

ШУМЫ КАК ФАКТОР УВЕЛИЧЕНИЯ ПРОПУСКНОЙ СПОСОБНОСТИ СТЕГАНОКАНАЛА

Рассмотрено влияние уровня шума в аналоговом сигнале, который после аналого-цифрового преобразования используется как контейнер для передачи скрытого сообщения. Получены аналитические оценки, приведен пример возможной аппаратной реализации.

Ключевые слова: пропускная способность, стеганоканал, квантование.

При использовании наиболее популярного и распространённого в настоящее время метода замены наименее значащего бита (НЗБ) максимальная информационная нагрузка контейнера с точки зрения возможности переноса скрытой от постороннего наблюдателя информации (скрытого вложения - СВ) определяется, по сути, только несовершенством зрительного либо слухового аппарата человека [1,2]. Исходя из экспериментальных данных, разрешающая способность слуха и зрения человека находится в границах $2^7 \dots 2^8$ различных значений яркости или громкости. Отсюда следует и усреднённая оценка пропускной способности стеганоканала: около 10% объема контейнера при использовании метода НЗБ. Однако при наличии помех в канале передачи либо наложения шума на файл контейнера число различных человеком уровней яркости или громкости существенно снижается, следовательно, и объем вложения теоретически может быть увеличен. Именно этот факт лежит в основе предлагаемого ниже подхода.

Вначале мысленно проделаем следующий эксперимент (который несложно провести и физически). Возьмем в качестве исходного некоторый графический файл (фотографию или рисунок) $K_{эм}$ объемом N бит. Отпечатаем эту фотографию (рисунок) на заведомо некачественной бумаге, отсканируем изображение и получим новый файл $K_{раб}$, объем которого, как можно ожидать, будет меньше. Сравнивая исходный и полученный после сканирования файлы, можно в явном виде найти файл-разницу ΔK как результат побитового сложения по модулю 2 соответствующих элементов. Логично предположить, что именно ΔK можно использовать для переноса СВ.

Аналогичные соображения могут возникнуть при рассмотрении в качестве контейнера аудиофайлов (например, оцифрованного звука). Если такой файл воспроизвести (озвучить) в помещении, где присутствует посторонний шум (улицы или других посторонних источников), зафиксировать его с помощью микрофона и записывающей аппаратуры, то после дискретизации и квантования, получим ситуацию точно такую же, как и в предыдущем случае. Поскольку рассмотренные информационные потоки являются составляющими видеофайлов, можно предполагать, что и этом случае шум может быть использован для увеличения той части контейнера, которая потенциально пригодна для переноса СВ.

Перейдем теперь к рассмотрению ситуации в общем случае.

Пусть по открытому каналу передаются в виде отдельных блоков длиной n бит двоичные слова, соответствующие отсчетам оцифрованного звука или пикселям графической информации. В канале присутствует шум (помеха), например, с равномерным распределением. Уровень шума, как принято, будем оценивать отношением сигнал/шум

$$A = 20 \lg \frac{U_s}{U_n}, \quad (1)$$

где U_s – среднее значение напряжения сигнала; U_n – среднее значение напряжения шума (помехи).

В первом приближении можно считать, что после аналого-цифрового преобразования уровень U_s отображается n битами, а $U_n - k$ битами в кодовых словах. Поэтому для грубой оценки уровня шума отношение (1) можно переписать как $A \approx 20 \lg \frac{n}{k}$, или $\frac{A}{20} \approx \lg \frac{n}{k}$.

После несложных преобразований получим $\lg k \approx \lg n - \frac{A}{20}$ и для количества "зашумленных" бит блока

$$k \approx \frac{n}{10^{A/20}}. \quad (2)$$

Уже на качественном уровне видно, что чем выше уровень шума (меньше значение A), тем больше k . Более наглядно это видно из рис. 1.

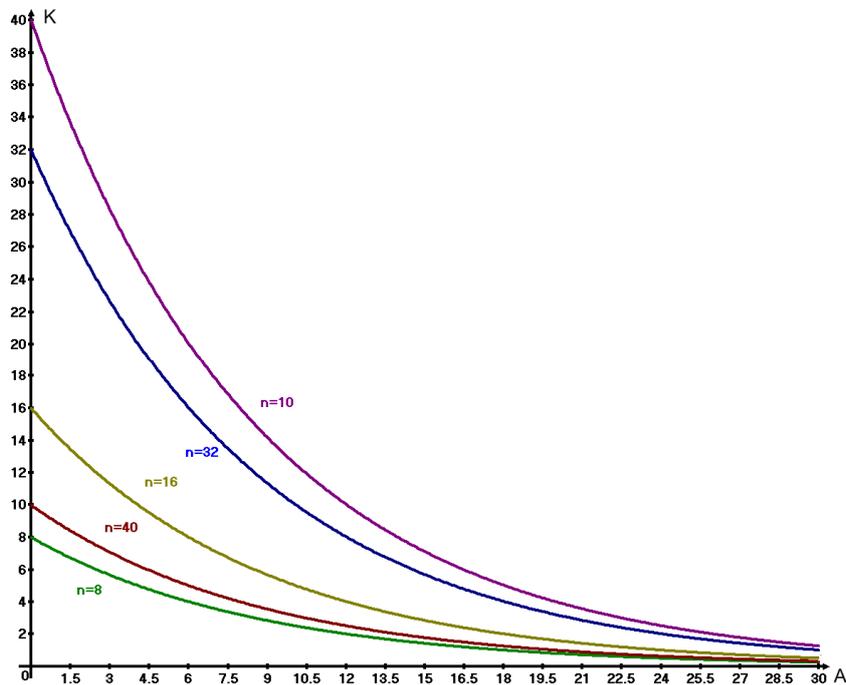


Рис. 1. Зависимость величины k от отношения сигнал/шум

Однако следует иметь в виду, что полученные оценки являются лишь некоторым достаточно грубым приближением к точным значениям (если они в принципе существуют). Это связано с тем, что:

- средние значения напряжений сигнала и шума зависят от интервала времени, за который проведено усреднение, поэтому возможны значительные колебания вокруг средних значений, полученных за разные интервалы;

- шаг квантования $\delta = \frac{1}{2^n}$ и соответственно, разрядность представления каждого отсчета после аналого-цифрового преобразования далеко не всегда выбирается исходя из среднего (или максимального) уровня сигнала. Так, например, при передаче аудиосигналов высокого качества учитывается и требуемый динамический диапазон и другие факторы.

Кроме того, при использовании неравномерного квантования (например, в реальных системах многоканальной связи) малые значения сигнала квантуются с меньшим шагом, а уровень шума при этом остается примерно тем же. То есть, отношение сигнал/шум в этом случае снижается. Следует отметить, что здесь речь идет не о шуме квантования, который уменьшается, а о связанном с помехами шуме канала передачи. Существуют и другие факторы, не позволяющие в общем случае дать точную оценку влияния шума.

Таким образом, можно принять, по крайней мере, как факт, что наличие шума увеличивает потенциальные возможности передачи СВ. Рассмотрим теперь, как можно реализовать такую возможность.

При любой реализации необходимо обеспечить статистическую близость СВ характеристикам шума с тем, чтобы максимально усложнить задачу криптоанализа (выявления самого факта наличия СВ). С этой целью, очевидно, СВ необходимо перед загрузкой в контейнер предварительно преобразовать, приблизив вероятности (частоты) появления 0 и 1 к равномерному распределению (либо к другому распределению, соответствующему характеристикам шума). Здесь возникает закономерный вопрос, что считать равномерным распределением для битового потока. Если в случае аналоговых сигналов наиболее близким к равномерному распределению можно считать так называемый "белый" шум, когда вероятности появления сигналов с разными значениями одинаковы и не зависят от значений сигнала в предыдущие моменты времени, то, по аналогии, к битовому потоку должны предъявляться такие же требования.

Это означает, что, во-первых, вероятности 0 и 1 должны быть равны $1/2$ (что сравнительно легко выполнить) и вероятности появления любых 2-битовых (00,01,10,11), 3-битовых (000,001,010,011,...), 4-битовых и т.д. комбинаций также должны быть одинаковыми. Кроме того, должна быть обеспечена независимость перечисленных вероятностей от "предыстории", т.е. от того, какие комбинации были в потоке в предшествующие моменты времени. Строго говоря, перечисленным требованиям отвечают лишь оцифрованные значения напряжений на выходах физического датчика случайных чисел (например, усиленного напряжения теплового шума на резисторе). Алгоритмическая генерация случайных чисел может дать лишь псевдослучайную последовательность значений, поскольку любой алгоритм (по сути, регулярное правило), уже по определению исключает принцип случайности.

В настоящее время существует большое количество разнообразных аппаратных и программных реализаций генераторов псевдослучайных чисел (ГПСЧ). По-видимому, наиболее распространенной реализацией ГПСЧ являются сдвигающие регистры с обратными связями и суммированием по модулю 2 [3]. Строго говоря, последовательности, генерируемые такими регистрами, далеки от случайных, но их частотные характеристики очень похожи на случайные. Вопрос качества генерируемых последовательностей важен с точки зрения криптографической стойкости информационного обмена. По сути, речь идет о возможности предсказать следующий бит в последовательности на основе знания значений всех (или части) предыдущих битов. Эта задача может быть сведена к определению структуры генератора (конкретно, стартового состояния и вида обратных связей) [4]. Не вдаваясь в детали этой достаточно сложной математической задачи, отметим лишь, что сдвигающие регистры с обратными связями на сегодня являются основным средством генерации псевдослучайных двоичных чисел. Не в последнюю очередь это можно объяснить исключительной простотой реализации таких генераторов и достаточной практической стойкостью информационного обмена к несанкционированному доступу при их использовании.

Таким образом, используя тот или иной ГПСЧ, можно преобразовать битовую последовательность с произвольными статистическими характеристиками (частотами появления 0 и 1, а также более длинных комбинаций) в последовательность с характеристиками, близкими к равномерному распределению в оговоренном выше смысле. Для этого достаточно каждый бит последовательности с произвольным (но конкретным) распределением сложить по модулю 2 с битом, генерируемым ГПСЧ. Используя такой подход, общую схему информационного обмена можно представить в виде, показанном на рис. 2.

В этой схеме битовый поток, поступающий от ГПСЧ, несет двойную функциональную нагрузку. С одной стороны, он используется для "забеливания" СВ с тем, чтобы соответствующая двоичная последовательность, полученная в результате побитового

сложения по модулю 2, была по возможности неотличима от случайной (в данном случае, последовательности с равномерным распределением). С другой, – конкретный вид генерируемой последовательности является, по сути, ключом в криптографическом смысле. Для этого отправитель и получатель информации должны иметь в составе своего оборудования (передатчика и приемника) одинаковые ГПСЧ, стартовые состояния которых и вид обратных связей в моменты начала каждого сеанса связи должны быть строго согласованы.

Защищенность стеганоканала при необходимости может быть существенно увеличена путем избирательного размещения битов СВ. Речь идет либо о частичном использовании потенциальной пропускной способности стеганоканала, образующейся за счет шума, либо выборе порядка заполнения (модификации) позиций контейнера битами СВ в соответствии с адресами, которые задаются тем же ГПСЧ. Такой механизм может быть реализован схемой, представленной на рис 3. Очевидно, такой же механизм должен быть задействован и при "выемке" СВ.

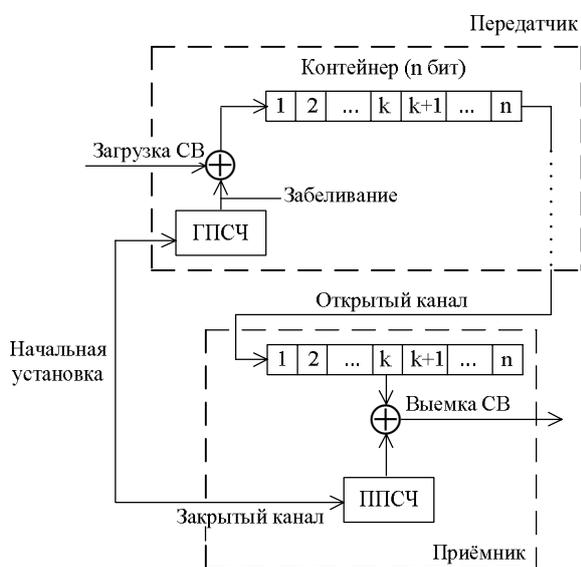


Рис. 2. Общая схема информационного обмена

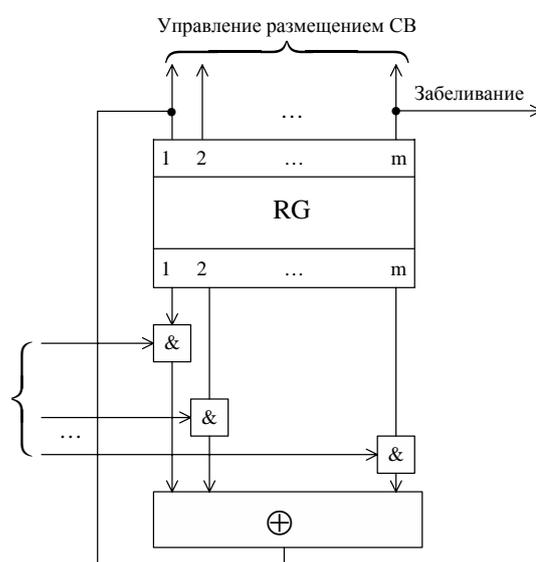


Рис. 3. Вариант аппаратной реализации стеганоканала

Выводы. Таким образом, предлагаемый подход иллюстрирует, достаточно широкие, с нашей точки зрения, возможности изменения пропускной способности стеганоканала (в сторону увеличения) за счет предварительного наложения шума на файл контейнера. Естественно, это потребует внесения изменений в организацию сеанса информационного обмена и согласования деталей этих изменений между отправителем и получателем.

ЛИТЕРАТУРА

1. Грибунин В.Г. Цифровая стеганография / В.Г. Грибунин, И.Н. Оков, И.В. Туринцев - М.: Солон-Пресс, 2002. – 272 с.
2. Коначович Г.Ф. Компьютерная стеганография / Г.Ф. Коначович, Пузыренко А.Ю. - К.: МК-Пресс, 2006. - 284 с.
3. Schneir B. Applied Cryptography: Protocols, Algorithms, and Source Code in C, 2nd ed New York // John Wiley and Sons, 1996.
4. Пометун С.О. Алгебраїчні атаки на потокові шифратори як узагальнення кореляційних атак / С.О. Пометун // Системні дослідження та інформаційні технології - 2008. - №2. - С. 29-40.

Надійшла: 20.12.2012

Рецензент: д.т.н., проф. Шелест М.Є.