О допустимых пределах искажений электроакустических речевых сигналов при скрытом встраивании данных

М.О. Пономарь

Предлагаемый метод сокрытия данных в речевых сигналах основан на использовании стеганографии под прикрытием поточной криптозащиты.

Особый интерес для систем скрытой связи по открытым речевым каналам представляют те методы, в которых скрываемые данные внедряются в значения непрерывных несущих параметров: время запаздывания эхо-сигнала, значения фазы спектральной составляющей, значения частоты основного тона и длительности вокализованных сегментов речи. При этом скрываемые данные оказываются достаточно стойкими к воздействию шумов, фильтрованию, сжатию с потерями, вокодерному, аналого-цифровому, цифро-аналоговому преобразованиям и для их извлечения не требуется исходный аудиосигнал [1].

При внедрении дискретных данных в непрерывные характеристики речевого сигнала требуется использовать искусственное квантование его по времени и уровню. Сегментация сигнала на естественные однородные вокализованные стационарные участки является аналогом его квантования по времени, а для квантования значений несущих параметров по уровню наиболее простым в поточной реализации является метод кодирования с модуляцией индекса квантования (Quantization Index Modulatin — QIM) [2].

Физически результат преобразования кодером QIM, например, частоты основного тона (ЧОТ) на передающей стороне, состоит в том, что из естественных, произвольных по частоте сегментов речи, поступающих на вход кодера, на выходе кодера формируются сегменты речи с нормированными частотами, соответствующие центрам интервалов квантования. На приёмном конце канала связи декодер извлекает из них скрытые данные на основе определения значений принятых ЧОТ сегментов и сопоставления их с общей для передающей и приёмной сторон кодовой таблицей [3].

Совершенно очевидно, что нарушитель, обнаружив в речи абонента нормированные частоты (нормированные значения эхо-сигнала, фазы или дли-

Рис. 1. Схема защиты скрываемых данных путём встраивания их с помощью искусственной модификации речи QIM-кодером и передачи вместе с речью в открытый канал под прикрытием криптографического преобразования

тельности вокализованных сегментов речи), легко определяет наличие в канале связи скрытых данных. Он даже сможет сразу прочесть их, в случае если они представляют собой сообщение и переданы открытым текстом в какой-либо из известных ему кодировок. В случае если сообщение зашифровано, в дело вступает криптоаналитик. Как известно, поточные шифры значительно менее стойки к дешифрации, чем блочные, а это значит, что есть шансы и у криптоаналитика. Но главное то, что стегоканал им обнаружен. В зависимости от результатов криптоанализа и цели нарушителя он может продолжать прослушивать канал или воздействовать на него, например, с целью разрушения скрытого сообщения или навязывания получателю ложной информации — то есть превратиться в активного нарушителя.

Из этого следует, что проектируемый стегоканал должен быть защищён от обнаружения и единственным практическим способом его защиты является криптозащита. В данном методе сокрытия демаскирующим признаком скрываемого сообщения являются нормированные значения несущих параметров, то есть необычные, неестественные статистические свойства заполненного контейнера по сравнению с пустым. Можно применить дизеринг — добавление небольшого шумового сигнала, делающего основной сигнал более естественным, но это затруднит его декодирование.

Единственным радикальным решением является криптозащита. Это значит, что на выходе кодера QIM необходимо иметь криптографический преобразователь, который формирует из нормированных параметров новые параметры, похожие на произвольные, естественные для человеческой речи, но зашифрованные. На приёмном конце происходит сначала расшифровывание каждого параметра, а затем QIM-декодирование его с целью извлечения скрытых данных. Нарушитель при этом не получает никаких сведений о наличии скрытого сообщения, а тем более не может его прочитать, так как он прослушивает полностью естественную речь, часть параметров которой

Таблица 1

Пример встраивания данных с использованием скрывающей модификации частоты основного тона сегментов речи с их криптозащитой

№ сегмента	1	2	3	4	5	6	7	8	9
ЧОТ пустого контейнера, container	114	114	114	130	206	115	115	159	206
Стего-коды пустого контейнера	14	14	14	30	06	15	15	59	06
Стего-вложение (симв/ дв/дес)	**рус/** 00000/ 0	**К/** НПО/ 30	**О/** 00011/ 03	**М/** 00111/ 39	**П/** 01101/ 13	**А/** 11000/ 24	**Н/** 00110/ 06	**И/** 01100/ 44	**Я/** 1110/ 29
ЧОТ заполненного стегокон тейнера, stego	100	130	103	139	213	124	106	144	229
Гамма (симв/дв)	**Лат./** 11111	**Е/** 10000	**М/** 00111	**Б/** 10011	**Е/** 10000	**Б/** 10010	**I/** 01100	**N/** 00110	**C/** 01011
Шифро-текст (симв/дв/ дес)	**Лат./** 11111/ 31	**C/** 01110/ 14	**Проб/** 00100/ 04	**S/** 10100/ 20	**Q/** 11101/ 18	**R/** 01010/ 10	**R/** 01010/ 10	**R/** 01010/ 10	**F/** 1011/ 22
ЧОТ заполненного крипто-стегокон-тейнера, stego + crypto	131	114	104	120	218	110	110	110	222

модифицирована в пределах психоакустической нормы с использованием криптозащиты.

Необходимо подчеркнуть, что это достаточно сложная задача, поскольку слуховое восприятие настолько совершенно, что позволяет опознать самые тонкие оттенки речевого сигнала. Человеческий слух, а тем более слух акустического стегоаналитика довольно точно определит признаки искусственности и естественности речи. И при встраивании данных необходимо учитывать два фактора: неслучайность характера сигналов незаполненного речевого контейнера и сохранение его качества при встраивании и шифровании данных.

Таким образом, речевой сигнал при встраивании в него данных и их извлечении должен претерпевать два прямых и два обратных стего- и криптопреобразования. Покажем, что эти преобразования алгоритмически реализуемы.

Воспользуемся для этого примером встраивания данных в модификацию ЧОТ, приведённым в работе [1]. Речевой контейнер со словами «Wow… Sound editing just…» длительностью 2 сек. разделён на 9 участков с приведёнными в первой строке таблицы 1. ЧОТ (в целых числах Гц, плотность вложения — 1 буква на сегмент речи в гомофоническом коде типа МТК-2 , интервал стегодекодера от –0,5 Гц до +05 Гц).

Достаточно длительные незашифрованные последовательности, подобные приведённому в строке 3 встроенному слову КОМПАНИЯ, будут легко обна-

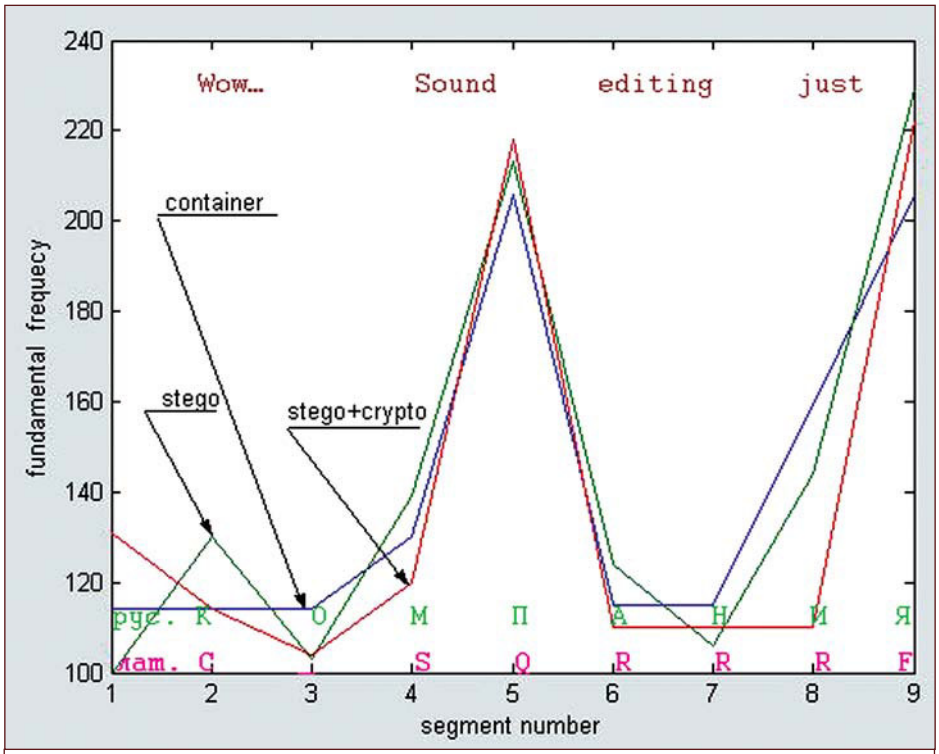

Рис. 2. Пример встраивания данных в речь путём стего- и криптопреобразования частоты основного тона сегментов речи.

ружены и прочтены нарушителем путём статистического анализа заполненного контейнера. Произведём шифрование этого слова. В поточных шифраторах каждый бит исходной информации шифруется с помощью гаммирования — наложения обратимым образом на открытые данные последовательности псевдослучайных чисел. В данном примере в качестве гаммы использовано слово EMBEDING с наложением побитовым «исключающим ИЛИ» (XOR). Получившийся шифротекст C_SQRRRF уже менее доступен криптоаналитику для прочтения, однако шифрование может повлиять на качество речевого контейнера (рис. 2), а значит, вызвать подозрение о наличии в нём стеговложения.

Но, как видно из рис. 2, в данном случае шифрование не вызвало деградирующих преобразований ЧОТ речи, сохранены и мелодический контур, и интонация фразы. Это объясняется тем, что используемый стегокод имеет диапазон значений от 0 до 99, а накладываемая гамма в коде МТК-2 только от 0 до 32. Можно назвать это криптографическим поточным преобразованием, сходным с дизерингом. Вполне очевидны метод расшифрования повторным наложением гаммы на текст и метод извлечения данных с использованием кодовой таблицы. Однако вопрос о допустимом пределе модификации ЧОТ и других параметров речи с учётом шифрования требует дальнейшего исследования.

Решение задачи встраивания данных в речь с шифрованием также тесно связано с непростой задачей синхронизации потокового криптопреобразования, которое требуется как при начале скрытой передачи данных, так и при её продолжении в случае временной потери связи. Трудность обеспечения синхронизации, по мнению некоторых авторов,

превращается в достоинство с точки зрения обеспечения скрытности передачи [4], поэтому использование искусственных средств синхронизации — синхронизирующих посылок, меток, заголовков и т.п. — является крайне нежелательным. Необходима оценка технической сложности задачи синхронизации и возможности её решения в реальном масштабе времени.

Заключение

В данной работе не рассматривались вопросы реализации и криптографической стойкости поточных шифров, это является предметом исследования других специалистов. Однако использование стегокодирования в сочетании с шифрованием вносит дополнительные искажения в речевой сигнал и задержки при его передаче. В связи с этим при разработке технических средств скрытной передачи данных потребуется определить их допустимые пределы, как с точки зрения скрытности, так и с точки зрения возможности поддержания речевой коммуникации.

Литература

1. Ponomar, Marina. Data hiding in speech signals on the basis of the modification of segment pitch and duration. 19th International Congress on Acoustics ICA2007MADRID, 2–7 Sept. 2007, Madrid, Spain, 2007, CAS-03-023, p.46.

2. Chen B., Wornel G.W. System, method, and product for information embedding using an ensemble of non-intersecting embedding generators. U.S. patent pending. Licensing info.: MIT Technology Lic. Office. 1996.

3. M.O.Ponomar. Coding with the Quantization of Speech Signal Carrier Features for Data Hiding. XX Session of the Russian Acoustic Sociaety. T.3. M.: GEOS, 2008, p.645–648.

4. Грибунин В.Г., Оков И.Н., Туринцев И.В. Цифровая стеганография. «Методы и технические средства обеспечения безопасности информации». — СПб.: ГТУ, 2001.

Пономарь Марина Олеговна,
аспирантка кафедры прикладной и экспериментальной лингвистики Московского государственного лингвистического университета (Москва, ул. Остоженка, д.38).
E-mail: oponomar@inbox.ru.