Лекция 5. Идеальные и почти идеальные СГС.

Определение. СГС называется *идеальной* (*совершенной, безусловно необнаруживаемой*), если ее обнаружение, при использовании наилучших статистических методов, равносильно случайному угадыванию ее наличия или отсутствия.

Определение: СГС называется *почти идеальной* (*σ-необнаруживаемой*), если при использовании наилучших статистических методов $min\{P_m, P_{fa}\} \geq \sigma$, где P_m – вероятность пропуска СГС, P_{fa} – вероятность ложного обнаружения СГС.

Можно ли, вообще, построить идеальные (ИСГС) или почти идеальные (ПИСГС)?

Ответ на этот вопрос зависит от следующих условий:

1. СГС с заданным ПО или с выбираемым.

2. Требуется ли относительно высокая скорость вложения или достаточно обеспечить низкую скорость (малое количество вкладываемых бит).

3. Требуется автоматическая СГС или с участием человека.

4. Требуется ли устойчивость СГС к удалению вложенного сообщения.

5. Имеется ли естественный шум в канале обнаружения СГС.

6. Задана ли в точности статистическая модель ПО (*теоретическая*, *модельно-обусловленная* СГС) или речь идет о практической СГС, где статистические модели ПО, обычно, известны лишь частично.

Существование ИСГС (или ПИСГС) при перечисленных выше условиях:

1. Если ПО может выбираться и допустимо участие человека, то ИСГС реализуема в Л-СГС с использованием хеш-функций (х.ф.), (см. лекцию 4).

2. Если допустимо вложение даже малого количества бит, то ИСГС реализуема для любых выбираемых ПО автоматически с использованием х.ф. (такой метод называется еще *“rejection-sampling” – см. лекцию 4*).

3. Если в канале имеется естественный шум, то построение ИСГС, устойчивых к атаке удаления, возможно при малой скорости вложения (см. лекцию 6. “СГС в каналах с шумами”).

4. Если задана точная статистическая модель ПО, то построение ИСГС, устойчивой к атаке удаления, возможно (см. следующие разделы лекции).

5. Если статистическая модель ПО в точности не известна, то возможно построение ПИСГС с малой скоростью вложения (см.следующие разделы лекции).

Уточнение критерия ИСГС и ПИСГС:

P_c – статистическое распределение, полностью описывающее ПО *C(n), n =1,2…N*.

P_w – статистическое распределение, полностью описывающее СГ $C_w(n)$ при выбранном методе вложения.

Определение. Относительной энтропией СГС называется

$$D(P_w \| P_c) = \begin{cases} \int_{x \in X} P_w(x) \log\left(\frac{P_w(x)}{P_c(x)}\right) dx, \\ \sum_{x \in X} P_w(x) \log\left(\frac{P_w(x)}{P_c(x)}\right), \end{cases} \quad (1)$$

(если X – непрерывное множество наблюдения СГС)

(если X – дискретное множество наблюдения СГС)

При любых статистических методах обнаружения СГС справедливы неравенства:

$$P_{fa} \log\left(\frac{P_{fa}}{1-P_m}\right) + (1-P_{fa}) \log\left(\frac{1-P_{fa}}{P_m}\right) \le D(P_w \| P_c), \quad (2)$$

$$P_m \log\left(\frac{P_m}{1-P_{fa}}\right) + (1-P_m) \log\left(\frac{1-P_m}{P_{fa}}\right) \le D(P_c \| P_w), \quad (3)$$

где

$$P_e > \pi_0 \pi_1 \exp(-J/2), \qquad J = D(P_w \| P_c) + D(P_c \| P_w), \quad (4)$$

$P_e = \pi_0 P_{fa} + \pi_1 P_m, \pi_0, \pi_1 -$ априорные вероятности отсутствия и присутствия СГС.

Если $P_{fa} = 0$, то из (2) следует, что

$$P_m \le 2^{-D(P_w||P_c)}, \qquad (5)$$

СГС будет ИСГС, если

$$D(P_w \,||\, P_c) = D(P_c \,||\, P_w) = 0 \Rightarrow P_{fa} = P_m = 1/2, \qquad (6)$$

СГС будет ПИСГС, (или ε-ИСГС), если

$$D(P_w \,||\, P_c) \le \varepsilon \quad \text{или} \quad D(P_c \,||\, P_w) \le \varepsilon. \qquad (7)$$

Пример. Положим $D(P_w||P_c) = 0.1$, тогда, если $P_{fa} = 0$, то $P_m \ge 2^{-D} \approx 0.933$.

Определение. Расстоянием Бхаттачариа (Bhattacharyy) между двумя вероятностными распределениями P_w и P_c называется

$$D_B(P_c, P_w) = -\ln(\rho_B(P_c, P_w)), \qquad (8)$$

где

$$\rho_B(P_c, P_w) = \int_{x \in X} (P_c(x) P_w(x))^{1/2} dx = \sum_{x \in X} (P_c(x) P_w(x))^{1/2}.$$

При любых статистических методах обнаружения СГС справедливы неравенства:

$$0.25\rho_B^2(P_c,P_w) \le P_e \le 0.5\rho_B(P_c,P_w), \qquad (9)$$

где $P_e = \pi_0 P_{fa} + \pi_1 P_m$, если $\pi_0 = \pi_1 = 1/2$.

$$P_e = 1/2(P_{fa} + P_m), \qquad (10)$$

$$P_e \ge 0.25e^{-2D_B(P_c,P_w)}.$$

СГС будет ИСГС, если

$$D_B(P_c,P_w) = 0 \Rightarrow \rho_B(P_c,P_w) = 1. \qquad (11)$$

СГС будет ПИСГС, (или ε-ИСГС), если

$$D_B(P_c,P_w) \le \varepsilon. \qquad (12)$$

Замечание. Расстояние Бхаттачариа (РБ) - симметричная функция, т.е. $D_B(P_c,P_w) = D_B(P_w,P_c)$, в отличие от относительной энтропии (ОЭ), которая несимметрична $(D(P_c \,||\, P_w) \ne D(P_w \,||\, P_c))$.

РБ не удовлетворяет свойству “треугольника” $(D_B(P_1,P_2) < D_B(P_2,P_3) + D_B(P_1,P_3))$, однако его модификация $(\sqrt{1-\rho_B^2(P_c,P_w)})$ этому свойству удовлетворяет.

Модельно-обусловленная (model-based) СГС.

В этом случае предполагается, что при разработке и обнаружении СГС статистические свойства ПО известны в точности.

1. $C(n) \in N(0,\sigma_c^2), E\{C(n)C(n')\} = 0,$ при $n \neq n'$.
Тогда можно построить ИСГС при вложении:

$$C_w(n) = \beta C(n) + \alpha(-1)^b \pi(n), \quad (13)$$

где

$$\beta = \sqrt{1-(\alpha^2/\sigma_c^2)}, \pi(n) \in N(0,1), E\{\pi(n)\pi(n')\} = 0, n \neq n'.$$

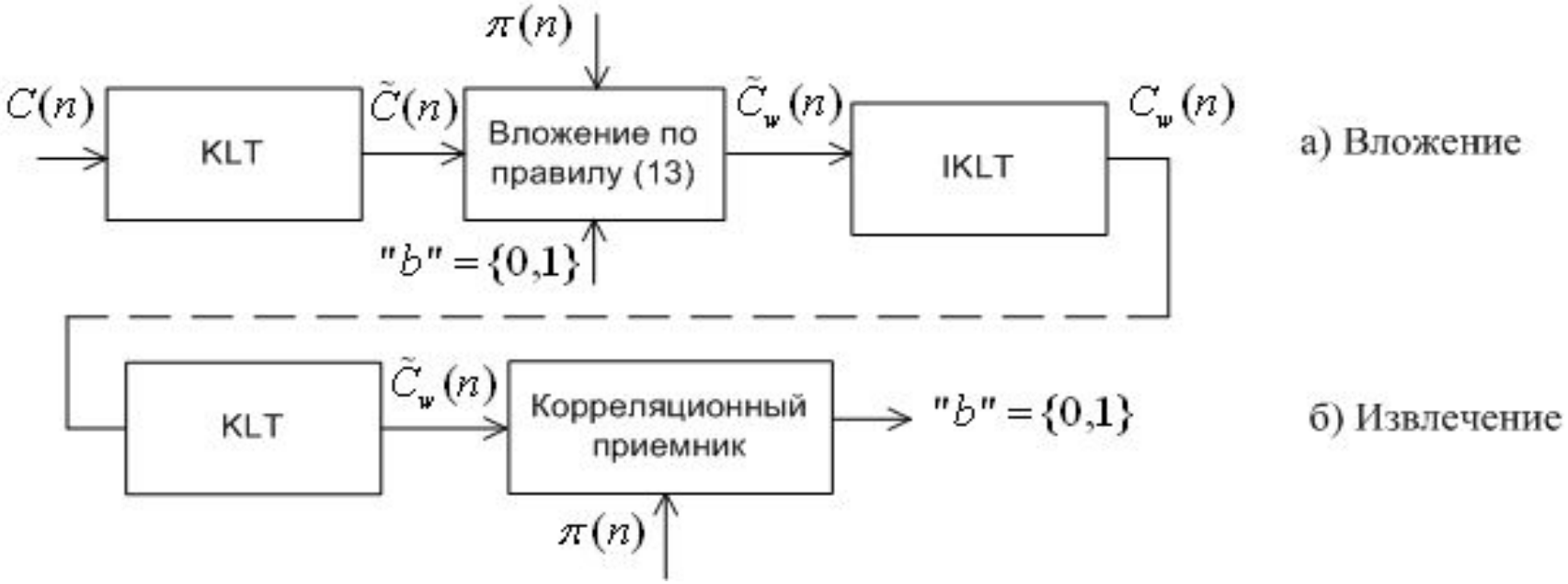
Действительно, $C_w(n) \in N(0,\sigma_w^2), E\{C_w(n)C_w(n')\} = 0,$ при $n \neq n'$,

т.е. $P_w = P_c$,

$$\sigma_w^2 = \beta\sigma_c^2 + \alpha^2 = \sigma_c^2 - \alpha^2 + \alpha^2 = \sigma_c^2,$$

2. ПО является окрашенным гауссовским шумом.
$C(n) \in N(0,\sigma_w^2), E\{C(n)C(n')\} \neq 0,$ при $n \neq n'$, задается известной матрицей корреляций $\mathbf{R_c}$.

Тогда ИСГС обеспечивается при следующих способах вложения и извлечения информации:



Здесь введены обозначения:

KLT – Преобразование Карунена-Лоэва, которое преобразует “окрашенный” гауссовский шум в “белый” (с независимыми отсчетами) гауссовский шум . $\tilde{C}(n)$

IKLT – обратное преобразование Карунена-Лоэва.

3. *C(n)* имеет произвольное статистическое (не гауссовское) распределение P_c с независимыми отсчетами. Тогда можно обеспечить построение ИСГС при вложении и извлечении по правилам:

$$C_w(n) = F^{-1}(\beta F(C(n)) + \alpha(-1)^b \pi(n)), \quad (14)$$

$$b = \begin{cases} 0, & \text{если } \sum_{n=1}^{N} F(C_w(n)\pi(n)) \geq 0, \\ 1, & \text{если } \sum_{n=1}^{N} F(C_w(n)\pi(n)) < 0, \end{cases} \quad (15)$$

где *F(.)* – преобразование случайной негауссовской величины *C(n)* с $E\{C(n)\} = 0, Var\{C(n)\} = \sigma_c^2$ к гауссовской случайной величине *N (0, σ^2_c).*

F^{-1}(.) – обратное преобразование к преобразованию *F(.).*

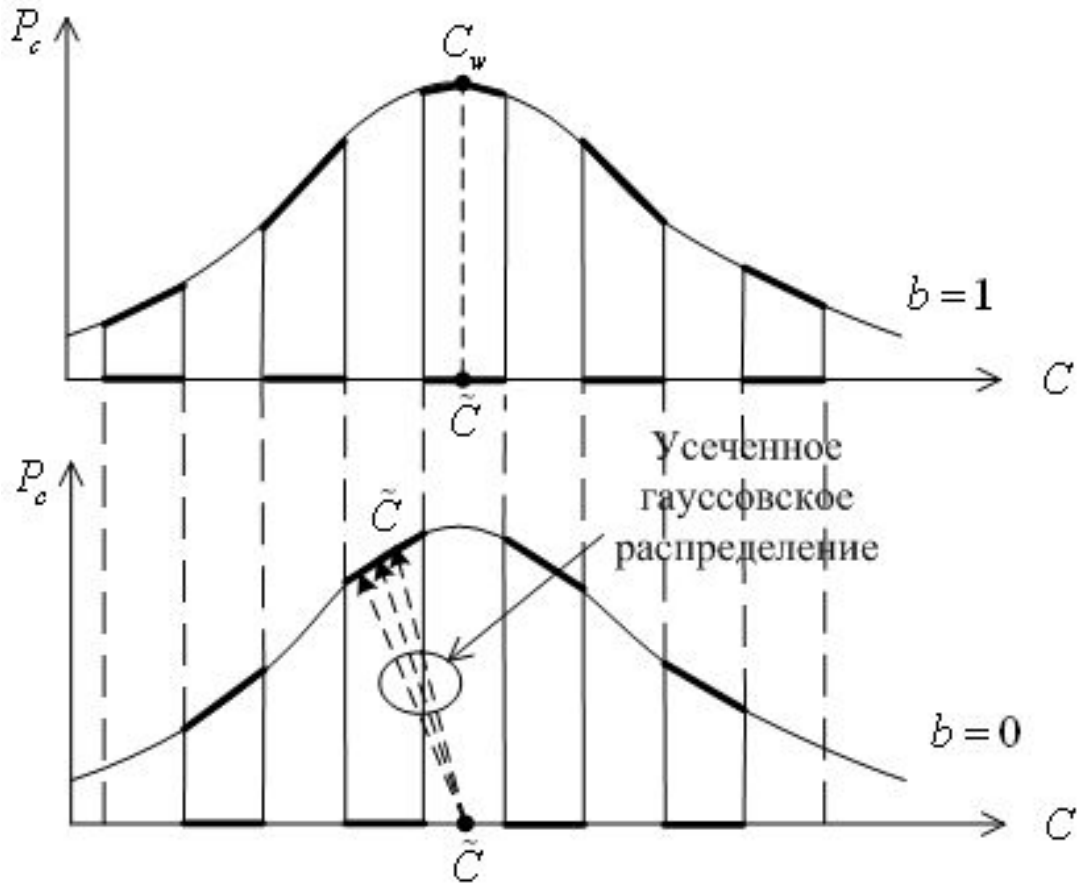
Замечание 1. Во всех случаях предполагается, что для вложения и извлечения необходимо в точности знать σ_c^2, R_c, а в последнем случае P_c.

Замечание 2. Во всех случаях СГС является “робастной”, т.е. может противостоять атаке о удалению возможно скрытого сообщения. Это достигается при помощи ПОП $\pi(n)$ длиной N, выбираемой по секретному ключу, и корреляционного приемника (15).

Однако, чем больше N, тем меньше скорость вложения. (Возможны и более изощренные атаки по удалению вложенной информации – см. лекции далее).

Замечание 3. Расчет вероятностей ошибок можно для различных типов детекторов (“слепой” и информированный) производить по формулам (12) и (19) из лекции 3.

4. *C(n)* є *N* $(0,\sigma_c^2), E\{C(n)C(n')\} = 0$,при *n* ≠ *n`.*
Метод погружения с квантованием СМ. (В дальнейшем этот метод будет изучаться более подробно в разделе 2. “ЦВЗ”.)



Метод погружения и извлечения бита “*b*” в отсчет *C(n)* очевиден из рисунка, причем извлечение происходит без ошибок при любом ПО. Скорость равна 1 бит/отсчет. СГС оказывается ИСГС и робастной относительно небольшого аддитивного шума.

СГС, построенные при неполном знании распределений ПО.
(Это дает подход к практической реализации СГС для типичных ПО (аудио и видео)).
1. Аддитивное вложение в гауссовские ПО с нулевыми средними и неизвестной, при построении СГС, матрицей корреляций.

Тогда $$D_B(P_c, P_w) = \frac{1}{2}\ln\frac{\det R}{\sqrt{\det R_w \det R_c}}, \quad (15)$$

где R_c – матрица корреляций ПО, R_w – матрица корреляций СГС,
$R = (R_w + R_w)/2$.
Для метода погружения по (13) и экспоненциальной матрице

$$R_c = \sigma_c^2 \begin{bmatrix} 1 & r & r^2 & \dots & r^{N-1} \\ r & 1 & r^2 & \dots & r^{N-2} \\ \dots & \dots & \dots & \dots & \dots \\ r^{N-1} & \dots & \dots & \dots & 1 \end{bmatrix}$$

Можно доказать [ ], что

$$\rho_\beta^2(P_c, P_w) \approx \left(\frac{\sqrt{(1-r^2)(1-2r^2\delta^{-1}+r^2)}}{1-2r^2\tau+r^2} \right)^{N-1}, \quad (16)$$

где $$\delta = \frac{1}{\left(1-\frac{\alpha^2}{2\sigma_c^2}\right)^2}; \tau = (1+\delta^{-1})/2.$$

Пример. r = 0.5, δ = 1.01, N = 500, тогда по (16) $\rho_\beta^2(P_c, P_w) = 0.997.$

r = 0.9, δ = 1.01, N = 500, тогда по (14) $\rho_\beta^2(P_c, P_w) = 0.66.$

(Напомним, что $P_e \geq 0.25\rho_\beta^2(P_c, P_w)$).

Видно, что при малых корреляциях отсчетов это ПИСГС.

Данная СГС оказывается робастной относительно аддитивной помехи.

$P = Q\left(\sqrt{\frac{N_0}{\eta - 1}}\right)$ для информированного декодера,

$P = Q\left(\sqrt{\frac{N_0}{\eta_w}}\right)$ для "слепого" декодера,

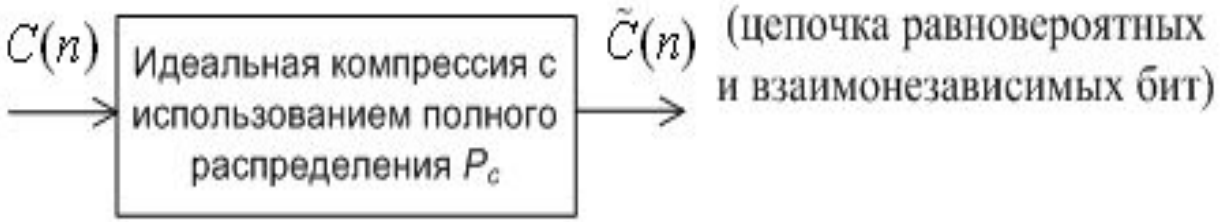
где $\eta_w = \frac{\sigma_0^2}{\alpha^2}$; $\eta_a = \frac{\sigma_\ell^2}{\alpha^2 + \sigma_\varepsilon^2}$ (см. лекцию 3).

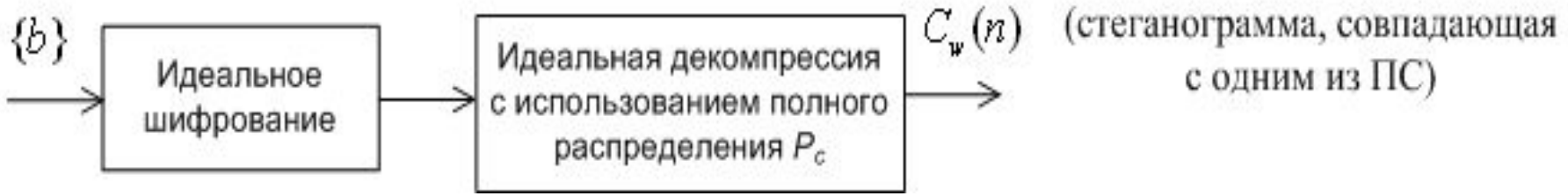
Пример. η_w = 100, η_a = 70, N_0 = 5, N = 500; тогда $P_I \geq 0.16$, $P \geq 3 \cdot 10^{-4}$ и на этом интервале можно секретно погрузить 100 бит.

При больших корреляциях между отсчетами ПО система перестает быть ПИСГС даже при малой скорости вложения. (Однако, это потенциальные условия обнаруживаемости СПО, а при использовании реальных одномерных или даже двумерных статистик, СГС может оставаться ПИСГС.)

2. СГС, реализованные с использованием методов сжатия ПО.

2.1. Идеальное сжатие.



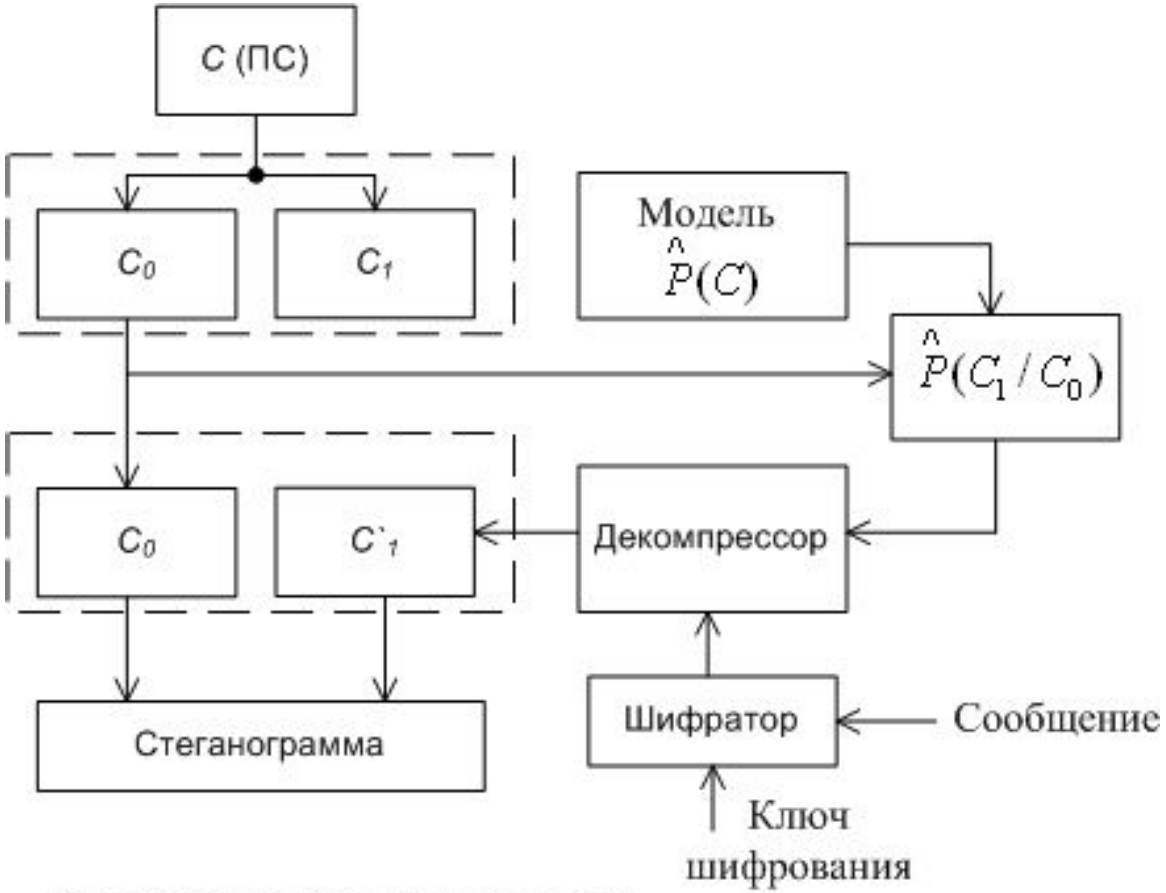


Такая СГС будет идеальной, но она практически нереализуема по двум причинам:

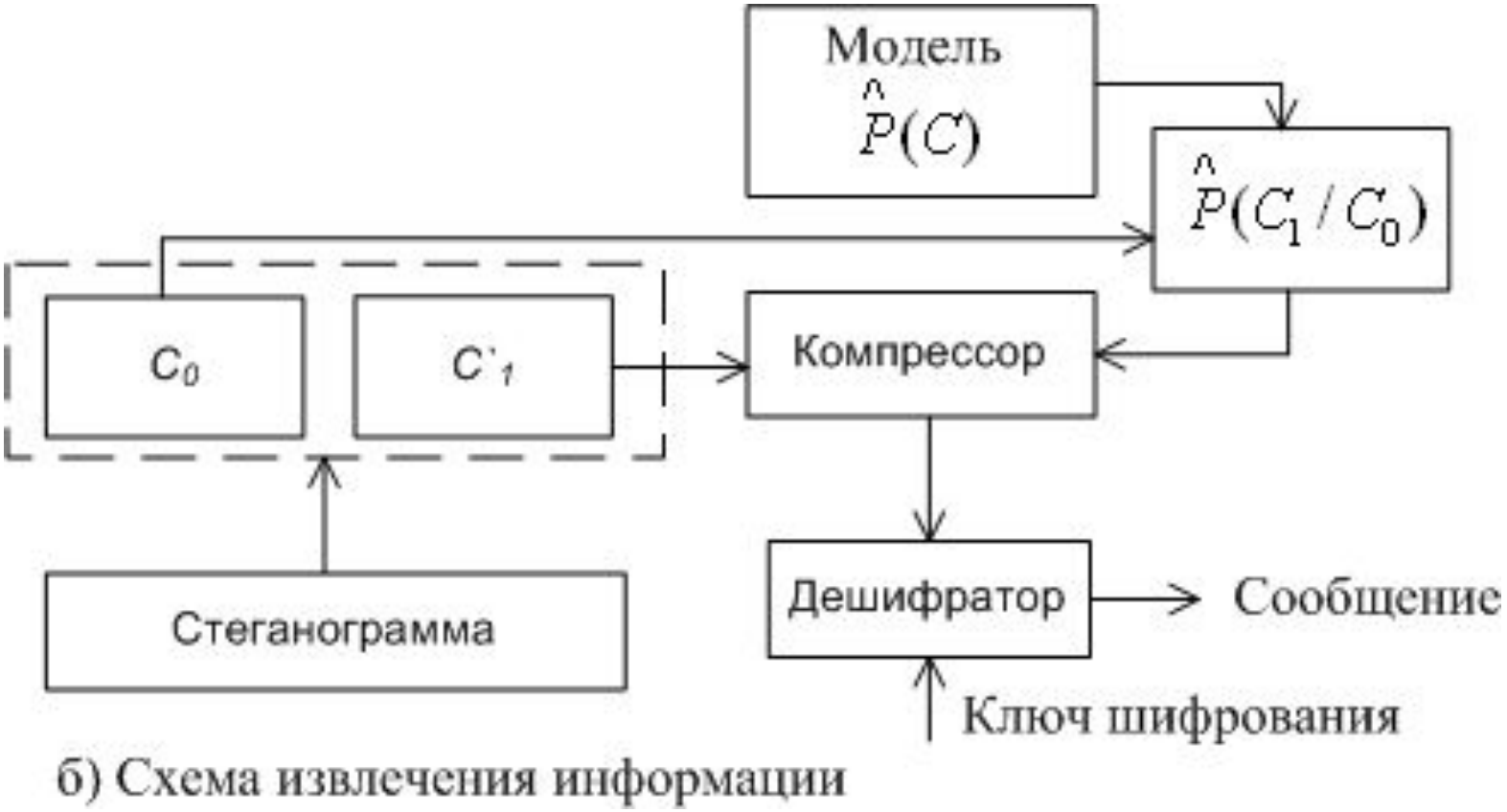
- распределение P_c для реальных ПО никогда не известно полностью,
- даже если бы P_c было в точности известно, реализовать практически идеальное сжатие невозможно.

2.2. Сжатие по частично известной статистике ПО [ ].

СМ делится на 2 части: C_0 и C_1, где C_0 сохраняется в СГ. Известная о ПО статистика используется для оценки условной вероятности $\hat{P}(C)$. $\hat{P}(C_1/C_0)$ Зашифрованное сообщение поступает в виде криптограммы (ПОевдослучайной последовательности) на вход декомпрессора, который использует условное распределение $\hat{P}(C_1/C_0)$. Результат компрессии $C`_1$ заменяет часть ПО - C_1. Объединение C_0 и C_1 дает стеганограмму.



а) Схема вложения информации



б) Схема извлечения информации

Частный случай. В качестве C_1 используются НЗБ некоторых коэффициентов DCT изображения. Моделирование такой СГС показало достаточно высокую скорость вложения информации и распределения НЗБ DCT коэффициентов в стеганограмме близкие к их распределениям в ПО.

Замечание. Фактически, данный метод состоит в более точной оценке распределения НЗБ DCT коэффициентов по реальному ПО и использованию этой статистики при вложении информации.

2.3. Адаптивная модуляция процедуры квантования (АК) (Perturbed Quantization Steganography (PQS) [ ]).

2.3.1. Вложение при помощи обычной модуляции квантования.

C(n) – непрерывная по амплитуде последовательность отсчетов или квантованная с повышенной точностью (например, 16-битовые отсчеты WAV или при передаче медицинских изображений).

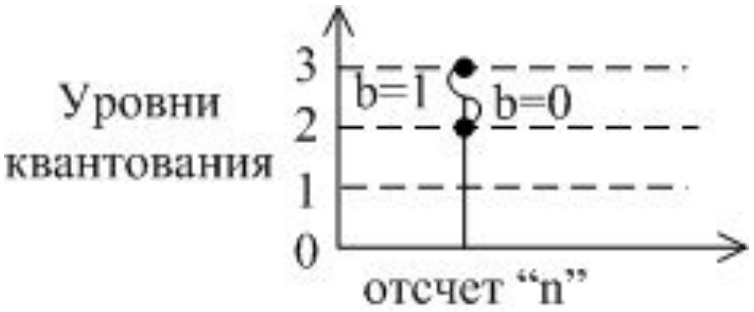
Метод вложения

$$C_w(n) = \begin{cases} C_{qe}(n), b = 0, \\ C_{qo}(n), b = 1, n = 1, 2..., \end{cases} \quad (17)$$

где $C_{qe}(n)$ - квантование отсчета *C(n)* до ближайшего четного уровня,

$C_{qo}(n)$ - квантование отсчета *C(n)* до ближайшего нечетного уровня.

Иллюстрирующие примеры:



Свойства данной СГС:

- скорость вложения 1 бит/отсчет,
- для декодирования не требует знания ПО,
- подвержена атаке по удалению вложения (рандомизация квантования),
- легко обнаруживается.

Для преодоления последнего свойства и была предложена СГ-АК.

Основная идея: использовать при вложении в квантователе тот факт, что атакующему никогда не известны отсчеты ПО до квантования.

Метод погружения иллюстрируется рисунком:

изводится только в те

интервал шириной *ε*, вокруг

ежутков) между уровнями. В

тся в

тный уровень, если

гный, если *b = 1*.

аномалию” квантования, поскольку ему никогда не известны значения отсчетов ПО до квантования. Кроме того, распределение амплитуд отсчетов на *ε*–интервале близко к равномерному и поэтому после вложения статистика отсчетов изменяется мало.

Проблема. Как определить на приеме (при декодировании) отсчеты с вложением?

Решается при помощи использования “*wet paper codes (WPC)*”:
b = $(b_1, b_2, \ldots b_k)^T$, $b_i = \{0,1\}^k$, “T” – транспонирование, n_0 – количество отсчетов, куда надо вложить эти *k* бит.
H – kxn_0 двоичная матрица, которая генерируется легальными пользователями по заранее распределенному стегоключу.
$S_p(n) = C_q(n) mod 2$, $n = 1, 2\ldots$, где $C_q(n)$ – величина $C(n)$, квантованная до ближайшего уровня при выборе интервала квантования Δ и числа уровней квантования *L*. (Видно, что $C_q(n) mod 2$ принимает значения 0 и 1, соответствующие четным и нечетным уровням квантования.)

- множество отсчетов, куда допустимо погружение по методу АК.

$E_\varepsilon \subseteq \{1, 2 \ldots n_0\}$

Метод погружения:

(18)

$$\mathbf{S_p} \to \mathbf{S'}_p = [S'_p(n)]_{n=1}^{n_0} : \mathbf{HS'_p} = \mathbf{b}, S'_p(n) \neq S_p(n), n \in E_\varepsilon.$$

Необходимое условие решения матричного уравнения (18)

(19)

$$ran\tilde{k}\mathbf{H} = k,$$

где $\tilde{\mathbf{H}}$ - двоичная (*kxm*) подматрица ***H***, полученная при помощи удаления из *H* всех столбцов с номерами, не принадлежащими E_ε, *m* – количество отсчетов среди n_0, куда можно погружать информацию по методу PS.

Можно доказать [7], что $m \approx k$, когда $n_0 \to \infty$, т.е., в среднем, при больших длинах n_0 блоков, в каждый из них можно будет погрузить $k = m$ бит информации.

Декодирование:

$$\mathbf{b} = \mathbf{H}\mathbf{S}'_{\mathbf{p}}. \quad (20)$$

Видно, что в отличие от обычных кодов с исправлением ошибок, WPC имеют значительно более сложную процедуру *кодирования* (нужно решать линейную систему), чем декодирования (умножение матрицы на вектор).

Замечание 1. Поскольку длина блоков n_0 может быть велика (сотни и тысячи бит), то для решения уравнения (18) нужно использовать специальные вычислительные методы [ ].

Замечание 2. В СГ должна содержаться дополнительная информация о числе бит, вкладываемых на этом блоке k (она вкладывается в заголовок длиной d бит).

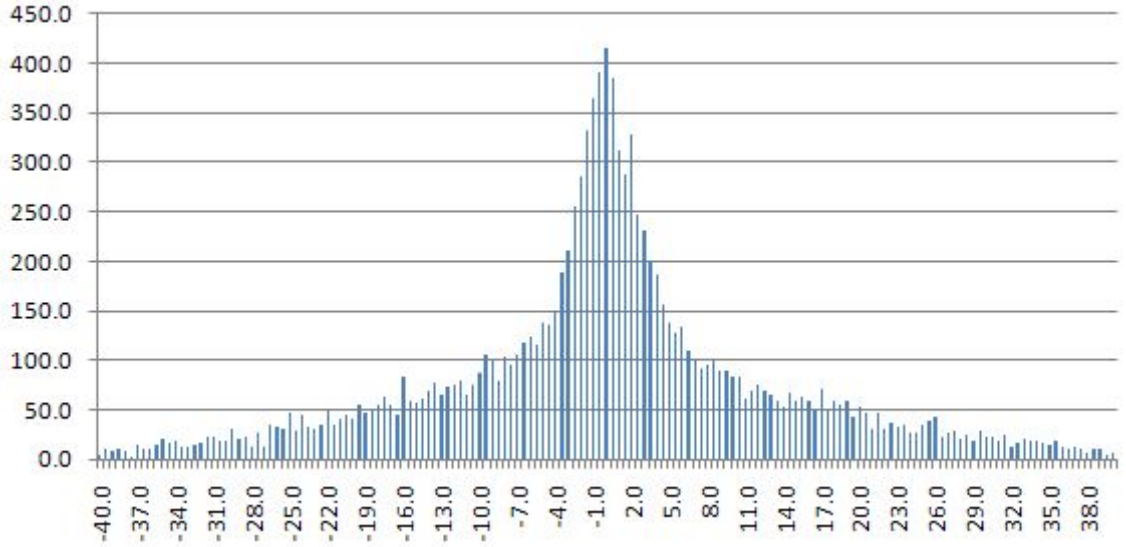
Доказывается [ ] граница: для вероятности $P_e(k, d, n_0, P_w)$ такого события, что в блок из n_0 отсчетов не удастся вложить k бит, при заголовке длиной d и при вероятности погружения в отсчет P_w:

$$P_e(k,d,n_0,P_w) \le \left[\left(\frac{P_w}{\mu} \right)^{\mu} \left(\frac{1-P_w}{1-\mu} \right)^{1-\mu} \right], \quad (21)$$

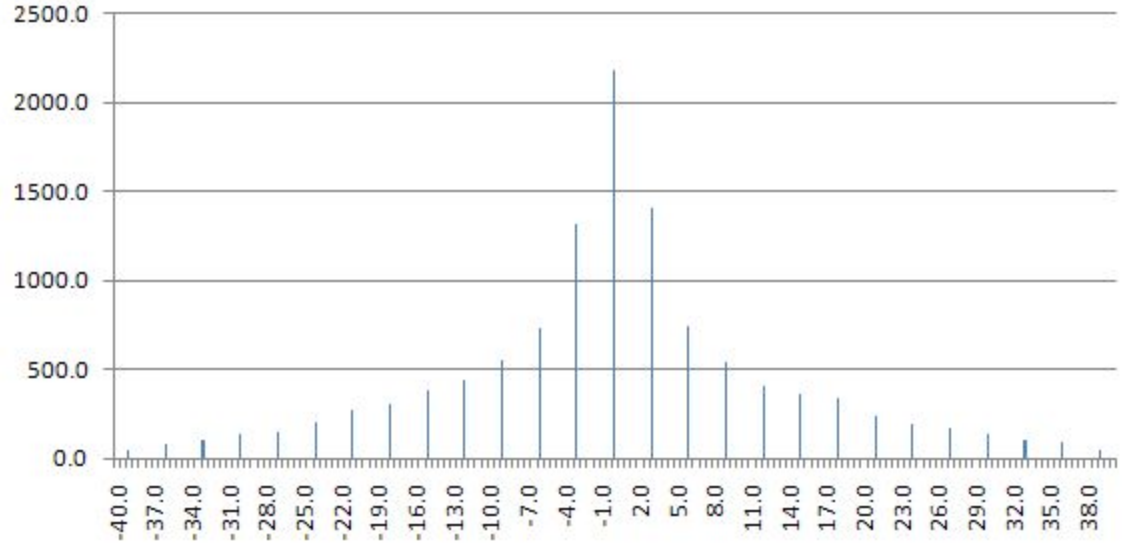
где $\mu = \dfrac{k+d-1}{n_0}$.

2.4 СГ-АК после двойной компрессии[12].

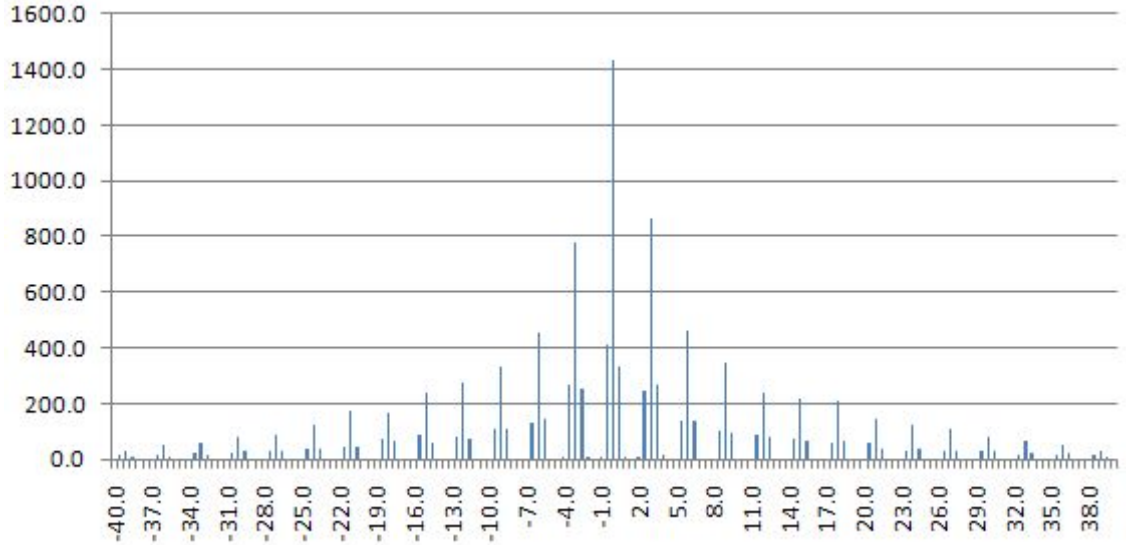
Метод, использующий двойное сжатие изображений в формате JPEG



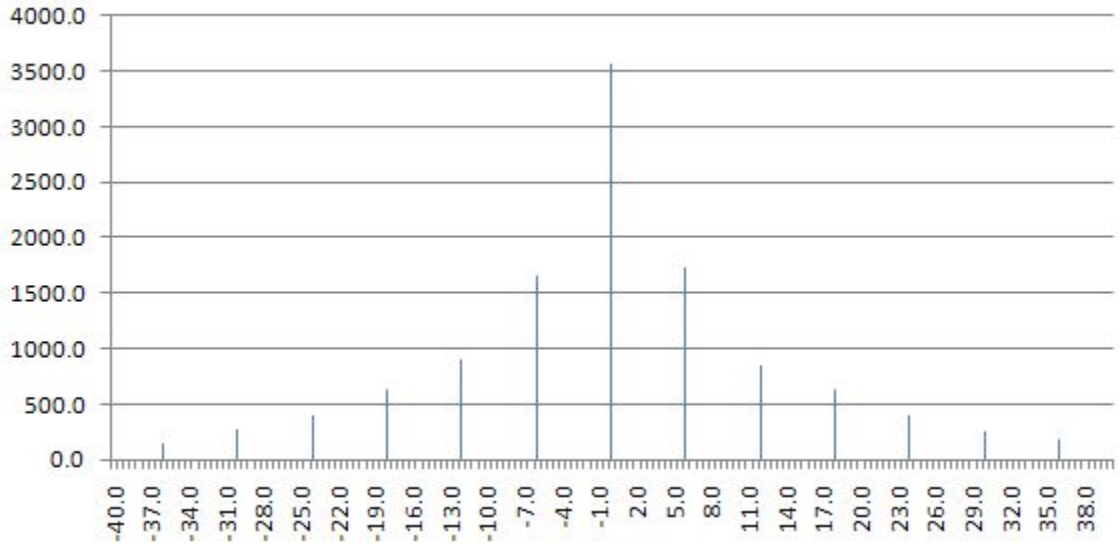
Гистограмма DCT коэффициента C21 до первого квантования



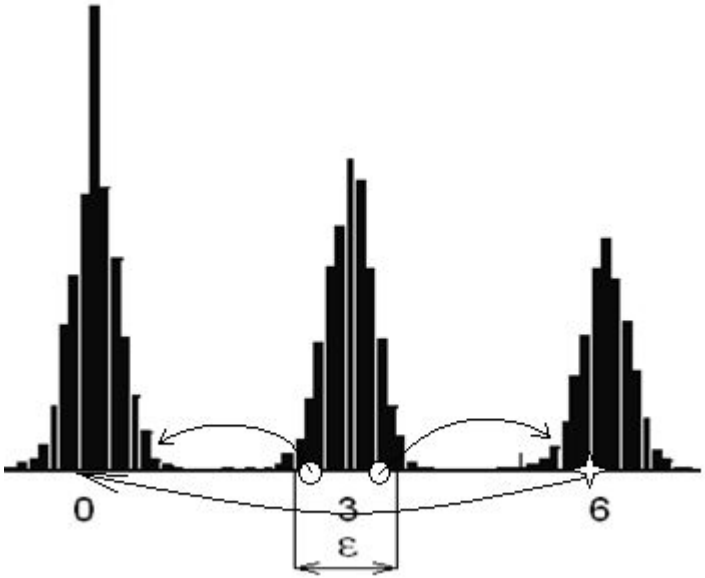
Гистограмма DCT коэффициента C21 , Q = 88%, q21 = 3 (после квантования).



Гистограмма DCT коэффициента C21 после преобразования JPEG изображения, выполненного с показателями качества Q=88% к формату bmp



Гистограмма DCT коэффициента C21 , после BMP-Jpeg-BMP
Q = 76%, q21 = 6



Участок гистограммы коэффициента C21 до второго квантования

АК (при двойном сжатии) вложение

Правило выбора коэффициента: $kq_{ij}^{(1)} = lq_{ij}^{(2)} + q_{ij}^{(2)} / 2$,
где k и l – целые числа, $q_{ij}^{(1,2)}$ - шаги квантования .

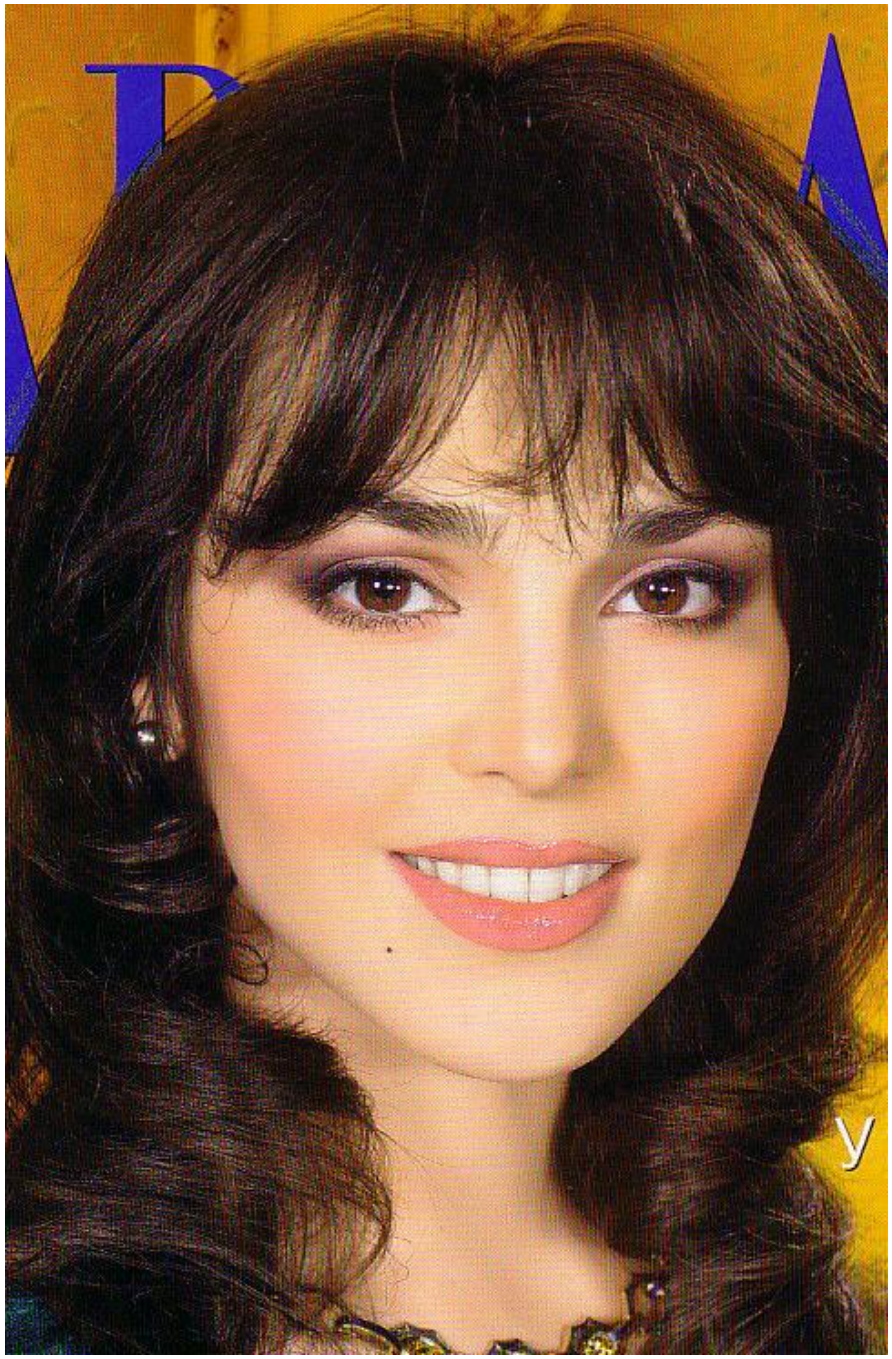
Общее количество изменяемых коэффициентов: :

$$|S| = \sum_{i,j=0}^{7} \sum_{m} z_{ij} h_{ij} ((2m+1) \frac{q_{ij}^{(2)}}{2g})$$

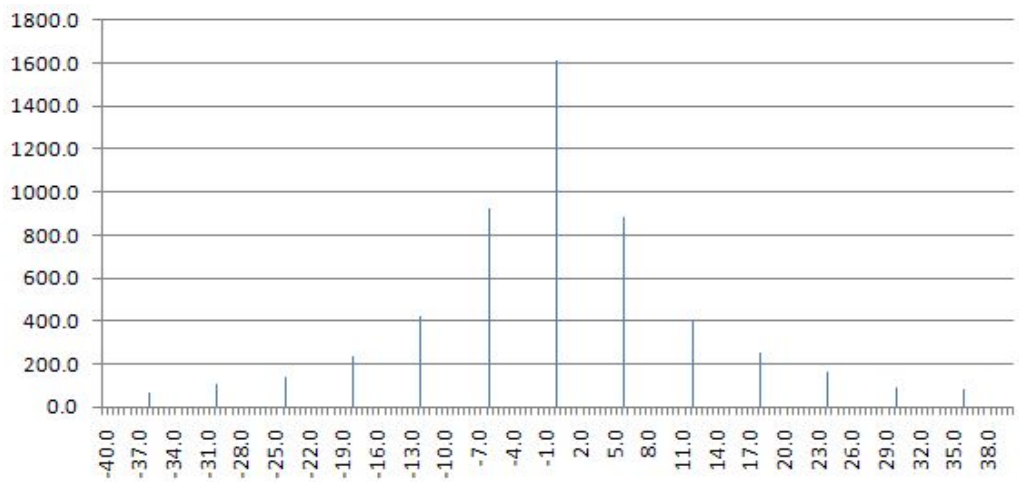
где z_{ij} = 1, если ($q_{ij}^{(1)}$, $q_{ij}^{(2)}$) являются «сотрудничающей парой» и z_{ij} = 0, в противоположном случае; $g = \gcd(q_{ij}^{(1)}, q_{ij}^{(2)})$

Исходное BMP ПО

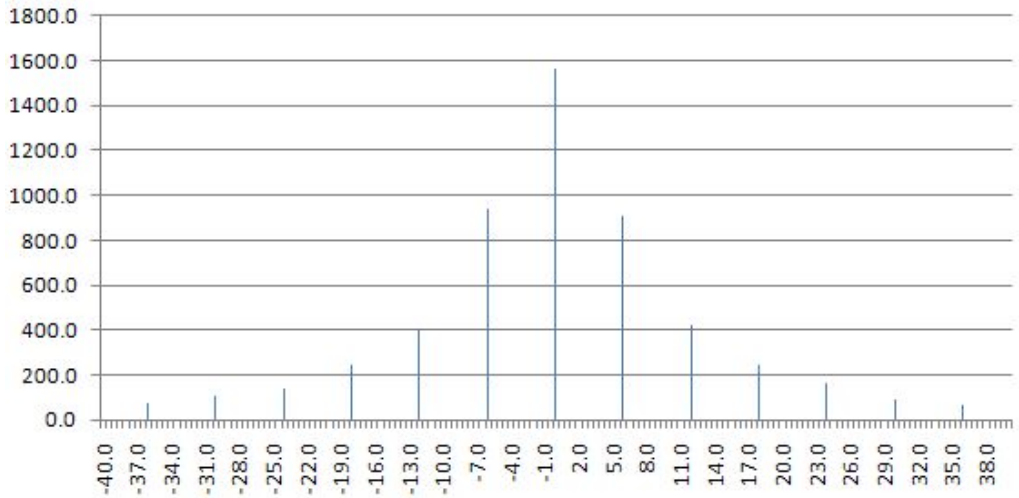
Изображение после жвоейного сжатия $Q_1 = 88\%$, $Q_2 = 76\%$



Изображение после двойного сжатия и вложения 65622 бит, $Q_1 = 88\%$, $Q_2 = 76\%$



Гистограмма DCT C_{21} изображения после двойного сжатия $Q_1 = 88\%$, $Q_2 = 76\%$



Гистограмма DCT C_{21} изображения после двойного сжатия и вложением 65622 bits, $Q_1 = 88\%$, $Q_2 = 76\%$

Видно, что гистограммы практически совпадают, что говорит о высокой стойкости СГ-АК.

Секретность СГС – PQS.

Воспользуемся критерием относительной энтропии для одномерной статистики:

$$D(\mathbf{P}_c \,||\, \mathbf{P}_w) = N\sum_{\ell=1}^{L} P_\ell \log\frac{P_\ell}{\tilde{P}_\ell} = ND_1(P_c \,||\, P_w), \qquad (22)$$

где $P_c = (P_\ell)_{\ell=1}^{L}$ - одномерное распределение, соответствующее ПО после квантования,

$P_w = (\tilde{P}_\ell)_{\ell=1}^{L}$ - одномерное распределение, соответствующее СГС – PS,

N - общее число наблюдаемых отсчетов,

L - количество уровней квантования.

Легко проверить, что

$$P_\ell = \int_{J_\ell} P_c(x)dx, \qquad J_\ell = [\Delta(\ell - 1/2), \Delta(\ell + 1/2)], \Delta -$$

где интервал квантования.

$$\tilde{P}_\ell = \frac{1}{2}\int_{I_{\ell,\varepsilon}} P_c(x)dx + \int_{M_{\ell,\varepsilon}} P_c(x)dx + \frac{1}{2}\int_{K_{\ell,\varepsilon}} P_c(x)dx,$$

где

$$I_{\ell,\varepsilon} = [\Delta(\ell - 1/2) - \varepsilon/2, \Delta(\ell - 1/2) + \varepsilon/2], K_{\ell,\varepsilon} = [\Delta(\ell + 1/2) - \varepsilon/2, \Delta(\ell + 1/2) + \varepsilon/2],$$

$$M_{\ell,\varepsilon} = [\Delta(\ell - 1/2) + \varepsilon/2, \Delta(\ell + 1/2) - \varepsilon/2].$$

$$P_w = \sum_{\ell=1}^{L}\int_{H_{\ell,\varepsilon}} P_c(x)dx, \qquad H_{\ell,\varepsilon} = [\Delta(\ell + 1/2) - \varepsilon/2, \Delta(\ell + 1/2) + \varepsilon/2].$$

где

Экспериментальные результаты для гауссовского ПО.

$v = \varepsilon/\Delta$	D_1	P_w	m
0.005	$6.588 \cdot 10^{-11}$	0.00498	$7.56 \cdot 10^{6}$
0.01	$2.6 \cdot 10^{-11}$	0.00973	$3.78 \cdot 10^{6}$
0.025	$1.647 \cdot 10^{-9}$	0.025	$2.5 \cdot 10^{6}$
0.05	$6.58 \cdot 10^{-9}$	0.05	$7.5 \cdot 10^{5}$
0.1	$2.6 \cdot 10^{-8}$	0.1	$3.8 \cdot 10^{5}$
0.25	$1.6 \cdot 10^{-7}$	0.249	$1.5 \cdot 10^{5}$
0.3	$2.372 \cdot 10^{-7}$	0.299	$1.26 \cdot 10^{5}$
0.5	$6.592 \cdot 10^{-7}$	0.499	$7.56 \cdot 10^{4}$
1.0	$2.641 \cdot 10^{6}$	0.997	$3.77 \cdot 10^{4}$

секретно

меньше параметр v =

сть вложения).

эти соотношения,

Одномерная статистика

очна для обнаружения PS.

обнаруживаться

малых v) для реальных аудио

Если мы положим, что в среднем число погружаемых бит равно $N P_w$, то число секретно погружаемых бит m (при $D(P_c||P_w) = 0.1$) будет равно

$$m = N P_w = 0.1 \frac{P_w}{D_1(P_c \,\|\, P_w)}.$$

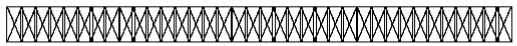
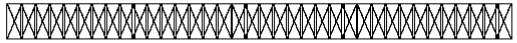
2.5. СГС с сохранением одномерной статистики (СГ-СОС).

ПО – двоичные (a,b) последовательность с взаимонезависимыми, но не обязательно равновероятными символами.
Вкладываемая информация – (0,1)

Алгоритм погружения:
1. Разделить последовательность ПО на пары.
2. Выполнить отображение: $aa \to u, bb \to u, ab \to v_0, ba \to v_1$
3. Изменить (если необходимо) символы $v_k, k = 0{,}1$ на символы v_y, где *y* – текущий бит вкладываемой информации,
4. Изменить символы v_y на соответствующие им пары двоичных символов, а вместо символов *u* восстановить ранее соответствовавшие им двоичные символы

Алгоритм извлечения:
1. Разделить последовательность двоичных символов на пары.
2. Выделить пары ab и ba, исключить из рассмотренных все u -пары.
3. Произвести последовательное декодирование вложенной последовательности по правилу: $ab \to 0, ba \to 1$

Пример:

Вкладываемая информация:	0110
Пусть ПО =	
Деление на пары:	
Первичное кодирование	
Вложение информации	
Двоичное кодирование	
Извлечение информации	0110

Свойства

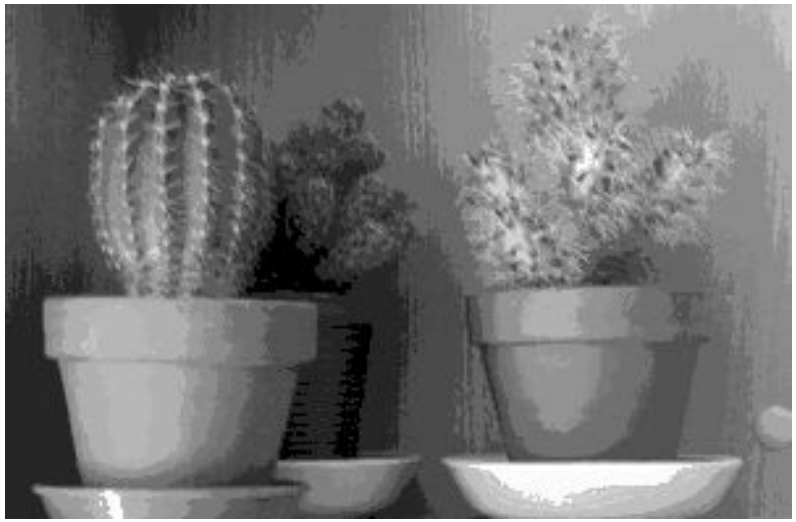
- При отсутствии помех извлечение производится без ошибок.
- Длина СГС равна длине ПО.
- Одномерная статистика (т.е. вероятности появления «a» и «b») в СГС сохраняются такими же, какими они были в ПО.
- Вложение информации существенно искажает последовательность бит, представляющую ПО. Более того, даже если вложенная информация извлечена верно, ПО не может быть всегда восстановлена такой же, какой она была на самом деле. (Действительно, если в СГС принята пара «ab», то не известно, была ли эта пара равна «ab» или «ba», а известно лишь, что они не могли быть равной «aa» или «bb»).
- Количество бит вложенных в ПО длины «n» всегда будет меньше, чем $n/2$ и данная величина будет зависеть от статистики ПО (P(a),P(b)).
- Свойство 4 приводит к тому, что данный метод может применяться не ко всему ПО, а лишь к некоторой его части, которая слабо влияет на восприятие всего ПО, например, к НЗБ.

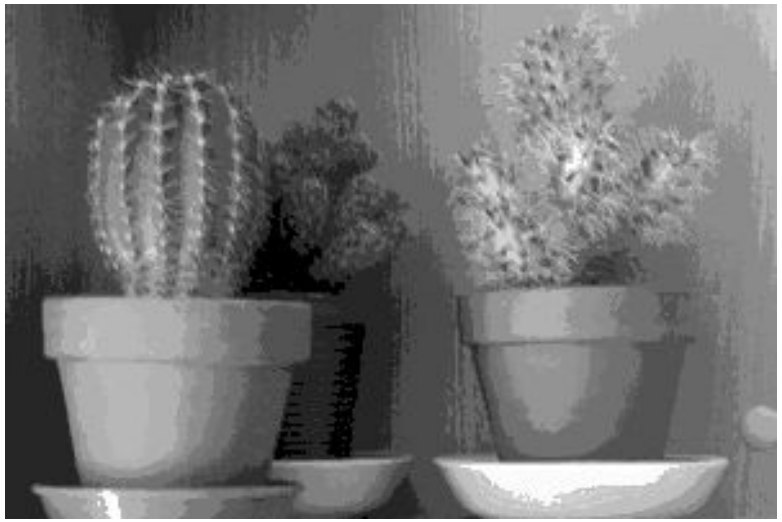
На рисунках показано изображение без вложения и тоже самое изображение с вложением 3 000 бит, тогда как потенциальные возможности вложения в НЗБ данного изображения составляют около 4 740 бит.

Пример 100% вложения информации СГ-СОС

Поскольку изображения имеют различные распределения яркости, пропускная возможность будет меняться по мере изменения ПО. Рассмотрим несколько изображений и получим возможный объем дополнительной информации, которую можно было бы вложить в них..

1. Размер изображения 200x300px. N_{max} = 6,330 bits.





Слева – исходное изображение, справа – изображение с вложением.

2. Размер изображения 200x300px. Nmax = 12,057 bits.

3. Размер изображения 199x300px. Nmax = 11,443 bits.

4. Размер изображения 200x300px. Nmax = 4,051 bits.

5. Размер изображения 200x300px. Nmax = 13,747 bits.

Полученные результаты представлены в таблице ниже

№	car	1	2	3	4	5
N_{max}	4,740	6,330	12,057	11,443	4,051	13,747

Представленные данные позволяют сделать вывод о том, что в среднем 5000-10000 бит дополнительной информации могут быть встроены в 8-битное изображение в градациях серого, содержащее 200x300px.

В этом методе возможный объем встраивания зависит от размера и распределения яркости изображения, которое изменяется при изменении глубины цвета изображения.

Steganalysis of SG-LSB-R method:
Сравнение методов стегоанализа СГ-НЗБ и СГ-СОС

Visual attack

Визуальная атака означает извлечение НЗБ из изображения. Четкий контур содержания изображения означает отсутствие встраивания, иначе мы увидим шумы, которые будут зависеть от размера встраивания.

Визуальная атака на исходное изображением

.

СГ-НЗБ

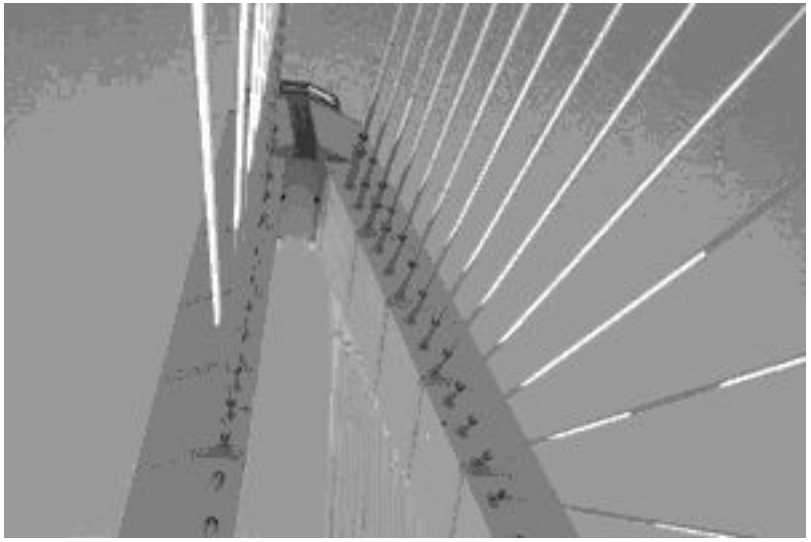
СГ-СОС

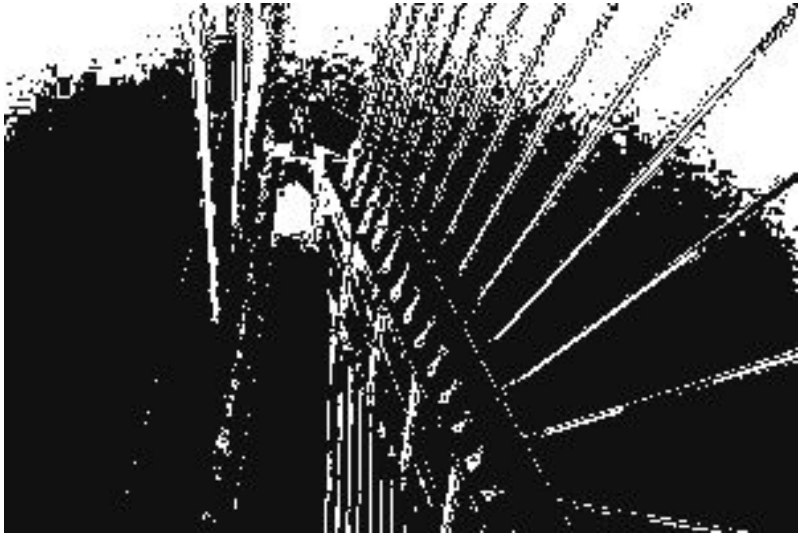
3,000 бит информации может быть вложено. В случае СГ-НЗБ мы получаем малошумную картину. В то время как в случае метода СГ-СОС факт встраивания не может быть обнаружен достаточно просто и необходимо сравнение с исходным изображением. При уменьшении вложенной информации (например до 1000 бит и ниже) факт встраивания становится неочевидным.

Другие методы стегоанализа на СГ-СОС

- Исследованы следующие атаки:
- Статистическая атака первого порядка (FOSA)
- Статистическая атака второго порядка (SOSA)
- Атака, основанная на подсчете нулей в гистограмме (ZHA)
- Атака, основанная на сравнении соседних значений гистограммы (NVA)

Результаты атак

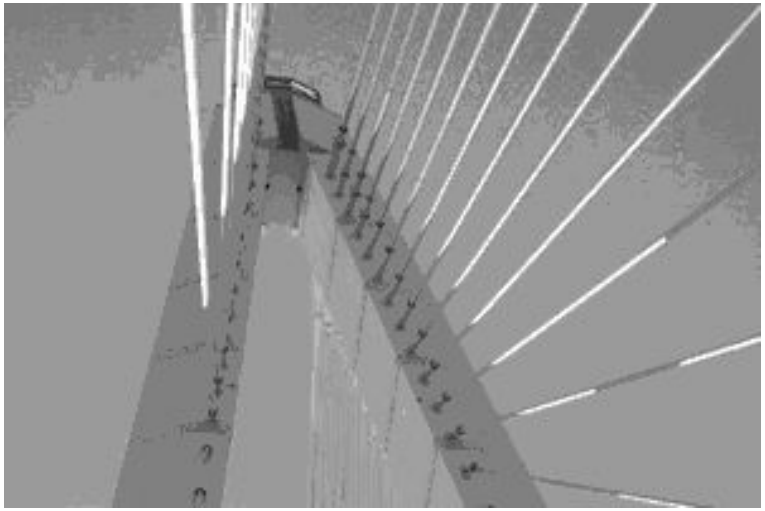


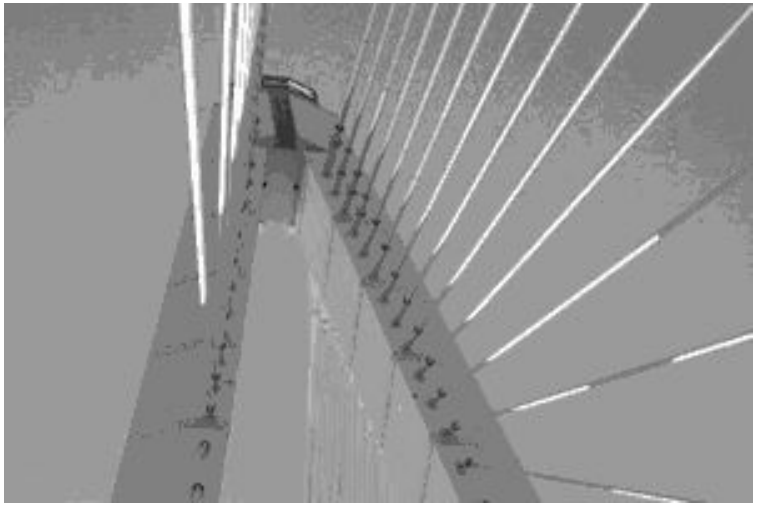


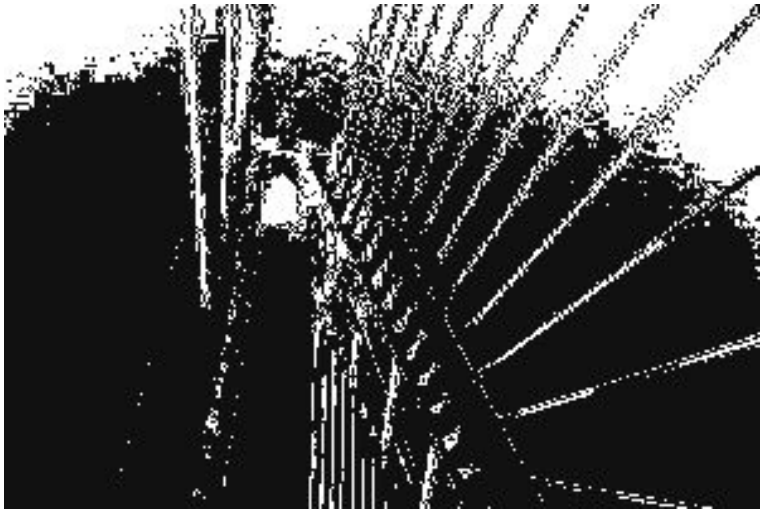
method	FOSA	SOSA	ZHA	NVA
СГ-СОС	detected	not detected	not detected	not detected
СГ-НЗБ	detected	detected	not detected	not detected

3,000 были вложены, однако SOSA не обнаружил этот факт.

Визуальная атака СГ-НЗБ (слева) и СГ-СОС (справа), вложено 3,000 бит







Выводы

- прямое использование метода СГ-СОС приводит к недопустимым искажениям ПО.
- Данный метод может быть использован как усовершенствоанный метод СГ-НЗБ