**Введение**

Проблема защиты информации от несанкционированного (самовольного) доступа заметно обострилась в связи с широким распространением локальных и особенно глобальных компьютерных сетей. Защита информации необходима для уменьшения вероятности утечки (разглашения), модификации (умышленного искажения) или утраты (уничтожения) информации, представляющей определенную ценность для ее владельца.

Задача защиты информации от несанкционированного доступа решалась во все времена на протяжении истории человечества. Уже в древнем мире выделилось два основных направления решения этой задачи, существующие и по сегодняшний день: криптография и стеганография. Целью криптографии является скрытие содержимого сообщений за счет их шифрования. В отличие от этого, при стеганографии скрывается сам факт существования тайного сообщения.

Слово «стеганография» имеет греческие корни и буквально означает «тайнопись». Исторически это направление появилось первым, но затем во многом было вытеснено криптографией. Тайнопись осуществляется самыми различными способами. Общей чертой этих способов является то, что скрываемое сообщение встраивается в некоторый безобидный, не привлекающий внимание объект. Затем этот объект открыто транспортируется адресату. При криптографии наличие шифрованного сообщения само по себе привлекает внимание противников, при стеганографии же наличие скрытой связи остается незаметным.

Под цифровой стеганографией понимается скрытие одной информации в другой. Причем сокрытие это должно реализоваться таким образом, чтобы, во-первых, не были утрачены свойства и некоторая ценность скрываемой информации, а во-вторых, неизбежная модификация информационного носителя, не только не уничтожила смысловые функции, но и на определенном уровне абстракции даже не меняла их. Тем самым факт передачи одного сообщения внутри другого не выявляется традиционными методами.

В качестве носителя скрытой информации должен выступать объект (файл), допускающий искажения собственной информации, не нарушающие его функциональность. Внесенные искажения должны быть ниже уровня чувствительности средств распознавания.

В качестве носителя обычно используются файлы изображений или звуковые файлы. Такие файлы обладают большой избыточностью и, кроме того, обычно велики по размеру, обеспечивая достаточно места для сокрытия простого или форматированного текста. Скрываемое сообщение может быть простым набором чисел, изображением, простым или зашифрованным текстом.

Многие мультимедийные форматы имеют поля расширения, которые могут заполняться пользовательской информацией, а могут быть забиты нулями – в последнем случае их также можно использовать для хранения и передачи информации. Однако этот наивный способ не только не обеспечивает требуемого уровня секретности, но и не может прятать значительные объемы данных. Решение этих проблем нашлось в следующем подходе.

В настоящее время можно наблюдать повсеместное распространение технологий VoIP(Skype) и рост популярности IP-телефонии. Что повышает интерес к ней, и становится объектом исследования ученых. Это направление в области стеганографии является наиболее актуальной. Технология VoIP весьма уязвима для стеганографических атак, из-за своей высокой пропускной способности, позволяющей относительно легко вставлять скрытые сообщения в трафик и доставлять их.

VoIP (англ. Voice over IP; IP-телефония) — общее название коммуникационных протоколов, технологий и методов, обеспечивающих передачу речевого сигнала по сети Интернет или по любым другим IP-сетям. Сигнал по каналу связи передаётся в цифровом виде и, как правило, перед передачей преобразовывается (сжимается) с тем, чтобы удалить избыточность.

Стеганография реального времени – использование стеганографических методов с каналом активного сетевого соединение. В частности технологии Skype, которая позволяет Бла-бла.

**ГЛАВА 1. Современная стеганография**

* 1. **Основные понятия и термины стеганографии и области ее применения**

За последние 15 лет, благодаря массовому распространению мультимедийных технологий и средств телекоммуникаций, развитие стеганографии вышло на принципиально новый этап, который специалисты называют компьютерной стеганографией. Среди основных областей использование цифровая стеганографии – скрытие (путем встраивания) сообщений в цифровые данных, которые, как правило, имеют аналоговую природу (речь, изображения, аудио или видеозаписи). В качестве контейнеров (носителей) возможно также использование текстовых файлов или исполняемых файлов. Так, например, наименее значимые биты цифрового изображения или аудиофайла могут быть заменены данными из текстового файла таким образом, что посторонний независимый наблюдатель не обнаружит никакой потери в качестве изображений или звука.

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 т.д. 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. Один 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.

Отсылать скрытые сообщения в VoIP трафике – последняя стадия эволюции стеганографии. Такой вид стеганографии называют – сетевой стеганографией. Сетевая стеганография - сокрытие информации в сетевом трафике, уже получила некоторое распространение, для сокрытия обычно использовался стандартный поток данных. В качестве носителей возможно использование сетевых пакетов и сетевого трафика. В VoIP трафике невозможно обнаружить посторонние данные, если не знать наверняка, что они есть и если не применить специальное программное обеспечение, предназначенное для расшифровки сигналов.

Так как цифровая стеганография является молодой наукой, то ее терминология не до конца устоялась. Основные понятия стеганографии были согласованы на первой международной конференции по скрытию данных. 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, метеорную радиосвязь, радиосвязь с псевдослучайной перестройкой радиочастоты, широкополосную радиосвязь.

Стеганографическая система (стегосистема) — объединение методов и средств используемых для создания скрытого канала для передачи информации. При построении такой системы условились о том, что:

1) враг представляет работу стеганографической системы. Неизвестным для противника является ключ с помощью которого можно узнать о факте существования и содержания тайного сообщения.

2) При обнаружении противником наличия скрытого сообщения он не должен смочь извлечь сообщение до тех пор пока он не будет владеть ключом.

3) Противник не имеет технических и прочих преимуществ.

Сообщение — это термин, используемый для общего названия передаваемой скрытой информации, будь то лист с надписями молоком, голова раба или цифровой файл.

Контейнер — так называется любая информация, используемая для сокрытия тайного сообщения. Пустой контейнер — контейнер, не содержащий секретного послания. Заполненный контейнер (стегоконтейнер) — контейнер, содержащий секретное послание.

Стеганографический канал (стегоканал) — канал передачи стегоконтейнера.

Ключ (стегоключ) — секретный ключ, нужный для сокрытия стегоконтейнера. Ключи в стегосистемах бывают двух типов: секретные и открытые. Если стегосистема использует секретный ключ, то он должен быть создан или до начала обмена сообщениями, или передан по защищённому каналу. Стегосистема, использующая открытый ключ, должна быть устроена таким образом, чтобы было невозможно получить из него закрытый ключ. В этом случае открытый ключ мы можем передавать по незащищённому каналу.

**Области применения стеганографии**

*Скрытая аннотация документов*

Медицинские снимки, картография, мультимедийные базы данных

*Защита от копирования*

Электронная коммерция, контроль за копированием, распространение мультимедийной информации

Рис.1. Области применения

стеганографии

*Аутентификация*

Системы видеонаблюдения, голосовой почты, электронное конфиденциальное делопроизводство

*Скрытая связь*

Военные и разведывательные приложения, а также в применение случаях, когда криптография использовать нельзя. В личных целях.

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, эти средства 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. Кроме того, важным недостатком цифровой подписи является 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, 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.

* 1. **Классификация методов стегоалгоритмов встраивания информации**

Большинство методов компьютерной стеганографии основываются на двух ключевых принципах:

* данные, которые не требуют абсолютной точности, могут быть видоизменены без потери своей функциональности
* органы чувств человека неспособны надежно различать незначительные изменения в модифицированных таким образом файлах и/или отсутствует специальный инструментарий, который был бы способен выполнять данную задачу.

В основе базовых подходов к реализации методов компьютерной стеганографии (КС) в рамках той или иной информационной среды лежит выделение малозначительных фрагментов этой среды и замена существующей в них информации информацией, которую необходимо скрыть. Поскольку в КС рассматриваются среды, поддерживаемые средствами вычислительной техники и компьютерных сетей, то вся информационная среда в результате может быть представлена в цифровом виде.

Таким образом, незначительные для кадра информационной среды фрагменты относительно того или иного алгоритма или методики заменяются фрагментами скрываемой информации. Под кадром информационной среды в данном случае подразумеваются определенная его часть, выделенная по характерным признакам. Такими признаками зачастую являются семантические характеристики выделяемой части информации. Например, кадром может быть избрано какое-нибудь отдельное изображение, звуковой файл.

Для существующих методов компьютерной стеганографии вводят следующую классификацию:

* по способу выбора контейнера
* по способу доступа к информации
* по способу организации контейнера
* по способу извлечения сообщения
* по принципу скрытия
* по формату контейнера
* по назначению.

По способу выбора контейнера различают суррогатные, селективные и конструирующие методы стеганографии. В суррогатных методах полностью отсутствуют возможность выбора контейнера, и для скрытия сообщения избирается первый попавшийся контейнер, который в большинстве случаев не оптимален для скрытия сообщения заданного формата. В селективных методах предусматривается, что скрытое сообщение должно воспроизводить специальные статистические характеристики шума контейнера. Для этого генерируют большое количество альтернативных контейнеров с последующим выбором наиболее оптимального из них для корректного сообщения. Особым случаем такого подхода является вычисление некоторой хеш-функции для каждого контейнера. При этом для скрытия сообщения избирается тот контейнер, хеш-функция которого совпадает со значением хеш-функции сообщения. В конструирующих методах стеганографии контейнер генерируется самой стеганосистемой. При этом существует несколько вариантов реализации. Так, например, шум контейнера может имитироваться скрытым сообщением. Это реализуется с помощью процедур, не только кодируют скрываемое сообщение под шум, но и сохраняют модель изначального шума. В предельном случае по модели шума может строится целое сообщение.

По способу доступа к скрываемой информации различают методы для потоковых (беспрерывных) контейнеров и методы для фиксированных контейнеров.

По способу организации контейнеры, помехоустойчивым кодам, могут быть систематическими и несистематическими. В первых можно указать конкретные места стеганограммы, где находятся информационные биты собственно контейнера, а где – шумовые биты, предназначенные для скрытия информации. Во втором случае организация контейнера такое разделение невозможно и для выделения скрытой информации необходимо обрабатывать содержимое всей стеганограммы.

По используемому принципу скрытия методы компьютерной стеганографии делятся на два основных класса: методы непосредственной замены и спектральные методы. Если первые, используя избыток информационной среды в пространственной или временной области, заключается в замене малозначительной части контейнера битами секретного сообщения, то другие для скрытия данных используют спектральные представления элементов среды, в которую встраиваются скрываемые данные (например, в разные коэффициенты массивов дискретно-косинусных преобразований, преобразований Фурье).

Основным направлением компьютерной стеганографии является использование свойств именно избыточности контейнера-оригинала, но при этом следует принимать во внимание то, что в результате скрытия информации происходит искажение некоторых статистических свойств контейнера или же нарушение его структуры. Это необходимо учитывать для уменьшения демаскирующих признаков.

В особую группу можно также выделить методы, которые используют специальные свойства форматов представления файлов:

* зарезервированные для расширения поля файлов, которые зачастую заполняются нулями и не учитываются программой
* специальное форматирование данных
* использование незадействованных участков на магнитных и оптических носителях
* удаление файловых заголовков-идентификаторов.

В основном, для таких методов характерны низкая степень скрытности, низкая пропускная способность и слабая производительность.

По назначению различают методы собственно для скрытой передачи данных и методы для скрытия данных в цифровых объектах с целью защиты авторских прав на них.

По типам контейнера выделяют методы с контейнерами вида текста, аудиофайла, изображения, видео, сетевые пакеты.

* 1. **Скрытие данных в изображениях**

Большинство исследований посвящено использованию в качестве стегоконтейнеров изображений. Это обусловлено следующими причинами:

* существованием практически значимой задачей защиты фотографий, картин, видео от незаконного тиражирования и распространения
* относительно большим объемом цифрового представления изображений, что позволяет внедрять ЦВЗ большого объема либо повышать робастность внедрения
* заранее известным размером контейнера, отсутствием ограничений, накладываемых требованиями реального времени
* наличием в большинстве реальных изображений текстурных областей, имеющих шумовую структуру и хорошо подходящих для встраивания информации
* слабой чувствительностью человеческого глаза к незначительным изменениям цветов изображения, его яркости, контрастности, содержанию в нем шума, искажениям вблизи контуров
* хорошо разработанными в последнее время методами цифровой обработки изображений.

Надо отметить, что последняя причина вызывает и значительные трудности в обеспечении робастности ЦВЗ: чем более совершенными становятся методы сжатия, тем меньше остается возможностей для встраивания посторонней информации. Развитие теории и практики алгоритмов сжатия изображений привело к изменению представлений о технике внедрения ЦВЗ. Если первоначально предлагалось вкладывать информацию в незначащие биты для уменьшения визуальной заметности, то современный подход заключается во встраивании ЦВЗ в наиболее существенные области изображений, разрушение которых приведет к полной деградации самого изображения. Не случайно поэтому стегоалгоритмы учитывают свойства системы человеческого зрения (СЧЗ), аналогично алгоритмам сжатия изображений. В стегоалгоритмах зачастую используются те же преобразования, что и в современных алгоритмах сжатия (дискретное косинусное преобразование в JPEG, вейвлет-преобразование в JPEG). При этом существуют, очевидно, три возможности. Вложение информации может производиться в исходное изображение, либо одновременно с осуществлением сжатия изображения-контейнера, либо в уже сжатое алгоритмом JPEG изображение. Поэтому рассмотрены свойства человеческого зрения и их учет в алгоритмах сжатия изображений.

* + 1. **Скрытие данных в пространственной области**

Алгоритмы скрытие данных в пространственной области, встраивают скрываемые данные в области первичного изображения. Их преимущество заключается в том, что для встраивания нет необходимости выполнять вычислительно сложные и длительные преобразование изображений.

Цветное изображение C представим через дискретную функцию, которая определяет вектор цвета для каждого пикселя изображения , где значение цвета задает трехкомпонентный вектор в цветовом пространстве. Наиболее распространенный способ передачи цвета – это модель RGB, в которой основные цвета – красный, зеленый и синий, голубой цвет может быть представлен в виде суммы основных цветов.

Вектор цвета в RGB-пространстве представляет интенсивность основных цветов. Сообщения встраиваются за счет манипуляций цветовыми составляющими или непосредственно яркостью .

Общий принцип этих методов заключается в замене избыточной, малозначимой части изображения битами секретного сообщения. Для извлечения сообщения необходимо знать алгоритм, по которому размещалась оп контейнеру скрытая информация.

* + - 1. **Метод замены наименее значащего бита (НЗБ или LSB)**

Такой подход типичен для стеганографии, и считается «простой системой». Обычно используются форматы без потерь, данные могут быть напрямую восстановлены. Безусловными достоинствами данного метода являются: простота, сравнительно большой объем встраиваемых данных, независимость от формата файла-контейнера (можно использовать любой объект, в котором допустима манипуляция младшими битами (LSB)). Однако он имеет два серьёзных недостатка:

* Скрытое сообщение легко разрушить.

Для этого необходимо просто записать в один или два младших бита каждого байта графического изображения нули или единицы, тогда, если картинка не содержала скрытого сообщения, то видимых искажений не появится, а если в картинке было скрыто сообщение, то оно будет испорчено. То есть, те же достоинства, которые используются для сокрытия информации, могут быть использованы и для незаметной борьбы со стеганографией. Таким образом, файл-контейнер критичен к любым модификациям.

* не обеспечена секретность встраивания информации.

Нарушителю точно известно местоположение всего сообщения. В случае перехвата информации, если у перехватчика возникнет подозрение, на то, что в изображении скрыто какое-то сообщение, ему относительно нетрудно будет извлечь эту информацию, так как количество возможных способов извлечения невелико.

Для преодоления последнего недостатка было предложено встраивать сообщение не во все пиксели изображения, а лишь в некоторые из них, определяемые по псевдослучайному закону в соответствии с ключом, известному только законному пользователю.

Рассмотрим суть метода на примере изображения формата *BMP*, т.е. полноцветные изображения, в которых под хранение информации о цвете каждой точки отводится по 3 байта. Каждый из байтов содержит информацию об интенсивности одной из трех составляющих цвета (палитра *RGB*): красной, синей, зеленой.

Так как каждый из трех цветов точки кодируется одним байтом, соответствующая ему интенсивность меняется в пределах от 0 до 255. Общее число возможных цветов равно 2563 ≈ 16⋅106. Известно, что человеческий глаз способен различить лишь порядка четырех тысяч цветов, а для их хранения достаточно всего четырех бит. Таким образом, *LSB*-методы предполагают использование для внедрения скрытой информации до четырех младших разрядов цветовых байтов.

Следует отметить, что это может привести к изменению интенсивности каждого цвета на *ΔI* = 42/256 = 16/256 = 0,0625 (на 6,25%). Согласно исследованиям, в действительности незаметными для человеческого глаза являются изменения в пределах 1-1,5%, а порог чувствительности глаза составляет 1-3%. Поэтому, при встраивании информации в изображение *RGB* рекомендуется модифицировать не более трех разрядов (искажение 3,5%) каждого из байтов, отвечающих за цвет точки.

Одной из стандартных схем является 3-2-3 *LSB*-метод. Как следует из названия, в ней используется по 3 младших разряда красной и синей составляющих точки и 2 младших разряда зеленой составляющей. Такой выбор обусловлен тем, что человеческий глаз лучше распознает градации зеленого цвета, чем других.

При использовании 3-2-3 схемы *LSB*-метода в каждом пикселе изображения скрывается 1 байт встраиваемого сообщения. Таким образом, максимальный объем скрываемой в контейнере информации равен количеству точек изображения:

*V* = *W*·*H* [байт], (1.1)

где *W* - ширина изображения в пикселях, *H* - высота изображения в пикселях.

Увеличение числа изменяемых бит повышает объем контейнера, но при этом сильнее искажается изображение, что становится заметно «невооруженному» глазу. При использовании только первого и второго разрядов (максимальное искажение 1,17%), либо только третьего разряда (искажение 1,56%) каждого из байтов обнаружить факт скрытой передачи данных возможно лишь с помощью *стегоанализа.*

**1.3.1.2.Метод псевдослучайного интервала**

В рассмотренном выше простейшем случае выполняется замена НЗБ всех последовательно размещенных пикселей изображения. Другой подход — метод случайного интервала, заключается в случайно распределении битов секретного сообщения по контейнеру, в результате чего расстояние между двумя встроенными битами определяется псевдослучайно. Эта методика особенно эффективна в случае, когда битовая длина секретного сообщения существенно меньше количества пикселей изображения.

* + - 1. **Метод псевдослучайной перестановки**

Недостатком метода псевдослучайного интервала является то, что биты сообщения в контейнере размещены в той же последовательности, что и в самом сообщении, и только интервал между ними изменяется псевдослучайно. Поэтому для контейнеров фиксированного размера более целесообразным является использование метода псевдослучайной перестановки (выбора), смысл которого заключается в том, что генератор ПСЧ образует последовательность индексов и сохраняет k-й бит сообщения в пикселе с индексом .

Пусть N – общее количество бит (самых младших) в имеющимся контейнере, – перестановка чисел . Тогда, если у нас имеется для скрытия конфиденциальное сообщение длиной n бит, то эти биты можно просто встроить вместо бит контейнера .

Функция перестановки должна быть псевдослучайной, иными словами, она должна обеспечивать выбор бит контейнера приблизительно случайным образом. Таким образом, секретные биты будут равномерно распределены по всему битовому пространству контейнера. Однако при этом индекс определенного бита контейнера может появиться в последовательности более одного раза и в этом случае может произойти «пересечение» — искажение уже встроенного бита. Если количество бит сообщения намного меньше количества младших бит изображения, то вероятность пересечения является незначительной, и искаженные биты в дальнейшем могут быть восстановлены с помощью корректирующих кодов.

Вероятность, по крайней мере, одного пересечения оценивается как

(1.2)

При увеличении и данная вероятность стремится к единице.

Для предотвращения пересечений можно запоминать все индексы использованных элементов и перед модификацией нового пикселя выполнять предварительную его проверку на повторяемость. Также можно применять генераторы ПСЧ без повторяемости чисел. Последний случай рассмотрим более подробно.

Для наших целей функция перестановки также зависит от секретного ключа К. При этом генератор псевдослучайной перестановки — это функция, которая для каждого значения K вырабатывает различные псевдослучайные перестановки чисел {1,2, …,N}.

Обозначим через генератор перестановок с соответствующим ключом К. Если перестановка защищена по вычислению (то есть, взлом алгоритма требует неоправданно больших затрат вычислительных ресурсов нарушителя), то возможность раскрытия содержания или предположение самого только вида перестановок без владения информацией о секретном ключе К практически равняется нулю.

Секретный генератор псевдослучайной перестановки (ГПСП) может быть эффективно реализован на основе генератора псевдослучайной функции (ГПСФ), который, как и ГПСП, вырабатывает различные, не подлежащие прогнозированию функции при каждом отдельном значении ключа; однако множество значений функций не должно равняться области ее определения. ГПСФ легко реализуется из секретной хэш-функции Н путем объединения аргумента с секретным ключом К и взятия от результирующей битовой строки функции Н:

(1.3)

Где - объединение (конкатенация) битовых строк и . - результирующая псевдослучайная функция от , которая зависит от параметра .

* + - 1. **Метод блочного скрытия**

Метод блочного скрытия — это еще один подход к реализации метода замены и заключается в следующем. Изображение-оригинал разбивается на непересекающихся блоков произвольной конфигурации, для каждого из которых вычисляется, бит четности :

В каждом блоке выполняется скрытие одного секретного бита , Если бит четности, то происходит инвертирование одного из НЗБ блока , в результате чего . Выбор блока может происходить псевдослучайно с использованием стегакоключа.

Хотя этот метод имеет такую же низкую устойчивость к искажениям, как и все предыдущие, у него есть ряд преимуществ. Во-первых, существует возможность модифицировать значение такого пикселя в блоке, изменение которого приведет к минимальному изменению статистики контейнера. Во-вторых, влияние последствий встраивания секретных данных в контейнер можно уменьшить за счет увеличения размера блока.

* + - 1. **Метод квантования изображения**

К методам скрытия в пространственной области можно также отнести метод квантования изображения, основанный на межпиксельной зависимости, которую можно описать некоторой функцией . В простейшем случае можно вычислить разницу между смежными пикселями и и задать ее как параметр функции, где  дискретная аппроксимация разницы сигналов .

Поскольку — целое число, а реальная разница — действительное число, то возникают ошибки квантования . Для сильно коррелированных сигналов эта ошибка близка к нулю:

При данном методе скрытие информации производится путем корректировки разностного сигнала . Стеганоключ представляет собой таблицу, которая каждому возможному значению ставит в соответствие определенный бит, например:

Таблица

|  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | -4 | -3 | -2 | -1 | 0 | 1 | 2 | 3 | 4 |
|  | 1 | 0 | 1 | 1 | 0 | 0 | 1 | 0 | 1 |

Для скрытия го бита сообщения вычисляется разница   
. Если при этом   
, не соответствует секретному биту, который необходимо скрыть, то значение заменяется ближайшим , для которого такое условие выполняется. При этом соответствующим образом корректируются значения интенсивностей пикселей, между которыми вычислялась разница . Извлечение секретного сообщения осуществляется согласно значению , соответствующему разнице .

* + - 1. **Метод Куттер-Джордана-Боссена**

Куттер (М. Kutter), Джордан (F. Jordan) и Боссен (F. Bossen) предложили алгоритм встраивания в канал синего цвета изображения, имеющего RGB-кодирование, поскольку к синему цвету ЗСЧ является наименее чувствительной. Рассмотрим алгоритм передачи одного бита секретной информации в предложенном методе.

Пусть — бит, который подлежит встраиванию, — изображение-контейнер, псевдослучайный пиксель контейнера, в который будет выполняться встраивание.

Секретный бит встраивается в канал синего цвета путем модификации яркости

Где - константа, определяющая энергию встраиваемого сигнала. Ее величина зависит от назначения стеганосистемы. Чем больше , тем выше устойчивость встроенной информации к искажениям, однако и тем сильнее ее заметность.

Получатель извлекает бит, не имея первичного изображения, то есть, «вслепую». Для этого выполняется предсказание значения первичного, не модифицированного пикселя на основе значений соседних пикселей. Для получения оценки пикселя предложено использовать значение нескольких пикселей, размещенных в том же столбце и в той же строке массива графического контейнера. Авторы использовали «крест» пикселей размером 7×7. Оценка получается в виде

где — количество пикселей сверху (снизу, слева, справа) от оцениваемого пикселя (в случае креста 7×7 =3).

При извлечении встроенного бита вычисляется разница между текущим и прогнозируемым значениями интенсивности пикселя :

Знак будет означать встроенный бит: если , то ; если , то .

Функции встраивания и извлечения в данном методе не симметричны, то есть, функция извлечения не является обратной функцией встраивания. Хотя, как указывают авторы метода, правильное распознание бита сообщения в случае применения описанных выше процедур является высоковероятным, однако не стопроцентным. Для уменьшения вероятности ошибок извлечения было предложено в процессе встраивания каждый бит повторять несколько раз (многократное встраивание). Поскольку при этом каждый бит был повторен раз, то получается оценок одного бита сообщения. Секретный бит извлекается по результатам усреднения разницы, между реальным и оцененным значениями интенсивности пикселя в полученном контейнере:

Знак усредненной разницы будет определять значение встроенного бита. Алгоритм устойчив ко многим известным видам атак: НЧ фильтрации изображения, его компрессии в соответствии с алгоритмом JPEG, обрезанию краев.

* + 1. **Скрытие данных в частотной области изображения**

Как уже было отмечено выше, стеганографические методы скрытия данных в пространственной области изображения являются нестойкими к большинству из известных видов искажений. Так, например, использование операции компрессии с потерями (относительно изображения, это может быть JPEG-компрессия) приводит к частичному или, что более вероятно, полному уничтожению встроенной в контейнер информации. Более стойкими к разнообразным искажениям, в том числе и компрессии, являются методы, использующие для скрытия данных не пространственную область контейнера, а частотную.

Существует несколько способов представления изображения в частотной области. При этом используется та или иная декомпозиция изображения, используемого в качестве контейнера. Например, существуют методы на основе использования дискретного косинусного преобразования (ДКП), дискретного преобразования Фурье (ДПФ), вейвлет-преобразования, преобразования Карунена-Лоева и некоторые другие. Подобные преобразования могут применяться либо к отдельным частям изображения, либо к изображению в целом.

Наибольшее распространение среди всех ортогональных преобразований в стеганографии получили вейвлет-преобразования и ДКП, что определенной мерой объясняется значительным распространением их использования при компрессии изображений. Кроме того, для скрытия данных целесообразно применять именно то преобразование изображения, которому последнее будет подвергаться со временем при возможной компрессии. Например, известно, что алгоритм ДКП является базовым в стандарте JPEG, а вейвлет-преобразования — в стандарте JPEG2000.

Стеганоалгоритм может быть достаточно стойким к последующей компрессии изображения, только если он будет учитывать особенности алгоритма перспективного сжатия. При этом, конечно, стеганоалгоритм, в основу которого заложено вейвлет-преобразование, совсем не обязательно будет стойким к дискретнокосинусному алгоритму сжатия, и наоборот. Еще большие трудности возникают при выборе метода стеганопреобразования во время скрытия данных в потоковом видео. Причина этого — одной из составляющих алгоритмов компрессии видеоинформации (в дополнение к компрессии неподвижного кадра), является кодирование векторов компенсации движения. При компрессии неподвижного изображения данная компенсация отсутствует за ненадобностью. Чтобы быть в достаточной степени стойким, стеганоалгоритм должен учитывать данный фактор.

Остается также открытым вопрос о существовании стойкого стеганопреобразования, которое было бы независимым от применяемого в дальнейшем алгоритма компрессии.

На сегодняшний день известно достаточно большое количество моделей, позволяющих оценить пропускную способность канала передачи скрытых данных. Ниже приведена одна из них.

Пусть — первичное изображение (контейнер-оригинал), — сообщение, которое подлежит скрытию. Тогда модифицированное изображение (стеганоконтейнер) . Также предполагается, что модифицированное изображение визуально неотличимо от первичного и может быть подвергнуто в стеганоканале компрессии с потерями: , где — оператор компрессии.

Задача адресата — извлечь из полученного контейнера встроенные на предыдущем этапе биты данных .

Основное, что будет при этом интересовать нас, — ответ на вопрос: какое количество бит можно эффективно встраивать в изображение и со временем извлечь из него при условии удовлетворительно низкой вероятности ошибок на последнем этапе. Другими словами, какова пропускная способность канала передачи скрытых данных при условии наличия в канале связи определенного алгоритма компрессии? Блок-схема такого стеганоканала представлена на рис.2.

S

M

C

Рис 2. Блок – схема стегоканала с атакой в виде компрессии

Сообщение передается по каналу, который имеет, два источника «шума»: — изображение-контейнер и «шум», возникающий в результате операций компрессии/декомпрессии. При этом   и — возможно искаженные стеганоконтейнер и, как результат, — оценка полезного сообщения.

Структурная схема стеганосистемы приведена на рис. 5.24.

S

Декодер

Прямое преобразование

Обратное преобразование

Прямое преобразование

Рис 3. Структурная схема

M

C

Изображение раскладывается на субполос (прямое преобразование), в каждую из которых встраивается скрываемая информация . После обратного преобразования получается модифицированное изображение . После компрессии/ декомпрессии   в канале связи получается изображение , которое на принимающей стороне вновь подвергается прямому преобразованию и из каждой субполосы независимо извлекается скрытое сообщение — оценка .

Реальные изображения не представляют собой случайные процессы с равномерно распределенными значениями величин. Известно, и данный факт используется в алгоритмах компрессии, что большая часть энергии изображений сосредоточена в низкочастотной (НЧ) области спектра. Отсюда и возникает необходимость в осуществлении декомпозиции изображения на субполосы, к которым прибавляется стеганособщение. НЧ субполосы содержат основную часть энергии изображения и, таким образом, носят шумовой характер. Высокочастотные (ВЧ) субполосы спектра Изображения наибольшим образом поддаются влиянию со стороны разнообразных алгоритмов обработки, таких как, например, компрессия или НЧ-фильтрация. Таким образом, можно сделать вывод, что для встраивания сообщения самыми оптимальными являются среднечастотные (СЧ) субполосы спектра изображения.

Стеганографический канал можно разложить на ряд независимых подканалов. Такое разложение совершается за счет выполнения прямого и обратного преобразований. В каждом из D подканалов существует по два источника шума. Пусть , при — дисперсия коэффициентов преобразования (шум изображения) в каждом из подканалов. Тогда выражение для пропускной способности канала стеганосистемы приобретет вид:

где — визуальный порог для субполосы (— максимально допустимая энергия стеганосообщения, исходя из требований сохранения визуального качества изображения); и — пиксельный размер изображения-контейнера.

Выбор значения визуального порога базируется на учете свойств ЗСЧ. Известно, что шум в ВЧ областях изображения более приемлемый, чем в НЧ областях.

Можно ввести некоторые весовые коэффициенты: , где , a — константа.

Случай, когда , отвечает равномерному распределению стеганограммы по всем субполосам. Случай отвечает распределению стеганограммы в соответствии с дисперсией субполос.

После некоторых упрощений можно получить выражение для пропускной способности канала передачи скрытых данных:

Знак приближения в выражении (1.10) является справедливым, для любого значения. Очевидно, что при декомпозиция никоим образом не будет влиять на пропускную способность стеганоканала. При пропускная способность будет возрастать за счет того, что в области с низкой дисперсией (высокочастотной области) стеганосигналу прибавляется относительно больше энергии.

Известно, что преобразования можно упорядочить по достижимым выигрышам от алгоритма кодирования рис.4.

Вейвлет

Адамара

Рис 4. Виды преобразований

ПКЛ

ДКП

Хаара

Единичное

Выигрыш

Под выигрышем от кодирования подразумевается степень перераспределения дисперсий коэффициентов преобразования. Наибольший выигрыш дает преобразование Карунена-Лоева (ПКЛ), наименьший — разложение, за базисом единичного импульса (то есть отсутствие преобразования).

Преобразования, которые характеризуются высокими значениями выигрыша от кодирования, такие как ДКП, вейвлет-преобразования, характеризуются резко неравномерным распределением дисперсий коэффициентов субполос. Высокочастотные субполосы не подходят для встраивания из-за большого шума обработки, а низкочастотные — из-за высокого шума изображения (рис. 3), Поэтому приходится ограничиваться среднечастотными полосами, у которых шум изображения приблизительно равен шуму обработки. Поскольку таких полос немного, то пропускная способность стеганоканала является сравнительно малой.

В случае применения преобразования с более низким выигрышем от кодирования, например, преобразования Адамара или Фурье, существует больше блоков, у которых шум изображения приблизительно равняется шуму обработки, а, следовательно, и пропускная способность выше. Вывод, к которому пришли авторы указанной работы, достаточно неожиданный: для повышения пропускной способности стеганографического канала целесообразно применять преобразования с меньшими выигрышами от кодирования, которые плохо подходят для компрессии сигналов.

Эффективность применения вейвлет-преобразования и ДКП для компрессии изображений объясняется тем, что они хорошо моделируют процесс обработки изображения в ЗСЧ, отделан существенные детали от второстепенных. Таким образом, данные преобразования более целесообразно использовать в случае присутствия активного нарушителя, поскольку модификация значимых, коэффициентов может привести к неприемлемому искажению изображения.

При применении преобразований с низкими значениями выигрыша от кодирования существует значительная опасность разрушения встроенных данных, по причине того, что коэффициенты преобразования менее стойки к модификациям. Однако при этом существует большая гибкость в выборе преобразования, и если последнее неизвестно нарушителю (хотя это и противоречит принципу Керхгофа) то модифицировать стеганограмму будет существенно сложнее.

Во время цифровой обработки изображения часто применяется двумерная версия дискретного косинусного преобразования:

(1.11)

(1.12)

где и — соответственно, элементы оригинального и восстановленного по коэффициентам ДКП изображения размерностью ; — пространственные координаты пикселей изображения; — массив коэффициентов ДКП; — координаты в частотной области; , если , и , если . Рассмотрим существующие методы, которые базируются на алгоритме ДКП.

**1.3.2.1. Метод относительной замены величин коэффициентов ДКП (метод Коха и Жао)**

Один из наиболее распространенных на сегодня методов скрытия конфиденциальной информации в частотной области изображения заключается в относительной замене величин коэффициентов ДКП, который в свое время описали Кох (Е. Koch) и Жао (J. Zhao).

На начальном этапе первичное изображение разбивается на блоки размерностью 8×8 пикселей. ДКП применяется к каждому блоку — формула (1.11), в результате чего получают матрицы 8×8 коэффициентов ДКП, которые зачастую обозначают , где b — номер блока контейнера , a — позиция коэффициента в этом блоке. Каждый блок при этом предназначен для скрытия одного бита данных.

Было предложено две реализации алгоритма: псевдослучайно могут выбираться два или три коэффициента ДКП. Рассмотрим первый вариант.

Во время организации секретного канала абоненты должны предварительно договориться о двух конкретных коэффициентах ДКП из каждого блока, которые будут использоваться для скрытия данных. Зададим данные коэффициенты их координатами в массивах коэффициентов ДКП: . Кроме этого, указанные коэффициенты должны отвечать косинус-функциям со средними частотами, что обеспечит скрытость информации в существенных для ЗСЧ областях сигнала, к тому же информация не будет искажаться при JPEG-компрессии с малым коэффициентом сжатия.

Непосредственно процесс скрытия начинается со случайного выбора блока изображения, предназначенного для кодирования b-го бита сообщения. Встраивание информации осуществляется таким образом: для передачи бит 0 стремятся, чтобы разница абсолютных значений коэффициентов ДКП превышала некоторую положительную величину, а для передачи бита 1 эта разница делается меньшей по сравнению с некоторой отрицательной величиной:

(1.13)

Таким образом, первичное изображение искажается за счет внесения изменений в коэффициенты ДКП, если их относительная величина не отвечает скрываемому биту. Чем больше значение , тем стеганосистема, созданная на основе данного метода, является более стойкой к компрессии, однако качество изображения при этом значительно ухудшается.

После соответствующего внесения коррекции в значения коэффициентов, которые должны удовлетворять неравенству (1.13), проводится обратное ДКП.

Для извлечения данных, в декодере выполняется аналогичная процедура выбора коэффициентов, а решение о переданном бите принимается в соответствии со следующим правилом:

(1.14)

* + 1. **Методы расширения спектра**

Изначально методы расширения спектра (PC или SS — Spread-Spectrum) использовались при разработке военных систем управления и связи. Во время второй мировой войны расширение спектра использовалось в радиолокации для борьбы с намеренными помехами. В последние годы развитие данной технологии объясняется желанием создать эффективные системы радиосвязи для обеспечения высокой помехоустойчивости при передаче узкополосных сигналов по каналам с шумами и осложнения их перехвата. Система связи является системой с расширенным спектром в следующих случаях.

* Полоса частот, которая используется при передаче, значительно шире минимально необходимой для передачи текущей информации. При этом энергия ин формационного сигнала расширяется на всю ширину полосы частот при низком соотношении сигнал/шум, в результате чего сигнал трудно обнаружить, пере хватить или воспрепятствовать его передаче путем внесения помех. Хотя суммарная мощность сигнала может быть большой, соотношение сигнал/шум в любом диапазоне частот является малым, что делает сигнал с расширенным спектром трудно определяемым при радиосвязи и, в контексте скрытия информации стеганографическими методами, трудно различимым человеком.
* Расширение спектра выполняется с помощью так называемого расширяющего (или кодового) сигнала, который не зависит от передаваемой информации. Присутствие энергии сигнала во всех частотных диапазонах делает радиосигнал с расширенным спектром стойким к внесению помех, а информацию, встроенную в контейнер методом расширения спектра, стойкой к ее устранению или извлечению из контейнера. Компрессия и другие виды атак на систему связи могут устранить энергию сигнала из некоторых участков спектра, но поскольку последняя была распространена по всему диапазону, в других полосах остается достаточное количество данных для восстановления информации. В результате, если, разумеется, не разглашать ключ, который использовался для генерации кодового сигнала, вероятность извлечения информации неавторизованными лицами существенно снижается.
* Восстановление первичной информации (то есть «сужение спектра») осуществляется путем сопоставления полученного сигнала и синхронизированной копии кодового сигнала.

В радиосвязи применяют три основных способа расширения спектра:

* с помощью прямой ПСП (РСПП)
* с помощью скачкообразного перестраивания частот
* с помощью компрессии с использованием линейной частотной модуляции(ЛЧМ).

При расширении спектра прямой последовательностью информационный сигнал модулируется функцией, которая принимает псевдослучайные значения в установленных пределах, и умножается на времени константу — частоту (скорость) следования элементарных посылок (элементов сигнала). Данный псевдослучайный сигнал содержит составляющие на всех частотах, которые, при их расширении, модулируют энергию сигнала в широком диапазоне.

В методе расширения спектра с помощью скачкообразного перестраивания частот передатчик мгновенно изменяет одну частоту несущего сигнала на другую. Секретным ключом при этом является псевдослучайный закон изменения частот.

При компрессии с использованием ЛЧМ сигнал модулируется функцией, частота которой изменяется во времени.

Очевидно, что любой из указанных методов может быть распространен на использование в пространственной области при построении стеганографических систем.

Рассмотрим один из вариантов реализации метода РСПП, авторами которого являются Смит (J.R. Smith) и Комиски (В.О. Comiskey). Алгоритм модуляции следующий: каждый бит сообщения , представляется некоторой базисной функцией , размерностью , умноженной, в зависимости от значения бита (1 или 0), на +1 или -1:

Модулированное сообщение , полученное при этом, попиксельно суммируется с изображением-контейнером в качестве которого используется полутоновое изображение размером . Результатом является стеганоизображение , при

Чтобы сделать невозможным искажение уже встроенного бита сообщения, базисные функции должны быть ортогональными:

где — количество значащих пикселей в базисной функции; — средняя мощность, приходящаяся на пиксель,; дельта-символ Кронекера.

В идеальном случае, все базисные функции должны быть некоррелированными с изображением-контейнером С, то есть должны быть ортогональными к нему: . Однако на практике тяжело найти контейнер, который был бы полностью ортогональным ко всем базисным функциям . В таком случае должна быть введена величина погрешности , которая учитывается увеличением мощности .

Для эффективного скрытия информации необходимо значительное количество базисных функций, ортогональных к типичным изображениям. Кодирование же изображении выдвигает противоположное требование: идеальным считается небольшое количество базисных функций, которые приблизительно перекрывает всю область изображения. Эти требования вступают в конфликт, когда изображение, содержащее скрытую информацию, подвергается компрессии: идеальная схема компрессии не способна полностью отобразить базисы, которые использовались для скрытия.

Базисные функции могут быть организованы и сравнимы в соответствии с такими свойствами как полная мощность, степень пространственного расширения (или локализации), а также степень пространственного частотного расширения (или локализации).

В случае РСПП модулирующая функция состоит из постоянного коэффициента усиления G (целое число), умноженного на псевдослучайный блок (массив) базисных функций , значений ±1. Каждый массив имеет индивидуальное расположение в массиве. Кроме того, массивы являются непересекающимися (то есть заведомо ортогональными друг к другу) и перекрывают - массив без промежутков.

Также будем считать, что все базисные функции имеют одинаковое количество значащих элементов . В этом случае полную мощность можно записать следующим образом:

(1.17)

На этапе извлечения данных нет необходимости владеть информацией о первичном контейнере С. Операция декодирования заключается в восстановлении скрытого сообщения путем проецирования полученного стеганоизображения на все базисные функции :

Значения могут быть легко восстановлены с помощью знаковой функции:

(1.19)

Если , то скрытая информация была утрачена. При малых значениях средней мощности возрастает вероятность извлечения ошибочного значения бита информации, однако контейнер при этом искажается в меньшей степени.

Основное преимущество стеганографических методов, основанных на расширении спектра — сравнительно высокая стойкость различного рода атакам на изображение, поскольку скрываемая информация распределена в широкой полосе частот и ее трудно удалить без полного разрушения контейнера. Искажения стеганоизображения увеличивают значение , однако до тех пор, пока выполняется условие , скрытое сообщение не пострадает.

* 1. **Скрытие данных в аудиосигналах**

Особое развитие получили цифровые методы стеганографии в аудиосреде. Скрытие данных в звуковых сигналах является особенно перспективным поскольку слуховая система человека (ССЧ) работает в сверхшироком динамическом диапазоне. ССЧ воспринимает более чем миллиард к одному в диапазон мощности и более чем тысяча к одному в частотном диапазоне. Кроме этого, высокой является и чувствительность к аддитивному флуктуационному (белому) шуму. Отклонения в звуковом файле могут быть выявлены вплоть до одной десяти миллионной (на 70 дБ ниже уровня внешних шумов).

Несмотря на это, существуют определенные возможности для скрытия информации и в аудиосреде. Хотя ССЧ и имеет широкий динамический диапазон, она характеризуется достаточно малым разностным диапазоном. Как следствие, громки звуки содействуют маскировке тихих звуков. Кроме того, ССЧ не способна различать абсолютную фазу, распознавая только относительную. Наконец, существует некоторые виды искажений, вызванных окружающей средой, которые настолько обычны для слушателя, что в большинстве случаев им игнорируются.

Подобные особенности слухового аппарата человека позволяют удачно использовать аудиосреду с целью стеганографической защиты конфиденциальной информации. Значительный вклад в развитие аудиостеганографии внесли Бендер (W. Bender), Моримото (N. Morimoto) и другие.

* + 1. **Кодирование наименее значащих бит**

Кодирование младших разрядов является простейшим способом внедрить конфиденциальные данные в иные структуры данных. Используя звуковой сигнал, путем замены НЗБ каждой точки осуществления выборки, представленной двоичной последовательностью, можно зашифровать значительный объем информации.

Теоретически, пропускная способность стеганоканала составляет 1 Кб/сек на 1 кГц в канале без помех, битовая скорость передачи данных составит 8 Кб/ сек в последовательности, которая оцифрована с частотой 8 кГц, и 44 Кб/сек в последовательности с частотой дискретизации 44 кГц. Платой за высокую пропускную способность канала является ощутимый на слух низкочастотный шум. Слышимость данного шума непосредственно зависит от содержимого сигнала-контейнера. Например, шум зрителей во время эфира спортивного соревнования в достаточной степени маскировал бы шум наименьших бит, модифицированных кодированием. Однако указанный шум будет ощутимым на слух при использовании в качестве контейнера аудиозаписи игры струнного квартета. Для компенсации внесенных искажений целесообразным будет использование адаптивной аттенюации данных.

* + 1. **Метод фазового кодирования**

Основная идея метода фазового кодирования состоит в замене фазы исходного звукового сегмента на опорную фазу, характер изменения которой отражает собой данные, которые необходимо скрыть. Для того чтобы сохранить разностную фазу между сегментами, фазы последних соответствующим образом согласовываются.

Фазовое кодирование, когда оно может быть использовано, является одним из наиболее эффективных методов по критерию отношения сигнал/воспринимаемый шум. Существенное изменение соотношения фаз между каждыми частотными составляющими приводит к значительному рассеиванию фазы. Тем не менее, до тех пор, пока модификация фазы в достаточной мере мала, может быть достигнуто скрытие, неощутимое на слух. Разумеется, модификация считается малой по отношению к конкретному наблюдателю, поскольку специалисты по спектральному анализу способны обнаружить те изменения, которые непрофессионалу могут показаться незначительными.

Процедура фазового кодирования заключается в следующем:

* Звуковая последовательность разбивается на серию N коротких сегментов (блоков)
* К n-му сегменту сигнала применяется K-точечное ДПФ, где K=I/N, и создаются массивы фаз и амплитуд для
* Запоминается разность фаз между каждыми соседними сегментами для :

(1.20)

* Двоичная последовательность данных представляется как или , отображая, соответственно, 1 или 0:
* С учетом разности фаз воссоздается новый массив фаз для
* Восстановление звукового сигнала осуществляется путем применения операции обратного ДПФ к исходной матрице амплитуд и модифицированной матрице фаз.

Перед процессом расшифровывания должна быть проведена синхронизация последовательности. Приемной стороне должны быть известны длина сегмента, точки ДПФ и интервал данных. Значение основной фазы первого сегмента определяется как 0 или 1, которые представляют закодированную двоичную последовательность.

Поскольку фаза является модифицированной, соответствующим образом будут изменены и абсолютные фазы последующих сегментов. Однако относительная разность фаз между каждыми смежными сегментами будет сохранена. Именно к этой относительной разности в фазе и обнаруживается наибольшая чувствительность человеческого слуха.

Разброс фаз — искажение, вызванное нарушением корреляции фаз между каждой из частотных составляющих. Уменьшение разброса фаз ограничивает скорость передачи данных при фазовом кодировании. Одной из причин разброса фаз можно назвать замещение фазы двоичным кодом. Для уменьшения искажений, значение модифицированной фазы должно быть близким к ее первичному значению. А для того чтобы снизить чувствительность встроенных данных к шуму, должна увеличиваться разность между структурами модифицированной фазы. Для этого, например, биту 0 может отвечать значение , а биту 1 — значение .

Еще одним источником искажения является скорость изменения модифицированной фазы. Если искажение применено к каждому элементу дискретизации ДПФ, это с большой вероятностью разрушит связи между фазами соседних частотных составляющих, что, в результате, приведет к наложению фонового биения. Путем более медленного изменения фазы и согласования переходов между изменениями фазы, достигается существенное снижение ощутимых на слух искажений.

Недостатком схемы, построенной на основе метода фазового кодирования, является низкая пропускная способность.

* + 1. **Метод расширения спектра**

В стандартном канале связи нередко бывает желательным сосредоточить информацию в как можно более узком диапазоне частотного спектра, например, для того чтобы сохранить имеющуюся полосу пропускания и уменьшить мощность сигнала. С другой стороны, основной метод расширения спектра предназначен для шифрования потока информации путем «рассеивания» кодированных данных по всему возможному частотному спектру. Последнее делает возможным прием сигнала даже при наличии помех на определенных частотах.

Как уже указывалось выше (см. методы скрытия данных в изображении путем расширения спектра), методы РСПП расширяют сигнал данных (сообщения), умножая его на элементарную посылку — ПСП максимальной длины, модулированную известной частотой.

Поскольку аудиосигналы, используемые в качестве контейнеров, имеют дискретный формат, то для кодирования в качестве частоты элементарной посылки можно использовать частоту дискретизации. Как следствие, дискретный характер сигнала устраняет наиболее сложную проблему, которая возникает при получении сигнала с расширенным прямой последовательностью спектром, — корректное определение начала и конца составляющих элементарной посылки с целью фазовой синхронизации. Следовательно, возникает возможность использования намного более высокой частоты следования элементарных посылок, и, таким образом, получения значительной связанной с ней скорости передачи данных. Кроме этого также могут применяться разнообразные алгоритмы блокирования сигнала, однако в вычислительном плане они являются достаточно сложными.

В РСПП для шифрования и дешифрования информации необходим один и тот же ключ - псевдослучайный шум, который в идеальном случае имеет плоскую частотную характеристику во всем диапазоне частот (так называемый белый шум). Ключ применяется к скрываемой информации и трансформирует ее последовательность в последовательность с расширенным спектром.

Метод РСПП по отношению к аудиосигналам заключается в следующем. Сигнал данных умножается на сигнал несущей и псевдослучайную шумовую последовательность, характеризующуюся широким частотным спектром. В результате этого спектр данных расширяется на всю доступную полосу. В дальнейшем последовательность расширенных данных ослабляется и прибавляется к исходному сигналу как аддитивный случайный шум.

РСПП использует двоичную фазовую манипуляцию, поскольку фаза сигнала ПСП поочередно чередуется с фазой модулированной двоичной последовательности сообщения.

На стадии извлечения фазовые значения и интерпретируются, соответственно, как биты 1 и 0, которыми кодировалась двоичная последовательность данных. При этом предусматривается следующее:

* Псевдослучайный ключ представляет собой М-последовательность (то есть он имеет максимально возможное количество комбинаций, которые равномерно распределены в заданном диапазоне, и максимально долго не повторяются). Следовательно, он имеет относительно плоский частотный спектр.
* Принимающей стороне известен поток ключей для шифрования. Выполнена синхронизация сигнала, а также известны точки начала и конца расширенных данных.
* Принимающей стороне также известны следующие параметры: частота следования элементарных посылок, скорость передачи данных и частота (вид) несущей.

В отличие от рассмотренного выше фазового кодирования, метод РСПП вводит в звук аддитивный случайный шум. Для того чтобы держать уровень шума низким и неощутимым на слух, расширенный код ослабляется (без адаптации) приблизительно до уровня 0,5% от динамического диапазона звукового файла-контейнера.

Объединение несложной техники повторения и кодирования с исправлением ошибок позволяет гарантировать целостность двоичной последовательности. Короткие сегменты двоичной кодовой комбинации объединяются и складываются с сигналом аудиоконтейнера таким образом, чтобы уменьшить шумы переходных процессов. Для этого в процессе декодирования проводится усреднение по всему сегменту.

* + 1. **Скрытие данных с использованием эхо-сигнала**

Данный метод подразумевает под собой встраивание данных в аудиосигнал-контейнер путем введения в него эхо-сигнала. Данные скрываются изменением трех параметров эхо-сигнала: начальной амплитуды, скорости затухания [(начальная амплитуда - затухание)/ ] и сдвига.

Когда сдвиг (задержка) между первичным и эхо-сигналом уменьшается, начиная с некоторого значения задержки, ССЧ становится не способной обнаружить разницу между двумя сигналами, а эхо-сигнал воспринимается только как дополнительный резонанс. Упомянутое значение трудно определить точно, поскольку оно зависит от качества первичной звукозаписи, типа звука, для которого формируется эхо-сигнал, и, в конечном итоге, — от слушателя.

В общем случае, что для большинства звуков и большинства слушателей смешивание происходит при задержке, соответствующей приблизительно одной миллисекунде.

Стеганокодер использует два времени задержки: одно для представления двоичного нуля («сдвиг»), а другое — для представления двоичной единицы (сдвиг + ). Оба времени задержки меньше того предельного времени, за которое ССЧ способна распознать эхо-сигнал. Кроме уменьшения времени задержки для обеспечения неощущаемости также можно установить уровни начальной амплитуды и времени затухания, которые бы не превышали порог чувствительности ССЧ.

Процесс встраивания данных может быть представлен в виде устройства, которое реализует одну из двух возможных системных функций. Во временной области системные функции — это дискретные во времени экспоненты, отличие между которыми состоит лишь в задержке между импульсами.

* 1. **Скрытие данных в тексте**

Для скрытия конфиденциальных сообщений в тексте (так называемая лингвистическая стеганография) используется или обычная избыточность письменной речи, или же форматы представления текста.

Наиболее сложным объектом для скрытия данных по многим причинам является электронная (файловая) версия текста. В отличие от текстового файла его «жесткая» копия (например, бумажная) может быть обработана как высокоструктурированное изображение и поэтому является относительно легко поддающейся разнообразным методам скрытия, таким как незначительные изменения формата текстовых шаблонов, регулирование расстояния между определенными парами символов (кернинг), расстояния между строками и т.п. В значительной степени такая ситуация вызвана относительным дефицитом в текстовом файле избыточной информации, особенно в сравнении с графическими или, например, звуковыми файлами. В то время как в большинстве случаев существует возможность внести незаметные глазу и неощутимые на слух модификации в изображение и звук, даже дополнительная буква или знак пунктуации в тексте могут быть легко распознаны случайным читателем.

Скрытие данных в тексте требует поиска таких модификаций, которые были бы незаметными подавляющему большинству читателей. В данной работе рассматриваются три группы методов, которые получили наибольшее распространение при встраивании скрываемых данных в текст:

* методы произвольного интервала, которые осуществляют встраивание путем манипуляции с пробельными символами (свободным местом на печатной полосе);
* синтаксические методы, которые работают с пунктуацией;
* семантические методы, в основу алгоритмов которых положено манипулирование словами, зависимое от скрываемых бит данных.
  + 1. **Метод произвольного интервала**

Существует, по меньшей мере, две причины, по которым манипулирование свободным местом в определенных случаях показывает довольно неплохие результаты. Во-первых, изменение количества пробелов в конце текстовой строки не вызывает существенных изменений в значении фразы или предложения. Во-вторых, среднестатистический читатель вряд ли заметит незначительные модификации свободного места страницы текста.

В данной работе рассмотрены три метода, которые для скрытия данных используют свободное место в тексте. Указанные методы оперируют с интервалами между предложениями, пропусками в конце текстовых строк и интервалами между словами в тексте, выровненном по ширине.

**1.5.1.1. Метод изменения интервала между предложениями**

Метод изменения интервала между предложениями позволяет встраивать в текст сообщение, имеющее двоичный формат, путем размещения одного или двух пробелов после каждого символа завершения предложения. В качестве символов окончания предложения могут служить, к примеру, точки в обычном тексте, точки с запятой для кода программ на языке C++ и т.п. При этом единичным пробелом может кодироваться бит 1, двойным — бит 0.

Кроме несомненной простоты, данный метод имеет и ряд недостатков. Во-первых, он не эффективен, поскольку для встраивания незначительного количества бит требуется текст значительного объема. В частности, один бит, который возможно скрыть в одном предложении, эквивалентен скорости передачи данных, соответствующей приблизительно одному биту на 160 байт текстового контейнера, при условии, что в среднем предложение представляет собой две строки по 80 символов каждая.

Во-вторых, возможность скрытия весьма зависит от структуры текстового контейнера (некоторые тексты, как например, верлибры или свободные стихи характеризуются отсутствием стабильных согласованных или однозначных знаков завершения строки).

В-третьих, существуют текстовые редакторы, которые автоматически устанавливают после точки в конце предложения один-два пробела (так называемое автозавершение). И, наконец, непоследовательное и противоречивое использование свободных мест может оказаться достаточно заметным для читателя.

**1.5.1.2. Метод изменения количества пробелов в конце текстовых строк**

Еще один метод использования свободных мест полосы текста для встраивания конфиденциальных данных заключается в добавлении пробелов в конец каждой текстовой строки. Количество добавляемых пробелов зависит от значения встраиваемого бита. Два пробела кодируют один бит на строку, четыре пробела — два бита, восемь — три и т.д. Такой подход позволяет существенно увеличить, по сравнению с предыдущим методом, количество информации, которую можно скрыть в тексте аналогичного объема.

Дополнительные преимущества указанного метода состоят в том, что он может быть применен к любому тексту. Изменения в формате последнего будут в достаточной степени незаметными, поскольку используемые при этом свободные места являются периферийными по отношению к основному тексту.

Недостатком данного (как, в конечном счете, и предыдущего) метода является то, что некоторые программы обработки текста могут непреднамеренно удалять дополнительно внесенные пробелы. Кроме того, характерным недостатком рассматриваемого метода является очевидная невозможность извлечения скрытых данных из бумажной копии текста (из-за невидимости пробелов).

* + - 1. **Метод изменения количество пробелов между словами выровненного по ширине текста**

Данный метод позволяет скрывать данные в свободных местах текста, выровненного по ширине. При этом биты данных встраиваются путем управляемого выбора позиций, в которых будут размещены дополнительные пробелы. Один пробел между словами интерпретируется как “0”. Два пробела — как “1”. В среднем метод позволяет встраивать по несколько бит в одну строку.

Из-за ограничений, которые накладываются выравниванием текста по ширине, не каждый пробел между словами может использоваться для встраивания данных. Для возможности принятия однозначного решения при определении принимающей стороной, какие же именно из пробелов между словами скрывают встроенную информацию, а какие являются частью оригинального текста, предложено использовать метод встраивания, аналогичный манчестерскому кодированию.

В результате упомянутого кодирования биты группируются попарно, причем последовательность “01” интерпретируется как “1”, “10” — как “0”, а пары “00” и “11” считаются пустыми. Например, извлеченное сообщение “1010000111” сводится к “001”, тогда как сообщение “0011110011” представляет пустую строку.

* + 1. **Синтаксические и семантические методы**

Тот факт, что свободное место для встраивания выбирается произвольно, является одновременно как преимуществом, так и недостатком с точки зрения скрытости данных. Обычный читатель может и не заметить манипуляции с текстом, тогда как текстовый редактор способен автоматически изменить количество и размещение пробелов, таким образом разрушая скрытые данные.

Низкая стойкость к атакам, в свете возможного переформатирования документа, выступает одной из причин поиска других методов встраивания данных в текстовые контейнеры. Кроме этого, синтаксические и семантические методы вообще никоим образом не используют свободные места в тексте, кардинально отличаясь от рассмотренных выше алгоритмов. Однако, все они могут использоваться одновременно, дублируя или же дополняя друг друга.

К синтаксическим методам текстовой стеганографии относятся методы изменения пунктуации и методы изменения структуры и стиля текста. Существует немало случаев, когда правила пунктуации являются неоднозначными и несоблюдение их не влияет существенно на общее содержание текста. Так, например, фразы «красный, зеленый, синий» и «красный, зеленый и синий» эквивалентны друг другу. Тот факт, что выбор подобных форм может быть произвольным , и используется при построении стеганосистем на основе синтаксических методов. Периодическое изменение форм при этом может быть поставлено в соответствие с двоичными данными. Например, появление в тексте формы перечисления с союзом «и» может подразумевать под собой встроенный бит “1”, в то время как отсутствие союза при перечислении будет говорить о том, что в данном случае встроен бит “0”. Другим примером может служить использование сокращений и аббревиатур. Средняя скорость передачи данных такими методами составляет несколько бит на один килобайт текста.

Однако в то время как письменный язык предоставляет достаточно возможностей для синтаксического скрытия данных, эти возможности исчезают в известных классических произведениях. Кроме того, хотя некоторые из правил пунктуации и считаются неоднозначными, их противоречивое использование может стать объектом внимания для цензора. Также возможны случаи, когда изменение пунктуации приводит к снижению воспринимаемости текста или же к приобретению текстом диаметрально противоположного смысла. Поэтому синтаксические методы рекомендуется применять с осмотрительностью.

К синтаксическим методам также относятся методы изменения стиля и структуры текста без значительного изменения его смысловой нагрузки. Например, предложение «Существует немало случаев, когда правила пунктуации являются неоднозначными» можно сформулировать как «Правила пунктуации являются неоднозначными во многих случаях». Такие методы являются более незаметными для посторонних, по сравнению с методами изменения пунктуации, однако возможность их использования ограничена сложностью автоматизирования процесса стеганографического встраивания и извлечения бит сообщения.

Семантические методы подобны синтаксическим. Наряду с этим, вместо того чтобы встраивать двоичные данные, используя двусмысленность грамматической формы, семантические методы определяют два синонима, которые отвечают значениям скрываемых бит. К примеру, слово “но”может быть поставлено в соответствие к “0”, а слово «однако» — к “1”.

Для проведения скрытия с использованием семантических методов необходимо наличие таблицы синонимов. Кроме того, как отмечается в, если слову, отвечает достаточно большое количество синонимов, возникает возможность одновременного кодирования большего количества бит. Скажем, выбор между синонимами «секретный», «тайный», «скрытый», «конфиденциальный», «негласный», «неизвестный», «засекреченный», «закрытый» дает возможность представить три бита данных за одно встраивание. Проблемы могут возникнуть, однако, когда желанию встроить бит информации препятствует нюанс значения слова.

**Глава 2. Стеганография с использованием VoIP. Теоретическая часть**

**2.1. Введение в VoIP стеганографию**

VoIP или IP-телефония это сервис работающий в режиме реального времени и позволяющий пользователям совершать звонки через IP сети. Это один из важнейших сервисов, основанных на IP сетях. С недавнего времени VoIP начал привлекать внимание ученых в области стеганографии. Особенности VoIP, которые делают его идеальным носителем для стеганографических целей:

* Популярность, поэтому ее использование не вызовет подозрений.
* Большой объем данных VoIP.
* Потенциально высокая стеганографическая пропускная способность.
* Использование различных протоколов.
* Это сервис в режиме реального времени, что вызывает дополнительные жесткие требования для обнаружения (раскрытие) и одновременно создает новые возможности для использования стеганографии.

В настоящее время методы стеганографии, которые могут использоваться в телекоммуникационных сетях, описывается термином сетевой стеганографии или, в частности, применительно к IP-телефонии, в соответствии с условиями VoIP стеганография или стеганофония. Эти термины относятся к методам сокрытия информации в любом слое TCP / IP стек протоколов (рис. 5).

Рис 5. VoIP стек и протоколы

Основное применение сетевой стеганографии это возможность тайного общение. Целью создания скрытой связи могут быть разными: использование в юридических или незаконных целях. Незаконное применение сетевой стеганографии начинается от криминальной связи, через конфиденциальные данные от проникновения охраняемой системы, вплоть до промышленного шпионажа. Законное использование включает обход Интернет цензуры и наблюдения, компьютерной/сетевой судебно-медицинской экспертизы, или защиты авторских прав. Также методы сетевой стеганографии могут быть использованы для улучшения VoIP, например, улучшить устойчивость к потерям пакетов и качество передачи голоса, предоставлять средства для безопасного распределения криптографических ключей.

**2.2. Методы VoIP стеганографии**

Большинство методов сетевой стеганографии покрывают все уровни ТСР/IP стека. Однако если посмотреть на VoIP с точки зрение трафика, то в общем, каждый метод стеганографии можно отнести к одной из трех групп (рис.6), основываясь на том, что используется в качестве носителя:

* Методы, основанные на модификации протокола PDU – заголовки сетевых протоколов или область полезной нагрузки. Например, модификация свободных или избыточных заголовков IP, UDP, RTP протоколов, в течение разговора, модификация области полезной нагрузки RTP пакетов или просто заменить пользовательские данные.
* Методы, которые меняют временные отношение PDU пакета, путем воздействия на порядок следования пакетов или путем преднамеренных потерь PDU пакетов. Эти методы не были предложены для VoIP, но они используют VoIP трафик.
* Смешанные методы, которые используют и заголовки и временные отношение.

Рис 6. Методы VoIP стеганографии

Последние десять лет активно уделялось внимание методам, основанных на модификации PDU и были опубликованы множества работ с использованием данного метода. Решение 2 группы методов довольно сложно реализовать практически, в связи с низкой пропускной способностью и необходимостью синхронизации. Кроме этого, методы 2 группы могут довольно значительно ухудшить качество связи. Последняя группа методов были разработаны недавно, и на сегодняшний день существуют небольшое количество работ на эту тему.

Польские ученые Мазурчик (Mazurczyk) и Котульски (Kotulski) в 2006[40] году впервые представили, возможность использование VoIP протокола как контейнер для скрытых сообщений. Авторы предложили использовать неиспользуемые поля в заголовках протокола RTP и цифровые водяные знаки для внедрения дополнительной информации в RTP трафик для проверки аутентификации и целостности. Необходимая информация была заложена в неиспользуемые поля в заголовках IP, UDP и RTP протоколов. Позже авторы улучшили им предложенный метод[39], путем включение RTCP (Real-time Transport Control Protocol) без необходимости использования отдельного протокола, что позволяет улучшить пропускную способность.

Внедрение водяных знаков в зашифрованном потоке (например, Skype), чтобы отслеживать его распространение через сеть, обеспечивая тем самым ее де-анонимность, такая идея была предложена Чен (Chen) [11] и соавторами. Водяные знаки внедряются в измененных пакетах в VoIP потоке. Авторы показали, что в зависимости от выбранных параметров водяных знаков, они способны достичь 99% положительных и 0 % ложных срабатываний.

Новые методы сетевой стеганографии, которые могут быть применены к VoIP: протокол установления связи, SIP c SDP(Session Description Protocol)[44] и RTCP были представлены польскими учеными Мазурчик и Шипиорски[45]. Они обнаружили, что комбинация сокрытия информации обеспечивает возможность скрытой передачи около 2000 бит в течении установлении связи и около 2,5 бит/секунду в течение разговора. В 2010 Лойд (Lloyd)[34] продолжил исследования[44], открыв новые методы с использованием SIP и SDP протоколов, и провел экспериментальный анализ на X-Lite (клиент) на работоспособность новых методов.

В работе [45] ,был объявлен новый метод под названием LACK (Lost Audio Packets Steganography), позднее был описан и проанализирован в работе [42]. Алгоритм LACK основан на модификации RTP пакета и его временного отношение. Этот метод использует тот факт, что в протоколе RTP задержанные пакеты не используются для перепостроение, то есть пакеты считаются бесполезными и, следовательно, отбрасываются. Таким образом, скрытая связь возможно путем введения преднамеренных задержек выбранных пакетов RTP и заменяя оригинальную полезную нагрузку в стенограмму. Практическая реализация и оценка нового метода LACK была также представлена Мазурчиком [41], где была проведены исследования на влияния этого метода на качество звука. Исследования метода LACK были продолжены Хамдага (Hamdaqa) и Тахвилдари (Tahvildari), ученые с помощью модифицированного (k,n) предела на основе интерполяции Лагранжа обеспечили надежность и отказоустойчивость системы. Результаты показали, что сложность стегоанализа увеличивается. «Цена» за высокую надежность – потеря стеганографической пропускной способности.

Форбс (Forbes) предложил метод, основанный на который изменяет значение временной метки заголовка RTP для отправки стеганограммы. Теоретически пропускная способность метода 350 бит/с.

Новый метод скрытие данных под названием «inference channel» был представлен Шахом (Shah) и Блейзом (Blaze) [53]. Метод создает внешнее вмешательство на общую коммуникационную среду (например, беспроводная сеть) для отправки скрытых данных. В данной работе [53], описывается внедрение помех в беспроводные каналы для сетей 802.11. Результаты показали, что данный метод обладает довольно низкой пропускной способностью 1бит/2,5с. Также этот метод отлично подходит для VoIP потоков.

Ученые Института телекоммуникаций Варшавского технологического университета предложили новую схему размещения скрытых данных в пакетах VoIP, что дает возможность во время обычных сеансов голосовой связи через интернет передавать из системы украденную информацию.

Метод TranSteg (от transcoding steganography, стеганография с перекодировкой)[43] состоит в повышении полезной нагрузки пакетов VoIP и использовании свободного места для передачи скрытых сообщений. Обычно при стеганографии сжимают скрытые данные, но TranSteg в этом не нуждается – здесь сжимается голосовой трафик, и таким образом высвобождается место для стеганограммы.

Основная новация TranSteg состоит в том, что для выбранного голосового потока находится кодек, который даст аналогичное качество голоса при заметно меньшем объеме передаваемых данных. Затем голосовой поток перекодируется (на этом этапе размер пакета уже известен, а выявить смену кодека невозможно) и в освободившийся блок помещается дополнительное скрытое сообщение. Правда, принимающая система должна быть специально сконфигурирована – ведь пакет содержит данные в одной кодировке, а маркирован как несущий данные другого кодека. Кроме того, она должна знать, как извлекать и собирать скрытое сообщение.

Ученые предложили тестовую реализацию нового метода, и в ходе экспериментов показали, что в течение обычного телефонного звонка длительностью 9 мин можно незаметно передать 2,2 Мб скрытых данных. При этом голосовые пакеты в G.711 перекодировались в G.726, качество передачи речи (MOS) упало с 4.46 до 3.834, что практически неразличимо для человека, а дополнительная задержка, которая потребовалась на переконфигурацию, составила всего 4 мс.

**2.3. Наш метод)**

В связи актуальностью защищенной связи, возникает необходимость разработки такого метода, который бы позволял передавать скрытые сообщения в режиме реального времени. Один из наиболее популярных VoIP сервисов на сегодняшний день это – Skype. Эта система и стала объектом нашего исследования. Данный сервис примечателен не только своей популярностью, но и безопасностью. Skype - программное обеспечение с закрытым кодом, обеспечивающее шифрованную голосовую связь и видеосвязь через Интернет между компьютерами (VoIP).

Предложенный нами алгоритм, заключается в передаче определенной последовательности нулевых бит в определенное время. Совокупность длины последовательности бит и расстояние между последовательностью представляют зашифрованные данные. Алгоритм включает в себя 3 варианта передаваемых данных:

1. Пакет, представляющий бит «0».
2. Пакет, представляющий бит «1».
3. Пакет, определяющий конец символа.

Каждый зашифрованный пакет состоит из двух последовательностей бит, первая последовательность идентифицируют зашифрованный пакет она начинается с 100-го бита и продолжается до 600-го. Данный идентификатор говорит о том, что передаваемый пакет содержит зашифрованную информацию. Вторая последовательность определяет конкретно какую информацию содержит пакет.

Если вторая последовательность имеет длину 400 бит, и начинается через 200 бит после идентификатора, то данный пакет содержит «0».

0

1600

800

100

600

1200

Если вторая последовательность длиной 400 бит и начинается через 600 бит после идентификатора, то пакет содержит «1».

0

1600

800

100

600

1200

Если вторая последовательность длиной 800 и начинается через 200 бит после идентификатора, то пакет несет информацию о конце передаваемого символа.

0

1600

800

100

600

**2.3. Описание проблемы**

Информация является одним из ценнейших предметов современной жизни. Получение доступа к ней с появлением глобальных компьютерных сетей стало невероятно простым.

Задача надежной защиты авторских прав, прав интеллектуальной собственности или конфиденциальных данных от несанкционированного доступа является одной из старейших и не решенных на сегодня проблем.

Использование сетевой стеганографии террористами появились с момента публикации в газете USA Today 5 февраля 2001 года двух статей — «Террористы прячут инструкции онлайн» и «Террористические группы прячутся за веб-шифрованием». 10 июля 2002 года в той же газете появилась статья «Боевики окутывают веб с помощью ссылок на джихад». В этой статье была опубликована информация о том, что террористы использовали фотографии на сайте eBay для передачи скрытых сообщений. Многие средства массовой информации перепечатывали данные сообщения, особенно после терактов 11 сентября.

Предполагаемое использование террористами сетевой стеганографии, ставит необходимость проведение исследование в этой области. Так же, необходимо решить задачу защищенной передачи данных по сети.

**Глава 3. Практическая реализация**

* 1. **Проектирование командной разработки**

Для совместной разработки нами был выбран набор инструментов от Microsoft – Visual Studio Team System (VSTS) и итеративная модель разработки. VSTS позволяет упростить совместную работу над проектами, предоставляет инструменты для тестирования и отладки разрабатываемых программ.

Основной методологией разработки в VSTS является Microsoft Solution Framework. Модель процессов MSF представляет собой общую методологию разработки и внедрения IT-решений. Благодаря своей гибкости данная модель может использоваться для разработки широкого круга IT-решений, она охватывает весь жизненный цикл создания решения, с самых ранних этапов до внедрения. Модель процессов MSF сочетает в себе качества двух классических моделей: каскадной и спиральной.

Процесс MSF ориентирован на «вехи» (milestones). Вехи – ключевые точки процесса разработки, которые характеризуют достижение какого–либо существенного результата.

Модель процессов MSF учитывает постоянные изменения требований к конечному продукту, процесс разработки состоит из коротких циклов и представляет собой поступательное движение от простейших ранних версий продукта к его окончательному виду.

* + 1. **Каскадная спиральная модель**

Модель процессов описывает последовательность действий при реализации проекта, по сути, модель процессов определяет жизненный цикл проекта.

В настоящее время существует множество различных моделей. Рассмотрим подробнее каскадную и спиральную модели, положения которых легли в основу модели процессов MSF.

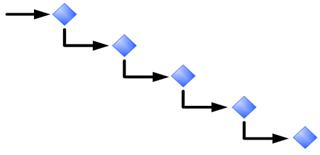


Рис 10. Каскадная модель (waterfall)

Как правило, данная модель применяется в случаях, когда требования к конечному продукту определены на ранних этапах и остаются неизменными, на протяжении всего процесса разработки. Каждый этап разработки заканчивается «вехой» - точкой оценки и перехода к следующему этапу. Каждый следующий этап может начаться только после завершения предыдущего. Фиксированные точки перехода между этапами облегчают распределение ответственности, создание отчетности и следование календарному графику.

На рисунке: Ромбы соответствуют вехам, стрелки – фазам.

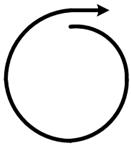


Рис. 11. Спиральная модель.

Данная модель предпочтительна для быстрой разработки небольших решений. В ней учитываются постоянные изменения и пересмотр проектных требований к конечному продукту. Имеет место постоянное и тесное взаимодействие разработчиков с заказчиком, который оценивает ход и результаты работы на протяжении всего процесса работы над проектом. К недостаткам спиральной модели можно отнести низкую степень формализации, порой тяжело понять на какой конкретной стадии (планирование, разработка т пр.) находится проект.

* + 1. **Модель процессов MSF**

Модель процессов MSF объединяет в себе упорядоченность каскадной модели с гибкостью спиральной.

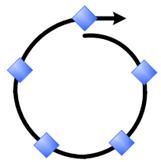


Рис. 12. Модель процессов MSF.

Базовые принципы:

Единое видение проекта. Изначально и у заказчика и у проектной группы имеется своё понимание целей и задач проекта, а также того, что должно быть достигнуто в ходе работы над проектом. Успех проекта невозможен без наличия единого видения и понимания проекта у заказчика и проектной группы, поскольку отсутствие ясности делает невозможным цели.

Поскольку выработка единого видения проекта и следование ему настолько важны, MSF выделяет для этих целей отдельную фазу (этап) разработки – «Выработка концепции».

Проявляйте гибкость – будьте готовы к переменам. В противоположность каскадной модели MSF основывается на принципе непрерывной изменяемости условий проекта при неизменной эффективности управления.

Концентрируйтесь на бизнес - приоритетах. Разрабатываемый продукт должен приносить определенную выгоду или отдачу конечным пользователям. В случае организаций – бизнес – отдачу.

Поскольку программный продукт способен привнести отдачу только после своего внедрения в среду организации, MSF включает в жизненный цикл создания решения фазу внедрения.

Поощряйте свободное общение. Исторически многие организации строили свою работу по принципу need-to-know, то есть сведения к минимуму информированности сотрудников.

Модель процессов MSF предлагает свободный и открытый обмен информацией между всеми участниками проекта, как внутри команды разработчиков, так и с ключевыми заинтересованными лицами. Это должно служить средством снижения недопонимания, неопределенности и необоснованных затрат.

Поэтому, MSF предлагает проводить анализ хода работ в определенных временных точках. Обязательное документирование результатов иллюстрирует прогресс работы над проектом, как для разработчиков, так и для заказчика.[1]

* 1. **Распределение ролей между разработчиками**

MSF основан на постулате о шести качественных целях, достижение которых определяет успешность проекта. Эти цели обуславливают модель проектной группы. В то время как за успех проекта ответственна вся команда, каждый из ее ролевых кластеров, определяемых моделью, ассоциирован с одной из упомянутых шести целей и работает над ее достижением. Шесть ролевых кластеров модели проектной группы – это:

* "Управление продуктом" (product management);
* "Управление программой" (program management);
* "Разработка" (development);
* "Тестирование" (test);
* "Удовлетворение потребителя" (user experience);
* "Управление выпуском" (release management).

Они ответственны за различные области компетенции (functional areas) и связанные с ними цели и задачи. Иногда ролевые кластеры называются просто ролями. Но в любом случае суть концепции остается той же – построить основу производственных отношений и связанную с ней модель команды такими, чтобы они были приспосабливаемыми (масштабируемыми) для удовлетворения нужд любого проекта. Одна роль (или один кластер) может быть представлена одним или несколькими сотрудниками, в зависимости от размера проекта, его сложности и профессиональных навыков, требуемых для реализации всех областей компетенции кластера.

Концепция абстрактных действующих лиц проекта, определяемых выполняемыми ими ролями, допускает при обсуждении управления проектами не рассматривать структуры коллективов на уровне персоналий. Это позволяет более точно представить сгруппированные в виде ролей функции разработчиков, что в свою очередь дает возможность проецировать ролевую структуру проекта на любую структуру коллектива разработчиков.[2]

В нашей работе, мы придерживаемся ролевой структуры проекта, и в нашей группе существуют следующие роли:

* Архитектор – Ролдугин Александр Николаевич
* Разработчики – Шантаев Нурсултан Омурбекович, Ролдугин Александр Николаевич
* Тестировщик – Шантаев Нурсултан Омурбекович, Ролдугин Александр Николаевич
  + 1. **Роль архитектора в проекте**

Архитектор — это член команды, который формирует структуру системы, будь то сеть, центр данных или прикладной программный комплекс. Архитектор обязательно должен быть экспертом в вопросах поддержки, функционирования, совместимости и защищенности продукта, поскольку результаты его работы отражаются на всех перечисленных областях. Кроме того, он должен уметь представить систему с использованием нотации, понятной всем пользователям его проектов – менеджеру, разработчикам, заказчику.

В Microsoft® Visual Studio® Team System определены две роли архитекторов: архитектор инфраструктуры, и архитектор приложений. В команде обе эти роли может выполнять один человек, но в Team System предусмотрена возможность разделения соответствующих функций.

**Архитектор инфраструктуры**

В обязанности архитекторов инфраструктуры входит проектирование центров данных и сетевой инфраструктуры, а также разработка соответствующей документации. Моделируя центр данных, эти архитекторы описывают его аппаратное обеспечение, типы коммуникаций, линии связи, брандмауэры и ограничения, определяемые требованиями безопасности.

Хотя модели инфраструктуры не связаны непосредственно с разработкой программного обеспечения, от них в значительной мере зависит успех приложения: его необходимо проверить на соответствие указанным ограничениям, дабы убедиться, что это приложение будет должным образом функционировать в целевом окружении. Такая стратегия называется проектированием с учетом последующего развертывания (design for deployment).

Для моделирования среды развертывания архитектор инфраструктуры (который привлекает к этой работе и специалистов из отдела эксплуатации), использует конструктор логического центра данных. Но сначала он должен составить себе четкое представление о целевом центре данных, в частности получить сведения о:

* типах хостов приложений (веб-сервер, сервер базы данных, произвольный сервер);
* версиях операционной системы и .NET на каждом сервере;
* наличии брандмауэров и ограничений коммуникационного потока между серверами (HTTP, TDS, универсальный сервер);
* входном и выходном протоколах каждого сервера и зоны;
* любых ограничениях, связанных с отдельными серверами и зонами.

**Архитектор приложений**

Архитектор приложений проектирует взаимосвязанные системы при­ложений, которые в результирующей системе будут представлены как комплексы взаимосвязанных веб-сервисов, веб-приложений, приложений Microsoft Windows®, внешних баз данных и прочих компонентов.

Прежде чем приступить к проектированию, архитектор приложений должен составить себе общее представление о таких параметрах архитектуры будущей системы, как:

количество и типы приложений (веб-сервисов, приложений баз данных, Windows-приложений);

версии операционных систем и .NET;

ограничения коммуникационных потоков между сервисами (HTTP, TDS и прочих);

другие параметры всех сервисов и зон (включая пользователь­ские).[3]

* + 1. **Роль разработчика в проекте**

Роль разработчика в проекте заключается в реализации (кодировании) идей архитектора проекта. На практике функции, выполняемые разработчиками, — понятие неформализованное. В разных проектах оно может обретать свое содержание. Тем не менее, типовые функции, которые предполагают практически все программные проекты, можно перечислить. Так, в любом программном проекте есть функции кодирования, т.е. записи программы на алгоритмическом языке по имеющимся спецификациям, анализа требований, т.е. выявления истинной потребности в создаваемой программе, тестирования и отладки.

Целью исследования в процессе дипломного проекта, было изучение методов командной работы при разработке приложений. Для достижения наиболее лучшего результата, необходимо соблюдение нескольких условий:

* ясное и четкое понимание каждым членом команды своей роли, что позволяет каждому выполнять свои задачи, не пересекаясь с работой других;
* спецификация проекта и график работ согласованы со всеми членами команды;
* все члены команды имеют четкое представление о модели процесса, которая будет использоваться в ходе выполнения проекта;
* каждый член команды должен основательно знать все аспекты плана проекта.

Руководитель группы - менеджер разработки руководит процессами создания архитектуры проекта, детального проектирования и созданием программных модулей.

Разработчики это программисты, которые создают необходимые программные модули. Поскольку создание приложений в архитектуре клиент-сервер требует знаний во многих областях: языки программирования, сети, коммуникации, управление базами данных, это должна быть группа высококвалифицированных специалистов по каждой из описанных областей.

Разработчик реализует проектируемые компоненты, владеет и создает специфичные классы и методы, осуществляет кодирование и автономное тестирование, строит продукт. Это широкое понятие, которое может подразделяться на специальные роли (например, разработчик классов). В зависимости от сложности проекта команда может включать различное число разработчиков.

Группа ролей, объединяемых названием разработчик, весьма многочисленна. Сюда входят и разработчики пользовательского интерфейса, специализирующиеся на создании визуальной части системы, и разработчики баз данных, роль которых в большинстве случаев берут на себя программисты, и интеграторы, отвечающие за построение всей системы из отдельных компонентов.[4]

* + 1. **Роль тестировщика в проекте**

Главная задача тестировщика — обнаружить и сообщить о проблемах программного продукта, которые могут сказаться на качестве. Ключевая задача тестировщика — найти и сообщить о существенных ошибках в продукте, протестировав его. Также в обязанности тестировщика входит точное сообщение о проявлениях ошибки и описание какого-либо обходного пути для уменьшения этих проявлений. Тестировшик описывает ошибки, а также простые в понимании и выполнении шаги для устранения этих ошибок. Он участвует в группе по разработке стандартов качества продукта. Цель тестирования состоит в том, чтобы доказать, что известные функции работают правильно, и обнаружить новые проблемы продукта.

В Visual Studio Team System есть средства для проведения тестирования вручную и автоматически. Также предоставляются возможности занесения результатов тестирования в базу данных, построения графиков, учета и анализа найденных ошибок и т. д.[5]

Для создания какого-либо теста нужно вначале создать тестировочный проект (test project). Впрочем, его можно не создавать отдельно - мастер создания тестов предложит сделать это автоматически. В Visual Studio Team System тестировочные проекты предназначены специально для хранения тестов различных типов. Каждый тестировочный проект связан с определенным языком из поддерживаемых в Visual Studio (Visual Basic, C#...), на котором будут писаться тестовые сценарии (test scripts). Чтобы создать новый тест, нужно выбрать пункт меню "Test – New Test..." и в открывшейся форме

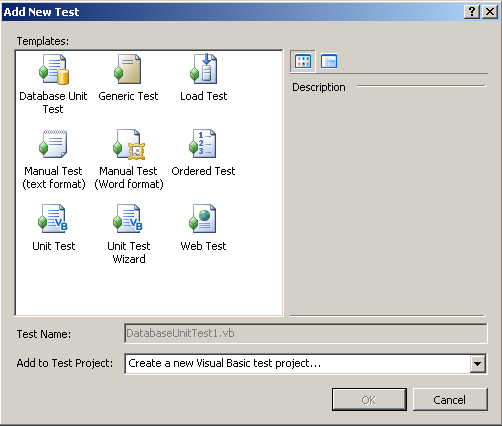


Рис. 13. Создание теста.

**Типы тестов в VSTS**

В VSTEST поддерживается шесть основных типов тестов:

* unit test (модульный тест) – программный тест, при выполнении которого вызываются методы класса и проверяются возвращаемые ими значения;
* manual test (тест, выполняемый вручную) – проводится тестировщиком, а не выполняется автоматически;
* generic test (пользовательский тест) – выполняется внешним приложением, вызываемым из Visual Studio;
* web test (веб-тест) – производятся вызовы http-приложения для проверки его функциональности;
* load test (нагрузочный тест) – одновременно и многократно выполняются модульные, пакетные, пользовательские или веб-тесты;
* ordered test (пакетный тест) – последовательно выполняются тесты из заданного списка.
* Database test (тест базы данных) – используется для тестирования функций, триггеров, индексов и любых других объектов БД.

Для каждого вида тестов используется свой конструктор, либо мастер, который проводит через процедуру создания теста.

* 1. **Взаимодействия между членами группы**

Для реализации взаимодействия между членами рабочей группы, был установлен Team Foundation Server. Он позволяет координировать работу отдельных членов команды, даже если их местоположение удалено друг от друга. В общем случае принцип работы с TFS(Team Foundation Server) выглядит следующим образом:

* Менеджер проекта производит постановку задачи, указывает сроки сдачи и т.д.
* Архитектор проекта разрабатывает архитектуру, реализовывает классы, формирует диаграммы
* Разработчик проекта при подключении к TFS, получает задание и реализует его, результаты работы помещает на сервер
* Тестер обращается к TFS, получает программный код программиста, и производит его тестирование

Для реализации проекта на рабочем компьютере программиста устанавливается пробная версия VS TFS 2010. Данная версия, позволяет реализовывать проекты практически любой сложности.