
MOBILE TRUST
TELECOMMUNICATIONS

Оглавление

Предисловие 2
1. Термины и определения 2
2. Стойкость шифрующих преобразований 3
2.1. Общие понятия о стойкости речевых преобразований. Числовые характеристики стойкости. 3
2.1.1 Маскировка речи 3
2.1.2. Разборчивость при прослушивании на ложном ключе. 4
2.1.3. Размеры переставляемых фрагментов речевого сигнала 4
2.1.4. Сложность шифрующих преобразований 5
2.1.5. Трудозатраты на дешифрование (восстановление речи по зашифрованному сигналу). 6
2.2. Метод аналогов при оценке стойкости. 7
3. Оценки стойкости разработанных шифрующих преобразований 8
3.1. Обоснование криптографической стойкости шифрующих речевых преобразований SCR4 8
3.1.1. Общая характеристика 8
3.1.2. Описание преобразований 8
3.1.3. Числовые характеристики стойкости SCR4 9
3.1.3.1.Маскировка речи. 9
3.1.3.2. Число вариантов зашифрования 11
3.1.3.3. Объем фрагмента речевого сигнала. 11
3.1.3.4. Прослушивание на ложном ключе. 11
3.1.4. Аналоги 12
3.2. Обоснование криптографической стойкости шифрующих речевых преобразований на основе однополосных временных перестановок с некратной коммутацией 13
3.2.1. Общая характеристика преобразований 13
3.2.2. Описание преобразований 13
3.2.3. Характеристики сложности преобразований 13
3.2.3.1. Число вариантов зашифрования 1 секунды речи. 13
3.2.3.2. Качественные характеристики сложности преобразований. 14
3.2.3.3. Маскировка речи. 14
3.2.3.4. Средний объем переставляемых фрагментов речевого сигнала. 15
3.2.4. Аналоги 15
4. Заключение 16

Предисловие

Данный материал содержит описание и обоснование стойкости шифрующих речевых преобразований, разработанных для использования в голосовом тракте сетей мобильной связи и Skype, Viber.

1. Термины и определения

В настоящем разделе содержатся основные термины и определения, которые используются по последующих разделах. Прочие термины поясняются непосредственно в самом тексте.

Фрагменты речевого сигнала — случайным образом выделенные звуки речи, или их части, ограниченные по времени и по частоте. Такие фрагменты могут сохраняться целиком в зашифрованном сигнале, однако они должны быть настолько малы, чтобы не позволить определить смысловое содержание речи.

Криптографическая стойкость шифрующих преобразований — устойчивость к возможному восстановлению злоумышленником зашифрованной информации. При высокой стойкости восстановление исходной речи невозможно или чрезвычайно затруднительно.

Дешифрование — восстановление зашифрованной информации. При дешифровании в отсутствие истинного ключа целью злоумышленника может являться восстановление смысла сказанного, при этом сам исходный речевой сигнал, вообще говоря, может восстанавливаться с искажениями.

Остаточная разборчивость речи — процент единиц речи, правильно идентифицируемых при прослушивании зашифрованного сигнала. Под единицами речи чаще всего понимают слова, предложения или слоги. При высокой стойкости шифрующих преобразований остаточная разборчивость близка к нулю.

Маскировка речи — понятие, употребляемое наряду с понятием остаточной разборчивости, имеющее противоположный смысл, то есть чем меньше

остаточная разборчивость, тем больше маскировка речи.

Объем фрагмента речевого сигнала — произведение длительности фрагмента речевого сигнала на ширину полосы его частот. Эта величина является безразмерной и характеризует сложность шифрующих преобразований. Чем меньше этот объем, то есть, мельче переставляемые фрагменты речевого сигнала, и чем больше число вариантов их перестановки, тем выше стойкость преобразований, то есть тем трудней восстановить исходный речевой сигнал и понять смысл сказанного.

2. Стойкость шифрующих преобразований

Настоящий раздел содержит общие сведения, необходимые для понимания обоснования стойкости преобразований.

2.1. Общие понятия о стойкости речевых преобразований. Числовые характеристики стойкости.

Мозаичные шифрующие преобразования осуществляют перестановки мелких фрагментов речевого сигнала по частоте и по времени.

Криптографическая стойкость мозаичных шифрующих преобразований характеризуется несколькими числовыми параметрами.

2.1.1 Маскировка речи

Некоторые преобразования слабо «перемешивают» фрагменты речевого сигнала, и для понимания смысла сказанного бывает достаточно просто внимательно один или несколько раз прослушать зашифрованную речь. В этом случае говорят о высокой остаточной разборчивости или низкой маскировке речи. Хорошие (стойкие) шифрующие преобразования характеризуются низкой остаточной разборчивостью, практически равной нулю.

2.1.2. Разборчивость при прослушивании на ложном ключе.

При наличии у злоумышленника аналогичного речевого шифратора он может включить его на прием и прослушать расшифрованную речь. Однако для этого ему необходимо знание ключа, так как при установке неверного ключа расшифрование произойдет неправильно, и вместо исходной речи будет звучать речь с переставленными фрагментами. Вообще говоря, если преобразования обладают низкой стойкостью, то при этом не исключается возможность разобрать отдельные слова. Стойкие преобразования характеризуются низкой разборчивостью при прослушивании на случайном ложном ключе, так как при этом происходит как бы двукратное перемешивание фрагментов речевого сигнала.

2.1.3. Размеры переставляемых фрагментов речевого сигнала

Размеры характеризуются длительностью переставляемых фрагментов и шириной полосы занимаемых ими частот. Эти размеры для повышения криптографической стойкости должны быть как можно меньше, однако при этом ухудшается качество восстановленной на приеме речи. Поэтому эти размеры выбираются исходя из компромиссных соображений — с одной стороны, чтобы обеспечить необходимую стойкость, а с другой — чтобы не слишком ухудшить качество речи. Обычно длительность переставляемых фрагментов не превосходит длительности звука речи (100 мс), а ширина полосы частот этих фрагментов может составлять 300 - 2000 Гц, то есть часть спектра речевого сигнала, который составляет не менее 3 КГц. Обобщенной характеристикой фрагмента речевого сигнала может служить его объем, то есть произведение длительности на ширину занимаемой полосы частот. Если переставляемые фрагменты различаются по этим параметрам, то можно рассматривать среднее значение объема.

2.1.4. Сложность шифрующих преобразований

Сложность шифрующих преобразований, непосредственно влияющая на

их стойкость, определяется некоторыми качественными и количественными характеристиками.

Число вариантов зашифрования 1 секунды речевого сигнала - основная числовая характеристика *сложности* преобразований. Для оценки числа вариантов зашифрования можно использовать расход шифрующей гаммы в единицу времени.

Следует, однако, отметить, что это число не связано напрямую со *стойкостью* преобразований.

Частотно-временная трансформация (ЧВТ) — преобразование речевого сигнала, при котором сигнал сжимается по времени и растягивается по частоте, или наоборот. На приеме производится обратная операция. При сильной трансформации речевого сигнала (в 2 – 4 раза) существенно снижается его разборчивость и пропадает узнаваемость говорящего.

Кадровая структура и скользящая шкала. Для перестановки фрагментов сигнала по времени существует два различных типа способов. В одном из них, называемых «перестановками с кадровой структурой», весь сигнал делится на кадры фиксированной длительности, и перестановка производится внутри кадра. При этом существует возможность перебора всех возможных вариантов перестановок внутри кадра, что является слабостью этого типа перестановок. При втором способе, называемом перестановками со «скользящей шкалой», нельзя выделить такие кадры, и это существенно усложняет восстановление речи по зашифрованному сигналу из-за увеличения числа перебираемых вариантов перестановок.

Различие длительностей переставляемых фрагментов сигнала. Переставляемые фрагменты речевого сигнала могут иметь одинаковые или различные длительности. В первом случае достаточно несложно произвести единую разметку зашифрованного сигнала, отметив границы переставляемых фрагментов, так как такая разметка в этом случае буде периодической. Во втором случае для дешифрования необходимо по отдельности определять границы каждого из переставляемых фрагментов, что существенно усложняет

процесс дешифрования и снижает его эффективность.

Маскировка моментов коммутаций при различных длительностях фрагментов создает дополнительные сложности в дешифровании, так как снижает надежность определения истинных вариантов перестановки фрагментов речи. Такая маскировка может производиться путем сглаживания переходных участков между соседними фрагментами зашифрованного речевого сигнала. Однако, такая операция ранее в известных шифраторах не использовалась, и исследований ее влияния на стойкость не проводилось.

2.1.5. Трудозатраты на дешифрование (восстановление речи по зашифрованному сигналу).

Ввиду того, что в переставляемых фрагментах структура речевого сигнала остается неизменной, существует принципиальная возможность дешифрования - восстановления зашифрованной речи, однако, чем сложней преобразования, тем более высокой квалификацией нужно обладать, и тем больше труда необходимо затратить на это. Для восстановления наиболее сложных преобразований необходимо иметь профессиональную квалификацию на уровне специалистов спецслужб индустриально развитых стран, располагать соответствующим оборудованием и тратить по несколько часов на восстановление одной секунды речи.

Оценка трудозатрат, необходимых для дешифрования единицы времени речевого сигнала, является основной характеристикой стойкости.

Однако, получение оценки трудозатрат на дешифрование представляет собой сложную задачу, решение которой в полном объеме само по себе весьма трудоемко, так как предполагает построение и реализацию метода дешифрования и проведение контрольного восстановления речевого сообщения с оценкой времени необходимого для восстановления смысла сообщения. Поэтому часто используется метод аналогов, изложенный в следующем разделе.

2.2. Метод аналогов при оценке стойкости.

Наиболее простым методом, который позволяет оценить стойкость

речевых преобразований, является сравнение с известными преобразованиями, стойкость которых уже оценена. Для этого необходимо подобрать наиболее близкий аналог преобразований, стойкость которых оценивается, выявить имеющиеся отличия рассматриваемых преобразований от аналога и оценить их влияние на стойкость преобразований.

3. Оценки стойкости разработанных шифрующих преобразований

3.1. Обоснование криптографической стойкости шифрующих речевых преобразований SCR4

3.1.1. Общая характеристика

Шифрующие преобразования SCR4 являются мозаичными 4-полосными частотно-временными перестановками с кратной коммутацией, частотно-временной трансформацией с постоянным коэффициентом и так называемой скользящей шкалой.

3.1.2. Описание преобразований

После предварительной фильтрации фильтром нижних частот с частотой среза 3 Кгц речевой сигнал, поступающий на вход речепреобразующего блока, преобразуется в цифровую форму с частотой дискретизации 6 КГц. С помощью фильтра нижних частот с частотой среза 1500 Гц входной сигнал расфильтровывается на 2 полосы, при этом сигнала нижней части спектра остается в прямом положении, а сигнал верхней части спектра инвертируется относительно частоты 1500 Гц. Таким образом обе частотных полосы сигнала оказываются в полосе частот от 0 до 1500 Гц. Производится двукратная децимация (прореживание отсчетов) сигнала, то есть из каждых двух подряд идущих отсчетов оставляется только 1, что приводит к сжатию сигнала по времени и растяжению по частоте в 2 раза. Сигнал каждого из полуканалов повторно подвергается такой же операции, в результате этого исходный сигнал расфильтровывается на 4 по-

лосы частот шириной 750 Гц, и сигнал в каждой полосе сжимается по времени и растягивается по частоте в 4 раза, то есть занимает полосу частот 3 КГц. Сигналы каждой из 4 частотных полос делится на фрагменты длительностью 60 мс, которые записываются в буферы задержки. Для каждого из подканалов используется по 2 буфера. Считывание из этих буферов производится в порядке, который задается шифрующей гаммой. При этом производится перестановка отрезков сигналов подканалов по времени. Считанный из буфера отрезок сигнала подканала поступает на выход блока речевых преобразований и передается в канал связи. На приемной стороне производится обратная перестановка фрагментов речевого сигнала с одновременной обратной частотно-временной трансформацией.

3.1.3. Числовые характеристики стойкости SCR4

3.1.3.1.Маскировка речи.

Ввиду четырехкратной частотно-временной трансформации зашифрованный сигнал напоминает на слух птичье щебетанье — из-за четырехкратного повышения частоты основного тона речевого сигнала. При этом полностью отсутствует как разборчивость, так и узнаваемость говорящего, то есть зашифрованная речь не содержит признаков индивидуальности и половой принадлежности.

Для иллюстрации преобразования можно привести спектрограмму исходного речевого сигнала (рис. 1) и зашифрованного сигнала (рис. 2).

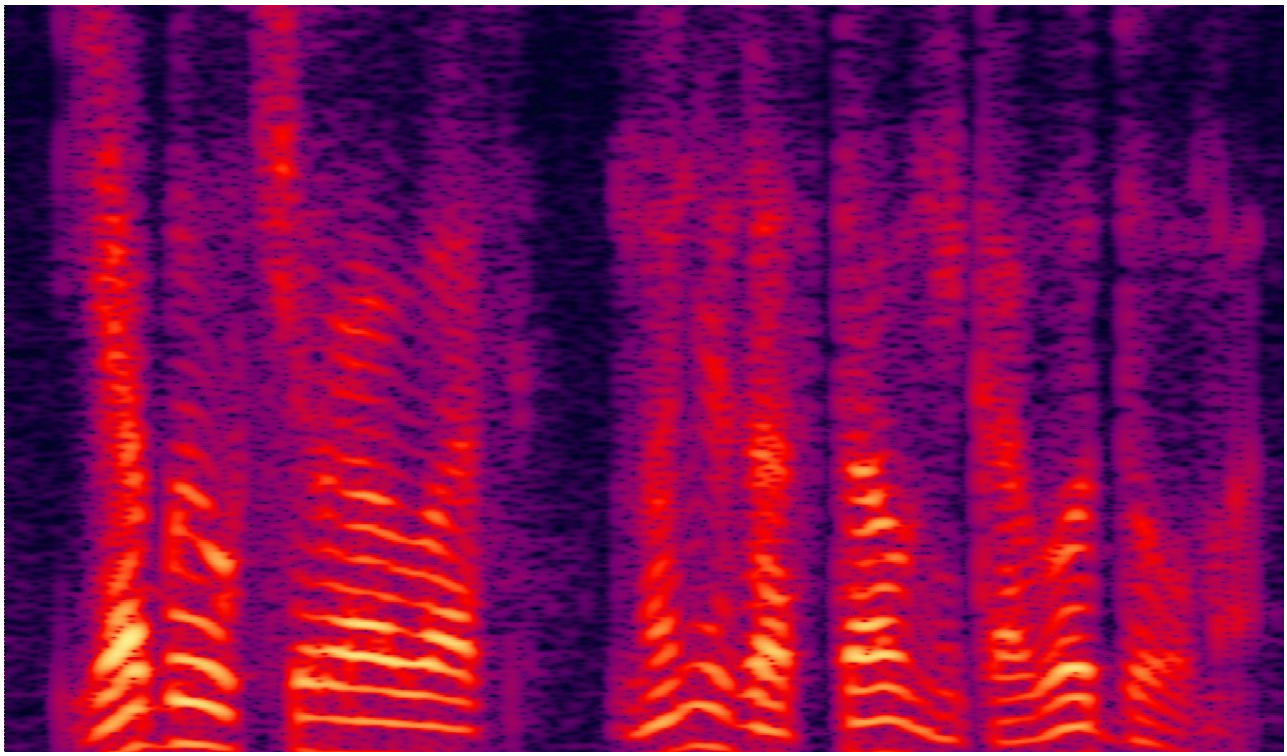

Рис. 1. Спектрограмма исходного речевого сигнала.

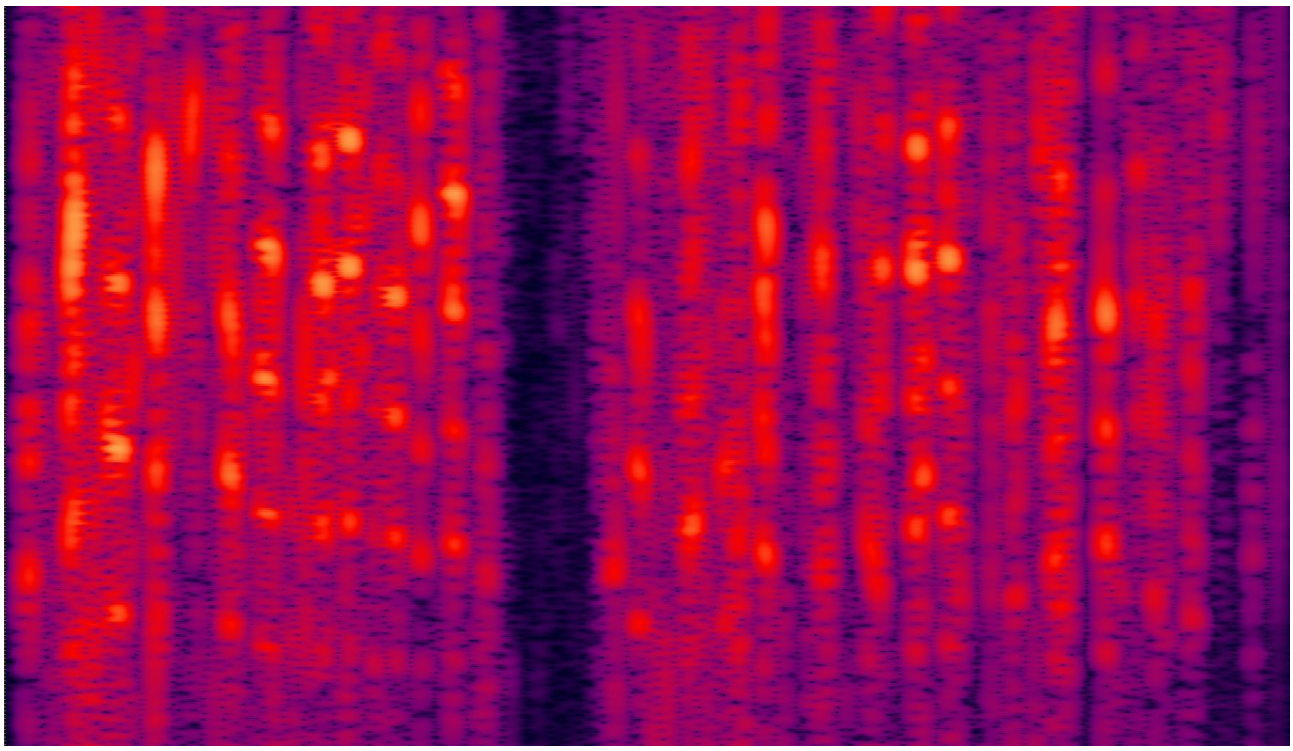

Рис. 2. Спектрограмма зашифрованного сигнала с помощью шифратора SCR4.

3.1.3.2. Число вариантов зашифрования

Один минимальный фрагмент речевого сигнала, имеющий длительность 60 миллисекунд и ширину полосы частот 750 Гц, может получить 5 вариантов временной задержки и занять одно из 4 возможных положений по частоте. При этом на 4 таких фрагмента при зашифровании расходуется 12 знаков двоичной гаммы. Таким образом, на зашифрование 1 секунды речевого сигнала расходуется 12/0,06=200 двоичных знаков гаммы. Исходя из этого, общее число вариантов зашифрования 1 секунды речи составляет $2^{200}=10^{65}$.

3.1.3.3. Объем фрагмента речевого сигнала.

Длительность отрезка составляет $\Delta t=60$ миллисекунд, а ширина полосы частот $\Delta f=$ 750 Гц. Таким образом, объем фрагмента речевого сигнала для рассматриваемых шифрующих преобразований составляет

$$V=\Delta t*\Delta f=0,06*750=45$$

3.1.3.4. Прослушивание на ложном ключе.

При прослушивании, как показывают экспертные оценки, разборчивость весьма низкая. Спектрограмма сигнала, зашифрованного с помощью шифратора SCR4 и расшифрованного на ложном ключе, изображена на рисунке 3. На ней можно видеть, что фрагменты речевого сигнала перемешаны по частоте и по времени.

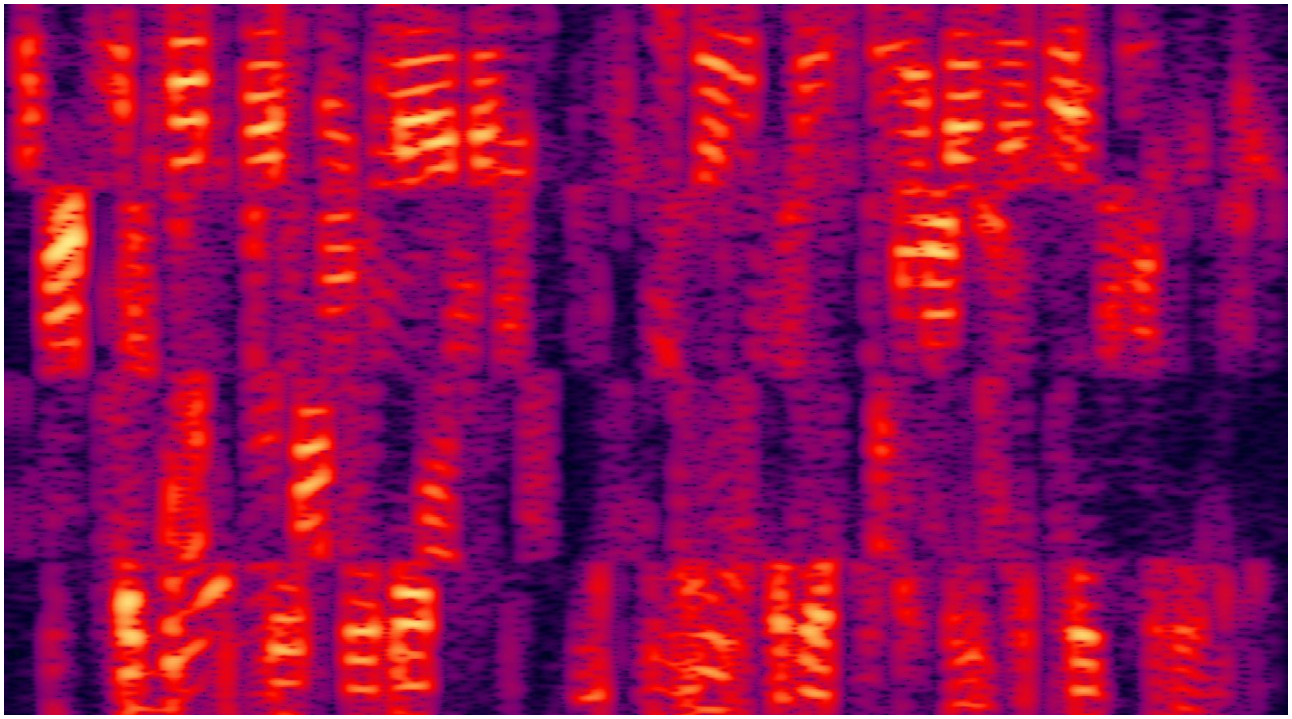

Рис. 3. Спектрограмма сигнала, зашифрованного с помощью шифратора SCR4 и расшифрованного на ложном ключе.

3.1.4. Аналоги

Преобразования SCR4 по принципу и параметрам (длительность и полоса частот некоммутируемых отрезков речи, расход шифрующей гаммы, число вариантов зашифрования 1 секунды речи, скользящая шкала, кратная коммутация) близки к известным шифрующим преобразованиям, для которых ранее были доказано, что они обладают высокой стойкостью.

Отличие заключается в наличии в SCR4 частотно-временной трансформации, повышающей уровень маскировки речи в зашифрованном сигнале по сравнению с известными преобразованиями.

3.2. Обоснование криптографической стойкости шифрующих речевых преобразований на основе однополосных временных перестановок с некратной коммутацией

3.2.1. Общая характеристика преобразований

Преобразования реализуют временные перестановки со «скользящей шкалой», различной длительностью переставляемых фрагментов сигнала и маскировкой границ переставляемых фрагментов.

3.2.2. Описание преобразований

Речевой сигнал, поступивший на вход, разбивается на фрагменты различной длительности, которая задается шифрующей гаммой. Длительности фрагментов составляют от 20 до 70 мс, средняя длительность составляет примерно 45 мс. На выработку значения длительности расходуется 3 или 4 двоичных знака гаммы. Перестановка фрагментов осуществляется с помощью буфера задержки в полной полосе частот, занимаемой сигналом, которая является параметром преобразования и ширина полосы частот составляет от 2,7 до 3,5 КГц. Порядок выборки фрагментов из буфера задается гаммой, на выбор каждого фрагмента расходуется также 3 или 4 знака гаммы. При этом фрагмент может инвертироваться по времени, или считываться в исходном положении, в зависимости от соответствующего знака гаммы. Таким образом, на зашифрование одного фрагмента речевого сигнала, средняя длительность которого составляет 45 миллисекунд, расходуется от 7 до 9 двоичных знаков гаммы.

3.2.3. Характеристики сложности преобразований

3.2.3.1. Число вариантов зашифрования 1 секунды речи.

Исходя из описания преобразований, на зашифрование одного фрагмента речевого сигнала расходуется от 7 до 9 знаков. С учетом того, что средняя

длительность фрагмента составляет 45 мс, на зашифрование 1 секунды речи расходуется

N=(7 …9)/0,045=156…200 бит гаммы шифратора.

Таким образом, число вариантов зашифрования 1 секунды речевого сигнала составляет от $2^{156} = 6{,}4*10^{46}$ до $2^{200}=10^{65}$.

3.2.3.2. Качественные характеристики сложности преобразований.

Различные длительности переставляемых фрагментов речевого сигнала наряду со скользящей шкалой и маскировкой границ переставляемых фрагментов характеризуют данное преобразование как весьма сложное.

С целью повышения криптографической стойкости преобразования осуществляется переход от кратной к некратной коммутации, т.е. к переменной длительности некоммутируемых отрезков.

Так же с целью повышения криптографической стойкости снижена средняя длительности переставляемых фрагментов речевого сигнала, что позволяет снизить остаточную разборчивость речи и увеличить число вариантов зашифрования.

З.2.3.3. Маскировка речи.

По предварительным оценкам, имеется узнаваемость диктора, и иногда возникает впечатление, что слышны знакомые слова. Вместе с тем, такое впечатление может складываться из-за речеподобности зашифрованного сигнала. Поэтому для достоверной оценки маскировки речи необходимо провести артикуляционные измерения, состоящие в экспертной оценке остаточной разборчивости речи по результатам прослушивания экспертами зашифрованного речевого сигнала. Спектрограмма зашифрованного сигнала приведена на рис. 4. Как можно видеть, спектрограмма имеет сходство со спектрограммой речевого сигнала (см. рис. 1), что также свидетельствует о речеподобности сигнала, зашифрованного с помощью данных преобразований.

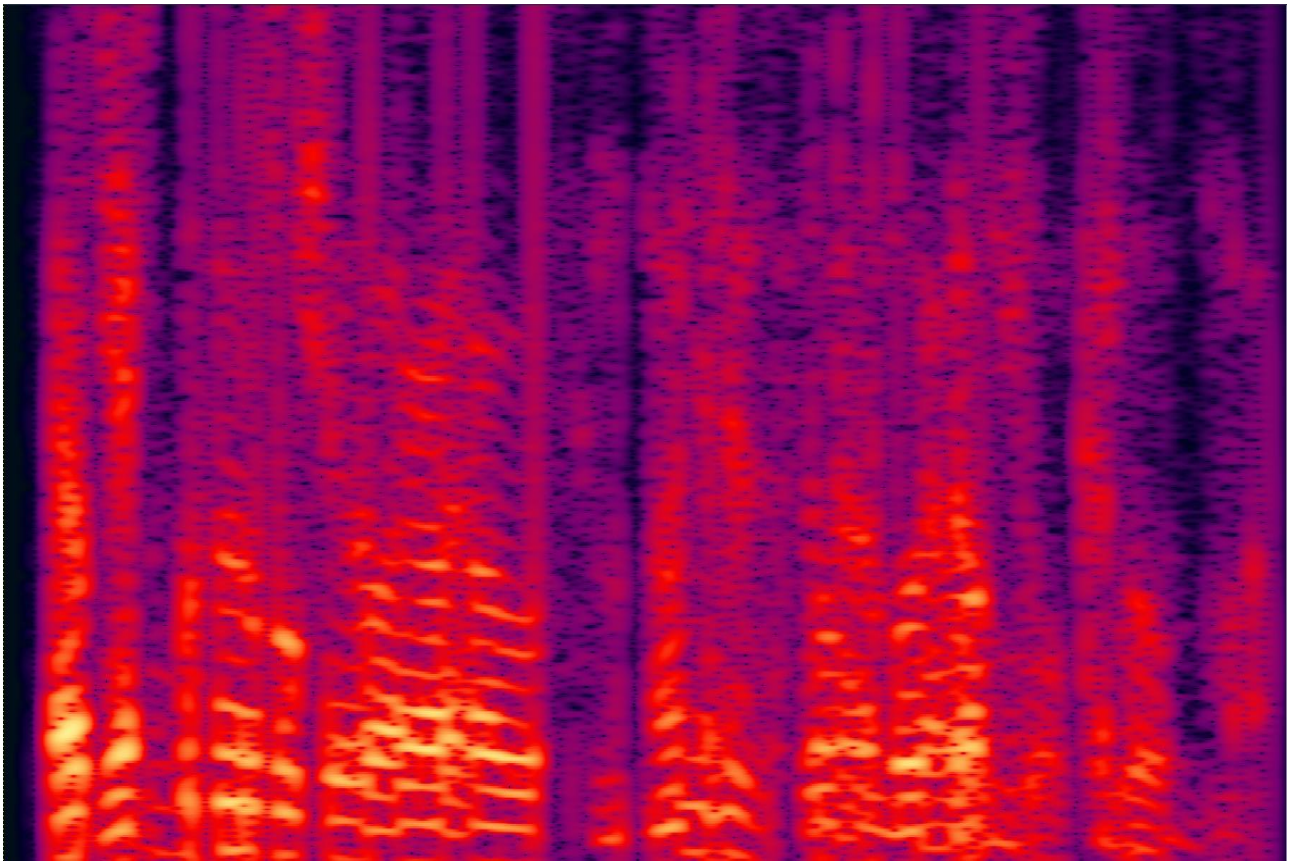

Рис. 4. Спектрограмма зашифрованного сигнала с помощью однополосных преобразований.

3.2.3.4. Средний объем переставляемых фрагментов речевого сигнала.

Средняя длительность фрагмента составляет $\Delta t=45$ миллисекунд, а ширина полосы частот примерно равна $\Delta f= 3000$ Гц. Таким образом, средний объем переставляемых фрагментов речевого сигнала для рассматриваемых шифрующих преобразований составляет

$$V=\Delta t*\Delta f=0{,}045*3000=135,$$

что в несколько раз выше, чем для преобразований SCR4, что косвенно характеризует эти преобразования как менее стойкие.

3.2.4. Аналоги

Наиболее близким аналогом является известная аппаратура, криптосхема которой реализует двухканальные частотно-временные перестановки со скользящей шкалой и различными длительностями переставляемых фрагментов.

Известно, что данные преобразования обладают довольно высокой стойкостью, и не могут быть дешифрованы автоматически.

4. Заключение

Шифратор Stealthphone Hard в режиме Crypto Voice over GSM осуществляет динамическое частотно-временное преобразование каждого отрезка речевого сигнала с последующей перестановкой сигналов во времени.

Прослушивание зашифрованного сигнала не позволяет восстановить смысл речи или узнать говорящего.

Шифрование осуществляется путем перестановки во времени фрагментов речевого сигнала с дополнительными спектральными преобразованиями. Особенностью алгоритма шифрования является невозможность временной сегментации зашифрованного речевого сигнала для перебора ключей.