анализа, целями которого являются определить пригодность данного контейнера для использования в системе скрытой передачи информации, а так же приблизительно оценить его информационную емкость. Если контейнер не отвечает, каким либо требованиям, он удаляется и вместо него запрашивается новый контейнер. Далее контейнер проходит статистический анализ, в ходе которого устанавливается: соответствуют ли параметры конкретного контейнера общим среднестатистическим параметрам остальных контейнеров. В случае если по какому-либо из параметров будет выявлено отклонение, контейнер отбраковывается и вместо него запрашивается другой контейнер. На данном шаге, к контейнеру также могут быть применены методы статистического стеганоанализа, позволяющие заранее отсечь контейнеры с исходно завышенной вероятностью наличия стегосообщения.

Запись данных в контейнер включает в себя несколько шагов. Сначала по типу и особенностям контейнера определяются параметры вложения. После выполняется процедура извлечения данных. Извлеченные данные представляют собой битовую строку, которая в последствии будет использована для записи данных передаваемого сообщения. Полученная битовая строка анализируется на предмет соответствия предъявляемым к ней алгоритмами согласования требованиям, кроме того, уточняется объем информации, которая может быть записана в данную строку. Если все необходимые требования выполняются, осуществляется процедура записи уже подготовленных и зашифрованных данных в извлеченную из контейнера битовую строку. На заключительном этапе осуществляется непосредственно

само стеганографическое сокрытие данных (битовой строки) в контейнере. Данный этап может также включать в себя дополнительный контроль вносимых искажений, отслеживание корреляционных связей внутри контейнера и статистическую коррекцию.

На заключительном этапе, полученное стегосообщение необходимо проверить на возможность выявления известными отправителю методами стеганоанализа. Если с помощью какого-либо из методов с некоторой вероятностью удастся обнаружить факт наличия посторонней информации в передаваемом контейнере, контейнер отбраковывается и весь процесс повторяется заново, но уже для другого контейнера. Если все тесты пройдены успешно и факт наличия посторонней информации в контейнере установлен не был, исходный контейнер удаляется из базы пустых контейнеров, а заполненный пересыпается адресату.

4.1.3 Общая схема криптостеганографической системы скрытой передачи ЭД на стороне адресата

Схема криптостеганографической системы на стороне адресата [114, 116] представлена на рисунке 21. Согласно предложенному протоколу на первом шаге полученный адресатом контейнер проверяется на целостность и соответствие типу используемых в системе контейнеров. Далее контейнер проверяется на предмет наличия его копии в базе ранее принятых контейнеров. Если идентичный контейнер в ходе проверки не был обнаружен, принятый контейнер добавляется в базу и передается на дальнейшую обработку. Здесь необходимо учитывать, что отправитель строго отслеживает повторяемость контейнеров. Получение одинаковых контейнеров сигнализирует о том, что скрытый канал мог быть подвергнут атаке со стороны злоумышленника.

Процедура извлечения информации из полученного контейнера включается в ряд шагов реализующих преобразования обратные процедуре встраивания данных. Из полученного контейнера стеганографическими алгоритмами извлекается битовая строка. Далее она декодируется, а

полученные в результате данные расшифровываются. Если процесс подготовки данных включал в себя наложение кодов коррекции ошибок, то выполняется процедура снятия этих кодов и выполняется коррекция ошибок. Затем из данных формируются блоки аналогичные блокам при отправке сообщения. После того, как блок данных был успешно восстановлен, проверяется его заголовок и контрольная сумма. Если заголовок содержит ошибки, или же не совпал результат вычисления контрольной суммы блока, считается, что контейнер был либо поврежден, либо изначально не содержал в себе скрытых данных. Если же целостность блока подтвердилась, выполняется процедура извлечения его данных – блока сжатого сообщения. На этом шаге работа с контейнером заканчивается.

Восстановление принятого сообщения осуществляется после того, как успешно приняты все отдельные блоки сжатого сообщения. Принятые блоки конкатенируются согласно их порядковым номерам. Полученные в результате данные распаковываются, проверяется общий заголовок, длина, целостность данных и ЭЦП. В случае успеха, получателю выдается восстановленное скрытое сообщение с подтверждением ЭЦП отправителя.

Возможно построение системы использующей несколько различных стеганографических алгоритмов с возможностью перенастройки параметров их работы. Отметим, что в предложенной схеме, выбор параметров вложения осуществляется на стороне отправителя, без информирования адресата. На стороне получателя такой подход потребует подбора параметров для извлечения информации из принятого контейнера. В данном случае контрольная сумма играет своего рода роль метки, если она совпала, то данные успешно извлечены, если же она не совпала (на схеме «Выход (2)»), то необходимо выбрать иные параметры извлечения информации (на схеме вход «(2)»). Если все возможные варианты перебраны и контрольная сумма ни разу не совпала, то контейнер содержит не скрытое сообщение, а «мусор». Очевидно, что в рамках предложенной схемы данная процедура может быть легко автоматизирована.

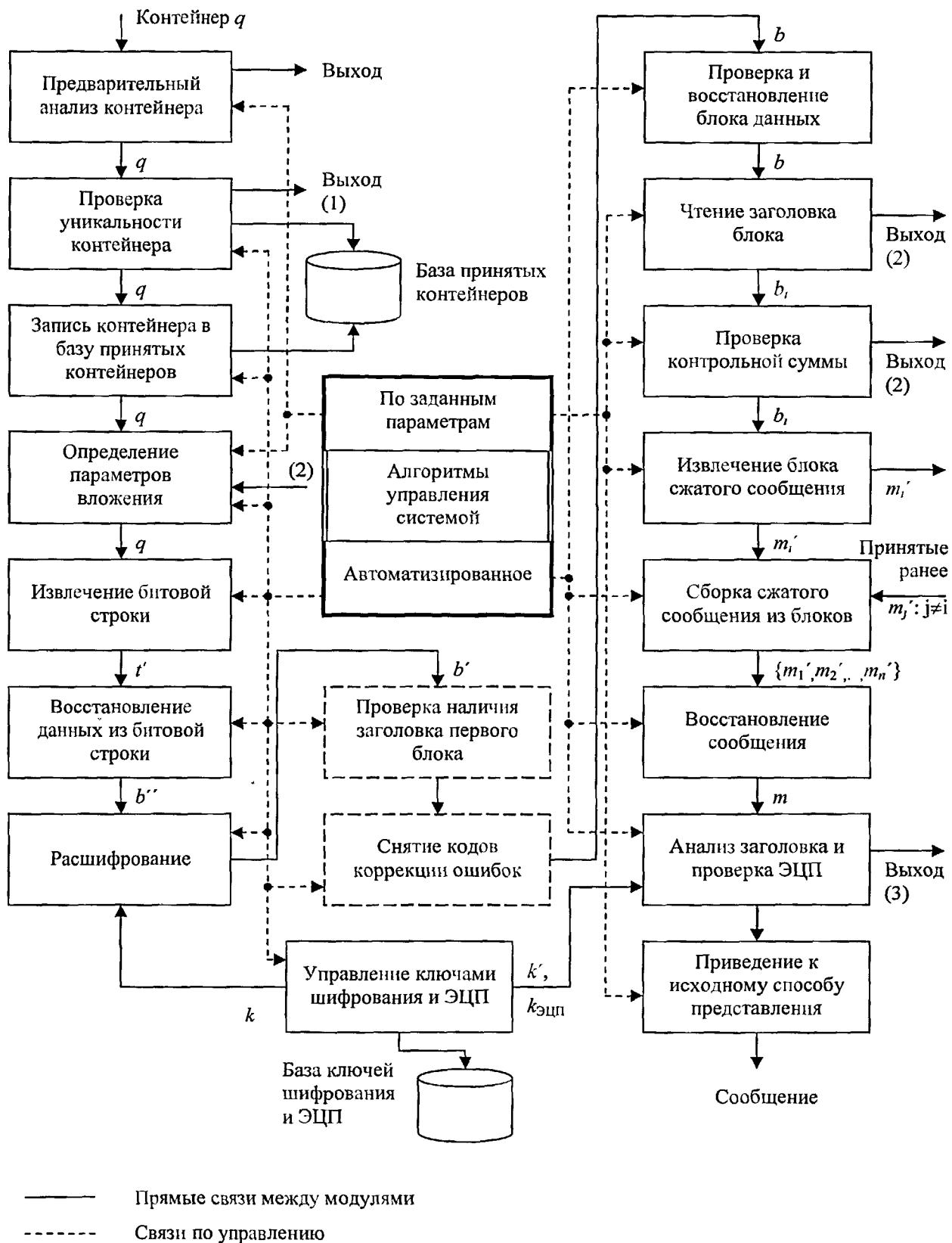


Рисунок 21 – Схема криптостеганографической системы на стороне адресата

4.2 Скрытая маркировка электронных документов в СЭД

Неочевидным применением криптостеганографических систем в электронном документообороте является возможность осуществления скрытой маркировки электронных документов [102, 119, 123].

На настоящее время задача отслеживания перемещения электронных документов как внутри организации, так и за его пределами остается нерешенной. Применение ЭЦП также не решает данной задачи. Основная причина состоит в том, что в отличие от бумажных документов получить идентичную оригиналу копию электронного документа не составляет труда. Для этого достаточно просто скопировать файл с электронным документом. Использование каких-либо обычных способов маркировки электронных документов будет очевидным для пользователей и при передаче документа маркировка может быть удалена.

4.2.1 Многопользовательская система скрытой уникальной маркировки электронных документов

Высокая гибкость криптостеганографических систем позволяет построить систему скрытой маркировки электронных документов. Очевидными преимуществами предлагаемой системы являются прозрачность процедуры маркировки и отсутствие визуально заметных для пользователей признаков. Работая с такой системой, пользователи не будут знать о том, что запрашиваемые ими документы содержат какую либо идентифицирующую их информацию. Рассмотрим общую идею системы скрытой маркировки электронных документов.

На рисунке 22 представлено электронное хранилище документов, реализованное на базе многопользовательской СУБД и системы аутентификации, авторизации и обработки запросов пользователей. В приведенном примере пользователи А и В обращаются к электронному хранилищу с запросом на получение копии одного общего документа. После прохождения процедур аутентификации пользователя системой

осуществляется проверка разрешений по доступу к запрашиваемому электронному документу. Если пользователь обладает необходимыми привилегиями, осуществляется запрос к электронному хранилищу. Полученный по запросу от электронного хранилища документ попадает на промежуточный сервер. На основе уникального ключа пользователя отправившего запрос соответствующего его идентификатору осуществляется маркировка электронного документа. После, маркированный документ передается пользователю. Для разных пользователей А и В один и тот же исходный документ будет иметь различную скрытую маркировку.

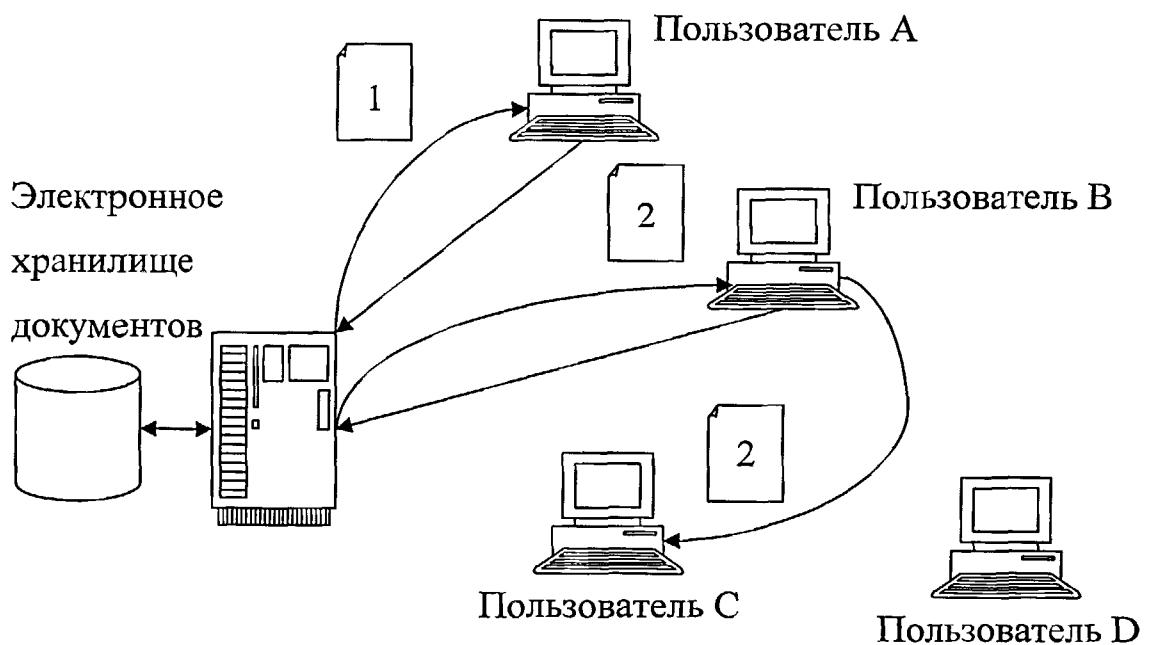


Рисунок 22 – Многопользовательская система скрытой маркировки электронных документов

Маркировка электронных документов позволяет однозначно установить вторичный источник – пользователя, передавшего документ третьей стороне в обход системы учета хранения и перемещения копий электронных документов. Допустим, пользователь В передал правомерно полученный им ранее документ пользователю С. В данном случае

пользователь С в действительности получит не точную копию документа хранящегося в электронном хранилище, а его промаркованную (под пользователя В) копию. Если каким либо образом, например, в ходе очередной проверки у пользователя С будет обнаружен электронный документ, к которому он не имел доступа, скрытая маркировка позволит однозначно установить вторичный источник. В нашем случае роль вторичного источника играет пользователь В, так или иначе причастный к разглашению полученного им электронного документа.

Использование в системе персональных уникальных ключей позволяет избежать ситуации, при которой пользователи А и В сговорившись, могли бы представить пользователя Д. Для каждой пары «документ» – «пользователь» или «документ» – «запрос» формируется уникальная маркировка, зависящая от идентификатора пользователя (уникального ключа), уникальных данных запроса и данных непосредственно маркируемого документа. Без знания секретной составляющей для пользователя Д, даже владея всей информацией об используемой технологии маркировки, пользователя А и В не могут сформировать маркировку документа, идентифицирующую пользователя Д.

Для решения представленной задачи успешно может быть использована криптостеганографическая система, построенная по базовой трехкомпонентной схеме. В данном случае криптографические алгоритмы отвечают за работу с ключевой информацией, алгоритмы согласования формируют данные ЦЗВЗ, а стеганографические алгоритмы обеспечивают скрытое встраивание данных ЦВЗ в электронные документы.

4.2.2 Модель системы скрытой маркировки ЭД на базе клиент-серверной архитектуры

Рассмотрим алгоритм взаимодействия пользователя с электронным хранилищем документов, представленный на рисунке 23. При подключении пользователя к системе хранения электронных документов выполняется процедура аутентификации, позволяющая идентифицировать пользователя и назначить ему уникальный цифровой идентификатор.

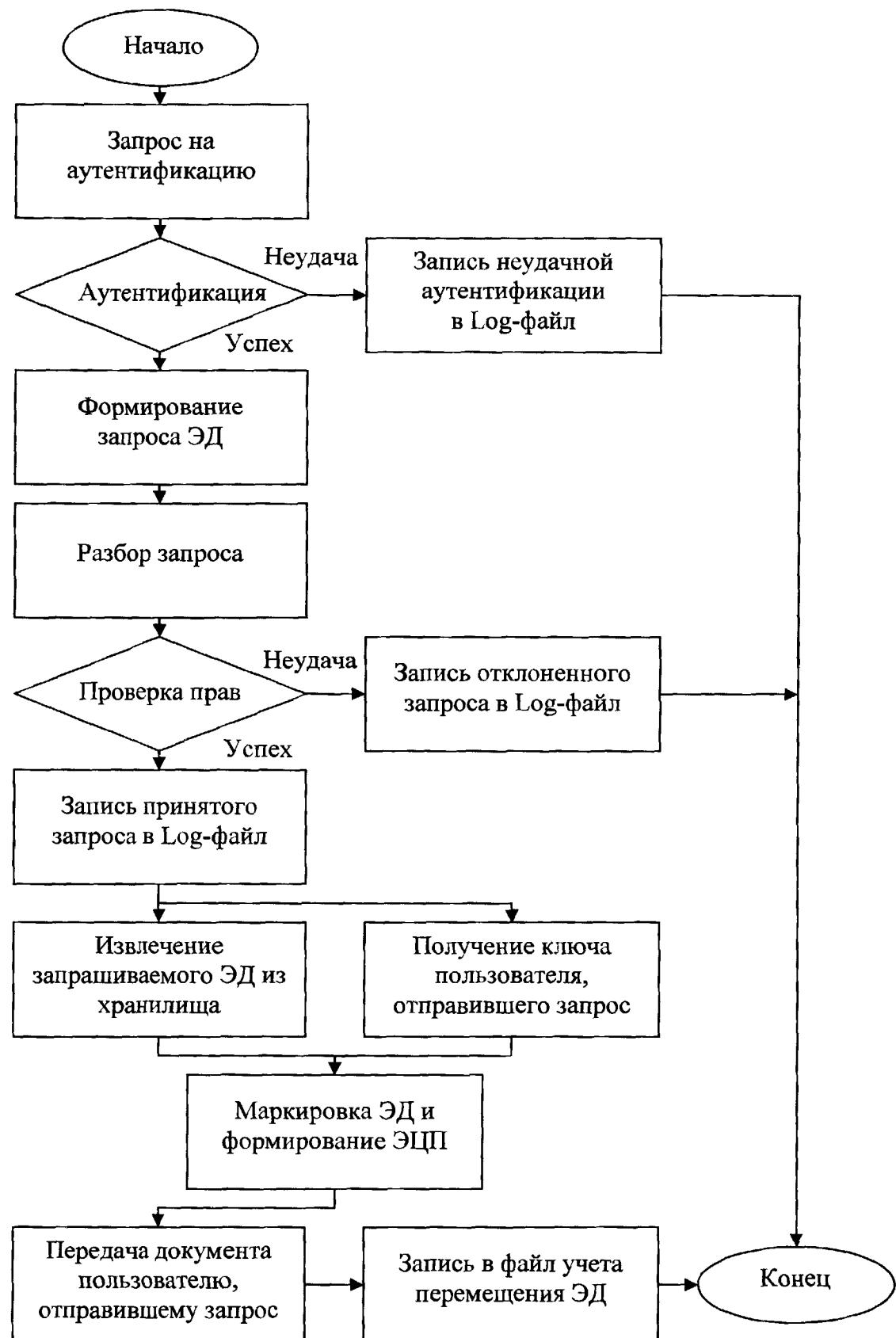


Рисунок 23 – Алгоритм взаимодействия пользователя с системой хранения ЭД

Информация о результате выполнения процедуры аутентификации пользователя вне зависимости от исхода последней записывается в журнал системы. После успешного прохождения процедуры аутентификации, пользователь может направить запрос на получение копии любого необходимого ему электронного документа, хранящегося в системе.

При получении запроса от пользователя системы выполняется процедура проверки прав доступа. Целью проверки является определить возможность предоставления пользователю запрашиваемого им документа. Если пользователь обладает необходимыми правами доступа на запрашиваемый электронный документ, системой осуществляется запрос к хранилищу на получение копии данного электронного документа. Одновременно выполняется процедура получения (формирования) уникальной ключевой последовательности для отправившего запрос пользователя согласно его уникальному цифровому идентификатору (UID). Полученный из хранилища электронный документ подвергается криптографическому преобразованию с использованием данной ключевой последовательности. Зашифрованные данные подаются на вход алгоритмов согласования выполняющих процедуру формирования данных уникального ЦВЗ. Далее стеганографическими алгоритмами осуществляется встраивание данных ЦВЗ в исходный документ.

В зависимости от алгоритмов встраивания ЦВЗ маркировка текстового документа может нарушить его целостность, но при этом суть и содержание документа, его соответствие оригиналу сохраняются. Нарушение целостности документа влечет появление несоответствия ЭЦП. В таких случаях требуется формирование второй подписи (ЭЦП) для копии документа. Для этой цели может быть использован секретный ключ ЭЦП для копий предоставляемых системой документов, оригиналы которых хранятся в системе с исходной подписью. Исходная ЭЦП проверяется системой единожды при поступлении документа в электронное хранилище, и далее

система выступает в роли доверенного посредника, гарантирующего достоверность и целостность предоставляемых ей копий документов.

На заключительном шаге, будучи промаркированной и подписанной, копия электронного документа передается пользователю, отправившему запрос. Информация о предоставлении копии документа с необходимыми для последующей проверки данными записывается в файл учета хранения и перемещения электронных документов.

Соответствующая описанной идеологии схема взаимодействия пользователя с электронным хранилищем документов на базе клиент-серверной архитектуры представлена на рисунке 24. На представленной схеме выделены следующие элементы: клиентская часть/терминал (1), модуль встраивания ЦВЗ и формирования ЭЦП копии документа (2), серверная часть – базовая СЭД (3), электронное хранилище/база данных электронных документов (4), модуль генерирования, хранения и предоставления ключей по заданному идентификатору пользователя (5), представление ЭД пользователю (6), модуль верификации ЦВЗ (7).

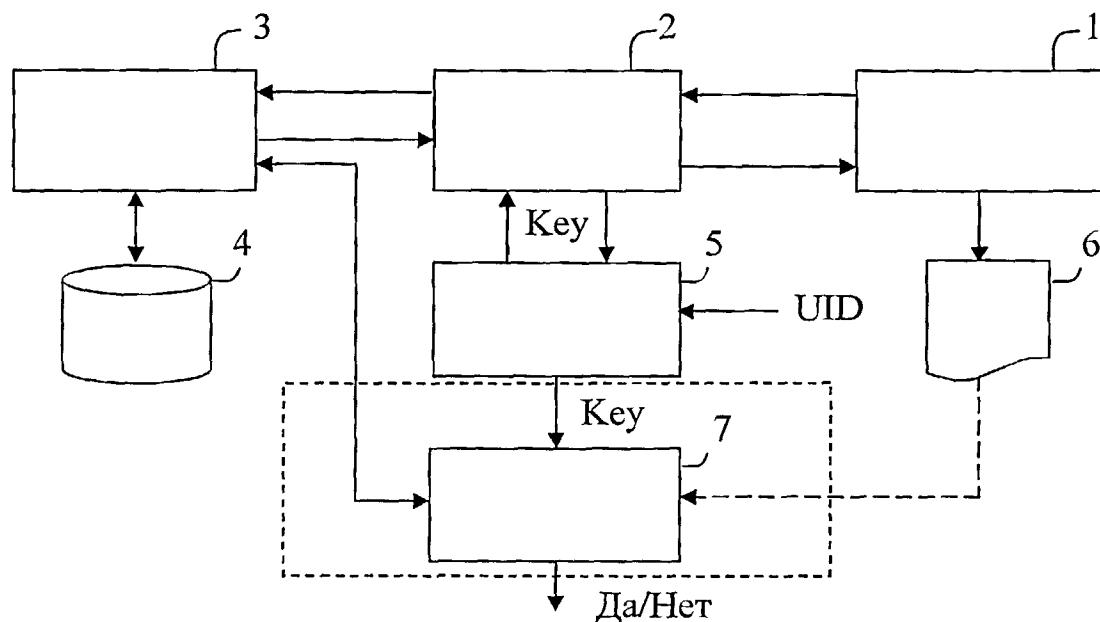


Рисунок 24 – Схема прозрачного взаимодействия клиентского терминала с электронным хранилищем через блок встраивания ЦВЗ

4.2.3 Методы встраивания данных ЦВЗ в системе маркировки ЭД

Маркировка текстовых документов может быть построена на базе различных стеганографических алгоритмов. Для целей скрытой маркировки могут эффективно использоваться методы на основе смещения строк и изменения межстрочного интервала, смещения слов и символов в строке, использования символов табуляции и пробела, добавляемых в начале, конце и середине текстовых строк. В большинстве современных, даже очень сложных редакторов текста, замена знака табуляции стоящего в начале абзаца пробельными символами будет неочевидна. Не более заметным является и добавление пробельных символов в конце абзаца. Оба предложенных варианта дают возможность максимально незаметно встроить один бит данных ЦВЗ на абзац.

Применение методов на основе изменения межстрочного интервала в действительности является более сложным, так как в большинстве современных текстовых процессорах межстрочный интервал настраивается только для абзаца, а не отдельной текстовой строки. В результате объемы скрываемой информации будут невелики. Кроме того, практически любая попытка изменить форматирование текста приведет к разрушению ЦВЗ. Так же такой подход, как и подход, основанный на относительном смещении символов в строке, не может быть применен для неформатированных текстовых данных (текстовые файлы формата *.txt). В тоже время метод смещения строк является наиболее эффективным для встраивания ЦВЗ, с точки зрения стойкости, в случае вывода ЭД на печать.

Наиболее стойкими к преднамеренному удалению данных ЦВЗ являются синтаксические и семантические стеганографические методы. В действительности, не имея на руках второй копии ЭД со встроенным ЦВЗ другого пользователя или же исходного ЭД, выявить такие методы практически невозможно. А попытка удаления данных ЦВЗ может привести к нарушению сути и содержания ЭД. Собственно этот же момент является и общим недостатком синтаксических и семантических методов при их

применении к ЭД. Именно по этой причине, применение данных методов разумно ограничить внешним документооборотом и материалами для открытой публикации. Снизить вероятность искажения сути ЭД при автоматизированной системе встраивания ЦВЗ позволяет более точное задание соответствующих правил и таблиц замен. Отметим также, что средний объем встраиваемых данных в этом случае может составить не один, а несколько бит на абзац. В тоже время, данные методы являются и одними из наиболее сложных в реализации.

В данной работе, в качестве достаточно простой в реализации и предполагающей большую информационную емкость альтернативы, предлагается метод, основанный на замене символов с идентичными по начертанию глифами (графическими образами символов). Все современные кодировки текстовых символов, согласно стандарту POSIX ISO/IEC 9945 содержат в себе базовый набор переносимых символов. Этот набор символов выделен особо, так как позволяет обеспечить совместимость всех современных операционных систем и переносимость исходных текстов программ на различных языках программирования. Набор переносимых символов включает в себя все символы латинского (английского) алфавита в малом и большом регистре, цифры, орфографические символы, знаки табуляции и пробела, а также ряд специальных символов. Таким образом, в любой локализованной под возможность использования русского языка кодировке уже изначально присутствуют и символы латинского алфавита. Как известно сравнительно большое количество символов в русском и латинском алфавитах имеют идентичное начертание. Именно этот факт позволяет, без каких либо визуально заметных искажений, подменять в текстовом документе символы одного алфавита другими. Практически во всех современных русифицированных кодировках можно установить соответствие символов 'A', 'B', 'C' английского алфавита символам 'А', 'В', 'С' русского алфавита. В приложении В представлены таблицы символов с одинаковым начертанием для различных, наиболее популярных кодировок.

Например, из таблицы В.1 видно, что русская буква «К» может быть представлена символом с кодом 138, принадлежащим русскому алфавиту, и символом с кодом 75, принадлежащим английскому алфавиту. Вследствие того, что символы имеют одинаковое начертание визуальное представление строк, в которых буква «К» представлена символом 'K' с кодом 138 будет идентичным визуальному представлению аналогичных строк, в которых буква «К» представлена символом с кодом 75. Аналогично и с другими символами. Более того, данный факт сохраняется и для большинства из так называемых TrueType шрифтов, которые используются в профессиональных и близких к профессиональным текстовых редакторах.

Как можно заключить из анализа представленных в приложении В таблиц, и учитывая также частоты встречаемости соответствующих символов в текстовых документах, предложенный метод обладает наибольшей информационной емкостью. А, учитывая полное отсутствие визуально заметных искажений, метод позволяет с абсолютной точностью сохранить не только суть и содержание маркируемых текстовых документов, но также и их структуру и форматирование. Однако, разумеется, данный метод не лишен и определенных недостатков ограничивающих область его применения. В частности серьезным ограничением является использование для просмотра и редактирования документов современных текстовых редакторов с автоматической проверкой орфографии. Данное ограничение касается просмотра и редактирования документа пользователями, а не процесса его создания автором документа.

Для целей встраивания данных ЦВЗ в электронные документы может быть использован любой из описанных выше методов. Выбор конкретного метода зависит от ограничений и конечных требований к системе в целом. Если требуется обеспечить максимальную стойкость маркировки к преднамеренному разрушению, то наиболее эффективными будут достаточно сложные синтаксические и семантические стеганографические методы. Если требуется обеспечить возможность маркировки коротких текстовых

документов, то выбор будет в пользу методов замены текстовых символов. Также эти методы весьма эффективны при предоставлении электронных документов в режиме только для чтения, например в формате Adobe Acrobat (*.pdf), XML или HTML. Если предполагается вывод электронных документов на печать, то предпочтение следует отдать методам на основе смещения текстовых строк. Таким образом, еще раз отметим, что выбор конкретного метода встраивания данных ЦВЗ определяется специфичными особенностями организуемой системы предоставления копий ЭД.

4.2.4 Формирование данных ЦВЗ

В предлагаемом решении цифровой водяной знак формируется непосредственно от данных электронного документа и уникальной ключевой последовательности пользователя. В связи с чем, в процессе маркировки ЭД ведется работа с текстовыми строками документа. Возможны два подхода, в первом случае весь текст рассматривается как единый контейнер (одна строка), во втором ЭД алгоритмами согласования разбивается на множество отдельных, независимых текстовых строк.

Первый подход предполагает поточную обработку текста, при которой формирование очередного бита данных ЦВЗ учитывается весь предшествующий текст. Такой подход достаточно прост в реализации и позволяет провести маркировку ЭД в один проход. Однако данному подходу присущ и один весьма существенный недостаток. Проверить результат маркировки можно только для неизмененных документов. Если хотя бы одна строка ЭД была изменена, целостность ЦВЗ будет нарушена. И все же данный подход может быть весьма эффективен в связке со стеганографическими методами на основе смещения текстовых строк.

Второй подход, основанный на разбиении исходного текстового документа на множество независимых текстовых строк, представляется более надежным. В данном случае проверка ЦВЗ возможна не только для модифицированных ЭД, но в ряде случаев и для отдельных абзацев. В зависимости от выбранного стеганографического метода единицей разбиения

могут быть абзацы, строки, несколько идущих подряд слов и отдельные слова. Так, абзацы необходимы для синтаксических методов, строки для семантических, а отдельные слова для методов замены символов с одинаковым начертанием.

Для определенности рассмотрим случай использования метода замены символов с одинаковым начертанием для группы слов. Пусть исходный текст алгоритмами согласования разбивается на группы из трех слов идущих последовательно в исходном тексте. Вероятность возможности внесения одного бита данных ЦВЗ в данном случае, учитывая среднюю длину слова в русском языке 5.28 символа, а также частоты появления однограмм в тексте [144], составляет более 0.999. Разумеется, существует вероятность встречи в исходном тексте группы слов, встраивание данных ЦВЗ в которую будет невозможно. В целях избежания таких ситуаций, непосредственно перед вычислением очередного бита ЦВЗ для данной группы слов, проводится анализ на возможность последующего встраивания данного бита в исходную группу. Так если группа слов не содержит ни одного символа из таблицы замен (таблицы символов двух и более алфавитов с одинаковым начертанием) она пропускается.

Непосредственно в процессе маркировки, для каждой выбранной группы слов содержащей символы из множества пар символов с одинаковым начертанием, вычисляют значение однобитовой хэш-функции. Структура возможной однобитовой хэш-функции представлена на рисунке 25 в контексте совместного применения с предварительным криптографическим преобразованием. На вход подают выбранную для маркировки текстовую строку (группу слов) и ключевую последовательность, соответствующую пользователю, запросившему документ. Текстовая строка и ключевая последовательность поступают на вход функции криптографического преобразования. В данном примере, на вход байтового сумматора. Байтовый сумматор представлен функцией битового сложения по модулю два (операция XOR) при достаточной длине ключевой последовательности или

функцией поточного шифрования текстовой строки. В реализации байтового сумматора с использованием функции битового сложения по модулю два, каждый бит очередного байта текстовой последовательности складывается по модулю два с соответствующим битом ключевой последовательности. Биты с выхода байтового сумматора последовательно поступают на вход регистра сдвига R_{\rightarrow} с линейной обратной связью [159]. Выбор конкретного примитивного многочлена для регистра сдвига зависит от особенностей реализации, в нашем случае выбран многочлен: $x^8 + x^4 + x^3 + x^2 + 1$. Биты с выхода регистра сдвига поступают во временный однобайтовый буфер WB, реализующий хранение последних восьми поступивших в него битов. По завершении циклической обработки всех байтов текстовой строки, последние восемь битов оставшиеся в однобайтовом буфере подают на вход конечного блока Σ реализующего свертку восьмибитного вектора в один бит данных. Конечная свертка реализована функцией сложения по модулю два всех битов вектора. В результате на выходе получают один бит данных цифровой подписи, значение которого зависит от всех символов текстовой строки и данных ключевой последовательности.

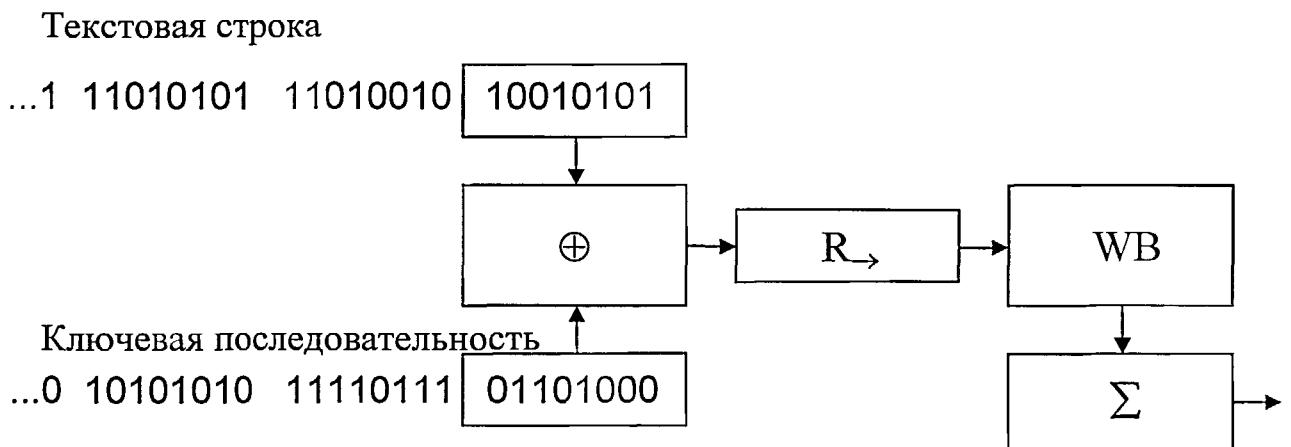


Рисунок 25 – Схема формирования данных ЦВЗ с применением однобитовой хэш-функции

После вычисления значения очередного бита данных ЦВЗ осуществляется встраивание бита в обрабатываемую текстовую строку. Встраивание данных в текстовую строку осуществляется путем замены одного из символов строки, соответствующим символом из таблицы замен. Замена осуществляется согласно следующему правилу:

Если значением бита на выходе однобитовой хэш-функции является 1, произвольный символ строки, принадлежащий множеству пар символов с одинаковым начертанием, заменяют символом идентичным по начертанию, но относящимся к другому алфавиту, если результатом вычисления хэш-функции является 0, то замену не осуществляют.

В том случае если текстовая строка обладает достаточно большой длиной и содержит в себе большое количество алфавитных символов принадлежащих множеству пар символов с одинаковым начертанием, для встраивания данных ЦВЗ может быть выбран не один, а несколько символов. В этом случае запись данных ЦВЗ осуществляется рекурсивно с последовательным переходом от текущего символа строки, к следующему символу строки входящим в таблицу замен. Порядок выбора символов в строке не принципиален и может варьироваться в зависимости от реализации. Так, в частности последовательно могут быть выбраны первый, третий, пятый и так далее символы, принадлежащий множеству пар символов с одинаковым начертанием, т.е. каждый нечетный символ. В любом случае процедуру вычисления однобитовой хэш-функции осуществляют для каждого символа отдельно. При этом для первого символа хэш-функцию вычисляют от исходной строки, а для каждого последующего от модифицированной на предыдущем шаге. Замену символов осуществляют согласно представленному выше правилу.

Если в тексте исходного документа в перемешку встречаются слова из разных языков (записанные с различными алфавитами) то является необходимым ввести дополнительные правила, которое в последствии позволили бы определить исходный алфавит каждого слова. Так в качестве

дополнительного правила можно ввести пропуск слов, состоящих из одного символа. Данное правило позволит избежать путаницы с неопределенным артиклем “а” и союзом «а». Другим правилом может быть непременное наличие в изменяемых словах алфавитных символов, не входящих в используемую таблицу замен. Альтернативным правилом может быть требование сохранения неизменным второго символа каждого слова. Все эти правила позволяют сделать процедуру маркировки текстовых документов обратимой. В свою очередь, это дает возможность проверки ЦВЗ даже без знания исходного документа.

4.2.5 Проверка ЦВЗ в электронных документах

Отличительной особенностью предложенного решения является возможность проверки скрытой маркировки ЭД даже по отдельным разрозненным фрагментам исходного документа. Проверка ЦВЗ может быть осуществлена не по целому документу, а по ограниченному не упорядоченному множеству строк. Серьезным ограничением является лишь исходная упорядоченность слов в пределах предложения.

Процедура проверки ЦВЗ начинается с того, что электронный документ разбивается на множество подстрок состоящих из трех последовательно идущих в исходном тексте слов. Из множества имеющихся подстрок случайным образом выбирают к попарно различных подстрок содержащих символы из множества пар символов с одинаковым начертанием. Строки, которые могли быть отбракованы при записи данных ЦВЗ согласно дополнительным правилам, пропускаются. Количество строк k выбранных для проверки зависит от числа символов $n \geq N$ относящихся к множеству символов с одинаковым начертанием, которые могли быть изменены в процессе записи данных ЦВЗ. Параметр N соответствует минимальному числу проверок и определяется исходя из необходимого уровня достоверности результатов проверки ЦВЗ. Другими словами параметр N задается по значению максимальной допустимой ошибки. Зависимость

вероятности ошибки от числа проведенных проверок представлена на графике – рисунок 26.

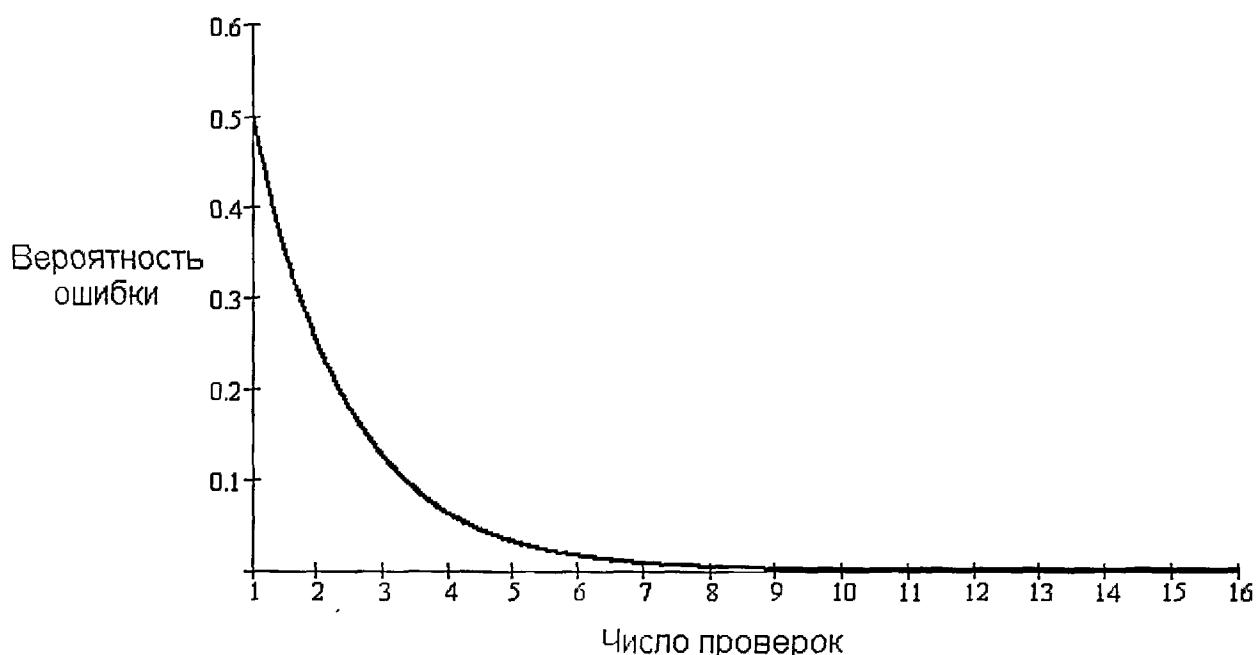


Рисунок 26 – Зависимость вероятности ошибки проверки ЦВЗ в зависимости от числа независимых проверок

Из представленного графика видно, что для обеспечения достоверности результатов проверки со значением более 0.99 достаточно проведения 8 независимых проверок отдельных символов. Чем заданное число проверок больше, тем больше достоверность полученных результатов. По результатам 16 независимых проверок вероятность ошибки составит уже менее 0.0001. Для обеспечения высокой степени достоверности результата проверки предлагается проведение порядка 32-256 независимых проверок, со случайным или близким к случайному порядком выбора подстрок. Это позволит свести к минимуму вероятность ошибки в случае близких ключевых последовательностей, вероятность появления которых существенно увеличивается с увеличением количества пользователей.

Алгоритм проверки ЦВЗ по множеству строк представлен на рисунке 27. В представленном алгоритме величина t характеризует общее количество символов использованных в процедуре маркировки текущей строки.

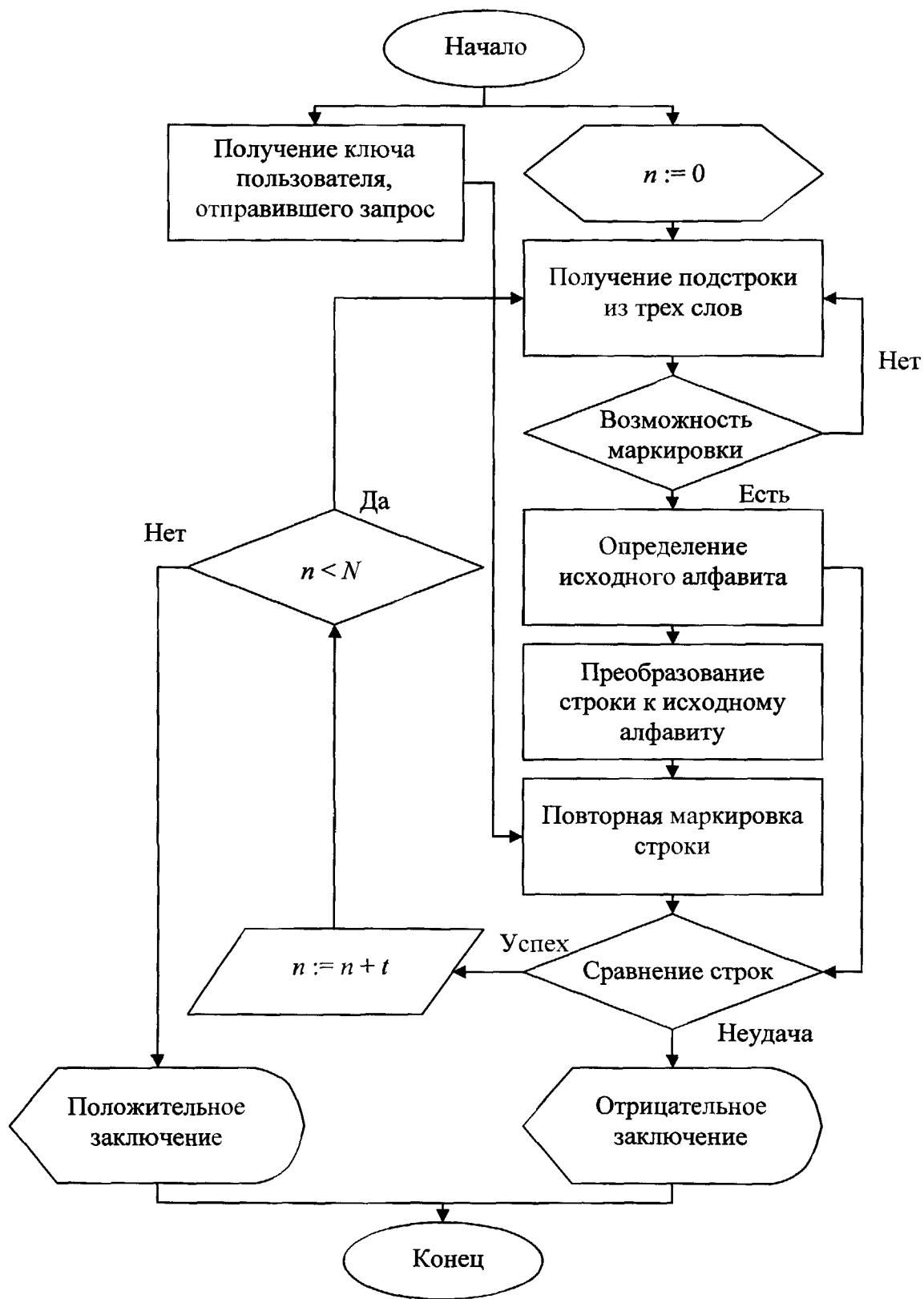


Рисунок 27 – Алгоритм проверки ЦВЗ

Согласно представленному алгоритму, на первом шаге для каждой строки из выбранного множества из k строк (подстрок из трех

последовательно идущих слов текста) определяют алфавит исходной строки. После определения алфавита исходной строки, на основе имеющейся строки формируют копию исходной строки путем преобразования всех символов к исходному алфавиту. По заданному идентификатору пользователя получают соответствующую ключевую последовательности и выполняют процедуру маркировки полученной на предыдущем шаге текстовой строки. Далее осуществляют сравнение полученной проверочной строки и строки из множества выбранных для проверки текстовых строк. Сравнение строк осуществляют по символам, выбираемым в ходе маркировки строк. Если все символы строки, выбранные в процессе маркировки, совпали с символами проверяемой строки, то счетчик проверок увеличивается на количество проверенных символов – t . Совпадшими считаются символы, относящиеся к одному алфавиту, т.е. имеющие один и тот же двоичный код. Если изначально, предполагается, что модификация электронного документа невозможна, то в случае появления первого же не совпадшего символа проверка может быть остановлена с выдачей отрицательного результата.

Возможен также случай раздельного подсчета совпавших S_c и не совпавших S_n символов. Если в итоге $S_n > S_c$, то результат проверки считается отрицательным. Если $S_n \leq S_c$, то результат проверки определяют по формуле: $D_{nc} = 1 - 2^{S_n}/2^{S_c}$. Если $D_{nc} > D$ (заданного уровня достоверности), то проверку считают прошедшей с положительным результатом. Такой подход позволяет эффективно противостоять возможному случайному или преднамеренному изменению отдельных строк электронного документа. В этом случае общее количество проверок также невелико. Сама же возможность проверки цифрового водяного знака за достаточно малое число шагов позволяет реализовать процедуру установления пользователя, чьи действия привели к несанкционированной передаче ЭД. С этой целью процедуру проверки последовательно повторяют для всех пользователей имевших доступ к данному ЭД и определяют максимальное значение D_{nc} .

4.3 Выводы по главе

1. Рассмотрена обобщенная архитектура многопользовательской распределенной системы скрытого электронного документооборота и предложена концепция модульного построения подобных систем на базе модели криптостеганографической системы связи. Детально проработаны основные моменты согласования различных модулей, протоколы взаимодействия, принципы, архитектурные и схематические решения. Приведены конкретные рекомендации по построению таких систем.

2. Показана возможность построения систем скрытой маркировки электронных документов на базе модели криптостеганографической системы связи и технологии цифровых водяных знаков. Разработаны методы и алгоритмы скрытой маркировки и проверки маркировки электронных документов, позволяющие отслеживать перемещение электронных документов, а также локализовать и выявлять каналы утечки информации. Представлена обобщенная архитектуры системы скрытой маркировки на базе многопользовательской распределенной СЭД. Проработаны алгоритмы взаимодействия пользователей с системой скрытой уникальной маркировки.

ЗАКЛЮЧЕНИЕ

Диссертационная работа посвящена разработке методов и средств защиты информации в системах электронного документооборота на базе использования современных технологий стеганографии, криптографии, цифровых водяных знаков и электронной цифровой подписи. Проведенные исследования и результаты работы могут быть использованы в качестве теоретической и практической базы при разработке новых и модернизации уже существующих программных средств защиты информации в системах электронного документооборота и скрытой передачи информации.

1. Проведен анализ существующих средств защиты информации в системах электронного документооборота и скрытой передачи электронных документов. Особое внимание уделено таким направлениям в защите информации, как стеганография и цифровые водяные знаки, так как данные направления предоставляют принципиально новые возможности защиты электронных документов. Исследованы уязвимости существующих систем скрытой передачи электронных документов, по результатам анализа уязвимостей введена классификация атак на системы скрытого электронного документооборота. Обоснована необходимость разработки нового подхода к построению средств защиты электронного документооборота на базе совместного использования методов и средств криптографической и стеганографической защиты информации.

2. Проведен анализ известных методов стеганоанализа на предмет границ их применимости и возможностей по противодействию анализу. На основе проведенного анализа предложена методика практической оценки защищенности информации в системах скрытой передачи электронных документов, определены соответствующие критерии оценки.

3. На основе теории совершенных стеганографических систем доказано существование совершенных стеганографических систем связи использующих в качестве контейнеров потенциально бесконечные битовые

строки, не отвечающие требованиям случайности и содержащие длинные серии нулей и единиц. Основанием для выбора таких строк в качестве контейнеров послужил тот факт, что они наиболее характерны для большинства типов контейнеров и известных стеганографических методов.

4. Введено понятие, предложена модель и сформулированы основные принципы построения крипстеганографических средств защиты информации, базирующихся на принципе совместного применения методов криптографии и стеганографии. В целях обеспечения корректного совмещения компонентов системы, на базе предложенной совершенной стеганографической системы связи на битовых строках содержащих длинные серии нулей и единиц, разработаны, построены и реализованы алгоритмы согласования. Разработана базовая архитектура системы, обладающая большой гибкостью и позволяющая строить системы скрытого электронного документооборота с использованием симметричных и открытых ключей. Показано, что в случае выполнения ряда требований к криптографическим и стеганографическим методам и алгоритмам, крипстеганографическая система связи будет обладать стойкостью к атакам со стороны пассивного противника эквивалентной стойкости ее криптографической части.

5. Определены требования к стеганографическим алгоритмам защиты информации в современных условиях функционирования. Разработаны новые стеганографические методы защиты информации, ориентированные на использование в крипстеганографических системах скрытой передачи электронных документов с использованием открытых каналов связи, отвечающие указанным требованиям. Базовыми принципами построения новых стеганографических методов были выбраны принцип использования для целей записи скрываемой информации наиболее трудно прогнозируемых составляющих мультимедиа контейнеров и принцип осуществления записи информации путем замены, изначально существующей в контейнере, информации данными скрываемого сообщения. Полученные в результате

стеганографические методы обладают высокой стойкостью ко всем известным методам стеганоанализа.

6. Предложена обобщенная архитектура многопользовательской распределенной криптостеганографической системы скрытого электронного документооборота. Детально проработаны принципы построения, протоколы и алгоритмы взаимодействия, а также схематические решения.

7. Впервые поставлена и решена задача скрытой маркировки электронных документов, созданы и конструктивно проработаны методы и алгоритмы маркировки и проверки маркировки текстовых документов. Разработанные методы и алгоритмы, позволяют отслеживать перемещение электронных документов, а также локализовать и выявлять каналы утечки информации. Разработана базовая архитектура системы скрытой маркировки ориентированная на многопользовательские электронные хранилища электронных документов. Проработаны алгоритмы взаимодействия пользователей с системой скрытой уникальной маркировки, разработана модель и схема системы на базе многопользовательской СЭД.

СПИСОК ЛИТЕРАТУРЫ

1. Adelson E. H., Digital Signal Encoding and Decoding Apparatus, US Patent, Patent Number: 4 939 515, Jul. 3, 1990.
2. Ahsan K., Covert Channel Analysis and Data Hiding in TCP/IP, M.A.Sc. thesis, Dept. of Electrical and Computer Engineering, University of Toronto, August 2002.
3. Baugher M., McGrew D., Naslund M., Carrara E., Norrman K., The secure real-time transport protocol (srtp), RFC 3711, Internet Society (IETF), March 2004.
4. Bender W., Gruhl D., Morimoto N., Lu A., Techniques for data hiding, IBM Systems Journal, vol. 35, nos. 3&4, pp. 313-336, 1996.
5. Bennett K., Linguistic Steganography: survey, analysis, and robustness concerns for hiding information in text, Center for Education and Research in Information Assurance and Security, CERIAS Tech Report 2004-13, 30 p.
6. Bhattacharjya A. K., Ancin H., Data Embedding in Text for a Copier System, in Proceedings of the ICIP, 2, 1999, pp. 245-249.
7. Bolshakov I. A., A method of linguistic steganography based on collocational-verified synonymy, Information Hiding: 6th International Workshop, Lecture Notes in Computer Science 3200, Springer, May 2004, pp. 180–191.
8. Bolshakov I. A., Gelbukh A., Synonymous paraphrasing using wordnet and internet, Natural Language Processing and Information Systems: 9th International Conference on Applications of Natural Language to Information Systems, NLDB 2004, Lecture Notes in Computer Science, Springer, June 2004, vol. 3136, pp. 312–323.
9. Brassil J., Low S., Maxemchuk N., O'Gorman L., Document marking and identification using both line and word shifting, Technical report, AT&T Bell Laboratories, 1994, pp. 853-860.
10. Calvo H., Bolshakov I. A., Using selectional preferences for extending a synonymous paraphrasing method in steganography, Avances en Ciencias de

- la Computacion e Ingenieria de Computo - CIC'2004: XIII Congreso Internacional de Computacion, October 2004, pp. 231–242.
11. Cedric T.M.M., Adi R. W., McLoughlin I., Data concealment in audio using a nonlinear frequency distribution of PRBS coded data and frequency-domain lsb insertion, In Proceeding: IEEE Region 10 International Conference on Electrical and Electronic Technology, Kuala Lumpur, Malaysia, vol. 1, 2000, pp. 275–278.
 12. Cha S.D., Park G.H., Lee H.K., A Solution to the Image Downgrading Problem, ACSAC, 1995, pp. 108-112.
 13. Chang K., Deng R. H., Feng B., Lee S., Kim H., Lim J., On Security Notions for Steganalysis, ICISC 2004, LNCS 3506, 2005, pp. 440–454.
 14. Chae J. J., Manjunath B. S., Data hiding in Video, Proceedings of 6th IEEE International Conference on Image Processing (ICIP'99), Volume 1, 1999, pp. 311-315.
 15. Christian Cachin, An Information-Theoretic Model for Steganography, In Proceeding of 2nd Workshop on Information Hiding (D. Aucsmith, ed.), Lecture Notes in Computer Science, Springer, 1525, pp. 306-318, 1998.
 16. Christian Cachin, An Information-Theoretic Model for Steganography, Information and Computation, 192(1):41-56, July 2004.
 17. Chotikakamthorn N., Document Image Data hiding Technique Using Character Spacing Width Sequence Coding, ICIP99, 1999, II:250-254.
 18. Costa M., Writing on dirty paper, IEEE Transactions on Information Theory, v.29(3), pp. 439- 441, 1983.
 19. Cox I. J., Public watermarks and resistance to tampering, International Conference on Image Processing (ICIP'97), Santa Barbara, California, U.S.A., 26–29 Oct. 1997, IEEE. ISBN 0-8186-8183-7.
 20. Cox I.J., Miller M.L., McKellips A.L., Watermarking as communication with side information, Proc. IEEE. v87. pp. 1127-1141.

21. Craver S., On Public-Key Steganography in the Presence of an Active Warden. in Information Hiding II, Springer Lecture Notes in Computer Science v 1525, pp. 355-368, April 1996.
22. Cvejic N., Seppanen T., Increasing Robustness of LSB Audio Steganography by Reduced Distortion LSB Coding, Journal of Universal Computer Science, vol. 11, no. 1 (2005), 56-65.
23. Dickman S. D., An Overview of Steganography, James Madison University, Department of Computer Science, Tech Reports, JMU-INFOSEC-TR-2007-002, July 2007, 10 p.
24. Diffie W., Hellman M., New Directions in Cryptography, IEEE Transactions on Information Theory, v. IT-22, n. 6, Nov 1976, pp. 644-654.
25. Fabien A. P. Petitcolas, Ross J. Anderson, Markus G. Kuhn, Information Hiding – A Survey, Proceedings of the IEEE, special issue on protection of multimedia content, 87(7):1062-1078, July 1999.
26. Franz E., Jerichow A., Möller S., Pfitzmann A., Stierand I., Computer based steganography: How it works and why therefore any restrictions on cryptography are nonsense, at best, Information Hiding, First International Workshop Cambridge, U.K., May 30 – June 1, 1996 Proceedings, Lecture Notes in Computer Science, Springer Berlin, Volume 1174, 1996, pp. 7-21.
27. Fridrich J., Goljan M., Rui Du, Reliable Detection of LSB Steganography in Grayscale and Color Images, Proc. of the ACM Workshop on Multimedia and Security, Ottawa, Canada, October 5, 2001, pp. 27-30.
28. Gang L., Akansu A. N., Ramkumar M., MP3 resistant oblivious Steganography, Proceedings of the Acoustics, Speech, and Signal Processing 2001, on IEEE International Conference - Volume 03, pp.1365-1368, 2001.
29. Girod B., The information theoretical significance of spatial and temporal masking in video signals, Proceedings of the SPIE Symposium on Electronic Imaging, vol. 1077, 1989, pp. 178-187.
30. Goldreich O., A note on computational indistinguishability, Information Processing Letters, v. 34, 1990, 277-281.

31. Gopalan K., Wenndt S., Audio Steganography for Covert Data Transmission by Imperceptible Tone Insertion, Proc. the IASTED International Conference on Communication Systems and Applications (CSA 2004), Banff, Canada, July, 2004, pp. 049-053.
32. Gruhl D., Lu A., Bender W., Echo hiding, Information Hiding, First International Workshop Cambridge, U.K., May 30 – June 1, 1996 Proceedings, Lecture Notes in Computer Science, Springer Berlin, Volume 1174, 1996, pp. 295-315.
33. Gutub A.A., Fattani M.M., A Novel Arabic Text Steganography Method Using Letter Points and Extensions, WASET International Conference on Computer, Information and Systems Science and Engineering (ICCISSE), Vienna, Austria, May 25-27, 2007, pp. 28-31.
34. Hartung F., Girod B., Digital Watermarking of Raw and Compressed Video, Proceedings of SPIE Volume 2952, Digital Compression Technologies and Systems for Video Communications, 1996, pp. 205-213.
35. Hartung F., Girod B., Watermarking of uncompressed and compressed video, Signal Processing, Vol. 66/3, pp.283--301, 1998.
36. Hartung F., Girod B., Watermarking of MPEG-2 encoded video without decoding and re-encoding, Proceeding of SPIE EI'97, Multimedia Computing and Networking, Vol. 3020, pp. 264-274, 1999.
37. Huang D., Yan H., Interword distance changes represented by sine waves for watermarking text images, IEEE transactions on circuits and systems for video technology, 2001, vol. 11, no12, pp. 1237-1245.
38. Huang D., Yeo T., Robust and inaudible multi-echo audio watermarking, In proc. IEEE Pacific-Rim Conference On Multimedia, Taipei, China, 2002, p 615–622.
39. Jie Song Liu, K.J.R., A data embedding scheme for H.263 compatible video coding, Circuits and Systems, 1999. ISCAS '99. Proceedings of the 1999 IEEE International Symposium, Volume 4, Issue , Jul 1999 pp.:390 - 393.

40. Jing Zhang, Anthony T. S., Gang Qiu, Pina Marziliano, Robust Video Watermarking of H.264/AVC, IEEE Transactions on Circuits and Systems II: Express Briefs, vol. 54, pp. 205-209, 2007.
41. Johnson N. F., An Introduction to Watermark Recovery from Images, SANS Intrusion Detection and Response (ID'99), Proceedings. San Diego, CA, February 9-13, 1999.
42. Johnson N. F., Jajodia S., Steganalysis: the investigation of hidden information, Information Technology Conference, 1998. IEEE 1-3, pp.113 – 116, Sep 1998.
43. Johnson N.F., Jajodia S., Steganalysis of Images Created Using Current Steganography Software, In Proceeding of 2nd Workshop on Information Hiding // Lecture Notes in Computer Science 1525, Springer, 1998. pp.273-289.
44. Kharrazi M., Sencar H. T., Memon N., Image Steganography: Concepts and Practice, WSPC, April 22, 2004.
45. Kirovski D., Malvar H., Robust Covert Communication over a Public Audio Channel Using Spread Spectrum, Information Hiding: 4th International Workshop, IHW 2001, Pittsburgh, PA, USA, April 25-27, 2001. Proceedings, Lecture Notes in Computer Science, Springer Berlin, Volume 2137, 2001, pp. 354-368.
46. Kirovski D., Malvar H., Spread-spectrum watermarking of audio signals, In proc. IEEE Transactions on Signal Processing, v. 51(4), 2003, p 1020–1033.
47. Ko B., Nishimura R., Suzuki Y., Time-spread echo method for digital audio watermarking using pn sequences, In proc. IEEE International Conference on Acoustics, Speech, and Signal Processing, Orlando, 2002, p 2001–2004.
48. Kundur D., Ahsan K., Practical Internet Steganography: Data Hiding in IP, Proc. Texas Workshop on Security of Information Systems, 5 pages, College Station, Texas, April 2003.
49. Kuo S., Johnston J., Turin W., Quackenbush S. Covert audio watermarking using perceptually tuned signal independent multiband phase modulation, In

- proc. IEEE International Conference on Acoustics, Speech, and Signal Processing, Orlando, 2002, p 1753–1756.
50. Kurosawa K., Watanabe O., Computational and Statistical Indistinguishability, Algorithms and Computation, Springer, ISSN 0302-9743, Volume 650/1992, 430-438.
 51. Kutter M., F. Petitcolas A.P., A fair benchmark for image watermarking systems, Electronic Imaging '99, Security and Watermarking of Multimedia Contents, vol.3657, San Jose, CA, USA, January, 1999, pp. 226–239.
 52. Langelaar G.C., van der Lubbe J.C.A., Biemond J., Copy Protection for Multimedia Data based on Labeling Techniques, 17th Symp. on Information Theory in the Benelux, pp. 33-40, 1996.
 53. Langelaar G. C., Lagendijk R. L., Biemond J., Removing Spatial Spread Spectrum Watermarks by Non-linear Filtering, 9th European Signal Processing Conference (EUSIPCO'98), Island of Rhodes, Greece, 8–11 Sept. 1998, pp. 2281–2284. ISBN 960-7620-05-4.
 54. Langelaar G.C., Lagendijk R.L., Biemond J., Real-Time Labeling of MPEG-2 Compressed Video, Journal of Visual Communication and Image Representation, Volume 9, Number 4, 1998, pp. 256-270.
 55. Low S. H., Maxemchuk N. F., Brassil J. T., O'Gorman L., Document marking and identification using both line and word shifting, IEEE INFOCOM, April, 1995, vol. 2, pp. 853-860.
 56. Lee K., Jung C., Lee S., Lim J., New Steganalysis Methodology: LR Cube Analysis for the Detection of LSB Steganography // Information Hiding, 2005. Volume 3727/2005, pp. 312-326.
 57. Lee K., Jung C., Lee S., Kim H., Lim J. Applying LR Cube Analysis to JSteg Detection // Communications and Multimedia Security, 2005. LNCS 3677, pp. 275-276.
 58. Low S. H., Maxemchuk N. F., Lapone A. M., Document Identification for Copyright Protection using Centroid Detection, IEEE Transactions on Communications, vol. 46/3, 1998, pp. 372-383.

59. Luo X., Liu B., Liu F., Improved RS Method for Detection of LSB Steganography // Computational Science and Its Applications – ICCSA 2005, 2005. pp. 508-516.
60. Matsuoka H., Spread Spectrum Audio Steganography Using Sub-band Phase Shifting, Intelligent Information Hiding and Multimedia Signal Processing, 2006. IIH-MSP '06., Dec. 2006, pp.3-6.
61. Maxemchuk F., Low S., Marking text documents, In IEEE-ICIP'97, vol. 3, Santa Barbara (Cal) 1997, pp. 13-16.
62. Mei Q., Wong E. K., Memon N., Data hiding in binary text documents, Proc. SPIE, Security and Watermarking of Multimedia Contents III, Aug. 2001, vol. 4314, pp. 369-375.
63. Meral, H.M., Sevinç, E., Ünkar, E.I., Sankur, B., Özsoy, S. and Güngör, T., Syntactic Tools for Text Watermarking, 19th Annual Symposium on Electronic Imaging: Security, Stenography, and Watermarking of Multimedia Contents IX, January 2007, Proceedings of the SPIE, Volume 6505, 12 p.
64. Mukherjee D., Chae J. J., Mitra S. K., Manjunath B. S., A source and channel-coding framework for vector-based data hiding in video, IEEE Trans. On Circuits and systems for video technology, vol. 10 (4), pp. 630-645 June 2000.
65. Mukherjee D., Chae J. J., S. K. Mitra, A source and channel coding approach to data hiding with application to hiding speech in video, in Proc. IEEE Int. Conf. Image Processing, vol. 1, Chicago, Oct. 1998, pp. 348–352.
66. Murphy B., The syntax of concealment: reliable methods for plain text information hiding, Proceedings of the SPIE International Conference on Security, Steganography, and Watermarking of Multimedia Contents, Volume 6505, January 2007.
67. Neubauer C., Herre J., Audio watermarking of MPEG2 AAC bitstreams, in Proc. of 108th Convention of Audio Engineering Society (AES), 19 p., Paris, 2000.
68. Oh H., Seok J., Hong J., Youn D., New echo embedding technique for robust and imperceptible audio watermarking, In proc. IEEE International

- Conference on Acoustic, Speech and Signal Processing, Salt Lake City, 2001, p 1341–1344.
69. F. Petitcolas A.P., Anderson R., Kuhn M., Attacks on Copyright Marking Systems // Lecture Notes in Computer Science, 1998, pp. 218-238.
 70. Pfitzmann B., Information Hiding Terminology // Information Hiding, First International Workshop, Cambridge, U.K., May 30 – June 1, 1996 Proceedings, Lecture Notes in Computer Science, Springer Berlin, Volume 1174, 1996, pp. 347-350.
 71. Provos N., Defending against on statistical Steganalysis // Proceeding of the 10 USENIX Security Symposium, 2001. p. 323-335.
 72. Reeds J., Solved: the Ciphers in Book III of Trithemius' Steganographia, Cryptologia v XXII no 4, pp 291–318, October 1998.
 73. Rivest R.L., Shamir A., Adleman L.M., A method for obtaining digital signatures and public-key cryptosystems // Communications of the ACM, v.21, n.2, Feb 1978, pp.120-126.
 74. Schulzrinne H., Casner S., Frederick R., Jacobson V., Rtp: A transport protocol for real-time applications // RFC 1889, Internet Society (IETF), January 1996.
 75. Shannon C.E., Communication theory of secrecy systems // Bell System Technical Journal, 28:657-715, 1949.
 76. Shirali-Shahreza M.H., Shirali-Shahreza M., A New Approach to Persian/Arabic Text Steganography // Computer and Information Science, 2006 and 2006 1st IEEE/ACIS International Workshop on Component-Based Software Engineering, Software Architecture and Reuse. ICIS-COMSAR 2006. 5th IEEE/ACIS International Conference, 10-12 July 2006, pp. 310-315.
 77. Simmons G.J., The Prisoners' Problem and Subliminal Channel, in Advances in Cryptology, Proceedings of CRYPTO '83, Plenum Press, 1984, pp. 51-67.

78. Sun Z., Ji Z., A Security Steganography Method for Lossy Compression Gray Scale Image // Springer-Verlag Berlin Heidelberg 2007, ICIC 2007, LNCS 4681, 2007, pp. 636–645.
79. Sung Min Kim, Sang Beom Kim, Youpyo Hong and Chee Sun Won, Data Hiding on H.264/AVC Compressed Video // Lecture Notes in Computer Science, Springer Berlin / Heidelberg, Volume 4633, 2007, pp.:698-707.
80. Swanson M.D., Zhu B., Tewfik A.H., Data hiding for video-in-video, International Conference on Image Processing, 1997 Proceedings, Volume 2, Issue , 26-29 Oct 1997, pp:676 - 679.
81. Swanson M. D., Zhu B., Chau B., Tewfik A. H., Object-based transparent video watermarking, in Proc. IEEE Workshop Multimedia Signal Processing, 1997, pp. 369–374.
82. Swanson M. D., Zhu B., Chau B., Tewfik A. H., Multiresolution video watermarking using perceptual models and scene segmentation, in Proc. IEEE Int. Conf. Image Processing, vol. 2, Santa Barbara, 1997, pp. 558–561.
83. Tachibana R. , Two-Dimensional Audio Watermark for MPEG AAC Audio, in Proc. of Security, Steganography and Watermarking of Multimedia Contents VI, SPIE vol. 5306, pp. 139-150, San Jose, USA, January 2004.
84. Takahashi T., Lee W., An Assessment of VoIP Covert Channel Threats // In Proc. of 3rd International Conference on Security and Privacy in Communication Networks (Secure-Comm'07), Nice, France, 2007.
85. Tao B., Dickenson B., Adaptive Watermarking in the DCT Domain, Proc. of Intl. Cond. Acoustics, Speech and Signal Processing (ICASSP '97), Vol. 4, pp. 2985-2988, April 1997.
86. Trabelsi Z., El-Sayed H., Frikha L., Rabie T., Traceroute Based IP Channel for Sending Hidden Short Messages, IWSEC 2006, Springer-Verlag Berlin Heidelberg, LNCS 4266, 2006, pp. 421 – 436.
87. Qiao L., Nahrstedt K., Watermarking Methods for MPEG Encoded Video: Towards Resolving Rightful Ownership, Proceedings of IEEE International Conference of Multimedia Computing and Systems, pp. 276-285, 1998.

88. Villán R., Voloshynovskiy S., Koval O., Vila-Forcén J.E., Topak E., Deguillaume F., Rytsar Y., Pun T., Text Data-Hiding for Digital and Printed Documents: Theoretical and Practical Considerations, in Proceedings of SPIE-IS&T Electronic Imaging 2006, Security, Steganography, and Watermarking of Multimedia Contents VIII, San Jose, USA, Jan.15-19 2006.
89. Voloshynovskiy, S., Pereira, S., Iquise, V., Pun, T., Attack Modelling: Towards a Second Generation Watermarking Benchmark, SP(81), No. 6, June 2001, pp. 1177-1214.
90. Wang Y., Izquierdo E., High-capacity data hiding in MPEG-2 compressed video, IWSSIP'02: international workshop on systems, signals and image processing No9, Manchester, 2002, pp. 212-218.
91. Wayner P., Disappearing Cryptography – Information Hiding: Steganography & Watermarking, Morgan Kaufmann Publishers, Los Altos, CA 94022, USA, second ed., 2002, 413 p.
92. Westfeld A., Pfitzmann A., Attacks on Steganographic Systems. Breaking the Steganographic Utilities EzStego, Jsteg, Steganos and S-Tools - and Some Lessons Learned, Proceedings of the Workshop on Information Hiding 1999, Lecture Notes in Computer Science. 2000. Vol. 1768. P. 61-75.
93. Wu T. L., Wu, S. F., Selective Encryption and Watermarking of {MPEG} Video, International Conference on Image Science, Systems, and Technology, {CISST}'97, June, 1997, 10p.
94. Yeh C., Kuo C., Digital watermarking through quasi m-arrays // Proceeding of IEEE Workshop on Signal Processing Systems, Taipei, Taiwan, 1999. pp. 456–461.
95. Zoran Duric, Neil F. Johnson, and Sushil Jajodia, Recovering Watermarks from Images, Information & Software Engineering Technical Report ISE-TR-99-04, April 1999.
96. Zou D., Shi Y.Q., Formatted text document data hiding robust to printing, copying and scanning, Circuits and Systems, 2005, ISCAS-2005, IEEE International Symposium, Vol. 5, pp. 4971-4974.

97. Аграновский А. В., Балакин А. В., Репалов С. А., Хади Р. А., Способ стеганографического преобразования блоков двоичных данных, Патент РФ № 2 257 010, РОСПАТЕНТ. - М., 20.07.2005, Бюл. №20, 9 с.
98. Аграновский А.В., Алиев А.Т., Селин С.Н., Хади Р.А., Современные запатентованные решения в области защиты информационных ресурсов корпоративных вычислительных сетей // Журнал "Информационные технологии" №10 '2005. с. 51 – 55.
99. Аграновский А.В., Алиев А.Т., Балакин А.В., Теоретико-множественный подход к оценке надежности многомодульных систем стеганографического анализа // Всероссийская научно-практическая конференция «Охрана, безопасность и связь – 2005»: Сборник материалов. Часть 3. – Воронеж: Воронежский институт МВД России, 2005. с. 22-25.
100. Аграновский А.В., Алиев А.Т., Иванков М.П., Уязвимости в системах безопасности современных систем электронных платежей // Всероссийская научно-практическая конференция «Современные проблемы борьбы с преступностью»: Сборник материалов (информационная безопасность). – Воронеж: Воронежский институт МВД России, 2005. с. 92-93.
101. Аграновский А.В., Алиев А.Т., Применение криптографических алгоритмов при построении стеганографических систем // Вестник Саровского Физтеха. Научно-популярный журнал. – Саров. Изд-во: ФГОУ ВПО "СарФТИ", 2006. №11, с. 55-56.
102. Аграновский А.В., Алиев А.Т., Балакин А.В., Защита прав интеллектуальной собственности при хранении, передаче и распространении информации в компьютерных сетях // Материалы X Международной научно-практической конференции «Информационная безопасность». - Таганрог: Изд-во ТТИ ЮФУ, 2008. Ч. 2. с.189-192.

103. Алиев А.Т., О применении стеганографического метода LSB к графическим файлам с большими областями монотонной заливки // Вестник ДГТУ. 2004. Т.4. №4(22), с. 454-460.
104. Алиев А.Т., Шагов Г.Н., О возможности обнаружения стеганографического канала в дискретных каналах передачи информации в случае использования несовершенных стегосистем // Обозрение прикладной и промышленной математики: Материалы шестого Всероссийского симпозиума по прикладной и промышленной математике (осенняя (открытая) сессия), Сочи - Дагомыс, 1 - 7 октября 2005 г.: т.12, выпуск 4, Редакция журнала «ОПиПМ» Москва 2005. – Ч. 2. с. 897-898.
105. Алиев А.Т., Балакин А.В., Колпаков Н.И. Основы построения стеганографической защиты мультимедиа-информации: Монография. Ростов-на-Дону: Изд-во СКНЦ ВШ, 2006. 204 с.: ил.
106. Алиев А.Т., О повышении стойкости стеганографического метода LSB // Теория, методы проектирования, программно-техническая платформа корпоративных информационных систем: Материалы Междунар. науч.-практ. конф., г. Новочеркасск, 16 мая 2003 г. / Юж.-Рос. гос. техн. ун-т (НПИ). – Новочеркасск: ЮРГТУ, 2003. с. 100-101.
107. Алиев А.Т., Способ внедрения дополнительной информации в цифровые изображения, Патент №2288544, РОСПАТЕНТ.-М. от 24.12.2004.
108. Алиев А.Т., Метод сдвига битовых последовательностей в стеганографии // Материалы VI Международной научно-практической конференции «Информационная безопасность», Таганрог: Изд-во ТРТУ, 2004. с.162-164.
109. Алиев А.Т., Балакин А.В., К вопросу оценки стойкости стегосистем // Методы и алгоритмы прикладной математики в технике, медицине и экономике: Материалы V Междунар. науч.-практ. конф., г. Новочеркасск, 21 янв. 2005 г.: В 3 ч./Юж.-Рос. гос. техн. ун-т (НПИ). – Новочеркасск: ЮРГТУ, 2004. Ч. 1. с. 51-53.

110. Алиев А.Т., Балакин А.В., Оценка стойкости систем скрытой передачи информации // Известия ТРТУ. Тематический выпуск. Материалы VII Международной научно-практической конференции «Информационная безопасность». Таганрог: Изд-во ТРТУ, 2005. №4 (48), с. 199-204.
111. Алиев А.Т., Построение стойких стеганографических систем на базе пространственных и пространственно-частотных фильтров // Материалы VII Международной научно-практической конференции «Информационная безопасность». Таганрог: Изд-во ТРТУ, 2005. с. 210-212.
112. Алиев А.Т., Построение стойких стеганографических систем на базе пространственных и пространственно-частотных фильтров // Информационное противодействие угрозам терроризма: науч.-практ. журн. / №4, 2005, Таганрог: ТРТУ, 2005. с. 159-164.
113. Алиев А.Т., Программное средство SBS BMP Hide организации скрытого хранения и передачи конфиденциальной информации // Свидетельство об официальной регистрации программы для ЭВМ №2006613532 РОСПАТЕНТ. - М., 19.08.2005.
114. Алиев А.Т., Аграновский А.В., Вопросы построения криптостеганографических систем. Модель стеганографического канала передачи данных // Известия ТРТУ. Тематический выпуск. «Информационная безопасность». Таганрог: Изд-во ТРТУ, 2006. № 7 (62), с. 185-192.
115. Алиев А.Т., Особенности организации оперативно-технических мероприятий по поиску скрытой информации // Международная научно-практическая конференция «Современные проблемы борьбы с преступностью»: Сборник материалов. – Пленарное заседание. – Воронеж: Воронежский институт МВД России, 2006, с. 7-8.
116. Алиев А.Т., Аграновский А.В., Вопросы построения крипто-стеганографических систем. Модель стеганографического канала передачи данных // Информационное противодействие угрозам терроризма: науч.-практ. Журн/№8, 2006, Таганрог: ТРТУ, 2006. с. 79-91.

117. Алиев А.Т., Мультиплексирование стеганографического канала // Материалы VIII Международной научно-практической конференции «Информационная безопасность». - Таганрог: Изд-во ТРТУ, 2006. Ч. 2. с.82-83.
118. Алиев А.Т., Полушкин Н.Ю., Лежнев А.В., Автосинхронизация в полудуплексных каналах с инъективными ошибками // Теория, методы проектирования, программно-техническая платформа корпоративных информационных систем: Материалы IV Междунар. науч.-практ. конф., г. Новочеркасск, 26 мая 2006 г./Юж.Рос. гос. техн. ун-т (НПИ). – Новочеркасск: ЮРГТУ, 2006. с. 80-83.
119. Алиев А.Т., Балакин А.В., Селин Р.Н., Способ маркировки и способ проверки маркировки строк ответов на запросы пользователей к базе данных с использованием цифровых водяных знаков, Заявка на изобретение №2007115462/09 (016791), РОСПАТЕНТ.-М. от 25.04.2007. Решение о выдаче патента от 16.05.2008.
120. Алиев А.Т., Сергеев Д.В., Криптостеганографические системы: теоретические основы, принципы построения и перспективы // Материалы IX Международной научно-технической конференции. «Информационная безопасность». Таганрог: Изд-во ТТИ ЮФУ, 2007. Ч.2., с. 49-54.
121. Алиев А.Т., Балакин А.В., Крассов С.А., Построение стеганографических систем на основе криптографических алгоритмов // Материалы XXXIV Международной конференции и дискуссионного научного клуба «Информационные технологии в науке, социологии, экономике и бизнесе» IT+SE'07. – Ялта-Гурзуф, 29 сентября – 8 октября 2007. с. 62-63.
122. Алиев А.Т., Крассов С.А., Шагов Г.Н., Построение стеганографических систем с симметричными и открытыми ключами на основе криптографических алгоритмов // Обозрение прикладной и промышленной математики: Материалы восьмого Всероссийского

- симпозиума по прикладной и промышленной математике (осенняя сессия), Сочи - Адлер, 29 сентября - 7 октября 2007 г.: т.15, выпуск 2, Редакция журнала «ОПиПМ» Москва 2008. Ч. 1. с. 253.
123. Алиев А.Т., Методы и механизмы защиты авторского права и смежных прав на электронные произведения в многопользовательских распределенных вычислительных системах / Высокопроизводительные вычислительные системы // Материалы Пятой Международной научной молодежной школы. Материалы Международной молодежной научно-технической конференции – Таганрог: Изд-во ТТИ ЮФУ, 2008, с. 402-404.
124. Алиев А.Т., Сергеев Д.В., Криптостеганографические системы: теоретические основы, принципы построения и перспективы // Информационное противодействие угрозам терроризма: науч.-практ. журн. / №11, 2008, Таганрог: ТРТУ, 2008. с. 23-31.
125. Бабаш А. В., История криптографии. Часть I./ Бабаш А. В., Шанкин Г. П. – М.: Гелиос, 2002. – 240с., ил.
126. Балакин А.В., Романцов А.П., Хади Р.А., Классификация современных методов стеганографического анализа // Вестник компьютерных и информационных технологий, №2, 2007г., с. 45-53.
127. Балакин А.В., Репалов С.А., Шагов Г.Н., Современная стеганография: модели и методы преобразования информации // Ростов-на-Дону, Изд-во СКНЦ ВШ, 2004. 240 с.
128. Быков С. Ф., Алгоритм сжатия JPEG с позиции компьютерной стеганографии // Защита информации. Конфидент, 2000. №3, с.26-33.
129. Генне О.В., Основные положения стеганографии // Защита информации. Конфидент, 2000. №3, с.20-24.
130. ГОСТ 28147-89. Системы обработки информации. Защита криптографическая. Алгоритм криптографического преобразования. – М.: ИПК Издательство стандартов, 1996.

131. ГОСТ Р 34.10-94. Информационная технология. Криптографическая защита информации. Процедуры выработки и проверки электронной цифровой подписи на базе асимметричного криптографического алгоритма. – М.: ИПК Издательство стандартов, 1994.
132. ГОСТ Р 34.10-2001. Информационная технология. Криптографическая защита информации. Процессы формирования и проверки электронной цифровой подписи. – М.: ИПК Издательство стандартов, 2001.
133. ГОСТ Р 34.11-94. Информационная технология. Криптографическая защита информации. Функция хэширования. – М.: ИПК Издательство стандартов, 1994.
134. ГОСТ Р 50739-95. Средства вычислительной техники. Защита от несанкционированного доступа к информации. Общие технические требования. – М.: Стандартинформ, 2006.
135. ГОСТ Р 50922-96. Защита информации. Основные термины и определения. – М.: ИПК Издательство стандартов, 1996.
136. ГОСТ Р 50922-2006. Защита информации. Основные термины и определения. – М.: Стандартинформ, 2008.
137. ГОСТ Р 51141-98 Делопроизводство и архивное дело. Термины и определения. – М.: Стандартинформ, 2006.
138. Грибунин В.Г., Оков И.Н., Туринцев И.В., Цифровая стеганография / М.: СОЛОН-Пресс, 2002. – 272 с.
139. Закон Российской Федерации «Об авторском праве и смежных правах» от 9 июля 1993г. № 5342-1 (с изменениями согласно Федеральному закону РФ от 20 июня 2004 г. № 72-ФЗ).
140. Запечников С.В., Криптографические протоколы и их применение в финансовой и коммерческой деятельности: Учебное пособие для вузов. – М. Горячая линия – Телеком, 2007. – 320 с.
141. Зубов А.Ю., Совершенные шифры. – М.: Гелиос АРВ, 2003. – 160с.
142. Кан Д., Взломщики кодов. – Пер. с англ. А. Ключевского. – «Секретная папка». – М.: ЗАО Изд-во Центрполиграф, 2000. – 473 с.

143. Кустов В. Н., Федчук А. А., Методы встраивания скрытых сообщений // Защита информации. Конфидент, 2000. №3, с.34-37.
144. Мартынов А.П., Фомченко В.П., Криптография и электроника / Под ред. А.И. Астайкина. Саров: ФГУП «РФЯЦ-ВНИИЭФ», 2006, 452 с.
145. Оков И.Н., Ковалев Р.М., Электронные водяные знаки как средство аутентификации передаваемых сообщений // Защита информации. Конфидент, 2001. №3, ст. 50-55.
146. Пашанин И., Исследование российских систем электронного документооборота [Электронный ресурс]: — Электрон. дан. — М.: Портал iTeam, 2008. — Режим доступа:
http://www.iteam.ru/publications/it/section_64/article_2886/, свободный. — Загл. с экрана.
147. Проект Федерального закона РФ «Об электронном документе» (от 16.04.2001).
148. РД «Безопасность информационных технологий. Критерии оценки безопасности информационных технологий». Приказ председателя Гостехкомиссии России от 19 июня 2002 года № 187.
149. Рублев Д.П., Макаревич О.Б., Федоров В.М., Архитектура сетевой системы обнаружения внедренных стеганографическим методом данных в речевых сообщениях и изображениях // Материалы IX Международной научно-технической конференции. «Информационная безопасность», Таганрог: Изд-во ТРТУ, 2007. Т.2., с. 72-78.
150. Рублев Д.П., Федоров В.М., Макаревич О.Б., Обнаружение скрытых сообщений в изображениях на основе вейвлет-декомпозиции // Материалы VII Международной научно-практической конференции «Информационная безопасность». – Таганрог: Изд-во ТРТУ, 2005. с.216-219.
151. Рублев Д.П., Федоров В.М., Макаревич О.Б., Бабенко Л.К., Метод встраивания данных в аудиопоток на основе преобразования фазовых составляющих // Материалы VII Международной научно-практической

- конференции «Информационная безопасность». – Таганрог: Изд-во ТРТУ, 2005. с.212-216.
152. Рублев Д.П., Макаревич О.Б., Федоров В.М., Обнаружение скрытых сообщений в изображениях на основе статистических моделей высших порядков // Материалы VI Международной научно-практической конференции «Информационная безопасность». Таганрог, 2004, стр. 287-289.
153. Сидоров М. А, Скрытые марковские модели и стегоанализ // Материалы VI Международной научно-практической конференции «Информационная безопасность». Таганрог, 2004, стр. 289-291.
154. Федеральный закон РФ «Об обязательном экземпляре документов» от 29 декабря 1994 г. N 77-ФЗ (в ред. Федерального закона от 11.02.2002 N 19-ФЗ, с изм., внесенными Федеральными законами от 27.12.2000 N 150-ФЗ, от 24.12.2002 N 176-ФЗ).
155. Федеральный закон РФ «Об электронной цифровой подписи» от 10 января 2002 г. N1-ФЗ.
156. Федоров В.М., Макаревич О.Б., Рублев Д.П., Метод стеганографии в аудиосигналах и изображениях, устойчивый к компрессии с потерями // Известия ТРТУ. Тематический выпуск. «Информационная безопасность». Таганрог: Изд-во ТРТУ, 2006. №7 (62), с.201-208.
157. Чмора А.Л., Современная прикладная криптография / М.: Гелиос АРВ, 2001. – 256 с.
158. Шенон К. Э. Теория связи в секретных системах / опубликована в «Работы по теории информации и кибернетике» // М.: Иностранная литература, 1963г, с. 333-369 (Перевод В.Ф.Писаренко).
159. Шнайер Б. Прикладная криптография. Протоколы, алгоритмы, исходные тексты на языке Си. – М.: Издательство ТРИУМФ, 2002. – 816 с.

ПРИЛОЖЕНИЯ

Приложение А. Термины и определения в теории стеганографии

Стеганография – направление в технологиях сокрытия информации, рассматривающее вопросы скрытого взаимодействия, обмена и передачи информации с использованием открытых каналов связи.

Следуя этому определению, целью стеганографии является организация защищенного обмена сообщениями между двумя сторонами, при котором от третьей стороны, осуществляющей контроль над используемым каналом связи, скрывается не только содержимое сообщений, но и факт их передачи.

Техническая (или специальная) стеганография – направление стеганографии сосредоточенное на задачах скрытой передачи информации с использованием достаточно сложных и неочевидных технических решений. Данное направление практически не ограничено в выборе носителей и методов для передачи скрытой информации.

В настоящее время в области технической стеганографии выделяют два относительно близких направления связанных с представлением данных в цифровом виде и их хранением и обработкой с использованием ЭВМ. Это направления компьютерной и цифровой стеганографии.

Компьютерная стеганография – изучает способы скрытого хранения и передачи информации на основе использования особенностей представления и обработки данных на ЭВМ. В компьютерной стеганографии для сокрытия информации широко используются особенности хранения данных в виде файлов и их записи на цифровые носители, нестандартное использование служебных полей и заголовков, особенности сетевых протоколов и операций копирования данных. Отличительной особенностью компьютерной стеганографии являются операции с представлением и хранением данных, сами данные при которых остаются неизмененными.

Цифровая стеганография – направление стеганографии связанное с цифровой обработкой сигналов, как правило, имеющих аналоговую природу

(фотоизображения, сканированные изображения, факсимильные сообщения, звуки, речь, аудио и видеозаписи) предполагающее непосредственное изменение данных и их обработку как в режиме реально времени, так и в отложенном режиме, предполагающем хранение данных в виде файлов на цифровых носителях.

Общая терминология для области технической стеганографии была принята в 1996 году по результатам открытых обсуждений на симпозиуме "Information Hiding – First information Workshop". В статье [70] подытожившей результаты обсуждений введены следующие определения:

1. *Встроенное <тип данных>* – то, что скрывается в чем-то другом.
 2. *Стего-<тип данных>* – нечто содержащее в себе встроенное сообщение.
 3. *Скрывающие <тип данных>* – исходное стегосообщение, которое еще не содержит в себе встроенного сообщения.
 4. *Стегоключ* или *ключ* – дополнительная секретная информация, которая может потребоваться в процессе встраивания/извлечения информации.

Основываясь на предложенной в [70] терминологии, а также терминологии используемой в отечественной открытой литературе [105, 129,

127, 138], поясним некоторые общие термины и определения, которые непосредственно использованы в данной работе.

Скрываемое/скрытое сообщение или *сообщение* – передаваемая секретная информация. В случае цифровой и компьютерной стеганографии скрываемое сообщение чаще всего представляет собой двоичную последовательность конечной длины.

Скрывающие сообщение или *контейнер* – любая информация или поток данных в которые можно записать скрываемое сообщение и пересылка которых по открытым каналам связи не вызывает подозрений.

Оригинальным контейнером называется контейнер, который не подвергался изменению с целью встраивания скрываемого сообщения.

Встроенное/внедренное сообщение – скрываемое сообщение, которое храниться или передается по открытому каналу связи, будучи тем или иным способом, записанным в контейнер.

Пустым контейнером называется контейнер, который не содержит встроенное сообщение, или признан таковым.

Соответственно контейнер, содержащий встроенное сообщение называется *заполненным*.

Так как запись информации в пустой контейнер приводит к изменению последнего, заполненный контейнер также называют *модифицированным*. Пустой или заполненный контейнер, передаваемый по открытому каналу связи в качестве обычного сообщения, называется *стегосообщением*.

Скрывающим преобразованием называется операция записи сообщения в контейнер, предполагающее изменение последнего.

Извлечение сообщения – операция восстановления скрытого сообщения в процессе обработки заполненного контейнера.

Стеганографическая система или *стегосистема* – совокупность средств и методов, которые используются для формирования скрытого канала передачи информации на основе открытого канала связи.

Канал передачи данных основанный на использовании той или иной стеганографической системы называется *стеганографическим каналом* или *стегоканалом*.

Стеганографическими методами принято называть методы сокрытия информации представленные скрывающими преобразованиями, а также методы извлечения сообщений из заполненных контейнеров.

Стеганографические методы сокрытия информации в области в цифровой и компьютерной стеганографии в зависимости от особенностей использования контейнеров делят на два класса: форматные и неформатные.

К *форматным стеганографическим методам* относят методы сокрытия информации в контейнерах, основанные на использовании тех или иных особенностей форматов хранения данных. Форматные методы являются наиболее простыми в реализации. Типичным примером форматного метода сокрытия информации является запись данных в конец графического файла. Наиболее же часто встречающимися подходами, используемыми форматными методами, применительно к графическим файлам являются:

- сокрытие информации в заголовках;
- использование зарезервированных полей;
- сокрытие информации между блоками данных;
- запись информации в конец файла;
- использование служебных блоков и блоков комментариев;
- сокрытие информации с использованием таблицы цветов.

Неформатные методы используют для сокрытия дополнительной информации непосредственно сами данные. Неформатные методы основаны на использовании таких свойств информации, как избыточность и неоднозначность. Запись информации с применением неформатных стеганографических методов неизбежно ведет к искажению данных. В сравнении с форматными стеганографическими методами, неформатные методы обладают более высокой надежностью и стойкостью к обнаружению.

Направление в технологиях противодействия защищенным системам связи нацеленное на выявление стеганографических вложений и каналов связи, включающее также методы извлечения скрытых сообщений, называется – *стеганоанализом*. Целью методов стеганоанализа является выявление скрытых сообщений, их извлечение, а также обеспечение возможности предоставления доказательств о наличии скрытого канала связи третьей стороне.

Приложение Б. Результаты экспериментального применения МСБП совместно с методом замены младших значащих битов

В целях испытания согласующего алгоритма на базе метода сдвига битовых последовательностей была рассмотрена его работа в связке со стеганографическим методом LSB. Метод замены младших значащих битов [1, 4, 26, 143], обозначаемый в зарубежной литературе аббревиатурой LSB (Least Significant Bits), является наиболее известным на сегодняшний день неформатным стеганографическим методом. К настоящему времени создано множество программных продуктов использующих различные модификации данного метода. К настоящему времени данный метод фактически стал считаться классикой цифровой стеганографии. Метод LSB ввиду своей простоты и сегодня пользуется большой популярностью среди разработчиков стеганографического программного обеспечения. Он также взят за основу большинства более современных методов сокрытия информации в растровых графических изображениях, аудио и виде потоках.

Коротко, идея метода заключается в замене младших битов в байтах цветового представления отдельных точек изображения битами скрываемого сообщения. Возможность замены объясняется избыточностью представления цвета, а также предположительно случайному поведением младших битов. Запись сообщения в классическом варианте метода LSB осуществляется путем последовательной замены всех младших битов байтового потока данных изображения битами скрываемого сообщения. Порог чувствительности глаза к изменению освещенности и интенсивности цвета при средних значениях составляет порядка 1~3 процента [29]. Таким образом, метод замены младших значащих битов в его базовой реализации в действительности не вносит перsepтивно заметных искажений [103].

Извлечение битовой строки из исходного графического изображения для последующей ее обработки с помощью МСБП осуществляется в пять этапов:

1. Разложение цветного изображения на цветовые составляющие или его преобразование к монотонному виду.
2. Выбор одной из не измененных цветовых составляющих.
3. Разложение выбранной на предыдущем шаге цветовой составляющей на битовые слои.
4. Выбор не модифицированного битового слоя.
5. Представление выбранного битового слоя в виде битовой последовательности построчным сканированием.

Для проведения экспериментов и наглядной демонстрации возможностей совместного применения методов (LSB и МСБП), а также перспективности предложенных решений была построена программная реализация с удобным графическим интерфейсом [113]. Главное окно программы представлено на рисунке Б.1.



Рисунок Б.1 – Интерфейс SBS BMP Hide: главное окно программы

Наглядным примером описанного процесса формирования битовой последовательности является окно задания параметров скрытия информации, представленное на рисунке Б.2.

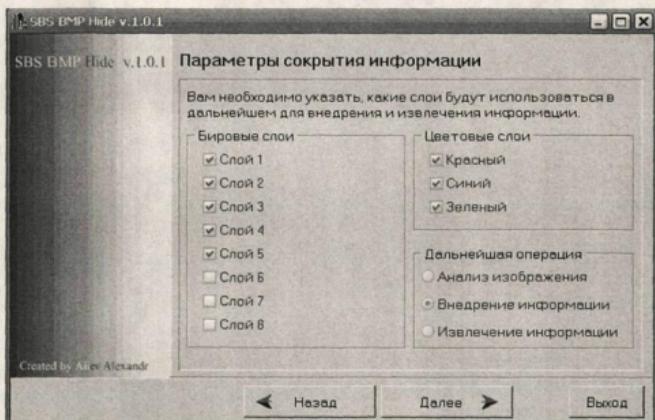


Рисунок Б.2 – Интерфейс SBS BMP Hide: задание параметров сокрытия

В примере интерфейса, приведенном на рисунке Б.2 видно, что для записи информации выбрано пять младших слоев изображения. Более того, в зависимости от изображения, используемого в качестве контейнера, допускается выбор для записи информации до шести из восьми возможных битовых слоев. В тоже время метод LSB в классическом варианте позволяет использовать только три младших слоя, и при этом одновременное использование всех трех слоев приведет к существенным искажениям исходного изображения. В нашем же случае визуальное искажение изображения не наблюдается.

Дело в том, что предложенный метод производит запись информации с учетом не только физических параметров человеческого зрения, но и с учетом высокоуровневых (психофизических) свойств человеческого зрения. Человеческий мозг проводит анализ и разделение поступающего видеосигнала на отдельные компоненты. Выделяемые компоненты имеют различные пространственные и частотные характеристики. В случае одновременного воздействия на глаз двух компонентов со сходными характеристиками возбуждаются одни и те же подканалы. Это приводит к

эффекту маскирования, заключающемуся в увеличении порога обнаружения одного сигнала в присутствии другого, обладающего аналогичными характеристиками. Поэтому аддитивный шум, хорошо заметный на гладких участках изображения, может оказаться незаметным на участках со сложной структурой. Заметим, что МСБП при записи информации подвергает изменению только граничные элементы серий одинаковых битов. Как следствие, для записи скрываемой информации используются наименее предсказуемые участки изображения, монотонные большие области, на которых заметны даже небольшие изменения, не затрагиваются.

Возможность использования большего по сравнению с базовым методом LSB битовых слоев дает значительный выигрыш в объемах скрываемой информации. В зависимости от изображения, объем скрываемых данных может достигать 60% от начального объема контейнера.



Рисунок Б.3 – Отношение объемов скрываемой информации для одного и того же изображения в методах LSB и LSB+МСБП

В п. 3.4.2 был приведен пример, показывающий высокую точность сохранения статистических свойств битовой строки, при записи в нее дополнительной информации с применением МСБП. На рисунке Б.4 приведен аналогичный график. На графике показаны частоты встречаемости пар и троек битов в битовых потоках, извлеченных из пустых и заполненных с применением МСБП контейнеров.

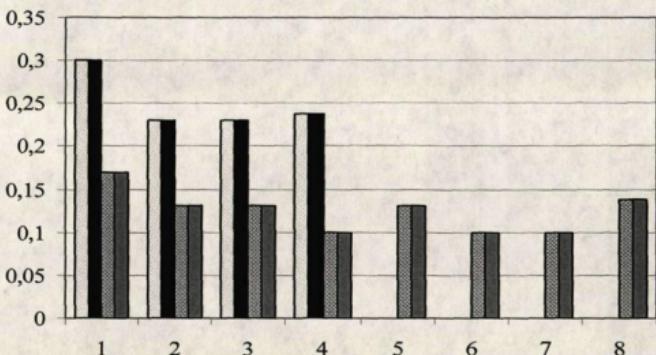
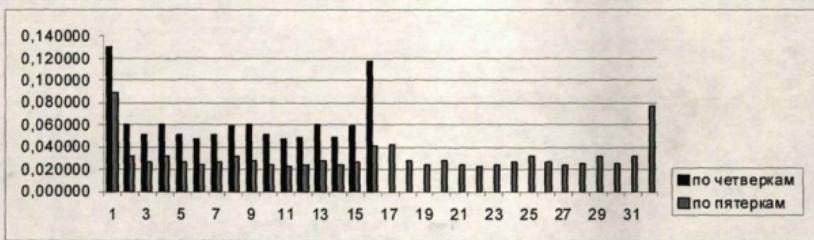
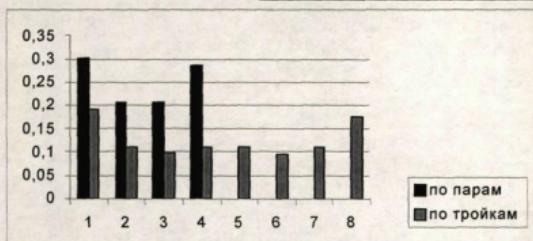


Рисунок Б.4 – Частоты встречаемости пар, троек битов в пустых и заполненных (LSB+МСБР) контейнерах

В ходе проведения эксперимента оценивались отклонения от исходных значений в частотах встречаемости векторов из двух-восьми битов. Все искажения ($<0,25\%$) оказались в допустимых пределах, гарантирующих невозможность выявления аномалии в данных частотах при анализе только заполненного контейнера. Оценка точности сохранения частот встречаемости пар и троек битов при применении МСБР проводилась на реальных цифровых фотoreалистичных изображениях. При этом были получены данные по частоте встречаемости битовых пар-восьмерок в 1-3 младших слоях растровых фотoreалистичных изображений. Объем выборки составил 520 различных изображений. Данные по средним, минимальным и максимальным значениям частот встречаемости троек, четверок и пятерок в 1-3 младших слоях растровых фотoreалистичных изображений представлены в таблице Б.1. Ввиду большого объема аналогичных таблиц для семерок и восьмерок битов, для них на рисунке Б.5 представлен сводный график вероятности появления пар, троек, четверок, пятерок, шестерок, семерок и восьмерок битов в 1-3 битовых слоях графических изображений.

Таблица Б.1

	по парам			по тройкам			по четверкам			по пятеркам		
	Average	Max	Min	Average	Max	Min	Average	Max	Min	Average	Max	Min
0	0,301427	0,842553	0,214824	0,190540	0,816664	0,113116	0,130643	0,800040	0,058886	0,089116	0,785506	0,027631
1	0,206772	0,250035	0,040362	0,110886	0,136183	0,025888	0,059897	0,081110	0,016822	0,032439	0,049096	0,010953
2	0,206772	0,250035	0,040362	0,097328	0,131620	0,012693	0,050566	0,075052	0,008290	0,026037	0,045023	0,004971
3	0,285026	0,546269	0,076722	0,109444	0,127085	0,024058	0,060319	0,071512	0,016390	0,032338	0,042620	0,010398
4				0,110886	0,136183	0,025888	0,050562	0,075199	0,008244	0,026176	0,045155	0,005211
5				0,095886	0,125353	0,012557	0,046766	0,062682	0,004449	0,023727	0,031668	0,002441
6				0,109444	0,127085	0,024058	0,051059	0,062736	0,009746	0,026704	0,033906	0,005902
7				0,175582	0,479336	0,052664	0,058385	0,067555	0,014312	0,031807	0,041545	0,009467
8							0,059897	0,081111	0,016822	0,027582	0,045069	0,006978
9							0,050989	0,062694	0,009266	0,024530	0,032508	0,003899
10							0,046761	0,062613	0,004403	0,023238	0,031799	0,002085
11							0,049124	0,064797	0,007667	0,023907	0,034141	0,003048
12							0,060324	0,071771	0,016379	0,028231	0,034381	0,007151
13							0,049120	0,064797	0,007679	0,024033	0,034429	0,003250
14							0,058385	0,067552	0,014312	0,026969	0,035557	0,006159
15							0,117196	0,441630	0,035536	0,040325	0,114480	0,012794
16										0,041526	0,130694	0,014533
17										0,027459	0,035818	0,005669
18										0,024529	0,034592	0,003319
19										0,027981	0,034740	0,005992
20										0,024385	0,034619	0,003033
21										0,023039	0,031827	0,002009
22										0,024354	0,032144	0,003788
23										0,026578	0,037840	0,004845
24										0,032315	0,041380	0,009645
25										0,026459	0,033442	0,005015
26										0,023524	0,032178	0,002318
27										0,025218	0,037852	0,004007
28										0,032092	0,042757	0,009228
29										0,025087	0,037261	0,003770
30										0,031416	0,041191	0,008152
31										0,076871	0,412081	0,018984



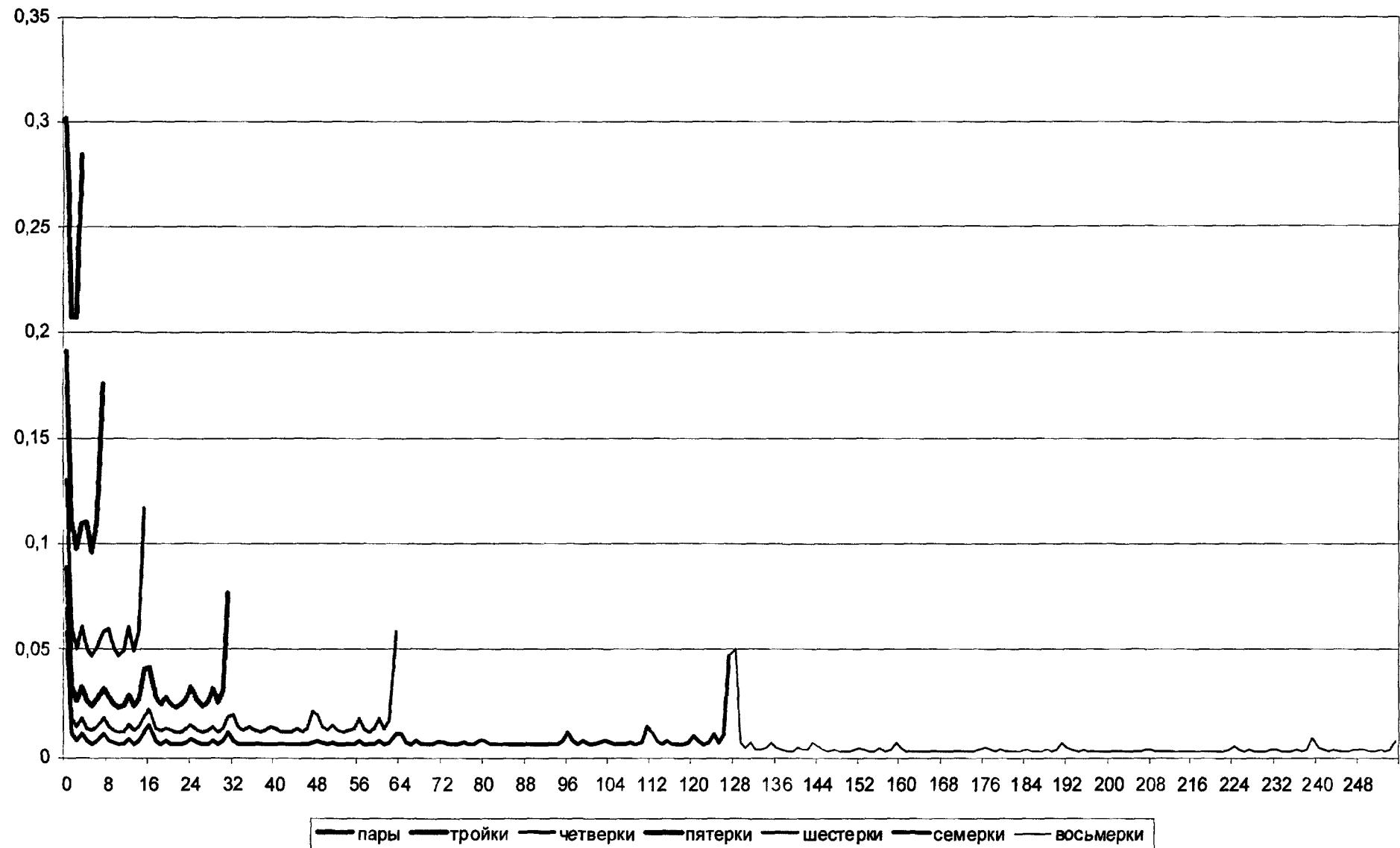


Рисунок Б.5 – Сводный график вероятности появления пар, троек, четверок, пятерок, шестерок, семерок и восьмерок битов в 1-3 битовых слоях графических изображений

Некоторые замены битов, произведенные согласно базовой кодовой таблице МСБП, приводят к изменению соотношения нулей и единиц. Однако элементы таблицы подобраны таким образом, чтобы взаимокомпенсировать вносимые искажения. В результате экспериментов было определено, что максимальное отклонение не превышает 1.45%. Но следует отметить, что указанные отклонения являются максимальным, и полученным путем использования в качестве скрываемых данных постоянных и строго периодических битовых строк. В реальных ситуациях указанные отклонения оказались значительно ниже, порядка 0.05%.

Вследствие того, что предложенный метод вносит изменения только в точки изображения расположенные на границах областей, искажения, вносимые в общую гистограмму изображения, оказываются малозаметными. В качестве примера на рисунке Б.6 приведены гистограммы исходного и заполненного контейнера. Для большей наглядности точности их совпадения на крайнем правом рисунке приведена разность двух гистограмм.

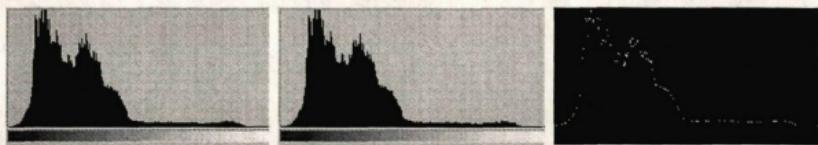


Рисунок Б.6 – Гистограммы исходного контейнера, заполненного контейнера и их разность.

На рисунке Б.7 представлены гистограммы разницы между распределением цветов исходного изображения и изображения содержащего дополнительную информацию в отношении к гистограмме исходного изображения. В первом случае внедрение информации осуществлялось с помощью стандартного метода замены младших битов, а во втором с помощью базового метода сдвига битовых последовательностей.

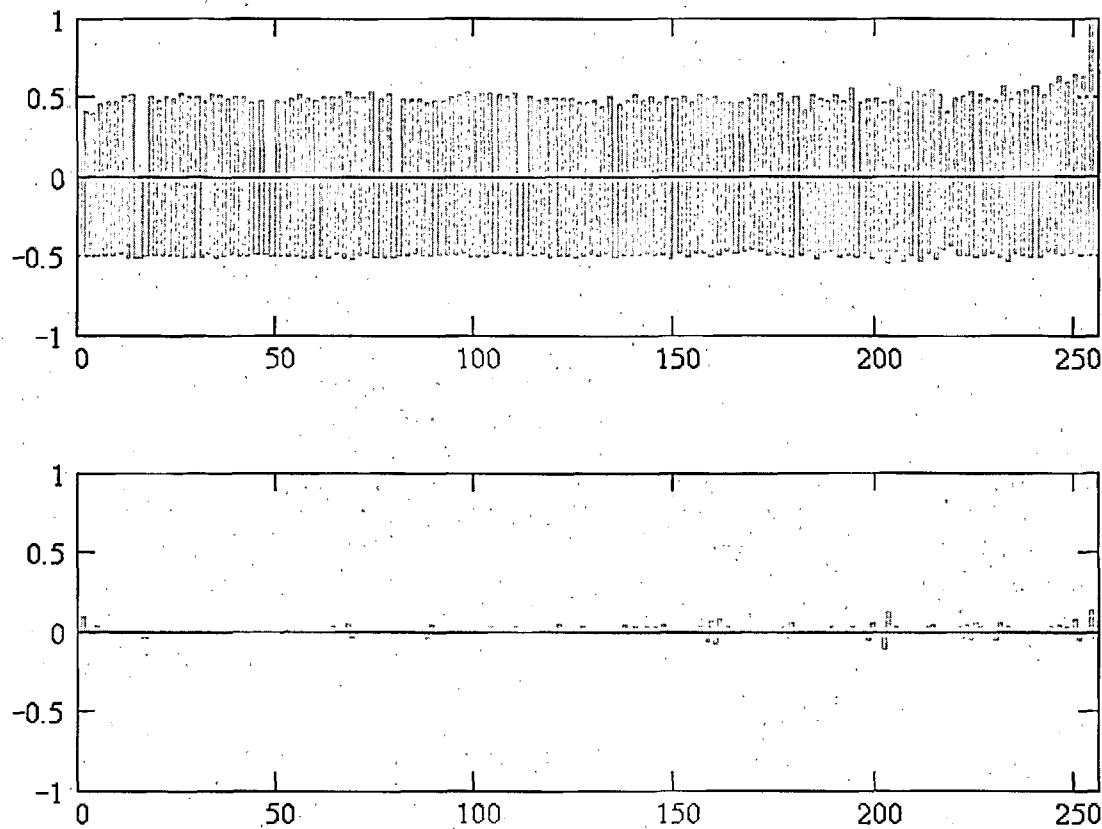


Рисунок Б.7 – Гистограммы разницы гистограмм исходного изображения и изображения содержащего дополнительную информацию. Верхний график – стандартный метод LSB.
Нижний график – метод LSB+МСБП

Из приведенного примера видно, что искажения, вносимые с использованием метода сдвига битовых последовательностей, оказывают гораздо меньшее воздействие на общую гистограмму распределения цветов обрабатываемого изображения. В проведенном эксперименте максимальное искажение гистограммы для метода LSB составило 130%, среднее 50.4%, для метода сдвига битовых последовательностей соответственно 13.5% и 1.9%. Таким образом, искажения, вносимые с использование МСБП, оказываются на порядок меньше искажений, вносимых в гистограмму при использовании стандартного метода LSB. Результаты проведенных экспериментов показывают, что искажения порядка нескольких процентов не выявляются методом χ^2 и, следовательно, факт записи информации методом сдвига

битовых последовательностей не может быть установлен. Было установлено, что LSB+МСБП не выявляется с помощью критерия χ^2 , вероятность использования стегосистемы согласно данному критерию стремиться к нулю с точностью до $3 \cdot 10^{-15}$, в то время, как метод замены младших битов был обнаружен во всех случаях с вероятностью близкой к единице.

В результате сильного сжатия графического изображения алгоритмом JPEG в младших битовых слоях могут возникать четкие прямоугольные области. Особенностью МСБП является то, что он осуществляет запись информации в точках перехода битовых последовательностей. В результате, если в младших слоях изображения будут встречены вертикальные линии, то они в результате встраивания данных покроются рябью. Пример подобной ситуации представлен на рисунке Б.8.



Рисунок Б.8. Слева – изображение сильно сжатое методом JPEG, в центре – его третий битовый срез, справа – битовый срез изображения после внедрения информации.

Как видно по нижней части срезов изображения нельзя сказать, в каком из них записана информация, и записана ли вообще. Однако на четких прямоугольных участках верхней части срезов появляется возможность выявления LSB+МСБП. Причина возможности выявления скрытой информации кроится в нарушении одного из главных требований, предъявляемых к битовым строкам при использовании МСБП. Данное требование заключается в том, что граничные биты в битовой строке должны

быть случайными. В приведенном же примере данное условие выполняется только для нижней части изображения.

В заключение, отметим, что метод замены младших значащих битов к настоящему времени можно считать уже отжившим свой век. Не так давно рядом исследователей были предложены новые революционные методы стеганоанализа, которые учитывают структуру изображения и связи между его элементами – корреляционные методы анализа. С их появлением дальнейшее развитие метода замены младших значащих битов приведет в тупик. Поэтому в данной работе вместо метода замены младших значащих битов, совместно с МСБП предлагается применять стеганографические методы нового поколения, такие как методы на основе пространственно частотных фильтров представленные в пункте 3.3.2 данной работы.

Приложение В. Таблицы кодов символов с одинаковым начертанием

Ниже приведены таблицы символов с идентичными глифами для кодовых таблиц различных наиболее популярных кириллических кодировок.

Таблица В.1 – 866 – MS-DOS

Символ английский	Код символа ASCII	Символ русский	Код символа 866 MS-DOS
A	65	А	128
B	66	В	130
C	67	С	145
E	69	Е	133
Н	72	Н	141
К	75	К	138
М	77	М	140
О	79	О	142
Р	80	Р	144
Т	84	Т	146
Х	88	Х	149
а	97	а	160
с	99	с	225
е	101	е	165
о	111	о	174
р	112	р	224
х	120	х	229

Таблица В.2 – MS Windows – 1251

Символ английский	Код символа ASCII	Символ русский	Код символа Win. 1251
A	65	А	192
B	66	В	194
C	67	С	209
E	69	Е	197
Н	72	Н	205
М	77	М	204
О	79	О	206
Р	80	Р	208
Т	84	Т	210

X	88	X	213
а	97	а	224
с	99	с	241
е	101	е	229
о	111	о	238
р	112	р	240
х	120	х	245
у	121	у	243

Таблица В.3 – KOI8-R

Символ английский	Код символа ASCII	Символ русский	Код символа KOI8-R
A	65	А	225
B	66	Б	247
C	67	С	243
E	69	Е	229
H	72	Н	238
M	77	М	237
O	79	О	239
P	80	Р	242
T	84	Т	244
X	88	Х	232
а	97	а	193
с	99	с	211
е	101	е	197
о	111	о	207
р	112	р	210
х	120	х	200
у	121	у	213

Таблица В.4 – ISO 8859-5

Символ английский	Код символа ASCII	Символ русский	Код символа ISO 8859-5
A	65	А	176
B	66	Б	178
C	67	С	193
E	69	Е	181
H	72	Н	189

M	77	M	188
O	79	O	190
P	80	P	192
T	84	T	194
X	88	X	197
a	97	a	208
c	99	c	225
e	101	e	213
o	111	o	222
p	112	p	224
x	120	x	229
y	121	y	227

Приложение Г. Способ борьбы с инъективными ошибками в каналах передачи данных малой пропускной способности

При организации скрытого канала связи с использованием стеганографических методов часто приходиться сталкиваться с ситуацией, когда контейнер передается только в одном направлении. Если используемый для передачи контейнера канал связи является каналом с гарантированной доставкой, то опасаться возможных искажений контейнера в ходе его передачи не приходится. Контейнеры на стороне получателя и отправителя будут идентичны, а, следовательно, сохраниться и скрытая в них информация. Использование же каналов допускающих определенный процент ошибок может полностью разрушить скрытую информацию.

В цифровых системах связи для борьбы с ошибками в канале используются специальные коды исправления ошибок. Существующие коды исправления ошибок позволяют эффективно бороться с аддитивными ошибками, т.е. с ошибками искажения отдельных битов передаваемого сообщения. Но эти коды не способны противостоять инъективным ошибкам в канале связи, при которых происходит добавление или потеря битов в битовом потоке при его передаче по каналу связи. Ошибки такого типа, в случае использования блочных или циклических кодов, приводят к ошибкам декодирования последующего потока битов, что в свою очередь приводит к необходимости повторной передачи практически всего сообщения. Следует отметить, что инъективные ошибки являются характерными для стеганографических каналов связи, так как информация часто передается на уровне шумовой составляющей несущего сигнала (контейнера).

Стандартным способом устранения данного недостатка является способ, при котором осуществляется периодическая вставка в битовый поток специальных меток синхронизации. Однако для обеспечения необходимой точности в обнаружении меток синхронизации их длина должна быть

достаточно большой. В случае малой пропускной способности канала такой подход приведет к большой загруженности канала.

В качестве альтернативы представленному решению предлагается способ построения самосинхронизирующегося канала связи. Основная идея заключается в использовании относительно небольших объемов дополнительной информации в каждом отдельном пакете данных. Каждый пакет данных снабжается короткой битовой строкой содержащей свертку или контрольную сумму от данных пакета и текущего номера пакета в передаваемой последовательности. Для борьбы с аддитивными ошибками, каждый пакет, после его формирования, подвергается кодированию блочным кодом коррекции ошибок. В проведенных экспериментах были использованы коды класса циклических кодов БЧХ (Боуза-Чоудхури-Хоквенгема) для двоичного алфавита с длиной кодового слова $n = 2^m - 1$, где $m = 5, 6, \dots$ и числом гарантированно исправляемых ошибок $t \geq 3$. После наложения кодов коррекции ошибок, на последнем шаге, блоки склеиваются в единую битовую последовательность, как показано на рисунке Г.1.

Исходная битовая последовательность:

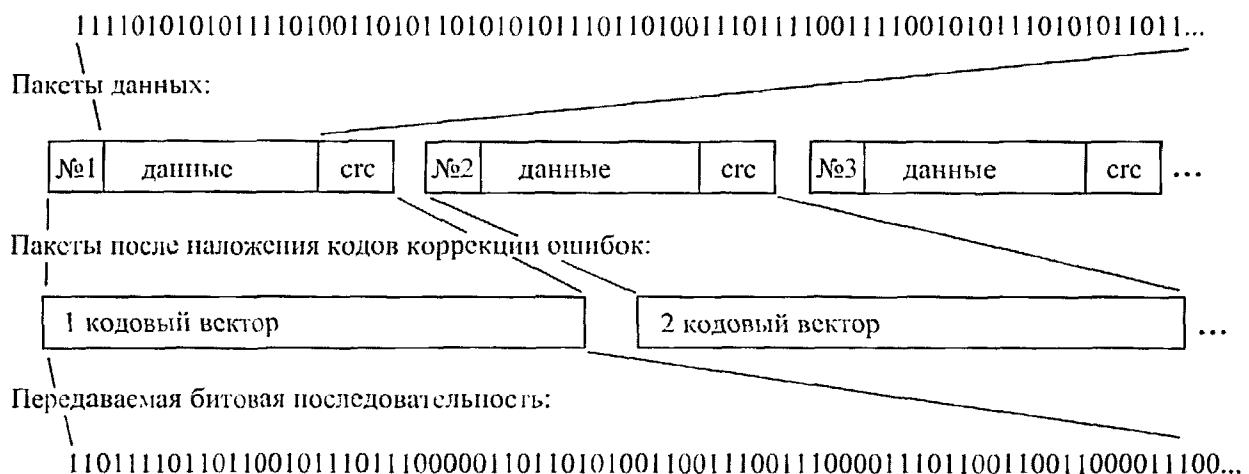


Рисунок Г.1 – Формирование передаваемой битовой последовательности

На принимающей стороне процесс восстановления передаваемых данных представляет собой итеративный процесс с последовательно выполняемыми шагами декодирования и проверки контрольной суммы.

На первом шаге, с позиции предполагаемого начала сообщения осуществляется декодирование первого блока сообщения длиной n бит. Если момент начала передачи данных известен с определенной точностью, то использование специальной метки синхронизации в начале передаваемой битовой последовательности не является необходимым, так как декодирование допускает определенное смещение начала серии информационных блоков. После декодирования блока осуществляется проверка контрольной суммы, и в случае, если контрольная сумма совпадает, то пакет считается принятым верно и с заданной вероятностью нами обнаружено начало потока передаваемых данных. Вероятность достоверного обнаружения начала передаваемых данных определяется размерностью контрольной суммы или длиной используемой в качестве ее свертки сообщения. Если проверка контрольной суммы дала отрицательный результат, то выбирается следующая наиболее вероятная точка начала потока данных и весь процесс повторяется заново. Так до тех пор, пока проверка контрольной суммы не закончится с положительным результатом, или не будет достигнут конец битового потока.

В случае успешного декодирования первого блока сообщения, по полу номер блока, определяется, является ли данный блок первым в передаваемом информационном сообщении. Если номер первого блока отличен от единицы и равен некоторому числу k , то считается что первые $k - 1$ блоков были повреждены и не подлежат декодированию.

После первого успешно декодированного блока переходят к обработке следующих блоков. Так как в канале передачи данных помимо аддитивных, возможны и инъективные ошибки начало следующего блока может оказаться смещенным относительно начала текущего блока на величину $n + \Delta$, где n – длина кодового вектора для выбранного кода коррекции ошибок, Δ – смещение, вызванное возможными инъективными ошибками. Величина смещения Δ зависит от вероятности и типа инъективной ошибки. С целью ускорения обработки и повышения достоверности декодирования, при малой

длине контрольной суммы, для текущего декодируемого блока строится массив смещений, упорядоченный по дереву Хаффмана в соответствии с вероятностями инъективной ошибки добавления и удаления символов, а также числом последних предшествующих поврежденных блоков. Пример демонстрирующий данный принцип приведен на рисунке Г.2. Как видно из примера, после каждого поврежденного блока фактически выполняется процедура автоматической синхронизации потока.

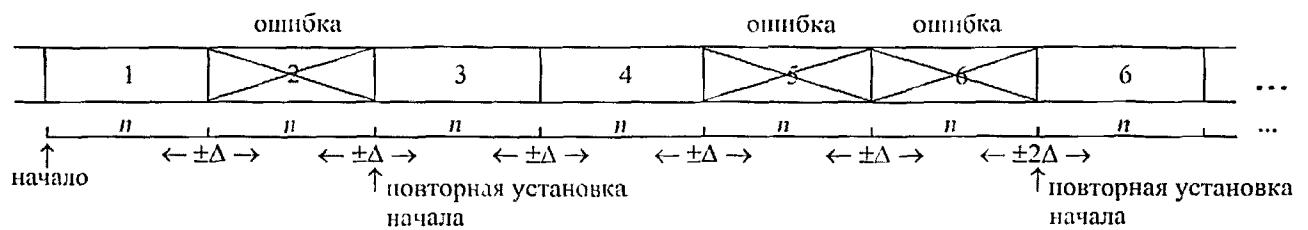


Рисунок Г.2 – Выделение блоков и автосинхронизация

Если в канале связи не было дублирования пакетов, то после декодирования всей принятой битовой последовательности формируется запрос на повторную передачу, содержащий список номеров поврежденных пакетов. Если же дублирование пакетов проводилось с учетом возможных искажений в канале связи, возможно полное восстановление передаваемого сообщения за одну итерацию. Предложенный подход может найти отражение в цифровых каналах связи. Использование данного подхода позволяет значительно сократить простаивание канала и временные затраты, связанные с ожиданием подтверждения доставки, как в стандартной системе ARQ. Кроме того, такой подход позволяет более эффективно использовать линию связи в полудуплексном режиме, когда канал передачи данных, монопольно занимается только одной передающей стороной, что в свою очередь позволяет расширить полосу пропускания и увеличить скорость передачи данных, с одновременным сокращением аддитивных ошибок.