# Оценка стойкости разнотипных стеганографических систем

Айк Казарян

Государственный Инженерный Университет Армении (Политехник)
Ереван, Армения
e-mail: ghazaryan.hayk@gmail.com

## РЕЗЮМЕ

В данной статье представлен теоретико-статистический метод оценки стойкости стеганографических систем, имеющих различную природу и организацию обеспечения скрытности.

## Ключевые слова

Защита информации, стеганография, скрытнопись, оценка стойкости.

## 1. ВВЕДЕНИЕ

В последние годы в связи с бурным развитием информационных технологий и их широким применением практически во всех областях человеческой деятельности все большее внимание уделяется средствам защиты информации. Вместе с тем, как показывает анализ, в ряде случаев современные криптографические средства защиты не полностью удовлетворяют потребностям пользователей, и на первый план выдвигаются стеганографические средства, которые призваны скрыть от посторонних сам факт присутствия объекта защиты.

Наряду с созданием новых средств защиты информации не менее актуальной остается разработка адекватных методов оценки и сравнения их эффективности. Существующие методы оценки стойкости стеганографических систем (син.: стеганосистема, система скрытнописи), отражая специфику применяемых принципов сокрытия информации, зачастую не позволяют достаточно полно сопоставлять средства защиты, имеющие различную природу и организацию обеспечения скрытности.

В данной статье предложен теоретико-статистический метод оценки стойкости стеганосистем, и, основанный на нем практических метод оценки стойкости разнотипных систем скрытнописи.

## 2. НЕДОСТАТКИ ТЕОРЕТИКО-ИНФОРМАЦИОННОГО КРИТЕРИЯ ОЦЕНКИ СТОЙКОСТИ

В работе [1] представлен теоретико-информационный метод оценки стойкости симметричных одноразовых стеганосистем с наличием пассивного противника. Метод основывается на известной модели стеганографического канала (рис. 1) и использует теорию информации и проверки гипотезы для анализа распределений вероятностей появления контейнеров и последующего получения оценки стойкости. Считаем необходимым перечислить основные недостатки данного метода.

- Предполагается, что противник знает точные вероятностные характеристики пустых контейнеров, стеганографических контейнеров, скрываемых сообщений и ключей. В итоге любое отклонение статистики при наблюдении противником трактуется как обнаружение стеганографического сообщения/канала. В итоге, любое незначительное отклонение наблюдаемой противником статистики от среднестатистических характеристик пустых контейнеров квалифицируется как обнаружение скрытого канала. Очевидно, что:
  - на практике любой канал вносит искажения в передаваемые сообщения, т.е. отклонения от ожидаемых значений неизбежны;
  - у противника в лучшем случае в наличии могут быть усредненные характеристики по вышеперечисленным сущностям;
  - отправитель волен подбирать или создавать такие контейнеры, для которых характеристики стеганографических контейнеров незначительно отличались от среднестатистических характеристик пустых контейнеров.
- Часто пользователю требуется оценить стойкость нескольких разнотипных методов сокрытия, чтобы сделать обоснованный выбор в пользу одного из них. Предложенный метод не предполагает использования при сравнении разнотипных стеганосистем.
- Трудно представить отправителя и получателя обменивающихся открытым текстом со случайным распределением символов.
- Трудно найти противника (надзирателя), который допускает передачу бессмысленных сообщений (со случайным распределением символов).

Очевидно, что такая идеализированная модель не адекватна реальным системам скрытнописи и не может быть применена на практике. Она представляет теоретический интерес и используется для

- доказательства абсолютной стойкости теоретических стегосистем, наподобие одноразового стеганографического блокнота [1];
- получения нижней оценки для вероятности ошибки второго рода (стеганографический контейнер был неправильно определен как пустой) используя теорему Неймана-Пирсона, имея верхнюю оценку для вероятности ошибки первого рода (пустой контейнер был неправильно определен как стеганографический) и стойкость стеганосистемы по данной модели.

## 3. ТЕОРЕТИКО-СТАТИСТИЧЕСКАЯ ОЦЕНКА СТОЙКОСТИ СТЕГАНОГРАФИЧЕСКИХ СИСТЕМ

При построении теоретико-статистической оценки мы основывались на следующих принципах:

- отталкиваться от известных методов организации атак на стеганосистемы, так как стойкость определяет возможность пассивных противников обнаружить

скрытый канал связи применением комплексных мер анализа (атак) системы скрытнописи;
- дать возможность сравнивать стойкость разнотипных стеганосистем;
- дать возможность построения практической оценки стойкости на основе разработанной теоретической оценки;
- уметь четко распознавать абсолютно стойкие стеганосистемы.

Итак, предположим, что имеется стеганосистема, работающая по схеме представленной на рис. 1. Отправитель отсылает контейнер получателю через канал общего пользования. Отсылаемые контейнеры могут быть либо пустыми, либо содержащими скрытое сообщение – стеганографическими контейнерами. Если передается стеганографический контейнер, то говорят, что существует скрытый канал связи между отправителем и получателем. Противник имеет доступ ко всем передаваемым сообщениям. Задачей противника является обнаружение факта передачи скрытого сообщения (выявление скрытого канала). Противник может применить собственные методы анализа контейнера, которые помогут ему решить стоящую перед ним задачу.

Нас будет интересовать случай, когда сообщение посланное отправителем содержит какую-то информацию, которая уменьшит энтропию на стороне получателя, как только последний извлечет и прочтет сообщение.

Получатель, при выявлении скрытого канала, используя закрытый ключ, применяет стеганографический алгоритм и извлекает сообщение. При построении данной модели нас не будет интересовать каким именно образом получатель узнает о существовании скрытого канала на данный момент времени. Будем предполагать, что существует *оракул*, при помощи которого получатель узнает о существовании скрытого канала.

Начнем с простого случая. По каналу общего пользования могут отсылаться контейнеры одного типа K и для этого типа контейнеров существует один-единственный параметр T, который имеет известное (в частности – известное всем трем участникам в нашей модели) среднестатистическое распределение значений:

$$\mathrm{Y} = \langle Y_1, Y_2, \ldots, Y_k \rangle,$$

где k-это количество категорий для параметра T (см. п. 6 для примеров). Случайным событием назовем определение значения параметра T противником и, соответственно, определение принадлежности к одной из $k$ категорий. Ясно, что для Y имеют место аксиомы Колмогорова. Противник, анализируя статистическую характеристику параметра T в контейнере, решает, либо он содержит скрытые данные, либо он пуст. Предполагается, что противник будет использовать множество статистических критериев, для оценки отклонений закономерностей в значениях параметра T для анализируемого контейнера. В настоящей статье предлагается использовать критерий Хи-квадрат для оценки стойкости стеганосистемы против атак пассивного противника, основываясь на следующих утверждениях:
- критерий Хи-квадрат один из самых известных статистических критериев [2, 3];
- успешно используется при анализе систем скрытнописи на выявление скрытого канала;
- является основным методом проверки нарушения закономерностей, используемым в сочетании с другими методами [2, 4];
- предполагает использование конечного числа категорий, что делает его использование на практике более предпочтительным, в отличии от непрерывных критериев, предполагающих бесконечное множество категорий;
- основан на общем принципе наименьших квадратов, что опять-таки делает его более предпочтительным с точки зрения применения на практике [4].

*Определение 3.1.* Стойкость представленной стеганосистемы к атакам пассивных противников является значение дистанции Хи-квадрат между ожидаемым и наблюдаемым распределением вероятностей параметра T:

$$V = \sum_{s=1}^{k} \frac{(Y_s - np_s)^2}{np_s}, \qquad (3.1)$$

где $Y_s$ - это наблюдаемое число попаданий значения параметра T в категорию $s$, $p_s$ - ожидаемая вероятность попадания значение параметра T в категорию $s$, $n$ – количество проведенных экспериментов.

При сравнении разнотипных стеганосистем следует учитывать тот факт, что значение критерия Хи-квадрат зависит от выбора количества категорий – чем больше категорий, тем больше значение критерия Хи-квадрат (см. таблицы по процентным точкам Хи-квадрат распределения) [2, 4, 5]. Те же таблицы выдают одинаковые значения для конкретной процентной точки распределения Хи-квадрат при равном количестве категорий. Отсюда следует, что при сравнении стеганосистем с разными параметрами $T_i$ следует выбирать одинаковое количество категорий.

# 4. СВОЙСТВА

Рассмотрим основные свойства критерия оценки стойкости стеганосистемы, определенной в 3.1.

*Лемма 4.1.* Стойкость согласно теоретико-статистическому методу всегда больше стойкости по теоретико-информационной модели. Другими словами дистанция Хи-квадрат всегда не меньше, чем относительная энтропия.
*Доказательство* получается применением теоремы Йенсена и можно найти в [3].

*Лемма 4.2.* Стойкость, определенная в 3.1, всегда неотрицательна и равна нулю тогда и только тогда, когда наблюдаемое распределение равно ожидаемому.

*Доказательство.* Первая часть леммы, а также достаточность второй части леммы, вытекают из леммы 4.1 и из аналогичного свойства относительной энтропии. Доказательство последнего можно найти в [6]. Необходимость второй части леммы очевидна – подставляя $Y_s = np_s$ в выражение 3.1 получим нулевое значение стойкости.

*Определение 4.1.* Случай, когда вероятности появления значения параметра T в любой из категорий равны, т.е.

$$p_1 = p_2 = \ldots = p_k = \tfrac{1}{k},$$

назовем случаем с равновероятными категориями.

*Лемма 4.3.* Стойкость стеганосистемы, определенной в 3.1, при равновероятных категориях, можно вычислить используя следующую упрощенную формулу:

$$V = \frac{k}{n}\sum_{s=1}^{k}\left(Y_s\right)^2 - n\,. \qquad (4.3)$$

*Доказательство.* Имея ввиду следующие равенства:

$$\sum_{s=1}^{k} Y_s = n \qquad \text{и} \qquad \sum_{s=1}^{k} p_s = 1\,.$$

очевидны следующие преобразования:

$$V = \sum_{s=1}^{k}\frac{\left(Y_s - np_s\right)^2}{np_s} = \sum_{s=1}^{k}\frac{\left(Y_s - \frac{n}{k}\right)^2}{\frac{n}{k}} =$$

$$= \frac{k}{n}\sum_{s=1}^{k}\left(Y_s\right)^2 - 2\sum_{s=1}^{k} Y_s + \sum_{s=1}^{k}\frac{n^2}{k^2} =$$

$$\frac{k}{n}\sum_{s=1}^{k}\left(Y_s\right)^2 - n\,.$$

*Лемма 4.4.* Стеганосистема, абсолютно стойкая по Качину, является также абсолютно стойкой по вышеопределенной теоретико-статистической оценке и наоборот, стеганосистема, которая не является абсолютно стойкой по Качину, также не является абсолютно стойкой по теоретико-статистической оценке.

*Доказательство.* Согласно модели стеганосистемы, представленной в [1], противнику известны оба распределения вероятностей появления конкретного контейнера в канале связи – и′ в случае обычного взаимодействия, и′ в случае создания скрытого канала. Различие этих двух распределений дает шанс противнику взломать систему. Согласно теоретико-информационной оценке, при одинаковых распределениях вероятностей пустых и стеганографических контейнеров противник не сумеет распознать стеганографических контейнер, если такой появится в канале связи, и, следовательно, такая стеганосистема будет абсолютно стойкой к атакам пассивных противников.

Выберем в качестве параметра для анализа появление конкретного контейнера в канале связи. Средняя статистика появления пустого контейнера является ничем иным, как ожидаемой статистикой появления контейнеров в нестеганографической системе. Наблюдаемой статистикой очевидно, что является вероятностное распределение появления стеганографических контейнеров в канале связи. Оценивая стойкость такой системы к атакам пассивных противников, используя теоретико-статистическую оценку, и учитывая лемму 4.2 получим необходимый результат.

# 5. ОБЩАЯ ПРАКТИЧЕСКАЯ ОЦЕНКА СТОЙКОСТИ ДЛЯ СЛОЖНЫХ СИСТЕМ СКРЫТНОПИСИ

Вышеизложенный метод оценки дает возможность оценивать стойкость стеганосистем, которые, с точки зрения разработчика системы, имеют один уязвимый параметр, отклонение которого может быть обнаружено при статистическом анализе контейнера. К таким системам в основном относятся классические методы сокрытия информации в цифровых контейнерах и протоколы передачи скрытых данных. Для более сложных методов и протоколов, а также для сложных подсистем стеганографической защиты информации, необходимо иметь более общий метод оценки стойкости, который не только даст возможность оценивать их стойкость, но и сравнивать со стеганосистемами другого типа. Для сложных систем существует множество взаимосвязанных параметров, существенное отклонение которых может иметь решающее значение для данной системы и провоцировать взлом такой системы. Пример для простых систем: пользователю необходимо скрыть данные в каком-либо часто используемом контейнере. Для него несущественно, будет ли этим контейнером JPEG файл или MP3 файл. Его интересует только стойкость выбранной системы и, соответственно, его предпочтение будет на стороне более стойкой системы. Пример для сложных систем: пользователю необходима система скрытнописи для сокрытия большого количества данных. Существует две альтернативы – либо это будет стеганографическая файловая система, либо стеганографическая база данных. Опять-таки, для пользователя существенна только стойкость и ему необходимо иметь под рукой метод сравнения стойкостей этих двух систем.

Перейдем к определению более общей и более удобной, с практической точки зрения, оценки стойкости, на основе вышеописанного метода оценки стойкости стеганосистем. Допустим, что для стеганосистемы $\Psi$ существует множество параметров

$$T = \left\{T_1, T_2, \ldots, T_m\right\},$$

которые могут быть анализированы противником на предмет выявления отклонений распределения их значений от ожидаемых значений. Целью разработчика системы является:

- выявление всех тех параметров, которые имеют существенные отклонения от среднестатистических значений при создании скрытого канала;
- разработка методов сокрытия, при которых у противника, при известном статистическом анализе контейнера, не появится серьезных оснований для предположения о существовании скрытого канала в системе.

Итак, предположим, что параметр $T_i$, как случайная величина, имеет функцию плотности распределения $f(T_i)$. Разобьем интервал $(-\infty, +\infty)$ на $k$ частей

$$(-\infty, a_{i1}), (a_{i1}, a_{i2}), \ldots, (a_{i,k-1}, +\infty)$$

таким образом, чтобы имели место следующие равенства (рис. 5.1.):

$$\int_{-\infty}^{a_{i1}} f(T_i) = \int_{a_{i1}}^{a_{i2}} f(T_i) = \ldots = \int_{a_{i,k-1}}^{+\infty} f(T_i) = \frac{1}{k}\,.$$

В этом случае мы имеем дело с равновероятными категориями (см. опред. 4.1) и согласно лемме 4.3 имеем право использовать упрощенную формулу (4.1) для определения стойкости по параметру $T_i$.

Для того чтобы получить оценку, которая будет учитывать все параметры $T_i$, будем следовать очевидной логике – просуммируем взвешенные значения стойкости по всем параметрам, т.е.:

$$V = \sum_{i=1}^{m}\lambda_i\left(\frac{k_i}{n}\sum_{s=1}^{k_i}\left(Y_{i,s}\right)^2 - n_i\right). \qquad (5.1)$$

Учитывая следующие зависимости между коэффициентами $\lambda_i$:

$$\lambda_i \geq 0 \qquad \text{и} \qquad \sum_{i=1}^{m}\lambda_i = m\,,$$

и подставляя в 5.1 получим конечный вид формулы оценки стойкости для сложных систем скрытописи:

$$V = \sum_{i=1}^{m} \lambda_i \frac{k_i}{n_i} \sum_{s=1}^{k_i} \left(Y_{i,s}\right)^2 - \sum_{i=1}^{m} \lambda_i n_i \,. \qquad (5.2)$$

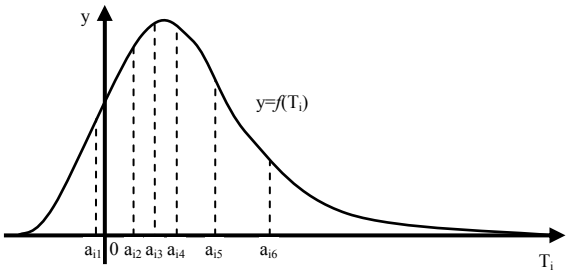


**Рис. 5.1.** Равновероятные категории (k=7)

# 6. ПРИМЕРЫ

Продемонстрируем применение метода при оценке стойкости реальных систем скрытописи. Как пример возьмем метод LSB [7] для сокрытия данных в графических контейнерах типа BMP [8]. В качестве параметра для анализа контейнеров выберем один из наиболее известных параметров, определяющий отклонение количества соседних номеров цвета в пикселах изображения [9]. В качестве контейнеров будем рассматривать BMP файлы с 24-битной глубиной цвета и разрешением 800x600 пикселей. Пространство сокрытия методом LSB для данного вида файлов равно 468.75 килобайт. Рассмотрим как меняется значение стойкости такой системы скрытописи, при изменении количества скрываемых данных.

Определим параметр $T_1$ как разницу количества пикселов с установленным младшим битом красной компоненты цвета от количества пикселов со сброшенным младшим битом красной компоненты цвета. Параметры $T_2$ и $T_3$ будут аналогичны $T_1$, но для зеленой и для синей компоненты цветов соответственно. Анализ примерно 10000 пустых контейнеров (www.desktopwallpapers.com, www.wallpapers.com) вышеопределенного вида дал изображенный на рис. 6.1 примерный вид функции плотности распределения разницы количества соседних цветовых компонент для 3-х выбранных цветов.

После получения ожидаемого распределения вероятностей по всем категориям и по всем 3-м цветовым компонентам, можно перейти к оценке разных модификаций метода LSB. Различие в основном представлялось в виде использованного пространства сокрытия (ИПС). Также были применены различные техники распределения скрываемой информации по всему контейнеру. В таблице 6.1 приведены оценки стойкости такой системы в зависимости от ИПС. В качестве контейнера было использовано одно и то же изображение. Как и ожидалось, чем меньше данных скрывается в контейнере, тем выше значение стойкости.

# 7. ЗАКЛЮЧЕНИЕ

В работе был представлен метод оценки стойкости разротипных стеганосистем. Полученная оценка имеет как важное теоретическое значение: возможность доказательства абсолютной стойкости, сравнение классических стеганосистем, и.т.д, так и практическое значение: использование метода при оценке реальных/сложных систем скрытописи.

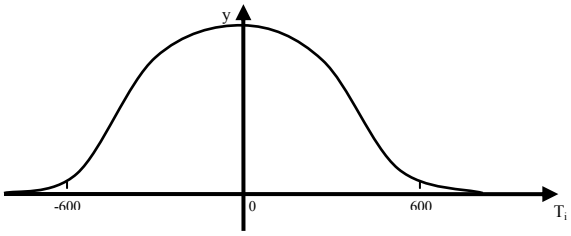


**Рис. 6.1.** Ожидаемое распределение рассматриваемого параметра при анализе LSB метода

**Таблица 6.1.**

| *ПС* | *ИПС* | *Стойкость (V)* |
|---|---|---|
| 468.75 кБ | 0% | 33 |
| 468.75 кБ | 100% | 1164 |
| 468.75 кБ | 50%. Сокрытие произведено в верхней части изображения. | 731 |
| 468.75 кБ | 50%. Сокрытие произведено в нижней части изображения. | 961 |
| 468.75 кБ | 50%. Сокрытие произведено последовательно в каждый второй пиксель изображения. | 1099 |
| 468.75 кБ | 25%. Сокрытие произведено последовательно в каждый четвертый пиксель изображения. | 696 |

Зависимость стойкости от используемого пространства сокрытия (ИПС)

# БИБЛИОГРАФИЯ


[1] Cachin C. “An Information-theoretic Model for Steganography”, *Proc. 2nd International Workshop on Information Hiding*. LNCS 1525. P. 306-318, 1998.

[2] Кнут Д., “Искусство программирования”, том 2. Получисленные алгоритмы, 3-е изд. : Пер. С англ. – М. : Издательский дом “Вильямс”, 2003. -832 с. : ил.

[3] Gibbs A., Su F. “On choosing and bounding probability metrics”, *Internaltional Statistical Review*, vol. 70, number 3, 419-435, 2002.

[4] Крамер Г., “Математические методы статистики” (пер. с англ.). Мир, 1975г.

[5] Abramowitz M., Stegun I. “Handbook of mathematical functions.”, *Washington, D.C.: U.S. Government Printing Office*, 1964.

[6] Казарян А., “Оценка стойкости стеганографических систем на модели с наличием пассивного противника.” *Вестник ГИУА. Серия “Моделирование, оптимизация, управление”*. –Вып. 10, -т. 1. -С. 60-65, –Ереван, 2007.

[7] Аграновский А.В., Девянин П.Н., Хади Р.А., Черемушкин А.В. “Основы компьютерной стеганографии.”, -152 с.: ил. –М.: Радио и связь, 2003.

[8] Swan T. Inside Windows File Formats. Sams Publishing, 1993.

[9] Westfeld A., Pfitzmann A. “Attacks on steganographic systems. Breaking the steganographic utilities EzStego, Jsteg, Steganos, and S-Tools and some lessons learned”, *Proceeding of the Workshop on Information Hiding.* 1999.