Реальные СГС-НЗБ и СГС-ШПС.

1. Реальные стегосистемы с вложением в НЗБ.
(программы для этих СГС распространяются свободно по Интернету)

1.1 Jsteg. В качестве ПО используется цветное изображение в формате JPEG.
Вложение производится в НЗБ частотных коэффициентов (за исключением нулевых и единичных), выбираемых по псевдослучайному пути, который задается стегоключом (паролем).

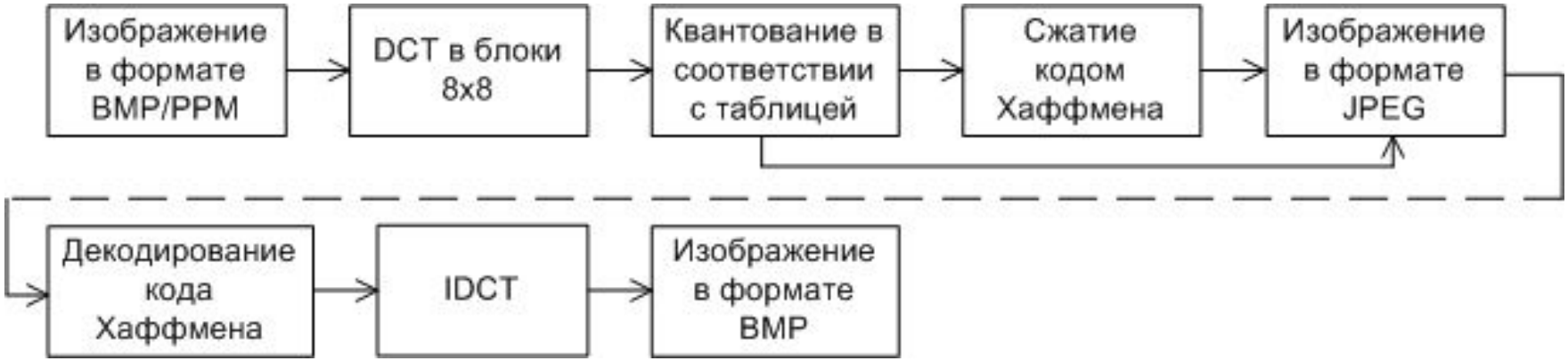


Рис. 1. Алгоритм сжатия изображения в формате JPEG.

Jsteg не обнаруживается визуальной атакой, но легко обнаруживается с использованием статистики χ^2 и анализа пар выборок.

1.2 Outguess. В качестве ПО используется цветное изображение в формате JPEG. Алгоритм реализован под операционную систему FreeBSD на языке C++. В лабораторной работе N1 портирован с помощью эмулятора cygwin. Работает из командной строки. Требует задания паролей (стегоключей) вложения и извлечения. Алгоритм вложения разработан для обеспечения защиты от атаки обнаружения χ^2.

Вложение происходит в два прохода: первый – по псевдослучайному пути, определяемому стегоключом (паролем), как в Jsteg, а второй – с изменением коэффициентов не затронутых первым проходом, с целью приближения гистограммы СГ-изображения к гистограмме ПО, что затрудняет χ^2-атаку.

Однако, обнаружение Outguess сказывается возможным, для чего используется факт увеличения “неоднородности” в блоках 8x8, которые сравниваются с неоднородностью исходного ПО, полученного при помощи оценки стегоизображения (см. далее: обнаружение СГС-ШПС и “слепой” стегоанализ).

1.3 F5. В качестве ПО используется цветное изображение в формате JPEG. Однако, в отличие от Jsteg и Outguess, это не чистая СГС-НЗБ. Основной принцип F5: при заданном числе вкладываемых бит информации минимизировать количество изменяемых бит ПО.

Пример. $x_1, x_2 \in \{0,1\}$ - биты вкладываемой информации. Обычное НЗБ требует изменение 2^x бит ПС. Модифицированное вложение (где a_1, a_2, a_3 – биты ПС, которые можно изменять):

$x_1 = a_1$ a_3, $x_2 = a_2$ a_3 => ничего не изменять,

$x_1 \neq a_1$ a_3, $x_2 = a_2$ a_3 => изменить a_1,

$x_1 = a_1$ a_3, $x_2 \neq a_2$ a_3 => изменить a_2,

$x_1 \neq a_1$ a_3, $x_2 \neq a_2$ a_3 => изменить a_3.

Во всех случаях изменяется не более одного бита. По заданным a_1, a_2, a_3 однозначно восстанавливаются x_1, x_2.

Правило извлечения:

$x_1 = a_1 \oplus a_3$, $x_2 = a_2 \oplus a_3$

Алгоритм F5 реализован с помощью JavaScript и использует обобщение данного подхода (матричный) *(1,n,k)*-код, где n – число позиций, которые могут меняться, *k* – число вкладываемых бит, 1 – максимальное число изменений при вложении *k* бит.

Параметры F5: $n = 2^k - 1$, - длина блоков, плотность изменений - $1/2^k$, скорость погружения – $k/n = k/2^k - 1$.

(Допустимые для вложения биты определяются ПСП, задаваемой стегоключом (паролем)). Уменьшение плотности изменений позволяет уменьшить вероятность обнаружения.

Однако СГС-F5 может быть обнаружена при помощи сравнения гистограмы выбранных DCT коэффициентов СГС и гистограммы таких же коэффициентов для оценки исходного ПС:

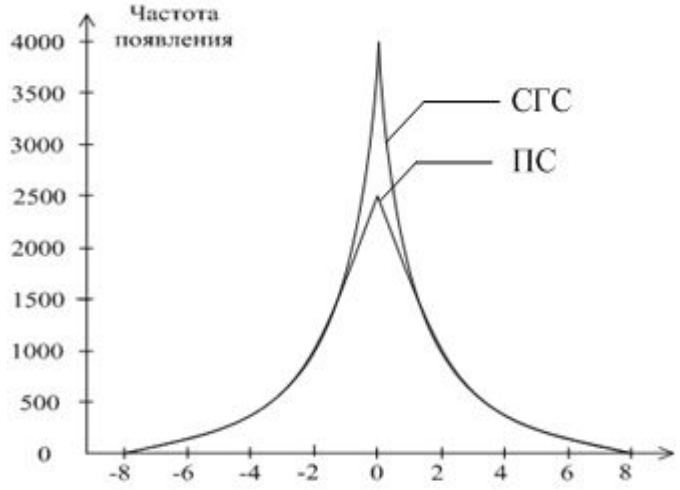


Рис. 2. Гистограмма DCT-(2,1) коэффициентов F5 и оценки действительного ПО.

Видно, что СГС-F5 может быть обнаружена.

2. СГС-ШПС.

Все СГС-НЗБ не выдерживают атаки по удалению вложенных сообщений даже при сохранении при этом высокого качества ПО.

Эта атака реализуется при помощи рандомизации НЗБ во временной или частотной области.

Для защиты от такой атаки необходимо использовать широкополосные сигналы (ШПС-СГ):

$$C_W(n) = C(n) + \alpha(-1)^b \pi(n), n = 1,2...N, \quad (1)$$

где α – коэффициент вложения, $\pi(n)$ – псевдослучайная (±1) последовательность (ПСП), вырабатываемая по секретному стегоключу, N – длина ПСП, на которой вкладывается один и тот же бит ($b=1$ или 0) информации.

Выделение информации при неизвестном ПО (“слепой” декодер):

$$\sum_{n=1}^{N} (C'_W(n) - m_c)\pi(n) \begin{cases} > 0 \rightarrow b = 0 \\ < 0 \rightarrow b = 1 \end{cases}, \quad (2)$$

где атака производится аддитивным шумом:

$$C'_W(n) = C_W(n) + \varepsilon(n), n = 1,2...N, \quad (3)$$

и где $m_c = E\{C(n)\}$.

Поскольку $\pi(n)$ при атаке неизвестна, то при выборе достаточно больших N и малых искажениях $C(n)$, атака не является успешной для любой статистики шума.

Действительно, рассмотрим вероятность ошибки для легитимного пользователя, который знает $\pi(n)$, $n = 1,2\ldots N$.

$$p(1/0) = P\{\Lambda \le 0 / b = 0\}, p(0/1) = P\{\Lambda > 0 / b = 1\} \quad (4)$$

$$\Lambda = \sum_{n=1}^{N} ((C_w^{'}(n) - m_c)\pi(n))$$

При $N \to \infty$, $\Lambda \sim N(E(\Lambda), Var(\Lambda))$ (ЦПТ теории вероятностей)

$$E\{\Lambda\} = E\{\sum_{n=1}^{N} (C(n) - m_c + \alpha(-1)^b \pi(n) + \varepsilon(n))\pi(n)\} = \alpha(-1)^b N \quad (5)$$

$$Var\{\Lambda\} = \sum_{n=1}^{N} E\{((C(n) - m_c + \varepsilon(n))\pi(n))^2\} =$$

$$= NE\{(C(n) - m_c)^2 + 2\varepsilon(n)\pi(n)(C(n) - m_c) + \varepsilon^2(n)\}) = N(\sigma_c^2 + \sigma_\varepsilon^2), \quad (6)$$

где $\sigma_c^2 = Var\{C(n)\}, \sigma_\varepsilon^2 = Var\{\varepsilon(n)\}$.

Если мы положим $m_c = 0$ в (3), то получим вместо (6):

$$Var\{\Lambda\} = N(E\{C^2(n)\} + \sigma_\varepsilon^2) = N(Var\{C(n)\} + m_c^2 + \sigma_c^2) = N(\sigma_c^2 + m_c^2 + \sigma_\varepsilon^2) \quad (7)$$

$$Var\{\Lambda\} \ge Var\{\Lambda\}$$

Положим сначала *b = 0.* Тогда

$$p(1/0) = p\{\Lambda \le 0 / b = 0\} = Q(E\{\Lambda\} / \sqrt{Var\{\Lambda\}}), \quad (8)$$

где $Q(x) = \frac{1}{\sqrt{2\pi}} \int_x^\infty e^{-t^2/2} dt.$

Подставляя (5) и (6) в (8), получим

$$p(1/0) = Q\left(\frac{\alpha}{\sqrt{N/(\sigma_c^2 + \sigma_\varepsilon^2)}}\right) \quad (9)$$

(Легко проверить, что аналогичное выражение получается и для случая *b = 1,* т.е. *p(1/0) = p(0/1) = p*).

Введем обозначения:

$\eta_w = \frac{\sigma_c^2}{\alpha^2}$ - (отношение сигнал/шум после погружения WM), (10)

$\eta_a = \frac{\sigma_c^2}{\alpha^2 + \sigma_\varepsilon^2}$ (отношение сигнал/шум после атаки). (11)

Подставляя (10) и (11) в (9), получим

$$p = Q(\sqrt{N\eta_a / (\eta_a \eta_w + \eta_w - \eta_a)}) \quad (12)$$

Типичным является случай, когда $\eta_w \ge \eta_a >> 1.$

Тогда для (12) получаем приближение

$$p = Q(\sqrt{N/\eta_w}) \quad (13)$$

Рассмотрим теперь случай *информированного декодера*, когда принятие решения о вложении информации выполняется по правилу:

$$\Lambda' \begin{cases} \geq 0 \rightarrow b = 0 \\ < 0 \rightarrow b = 1, \end{cases} \quad (14)$$

где

$$\Lambda' = \sum_{n=1}^{N} (C_w'(n) - C(n))\pi(n) \quad (15)$$

Используя ЦПТ получаем:

$$p' = P\{\Lambda' < 0 \mid b = 0\} = Q(\frac{E\{\Lambda'\}}{\sqrt{Var\{\Lambda'\}}}) \quad (16)$$

$$E\{\Lambda'\} = E\{\sum_{n=1}^{N} (C(n) + \alpha\pi(n) + \varepsilon(n) - C(n))\pi(n)\} = \alpha N \quad (17)$$

$$Var\{\Lambda'\} = Var\{\sum_{n=1}^{N} \varepsilon(n)\pi(n)\} = NVar\{\varepsilon(n)\}Var\{\pi(n)\} = N\sigma_\varepsilon^2 \quad (18)$$

Подставляя (17), (18) в (16) и используя (10), (11), получим

$$p' = Q(\frac{\alpha\sqrt{N}}{\sigma_\varepsilon}) = Q(\sqrt{\frac{N}{(\eta - 1)}}), \quad (19)$$

где

$$\eta = \eta_w / \eta_a$$

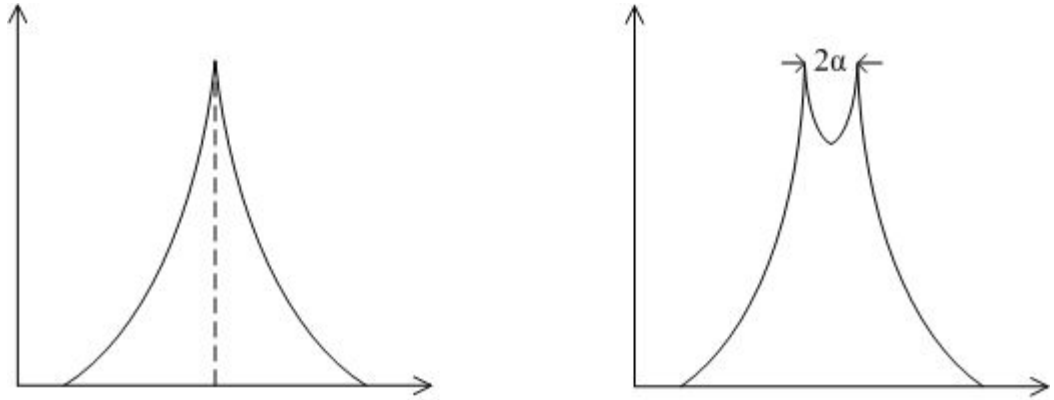
Сравнивая *p* по (12) и *p'* по (19) мы видим, что $p \geq p`$.
Действительно, выбирая $p = p`$ , но разные *N* и *N`* получаем

$$N' / (\eta - 1) = N / \eta_w => N / N' = \eta_w / (\eta - 1) \quad (20)$$

Пример. Положим $\eta_w = 120$, $\eta_a = 100$. Тогда $N/N' = 600$.
Это означает, что для "слепого" декодера скорость вложения будет в 600 раз меньше, чем для информированного!
Чтобы уменьшить эту разницу (для неизвестного у декодера ПС) используют *информированный кодер* (метод погружения), который отличается от (1). Однако это может привести к лучшему обнаружению СГС и поэтому он используется обычно для ЦВЗ (см. далее).

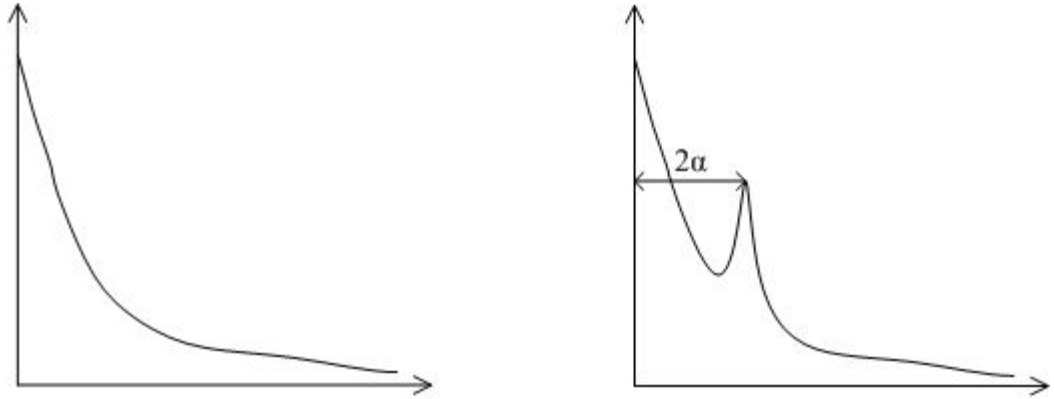
Обнаружение СГС-ШПС

1. По одномерной статистике (гистограмме)



а) ПС б) СГС-ШПС

2. По статистике второго порядка
(По гистограммам модулей разностей яркостей смежных пикселей, |C(n+1)-C(n)|)



а) ПС а) СГС-ШПС

α	№ изображения	P	X^2
1	1	1	51248
		0,5	15885
		0,1	45781
		0	58211
	2	1	40026
		0,5	3629
		0,1	43308
		0	60000
2	1	1	54824
		0,5	47791
		0,1	54739
		0	58211
	2	1	20386
		0,5	22046
		0,1	50453
		0	60000
3	1	1	49463
		0,5	50267
		0,1	55932
		0	58211
	2	1	39683
		0,5	44964
		0,1	55701
		0	60000

3. Использование критерия X^2 (см. Лекцию 2)

Можно сделать вывод, что этот метод работает для изображений высокого качества (без цифрового шума).
Лучшие результаты мы получаем для вероятности встраивания P = 0.5.

4. ПВА .

Экспериментальные результаты расчета $\widetilde{P}$ по методу ПВА ля метода СГ-ШПС

α	P	№ изображения				
		1	2	3	4	5
1	1	0,00306	0	0,03965	0,00525	0,00603
	0,5	0,03307	0	0,10544	0,01854	0,02475
	0,1	0	0	0,00580	0,00482	0,00468
2	1	0,00030	0	0	0,00005	0
	0,5	0	0,05735	0	0	0,00397
	0,1	0	0,00211	0	0	0,00050
3	1	0,00050	0,00335	0,01936	0,00215	0
	0,5	0,00061	0	0	0,00183	0
	0,1	0,00009	0,00397	0	0,00012	0
0		0	0	0	0	0,00005

Видно, что этот метод работает не очень хорошо, но он может быть использован в сочетании с другими методами.

5. Метод, основанный на подсчете нулей в гистограмме

Количество нулей в гистограмме СО всегда меньше, чем в ПО

Результаты подсчета количества нулей гистограммы для 5 различных изображений

α	P	№ изображения				
		1	2	3	4	5
1	1	165	95	154	138	161
	0,5	139	5	128	93	127
	0,1	141	6	134	100	131
2	1	153	68	133	131	157
	0,5	134	36	120	109	130
	0,1	138	36	127	120	136
3	1	165	95	154	138	161
	0,5	139	5	127	93	127
	0,1	140	5	132	95	132
0		201	162	193	174	200

Видно, что метод работает, однако не для всех изображений. Лучшие результаты при P=0.5.

6. По статистике суммы квадратов разностей яркостей соседних пикселей

$$\Gamma = \frac{1}{2N_0\sigma_c^2}\sum_{n=1}^{N_0}\left(C_W(n+1) - C_W(n)\right)^2, \qquad (21)$$

где $\sigma_c^2 = \frac{1}{N_0}\sum_{n=1}^{N_0} C_W^2(n).$

N_0 – общее число пикселей изображения.

Метод обнаружения СГС-ШПС:

$\Gamma > \gamma_0$ => СГС присутствует,

$\Gamma \le \gamma_0$ => СГС отсутствует. (22)

Действительно, для ПО:

$$E\{\Gamma\} = \frac{N_0}{2N_0\sigma_c^2}E\left\{C^2(n+1) + C^2(n) - 2C(n+1)C(n)\right\} = 1 - R_c(n, n+1), \qquad (23)$$

где $R_c(n,n+1)$ – нормированный коэффициент корреляции между яркостями соседних пикселей.

Замечание. Вложение ШПС-СГС по правилу (1) не обеспечит секретности, если при атаке известна $Var\{C(n)\} = \sigma_c^2$, поскольку тогда

$$Var\{C_w(n)\} = \sigma_c^2 + \alpha^2 > Var\{C(n)\}$$

Для секретной ШПС-СГС выполняется вложение по модифицированному правилу

$$C_w(n) = \beta C(n) + \alpha(-1)^b \pi(n), n = 1,2...N, \quad (24)$$

где $\beta = \sqrt{1 - \frac{\alpha^2}{\sigma_w^2}}.$

Тогда $Var\{C_w(n)\} = Var\{C(n)\} = \sigma_c^2$ (это можно легко проверить).

$$E\{\Gamma'\} = \frac{N_0}{2N_0\sigma_c^2} E\left\{(C_w(n+1) - C_w(n))^2\right\} =$$

$$= \frac{N_0}{2N_0\sigma_c^2} E\left\{\left[\beta C(n+1) + \alpha(-1)^b \pi(n+1) - (\beta C(n) + \alpha(-1)^b \pi(n))\right]^2\right\} \quad (25)$$

После преобразования (25) получим

$$E\{\Gamma'\} = 1 - \beta^2 R_c(n, n+1).$$

Поскольку $\beta < 1$, то $E\{\Gamma'\} \geq E\{\Gamma\}$, причем эта разница тем больше, чем больше $R_c(n, n+1)$, что и обуславливает возможность обнаружения ШПС-СГС.

Проверим обнаруживаемость СГС-ШПС для 20 различных изображений размером ~ 300x200 с градациями серого при $\alpha = 1$.

№	Изображение	Значение Γ для ПС	Значение Γ` для СГС-ШПС
1	17	447 **01720**	447 **20436**
2	20	706 **15072**	706 **82968**
3	24	73 **34841**	73 **30683**
4	25	73 **74376**	74 **12018**
5	27	428 **02232**	428 **67656**
6	29	235 **21896**	235 **48856**
7	32	822 **25568**	823 **40832**
8	35	628 **16864**	628 **25872**
9	37	123 **20428**	123 **74226**
10	38	303 **16216**	303 **88040**
11	41	643 **62788**	643 **89600**
12	43	200 **28484**	200 **27456**
13	44	135 **42418**	135 **94816**
14	45	738 **28064**	738 **52400**
15	47	1012 **11280**	1012 **68328**
16	49	746 **53152**	746 **84432**
17	50	51 **07258**	51 **53188**
18	51	217 **11080**	217 **88028**
19	52	1036 **45688**	1036 **95648**
20	53	376 **28880**	377 **21680**

Как видно из таблицы, типично изменяются последние пять цифр. Выберем для них порог $\lambda_0 = 46000$.
Тогда для выбранных ПС получаем:

- верно определена СПС-ШПС в 30 случаях
- получены ложные обнаружения в 3-х случаях
- пропущена СГС-ШПС в 7 случаях.

Из таблицы видно, что ПО и СГ различимы, но возникает проблема – как выбрать порог? Таким образом, данный подход применим в случае, когда необходимо различать, какой из двух образов ПО или СГ.

В действительности имеются и более эффективные методы обнаружения СГС-ШПС (см. далее “слепой” стегоанализ).