В.Г. ГРИБУНИН, И.Н. ОКОВ, И.В. ТУРИНЦЕВ

# ЦИФРОВАЯ СТЕГАНОГРАФИЯ

АСПЕКТЫ ЗАЩИТЫ

**В.Г. Грибунин, И.Н. Оков, И.В. Туринцев**

Цифровая стеганография / В.Г. Грибунин, И.Н. Оков, И.В. Туринцев — М.: СОЛОН-ПРЕСС, 2009— 272 с. (Серия «Аспекты защиты»)

ISBN 5-98003-011-5

Интерес к стеганографии появился в последнее десятилетие и вызван широким распространением мультимедийных технологий. Методы стеганографии позволяют не только скрытно передавать данные, но и решать задачи помехоустойчивой аутентификации, защиты информации от несанкционированного копирования, отслеживания распространения информации по сетям связи, поиска информации в мультимедийных базах данных.

На русском языке стеганографии было посвящено несколько обзорных статей в спецвыпуске журнала «Конфидент» — и только. Данная книга призвана восполнить существующий пробел. В ней обобщены самые последние результаты исследований зарубежных ученых, рассмотрены как теоретические, так и практические аспекты стеганографии, выполнена классификация стегосистем и методов встраивания, детально исследованы вопросы повышения пропускной способности стегоканала, обеспечения стойкости и незаметности внедрения, приведено более 50 алгоритмов встраивания данных.

Книга предназначена для студентов, аспирантов, научных работников, изучающих вопросы защиты информации, а также для инженеров-проектировщиков средств защиты информации. Несомненный интерес она вызовет также у специалистов в области теории информации и цифровой обработки сигналов.

ISBN 5-98003-011-5

# Оглавление

# 1. Введение в цифровую стеганографию

## 1.1. Цифровая стеганография. Предмет, терминология, области применения

Цифровая стеганография как наука родилась буквально в последние годы. По нашему мнению, она включает следующие направления:

1) встраивание информации с целью ее скрытой передачи;

2) встраивание цифровых водяных знаков (ЦВЗ) (watermarking);

3) встраивание идентификационных номеров (fingerprinting);

4) встраивание заголовков (captioning).

ЦВЗ могут применяться в основном для защиты от копирования и несанкционированного использования. В связи с бурным развитием технологий мультимедиа остро встал вопрос защиты авторских прав и интеллектуальной собственности, представленной в цифровом виде. Примерами могут являться фотографии, аудио- и видеозаписи и так далее. Преимущества, которые дают представление и передача сообщений в цифровом виде, могут оказаться перечеркнутыми легкостью, с которой возможно их воровство или модификация. Поэтому разрабатываются различные меры защиты информации организационного и технического характера. Одно из наиболее эффективных технических средств защиты мультимедийной информации и заключается во встраивании в защищаемый объект невидимых меток ЦВЗ. Разработки в этой области ведут крупнейшие фирмы во всем мире. Так как методы ЦВЗ начали разрабатываться совершенно недавно (первой статьей на эту тему была, видимо, работа [1]), то здесь имеется много неясных проблем, требующих своего разрешения.

Название этот метод получил от всем известного способа защиты ценных бумаг, в том числе и денег, от подделки. Термин «digital watermarking» был впервые применен в работе [2]. В отличие от обычных водяных знаков ЦВЗ могут быть не только видимыми, но и (как правило) невидимыми. Невидимые ЦВЗ анализируются специальным декодером, который выносит решение об их корректности. ЦВЗ могут содержать некоторый аутентичный код, информацию о собственнике либо какую-нибудь управляющую информацию. Наиболее подходящими объектами защиты при помощи ЦВЗ являются неподвижные изображения, файлы аудио- и видеоданных.

Технология встраивания идентификационных номеров производителей имеет много общего с технологией ЦВЗ. Отличие заключается в том, что в первом случае каждая защищенная копия имеет свой уникальный встраиваемый номер (отсюда и название — дословно «отпечатки пальцев»). Этот идентификационный номер позволяет производителю отслеживать дальнейшую судьбу своего детища: не занялся ли кто-нибудь из покупателей незаконным тиражированием. Если да, то «отпечатки пальцев» быстро укажут на виновного.

Встраивание заголовков (невидимое) может применяться, например, для подписи медицинских снимков, нанесения легенды на карту и в других случаях. Целью является хранение разнородно представленной информации в едином целом. Это, пожалуй, единственное приложение стеганографии, где в явном виде отсутствует потенциальный нарушитель.

Так как цифровая стеганография является молодой наукой, то ее терминология не до конца устоялась. Основные понятия были согласованы на Первой международной конференции по скрытию данных [3]. Тем не менее даже само понятие «стеганография» трактуется различно. Так, некоторые исследователи понимают под стеганографией только скрытую передачу информации. Другие относят к стеганографии такие приложения, как, например, метеорная радиосвязь, радиосвязь с псевдослучайной перестройкой радиочастоты, широкополосная радиосвязь. На наш взгляд, неформальное определение того, что такое цифровая стеганография, могло бы выглядеть следующим образом: «Наука о незаметном и надежном скрытии одних битовых последовательностей в других, имеющих аналоговую природу». Под это определение как раз подпадают все четыре вышеприведенных направления скрытия данных, а приложения радиосвязи нет. Кроме того, в определении содержится два главных требования к стеганографическому преобразованию: незаметность и надежность, то есть устойчивость к различного рода искажениям. Упоминание об аналоговой природе цифровых данных подчеркивает тот факт, что встраивание информации производится в оцифрованные непрерывные сигналы. Таким образом, в рамках цифровой стеганографии не рассматриваются вопросы внедрения данных в заголовки IP-пакетов и файлов различных форматов, в текстовые сообщения.

Как бы ни были различны направления стеганографии, предъявляемые ими требования во многом совпадают, что будет показано далее. Наиболее существенное отличие постановки задачи скрытой передачи данных от постановки задачи встраивания ЦВЗ состоит в том, что в первом случае нарушитель должен обнаружить скрытое сообщение, тогда как во втором случае о его существовании все знают. Более того, у нарушителя на законных основаниях может иметься устройство обнаружения ЦВЗ (например, в составе DVD-проигрывателя).

Слово «незаметном» в нашем определении цифровой стеганографии подразумевает обязательное включение человека в систему стеганографической передачи данных. Человек здесь может рассматриваться как дополнительный приемник данных, предъявляющий к системе передачи достаточно трудно формализуемые требования.

Задачу встраивания и выделения сообщений из другой информации выполняет стегосистема, состоящая из следующих основных элементов, представленных на рис. 1.1:

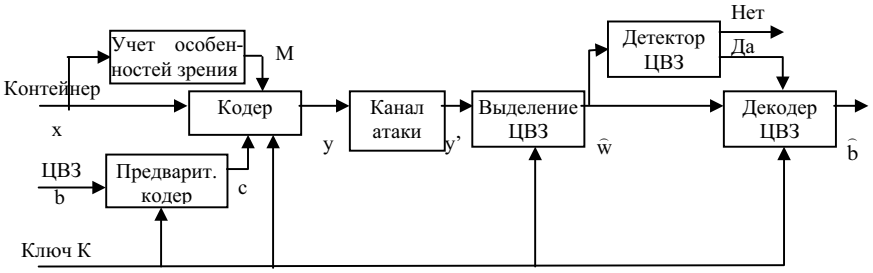


Рис. 1.1. Структурная схема типичной стегосистемы ЦВЗ

— прекодер — устройство, предназначенное для преобразования скрываемого сообщения к виду, удобному для встраивания в сигнал-контейнер (контейнером называется информационная последовательность, в которой прячется сообщение);

— стегокодер — устройство, предназначенное для осуществления вложения скрытого сообщения в другие данные с учетом их модели;

— устройство выделения встроенного сообщения;

— стегодетектор — устройство, предназначенное для определения наличия стегосообщения;

— декодер — устройство, восстанавливающее скрытое сообщение. Этот узел может отсутствовать.

Данные, содержащие скрытое сообщение, могут подвергаться преднамеренным атакам или случайным помехам, описание которых приведено в главе 3.

Как показано на рис. 1.1, в стегосистеме происходит объединение двух типов информации так, чтобы они могли быть различимы двумя принципиально разными детекторами. В качестве одного из детекторов выступает система выделения ЦВЗ, в качестве другого — человек.

Прежде чем осуществить вложение ЦВЗ в контейнер, ЦВЗ должен быть преобразован в некоторый подходящий вид. Например, если в качестве контейнера выступает изображение, то и последовательность ЦВЗ зачастую представляется как двумерный массив бит. Для того чтобы повысить устойчивость ЦВЗ к искажениям, нередко выполняют его помехоустойчивое кодирование либо применяют широкополосные сигналы. Первоначальную обработку скрытого сообщения выполняет показанный на рис. 1.1 прекодер.

В качестве важнейшей предварительной обработки ЦВЗ (а также и контейнера) назовем вычисление его обобщенного преобразования Фурье. Это позволяет осуществить встраивание ЦВЗ в спектральной области, что значительно повышает его устойчивость к искажениям. Предварительная обработка часто выполняется с использованием ключа $K$ для повышения секретно-

сти встраивания. Далее ЦВЗ «вкладывается» в контейнер, например путем модификации младших значащих бит коэффициентов. Этот процесс возможен благодаря особенностям системы восприятия человека. Хорошо известно, что изображения обладают большой психовизуальной избыточностью. Глаз человека подобен низкочастотному фильтру, пропускающему мелкие детали. Особенно незаметны искажения в высокочастотной области изображений. Эти особенности человеческого зрения используются, например, при разработке алгоритмов сжатия изображений и видео.

Процесс внедрения ЦВЗ также должен учитывать свойства системы восприятия человека. Стеганография использует имеющуюся в сигналах психовизуальную избыточность, но другим, чем при сжатии данных, образом. Приведем простой пример. Рассмотрим полутоновое изображение с 256 градациями серого, то есть с удельной скоростью кодирования 8 бит/пиксел. Хорошо известно, что глаз человека не способен заметить изменение младшего значащего бита. Еще в 1989 году был получен патент на способ скрытого вложения информации в изображение путем модификации младшего значащего бита. В данном случае детектор стего анализирует только значение этого бита для каждого пиксела, а глаз человека, напротив, воспринимает только старшие 7 бит. Этот метод прост в реализации и эффективен, но не удовлетворяет некоторым важным требованиям к ЦВЗ.

В большинстве стегосистем для внедрения и выделения ЦВЗ используется ключ. Ключ может быть предназначен для узкого круга лиц или же быть общедоступным. Например, ключ должен содержаться во всех DVD-плейерах, чтобы они могли прочесть содержащиеся на дисках ЦВЗ. Иногда по аналогии с криптографией стегосистемы делят на два класса: с открытым и с секретным ключом. На наш взгляд, аналогия неверна, так как понятие открытого ключа в данном случае в корне различно. Правильным выражением было бы «общедоступный ключ», причем ключ встраивания совпадает с ключом выделения. Не существует, насколько известно, стегосистемы, в которой бы при выделении ЦВЗ требовалась другая информация, чем при его вложении. Хотя и не доказана гипотеза о невозможности существования подобной системы. В системе с общедоступным ключом достаточно сложно противостоять возможным атакам со стороны злоумышленников. В самом деле, в данном случае нарушителю точно известны ключ и месторасположение ЦВЗ, а также его значение.

В стегодетекторе происходит обнаружение ЦВЗ в (возможно, измененном) защищенном ЦВЗ изображении. Это изменение может быть обусловлено влиянием ошибок в канале связи, операций обработки сигнала, преднамеренных атак нарушителей. Во многих моделях стегосистем сигнал-контейнер Выход детектора рассматривается как аддитивный шум [4]. Тогда задача обнаружения и выделения стегосообщения является классической для теории связи. Однако такой подход не учитывает двух факторов: неслучайного ха-

рактера сигнала контейнера и требований по сохранению его качества. Эти моменты не встречаются в известной теории обнаружения и выделения сигналов на фоне аддитивного шума. Их учет позволит построить более эффективные стегосистемы.

Различают стегодетекторы, предназначенные для обнаружения наличия ЦВЗ и устройства для выделения этого ЦВЗ (стегодекодеры). В первом случае возможны детекторы с жесткими (да/нет) или мягкими решениями. Для вынесения решения о наличии/отсутствии ЦВЗ удобно использовать такие меры, как расстояние по Хэммингу, либо взаимную корреляцию между имеющимся сигналом и оригиналом (при наличии последнего, разумеется). А что делать, если у нас нет исходного сигнала? Тогда в дело вступают более тонкие статистические методы, основанные на построении моделей исследуемого класса сигналов.

В зависимости от того, какая информация требуется детектору для обнаружения ЦВЗ, стегосистемы ЦВЗ делятся на три класса: открытые, полузакрытые и закрытые. Эта классификация приведена в табл. 1.1.

**Таблица 1.1. Классификация систем встраивания ЦВЗ**

| | | Что требуется детектору | | Выход детектора | |
|---|---|---|---|---|---|
| | | Исходный сигнал | Исходный ЦВЗ | Да/Нет | ЦВЗ |
| Закрытые | Тип I | + | + | + | – |
| | Тип II | + | – | – | + |
| Полузакрытые | | – | + | + | – |
| Открытые | | – | – | – | + |

Наибольшее применение могут иметь открытые стегосистемы ЦВЗ, которые аналогичны системам скрытой передачи данных. Наибольшую устойчивость по отношению к внешним воздействиям имеют закрытые стегосистемы I типа.

Рассмотрим подробнее понятие «контейнера». До стегокодера — это пустой контейнер, после него — заполненный контейнер, или стего. Стего должен быть визуально неотличим от пустого контейнера. Различают два основных типа контейнеров: потоковый и фиксированный.

Потоковый контейнер представляет собой непрерывно следующую последовательность бит. Сообщение вкладывается в него в реальном масштабе времени, так что в кодере неизвестно заранее, хватит ли размеров контейнера для передачи всего сообщения. В один контейнер большого размера может быть встроено и несколько сообщений. Интервалы между встраиваемыми битами определяются генератором псевдослучайной последовательности с

равномерным распределением интервалов между отсчетами. Основная трудность заключается в осуществлении синхронизации, определении начала и конца последовательности. Если в данных контейнера имеются биты синхронизации, заголовки пакетов и т. д., то скрываемая информация может идти сразу после них. Трудность обеспечения синхронизации превращается в достоинство с точки зрения обеспечения скрытности передачи. Кроме того, потоковый контейнер имеет большое практическое значение: представьте себе, например, стегоприставку к обычному телефону. Под прикрытием обычного, незначащего телефонного разговора можно было бы передавать другой разговор, данные и т. п., а не зная секретного ключа, нельзя было бы не только узнать содержание скрытой передачи, но и сам факт ее существования. Не случайно, что открытых работ, посвященных разработке стегосистем с потоковым контейнером, практически не встречается.

У фиксированного контейнера размеры и характеристики заранее известны. Это позволяет осуществлять вложение данных оптимальным в некотором смысле образом. В книге мы будем рассматривать в основном фиксированные контейнеры (далее — контейнеры).

Контейнер может быть выбранным, случайным или навязанным. Выбранный контейнер зависит от встраиваемого сообщения, а в предельном случае является его функцией. Этот тип контейнера больше характерен для стеганографии. Навязанный контейнер может появиться в сценарии, когда лицо, предоставляющее контейнер, подозревает о возможной скрытой переписке и желает предотвратить ее. На практике же чаще всего сталкиваются со случайным контейнером.

Встраивание сообщения в контейнер может производиться при помощи ключа — одного или нескольких. Ключ — псевдослучайная последовательность (ПСП) бит, порождаемая генератором, удовлетворяющим определенным требованиям (криптографически безопасный генератор). В качестве основы генератора может использоваться, например, регистр сдвига с линейной обратной связью. Тогда адресатам для обеспечения связи может сообщаться начальное заполнение этого регистра. Числа, порождаемые генератором ПСП, могут определять позиции модифицируемых отсчетов в случае фиксированного контейнера или интервалы между ними в случае потокового контейнера. Надо отметить, что метод случайного выбора величины интервала между встраиваемыми битами не особенно хорош. Причин — две. Во-первых, скрытые данные должны быть распределены по всему изображению. Поэтому равномерное распределение длин интервалов (от наименьшего до наибольшего) может быть достигнуто лишь приближенно: мы должны быть уверены в том, что все сообщение встроено, то есть «поместилось» в контейнер. Во-вторых, длины интервалов между отсчетами шума распределены не по равномерному, а по экспоненциальному закону. Генератор же ПСП с экспоненциально распределенными интервалами сложен в реализации.

Скрываемая информация внедряется в соответствии с ключом в те отсчеты, искажение которых не приводит к существенным искажениям контейнера. Эти биты образуют стегопуть. В зависимости от приложения под существенным можно понимать искажение, приводящее как к неприемлемости для человека-адресата заполненного контейнера, так и к возможности выявления скрытого сообщения после стегоанализа.

ЦВЗ могут быть трех типов: робастные, хрупкие и полухрупкие (semifragile). Под робастностью понимается устойчивость ЦВЗ к различного рода воздействиям на стего. Робастным ЦВЗ посвящено большинство исследований.

Хрупкие ЦВЗ разрушаются при незначительной модификации заполненного контейнера. Они применяются для аутентификации сигналов. Отличие от средств электронной цифровой подписи заключается в том, что хрупкие ЦВЗ все же допускают некоторую модификацию контента. Это важно для защиты мультимедийной информации, так как законный пользователь может, например, пожелать сжать изображение. Другое отличие заключается в том, что хрупкие ЦВЗ должны не только отразить факт модификации контейнера, но также вид и местоположение этого изменения.

Полухрупкие ЦВЗ устойчивы по отношению к одним воздействиям и не устойчивы — к другим. Вообще говоря, все ЦВЗ могут быть отнесены к этому типу. Однако полухрупкие ЦВЗ специально проектируются так, чтобы быть неустойчивыми по отношению к определенного рода операциям. Например, они могут позволять выполнять сжатие изображения, но запрещать вырезку из него или вставку в него фрагмента.

На рис. 1.2 представлена классификация систем цифровой стеганографии.

Стегосистема образует стегоканал, по которому передается заполненный контейнер. Этот канал считается подверженным воздействиям со стороны нарушителей. Следуя Симмонсу [5], в стеганографии обычно рассматривается такая постановка задачи («проблема заключенных»).

Двое заключенных, Алиса и Боб, желают конфиденциально обмениваться сообщениями, несмотря на то, что канал связи между ними контролирует охранник Вилли. Для того чтобы тайный обмен сообщениями был возможен, предполагается, что Алиса и Боб имеют известный обоим секретный ключ. Действия Вилли могут заключаться не только в попытке обнаружить скрытый канал связи, но и разрушить передаваемые сообщения, а также их модифицировать и создать новые, ложные сообщения. Соответственно, можно выделить три типа нарушителей, которым должна противостоять стегосистема: пассивный, активный и злоумышленный. Подробнее возможные действия нарушителей и защита от них рассмотрены во второй главе. Пока заметим лишь, что пассивный нарушитель может быть лишь в стегосистемах скрытой передачи данных. Для систем ЦВЗ характерны активные и злоумышленные нарушители.

Статья Симмонса [5], как он сам написал впоследствии [6], была вызвана желанием привлечь внимание научной общественности к закрытой в то время проблеме, связанной с контролем над ядерным оружием. Согласно Договору ОСВ СССР и США должны были разместить некие датчики на стратегических ракетах друг друга, которые передавали бы информацию о том, не подсоединена ли к ним ядерная боеголовка. Проблема, которой занимался Симмонс, заключалась в том, чтобы не допустить передачи какой-либо другой информации этими датчиками, например о местоположении ракет. Определение факта наличия скрытой информации — главная задача стегоанализа.


Рис. 1.2. Классификация систем цифровой стеганографии

Для того, чтобы стегосистема была надежной, при ее проектировании необходимо выполнение ряда требований.

— Безопасность системы должна полностью определяться секретностью ключа. Это означает, что нарушитель может знать все алгоритмы работы стегосистемы и статистические характеристики множества сообщений и контейнеров, что, однако, не даст ему никакой дополнительной информации о наличии или отсутствии сообщения в данном контейнере.

— Знание нарушителем факта наличия сообщения в каком-либо контейнере не должно помочь ему при обнаружении сообщений в других контейнерах.

— Заполненный контейнер должен быть визуально неотличим от незаполненного. Для этого надо, казалось бы, внедрять скрытое сообщение в визуально незначимые области сигнала. Однако эти же области используют и алгоритмы сжатия. Поэтому если изображение будет в дальнейшем подвергаться сжатию, то скрытое сообщение может разрушиться. Следовательно, биты должны встраиваться в визуально значимые области, а относительная незаметность может быть достигнута за счет использования специальных методов, например модуляции с расширением спектра.

— Стегосистема ЦВЗ должна иметь низкую вероятность ложного обнаружения скрытого сообщения в сигнале, его не содержащем. В некоторых приложениях такое обнаружение приводит к серьезным последствиям. Например, ложное обнаружение ЦВЗ на DVD-диске может вызвать отказ от его воспроизведения плейером.

— Должна обеспечиваться требуемая пропускная способность (это требование актуально в основном для стегосистем скрытой передачи информации). В третьей главе мы введем понятие скрытой пропускной способности и рассмотрим пути ее достижения.

— Стегосистема должна иметь приемлемую вычислительную сложность реализации. При этом возможна асимметричная по сложности реализации система ЦВЗ, то есть сложный стегокодер и простой стегодекодер.

К ЦВЗ предъявляются следующие требования.

— ЦВЗ должен легко (вычислительно) извлекаться законным пользователем.

— ЦВЗ должен быть устойчивым либо неустойчивым к преднамеренным и случайным воздействиям (в зависимости от приложения). Если ЦВЗ используется для подтверждения подлинности, то недопустимое изменение контейнера должно приводить к разрушению ЦВЗ (хрупкий ЦВЗ). Если же ЦВЗ содержит идентификационный код, логотип фирмы и т. п., то он должен сохраниться при максимальных искажениях контейнера, конечно, не приводящих к существенным искажениям исходного сигнала. Например, у изображения могут быть отредактированы цветовая гамма или яркость, у аудиозаписи — усилено звучание низких тонов и т. д. Кроме того, ЦВЗ должен быть робастным по отношению к аффинным преобразованиям изображения, то есть его поворотам, масштабированию. При этом надо различать устойчивость самого ЦВЗ и способность декодера верно его обнаружить. Скажем, при повороте изображения ЦВЗ не разрушится, а декодер может оказаться неспособным выделить его. Существуют приложения, когда ЦВЗ должен быть устойчивым по отношению к одним преобразованиям и неустойчивым — к другим. Например, может быть разрешено копирование изображения (ксерокс, сканер), но наложен запрет на внесение в него каких-либо изменений.

— Должна иметься возможность добавления к стего дополнительных ЦВЗ. Например, на DVD-диске имеется метка о допустимости однократного копирования. После осуществления копирования необходимо добавить метку о запрете дальнейшего копирования. Можно было бы, конечно, удалить первый ЦВЗ и записать на его место второй. Однако это противоречит предположению о трудноудалимости ЦВЗ. Лучшим выходом является добавление еще одного ЦВЗ, после которого первый не будет приниматься во внимание. Но наличие нескольких ЦВЗ на одном сообщении может облегчить атаку со стороны нарушителя, если не предпринять специальных мер.

В настоящее время технология ЦВЗ находится в самой начальной стадии своего развития. Практика показывает, что должно пройти лет 10–20 для того, чтобы новый криптографический метод начал широко использоваться в обществе. Наверное, аналогичная ситуация будет наблюдаться и со стеганографией. Одной из проблем, связанных с ЦВЗ, является многообразие требований к ним, в зависимости от приложения. Рассмотрим подробнее основные области применения ЦВЗ.

Вначале рассмотрим проблему пиратства, или неограниченного неавторизованного копирования. Алиса продает свое мультимедийное сообщение Питеру. Хотя информация могла быть зашифрована во время передачи, ничто не помешает Питеру заняться ее копированием после расшифровки. Следовательно, в данном случае требуется дополнительный уровень защиты от копирования, который не может быть обеспечен традиционными методами. Как будет показано далее, существует возможность внедрения ЦВЗ, разрешающего воспроизведение и запрещающего копирование информации.

Важной проблемой является определение подлинности полученной информации, то есть ее аутентификация. Обычно для аутентификации данных используются средства цифровой подписи. Однако они не совсем подходят для обеспечения аутентификации мультимедийной информации. Дело в том, что сообщение, снабженное электронной цифровой подписью, должно храниться и передаваться абсолютно точно, «бит в бит». Мультимедийная же информация может незначительно искажаться как при хранении (за счет сжатия), так и при передаче (влияние одиночных или пакетных ошибок в канале связи). При этом ее качество остается допустимым для пользователя, но цифровая подпись работать не будет. Получатель не сможет отличить истинное, хотя и несколько искаженное, сообщение от ложного. Кроме того, мультимедийные данные могут быть преобразованы из одного формата в другой. При этом традиционные средства защиты целостности работать также не будут. Можно сказать, что ЦВЗ способны защитить именно содержание аудио-, видеосообщения, а не его цифровое представление в виде последовательности бит. Важным недостатком цифровой подписи является и то, что ее легко удалить из заверенного ею сообщения, после чего приделать к нему новую подпись. Удаление подписи позволит нарушителю отказаться от авторства либо

ввести в заблуждение законного получателя относительно авторства сообщения. Система ЦВЗ проектируется таким образом, чтобы исключить возможность подобных нарушений.

Как видно из рис. 1.3, применение ЦВЗ не ограничивается приложениями безопасности информации. Основные области использования технологии ЦВЗ могут быть объединены в четыре группы: защита от копирования (использования), скрытая аннотация документов, доказательство аутентичности информации и скрытая связь.


Рис. 1.3. Потенциальные области применения стеганографии

Популярность мультимедиа-технологий вызвала множество исследований, связанных с разработкой алгоритмов ЦВЗ для использования в стандартах MP3, MPEG-4, JPEG2000, защиты DVD дисков от копирования.

### 1.2. Встраивание сообщений в незначащие элементы контейнера

Цифровые изображения представляют собой матрицу пикселов. Пиксел — это единичный элемент изображения. Он имеет фиксированную разрядность двоичного представления. Например, пикселы полутонового изображения кодируются 8 битами (значения яркости изменяются от 0 до 255).

Младший значащий бит (LSB) изображения несет в себе меньше всего информации. Известно, что человек обычно не способен заметить изменение в этом бите. Фактически он является шумом. Поэтому его можно использовать для встраивания информации. Таким образом, для полутонового изображения объем встраиваемых данных может составлять 1/8 объема контейнера. Напри-

мер, в изображение размером $512 \times 512$ можно встроить 32 килобайта информации. Если модифицировать два младших бита (что также почти незаметно), то можно скрытно передать вдвое больший объем данных.

Достоинства рассматриваемого метода заключаются в его простоте и сравнительно большом объеме встраиваемых данных. Однако он имеет серьезные недостатки. Во-первых, скрытое сообщение легко разрушить, как это показано в третьей главе. Во-вторых, не обеспечена секретность встраивания информации. Нарушителю точно известно местоположение всего ЦВЗ. Для преодоления последнего недостатка было предложено встраивать ЦВЗ не во все пикселы изображения, а лишь в некоторые из них, определяемые по псевдослучайному закону в соответствии с ключом, известному только законному пользователю. Пропускная способность при этом уменьшается.

Рассмотрим подробнее вопрос выбора пикселов изображения для встраивания в них скрытого сообщения.

В работе [7] отмечается неслучайный характер поведения младшего значащего бита изображений. Скрываемое сообщение не должно изменять статистики изображения. Для этого в принципе возможно, располагая достаточно большим количеством незаполненных контейнеров, подыскать наиболее подходящий и даже найти контейнер, уже содержащий в себе наше сообщение при данном ключе. Тогда изменять вообще ничего не надо и вскрыть факт передачи будет невозможно. Эту ситуацию уместно сравнить с применением одноразового блокнота в криптографии. Метод выбора подходящего контейнера требует выполнения большого количества вычислений и обладает малой пропускной способностью.

Альтернативным подходом является моделирование характеристик поведения LSB. Встраиваемое сообщение будет в этом случае частично или полностью зависеть от контейнера. Процесс моделирования является вычислительно трудоемким, кроме того, его надо повторять для каждого контейнера. Главный недостаток этого метода — повторение процесса моделирования нарушителем, возможно обладающим большим вычислительным ресурсом, создающим лучшие модели. Результат — обнаружение скрытого сообщения. Это противоречит требованию о независимости безопасности стегосистемы от вычислительной мощности сторон. Кроме того, для обеспечения скрытности необходимо держать используемую модель шума в тайне. А нарушителю не известен должен быть лишь ключ.

В силу указанных трудностей на практике обычно ограничиваются поиском пикселов, модификация которых не вносит заметных искажений в изображение. Затем из этих пикселов в соответствии с ключом выбираются те, которые будут модифицироваться. Скрываемое сообщение шифруется с применением другого ключа. Этот этап может быть дополнен предварительной компрессией для уменьшения объема сообщения.

### 1.3. Математическая модель стегосистемы

Стегосистема может быть рассмотрена как система связи [8].

Алгоритм встраивания ЦВЗ состоит из трех основных этапов: 1) генерации ЦВЗ, 2) встраивания ЦВЗ в кодере и 3) обнаружения ЦВЗ в детекторе.

1. Пусть $W^*, K^*, I^*, B^*$ есть множества возможных ЦВЗ, ключей, контейнеров и скрываемых сообщений, соответственно. Тогда генерация ЦВЗ может быть представлена в виде

$$F:\ I^* \times K^* \times B^* \to W^*, \quad W = F(I, K, B), \tag{1.1}$$

где $W, K, I, B$ – представители соответствующих множеств. Вообще говоря, функция $F$ может быть произвольной, но на практике требования робастности ЦВЗ накладывают на нее определенные ограничения. Так, в большинстве случаев, $F(I, K, B) \approx F(I + \varepsilon, K, B)$, то есть незначительно измененный контейнер не приводит к изменению ЦВЗ. Функция $F$ обычно является составной:

$$F = T \circ G, \text{ где } G: K^* \times B^* \to C^* \text{ и } T: C^* \times I^* \to W^*, \tag{1.2}$$

то есть ЦВЗ зависит от свойств контейнера. Функция $G$ реализуется при помощи криптографически безопасного генератора ПСП с $K$ в качестве начального значения.

Для повышения робастности ЦВЗ могут применяться помехоустойчивые коды, например коды БЧХ, сверточные коды [9]. В ряде публикаций отмечены хорошие результаты, достигаемые при встраивании ЦВЗ в области вейвлет-преобразования с использованием турбо-кодов. Отсчеты ЦВЗ принимают обычно значения из множества $\{-1,1\}$, при этом для отображения $\{0,1\} \to \{-1,1\}$ может применяться двоичная относительная фазовая модуляция (BPSK).

Оператор $T$ модифицирует кодовые слова $C^*$, в результате чего получается ЦВЗ $W^*$. На эту функцию можно не накладывать ограничения необратимости, так как соответствующий выбор $G$ уже гарантирует необратимость $F$. Функция $T$ должна быть выбрана так, чтобы незаполненный контейнер $I_0$, заполненный контейнер $I_W$ и незначительно модифицированный заполненный контейнер $I'_W$ порождали бы один и тот же ЦВЗ:

$$T(C, I_0) = T(C, I_W) = T(C, I'_W), \tag{1.3}$$

то есть она должна быть устойчивой к малым изменениям контейнера.

2. Процесс встраивания ЦВЗ $W(i,j)$ в исходное изображение $I_0(i,j)$ может быть описан как суперпозиция двух сигналов:

$$\varepsilon:\ I^* \times W^* \times L^* \to I_W^*,\quad I_w(i,j) = I_0(i,j) \oplus L(i,j)W(i,j)p(i,j), \tag{1.4}$$

где $L(i,j)$ — маска встраивания ЦВЗ, учитывающая характеристики зрительной системы человека; служит для уменьшения заметности ЦВЗ;

$p(i,j)$ — проектирующая функция, зависящая от ключа;

знаком $\oplus$ обозначен оператор суперпозиции, включающий в себя, помимо сложения, усечение и квантование.

Проектирующая функция осуществляет «распределение» ЦВЗ по области изображения. Ее использование может рассматриваться, как реализация разнесения информации по параллельным каналам. Кроме того, эта функция имеет определенную пространственную структуру и корреляционные свойства, использующиеся для противодействия геометрическим атакам (см. гл. 2).

Другое возможное описание процесса внедрения получим, представив стегосистему как систему связи с передачей дополнительной информации (рис. 1.4) [8]. В этой модели кодер и декодер имеют доступ, помимо ключа, к информации о канале (то есть о контейнере и о возможных атаках). В зависимости от положения переключателей А и Б выделяют четыре класса стегосистем (подразумевается, что ключ всегда известен кодеру и декодеру).

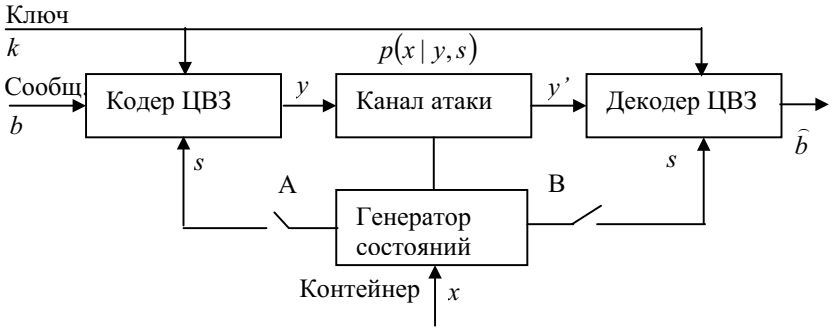


Рис. 1.4. Представление стегосистемы, как системы связи с передачей дополнительной информации

I класс: дополнительная информация отсутствует (переключатели разомкнуты) — «классические» стегосистемы. В ранних работах по стеганографии считалось, что информация о канале недоступна кодеку. Обнаружение ЦВЗ осуществлялось путем вычисления коэффициента корреляции между принятым стего и вычисленным по ключу ЦВЗ. Если коэффициент превышал

некоторый порог, выносилось решение о присутствии ЦВЗ. Известно, что корреляционный приемник оптимален лишь в случае аддитивной гауссовой помехи. При других атаках (например, геометрических искажениях) эти стегосистемы показывали удручающие результаты.

II класс: информация о канале известна только кодеру (А замкнут, Б разомкнут). Эта конструкция привлекла к себе внимание благодаря статье [10]. Интересной особенностью схемы является то, что, будучи слепой, она имеет ту же теоретическую пропускную способность, что и схема с наличием исходного контейнера в декодере. К недостаткам стегосистем II класса можно отнести высокую сложность кодера (необходимость построения кодовой книги для каждого изображения), а также отсутствие адаптации схемы к возможным атакам. В последнее время предложен ряд практических подходов, преодолевающих эти недостатки. В частности, для снижения сложности кодера предлагается использовать структурированные кодовые книги, а декодер рассчитывать на случай наихудшей атаки.

III класс: дополнительная информация известна только декодеру (А разомкнут, Б замкнут). В этих схемах декодер строится с учетом возможных атак. В результате получаются робастные к геометрическим атакам системы. Одним из методов достижения этой цели является использование так называемого опорного ЦВЗ (аналог пилот-сигнала в радиосвязи). Опорный ЦВЗ — небольшое число бит, внедряемых в инвариантные к преобразованиям коэффициенты сигнала. Например, можно выполнить встраивание в амплитудные коэффициенты преобразования Фурье, которые инвариантны к аффинным преобразованиям. Тогда опорный ЦВЗ «покажет», какое преобразование выполнил со стего атакующий. Другим назначением пилотного ЦВЗ является борьба с замираниями по аналогии с радиосвязью. Замираниями в данном случае можно считать изменение значений отсчетов сигнала при встраивании данных, атаках, добавлении негауссовского шума и т. д. В радиосвязи для борьбы с замираниями используется метод разнесенного приема (по частоте, времени, пространству, коду). В стеганографии же используется разнесение ЦВЗ по пространству контейнера. Пилотный ЦВЗ генерируется в декодере на основе ключа.

IV класс: дополнительная информация известна и в кодере и в декодере (оба ключа замкнуты). Как отмечено в [9], по всей видимости, все перспективные стегосистемы должны строиться по этому принципу. Оптимальность этой схемы достигается путем согласования кодера с сигналом-контейнером, а также адаптивным управлением декодером в условиях наблюдения канала атак.

3. Также как в радиосвязи наиболее важным устройством является приемник, в стегосистеме главным является стегодетектор. В зависимости от типа он может выдавать двоичные либо $M$-ичные решения о наличии/отсутствии ЦВЗ (в случае детектора с мягкими решениями). Рассмотрим вначале более