Белорусский Государственный Университет
Факультет Прикладной Математики и Информатики
Кафедра математического моделирования и анализа данных

Шимко Андрея Чеславовича

Обнаружение вкраплений в двоичную цепь Маркова на основе энтропийных характеристик

Отчет по преддипломной практике

студента 5 курса 9 группы

Научный руководитель:
старший преподаватель
Егор Валентинович Вечерко

Минск, 2017

1 Введение

Стеганография представляет собой специфическую область человеческой деятельности, связанной с разработкой и анализом методов сокрытия факта передачи информации. Подобно криптографии, стеганография известна со времен античности. Но на этом аналогии, по крайне мере в контексте теоретических исследований, заканчиваются. За последние четверть века возникла и успешно раз- вивается новая математическая дисциплина криптология, или, что то же самое, математическая криптография, изучающая математические модели криптографических схем. Попытки создания ма- тематической стеганографии (которую, быть может, следует именовать также стеганологией) пред- принимаются, но исследования здесь находятся лишь в зачаточном состоянии.

Такое положение дел обусловлено прежде всего сложностью возникающих в стеганографии задач. Всякая попытка построения математических моделей стеганографических систем сопряжена с необходимостью рассмотрения большого количества случаев и подслучаев, не допускающих простой и единообразной трактовки. Другими словами, внешняя среда, в которой должны функционировать стеганографические системы, имеет гораздо большее, по сравнению с внешней средой криптографических схем, количество степеней свободы.

Основными критериями для оценки и сравнения различных методов построения стеганографических систем являются их стойкость и емкость. В отличие от достаточно исследованных криптографических систем, оценки стойкости стегосистем более сложны и само понятие стойкости имеет большое число различных формулировок, что объясняется разнообразием задач стеганографи-ческой защиты данных. В настоящей работе исследуются методы построения стегосистем, предназначенных для скрытия факта передачи конфиденциальных сообщений. Говоря о стойкости криптографических систем, важно упомянуть о принципе Керкхоффса, который заключается в том, что система защиты информации должна обеспечивать свои функции даже при полной информированности противника о ее структуре и алгоритмах, и вся секретность системы должна заключаться в ключе. Этот принцип также можно соотнести с определением стойкости стегосистем. В данном случае, ключом может являться, например, секретная последовательность, определяющая порядок прохода, элементов контейнера при внедрении бит информации, что имеет место в алгоритмах рассеянного заполнения контейнеров. Второй критерий, емкость метода, определяет максимальное количество встраиваемой информации, и может выражаться в единицах бит на пиксель.

2 Основные понятия

Определение 2.1. *Стеганография - наука о способах передачи (хранения) скрытой информации, при которых скрытый канал организуется на базе и внутри открытого канала с использованием особенностей восприятия информации, причем для этой цели могут использоваться такие приемы, [2]как:*

1. *Полное сокрытие факта существования скрытого канала связи*
2. *Создание трудностй для обнаружения, извлечения и модификации передаваемых скрытых сообщений внутри открытых сообщений-контейнеров*
3. *Маскировки скрытой информации в протоколе*

Определение 2.2. *Контейнером $b \in B$ (носителем) называют несекретные данные, которые используют для сокрытия сообщения [2].*

Определение 2.3. *Сообщение $m \in M$ - секретная информация, наличие которой в контейнере необходимо скрыть [2].*

Определение 2.4. *Ключ $k \in K$ - секретная информация, известная только законному пользователю, которая определяет конкретный вид алгоритма сокрытия [2].*

Определение 2.5. *Пустой контейнер - контейнер, не содержащий сообщения [2].*

Определение 2.6. *Заполненный контейнер - контейнер, с внедренным в него сообщением [2].*

Определение 2.7. *Стеганографический алгоритм - два преобразования: прямое стеганографическое преобразование $F : M \times B \times K \to B$ и обратное стеганографическое преобразование $F^{-1} : B \times K \to M$, сопоставляющее соответственно тройке (сообщение, пустой контейнер, ключ) контейнер-результат и паре (заполненный контейнер, ключ) - исходное сообщение, [2]причем:*

$$F(m, b, k) = b_{m,k}, F^{-1}(b_{m,k}, k) = m, m \in M, b_{m,k}, b \in B, k \in K \quad (1)$$

Определение 2.8. *Под стеганографической системой будем понимать $S = (F, F^{-1}, M, B, K)$, представляющую собой совокупность сообщений, секретных ключей, контейнеров и связывающих их преобразований [2].*

Определение 2.9. *Внедрение (сокрытие) - применение прямого стегонаграфического преобразования к конкретным контейнеру, ключу и сообщению [2].*

Определение 2.10. *Извлечение сообщения - применение обатного стегонаграфического преобразованияю [2].*

Определение 2.11. *l-грамм - подпоследовательность из l подряд иущих элементов последовательности.*

Определение 2.12. *Под дискретным источником сообщений будем понимать устройство, порождающее последовательности, составленные из букв конечного алфавита A ($|A|=n<\infty$). При этом буквы последовательностей пораждаются в дискретный момент времени:$t = 0, 1, 2, ...; t = ..., -2, -1, 0, 1, 2, ...$; [1]*

Определение 2.13. *Если вероятность того, что источник порождает некоторую последовательность $a_{i_1}...a_{i_l}$, составленную из букв алфавита A, в момент времени 1, 2, ... l, равна вероятности того, что порождается точно такая же последовательность в момент времени j+1, ..., j+l для любых j,l;$a_{j_1}...a_{j_l}$, то источник называется стационарным.[1] Стационарность означает неизменность во времени всех конечномерных распределений соответствующего случайного процесса.*

Определение 2.14. *Энтропией источника назовем величину:*

$$H_\infty = \lim_{l\to\infty} \frac{H(C_l)}{l}; \tag{2}$$

если данный предел существует[1].

3 Теоретико-информационный подход к оценке стойкости систем

В настоящем разделе рассматривается теоретико-информационный подход к определению стойкости стеганосистем для стеганографических каналов без повторений в присутствии пассивного противника, обладающего неограниченными вычислительными возможностями.

В основе всех известных определений стойкости стеганосистем лежит требование неотличимости распределения вероятностей на множестве стего от распределения вероятностей на множестве пустых контейнеров. Рассматривается статистическая неотличимость, или, иначе говоря, неотличимость относительно произвольных алгоритмов.

Парадигма неотличимости распределений вероятностей заимствована из математической криптографии. Заметим, однако, что ее адекватность для стеганографии не очевидна. По крайней мере, в случае стеганографического канала без повторений не ясно, насколько оправданными будут усилия отправителя по имитации распределения вероятностей на множестве пустых контейнеров. Не следует ли вместо этого стремиться передать скрытое сообщение в одном из наиболее вероятных контейнеров?

Вполне очевидна идея создания стеганографического канала путем маскировки скрытого сообщения под шум, вносимый алгоритмом шифрования в исходный контейнер.

Проверка гипотезы о стойкости системы состоит в том, чтобы определить, какая из двух гипотез - H_0 или H_1 - является верным толкованием наблюдаемой величины Q. Есть два возможных распределения вероятностей, которые принято обозначать P_{Q_0}, P_{Q_1}, над пространством возможных наблюдений. Если верна гипотеза H_0, тогда Q была порождена согласно P_{Q_0}, если же верна гипотеза H_1, тогда Q была порождена согласно P_{Q_1}. Правило принятия решения - это двоичное отображение, заданное на пространстве возможных наблюдений, которое составляет одну из двух возможных гипотез для каждого возможного элемента q. Основной мерой проверки гипотезы является относительная энтропия или различие между двумя распределениями вероятности P_{Q_0} и P_{Q_1}, определяемое следующим выражением:

$$H(P_{Q_0}||P_{Q_1}) = \sum_q P_{Q_0}(q) log \frac{P_{Q_0}(q)}{P_{Q_1}(q)}; \tag{3}$$

Относительная энтропия между двумя распределениями всегда неотрицательна и равна нулю только тогда, когда распределения равномерны. Несмотря на то, что остиельная энтропия не является метрикой с точки зрения математики (так как не симметрична и не удовлетворяет аксиоме треугольника), полезно считать ее таковой. Двоичная относительная энтропия $d(\alpha, \beta)$ определяется как:

$$d(\alpha, \beta) = \alpha \log \frac{\alpha}{1-\beta} + (1-\alpha) \log \frac{1-\alpha}{\beta}; \tag{4}$$

где α - вероятность ошибки первого рода, β - вероятность ошибки второго рода.

4 Математическая модель вкраплений на основе схемы независимых испытаний

Контейнер представляет собой последовательность случайных величин распределенных по закону Бернулли с параметром p:

$$\mathcal{L}x_t = Bi(1,p), x_i \in V = 0,1, i = \overline{1,T}; \tag{5}$$

Вкрапляемое сообщение имеет вид:

$$\mathcal{L}m_t = Bi(1,\theta), m_i \in V = 0,1, i = \overline{1,\tau}; \tag{6}$$

Ключ γ_t определяет момент времени вкрапления i-того бита сообщения в исходный контейнер:

$$\mathcal{L}\gamma_t = Bi(1,\delta), \gamma_i \in V = 0,1, i = \overline{1,T}; \tag{7}$$

Вкрапление битов m_t производится по правилу, заданному следующим функциональным преобразованием:

$$y_t = (1-\gamma_t)x_t + \gamma_t m_{\tau_t}; \tag{8}$$

Лемма 4.1. *Для модели (5)-(8)*

$$P\{y_t = 1\} = (1-\delta)p + \delta\theta; \tag{9}$$

$$P\{y_t = 0\} = (1-\delta)(1-p) + \delta(1-\theta); \tag{10}$$

Доказательство. Воспользуемся формулой полной вероятности:
$P\{y_t = 1\} = P\{(1-\gamma_t)x_t + \gamma_t m_{\tau_t} = 1\} = \sum_{j\in V} P\{y_t = 1, \gamma_t = j\} = \sum_{j\in V} P\{\gamma_t = j\}P\{y_t = 1|\gamma_t = j\} = (1-\delta)P\{x_=1, \gamma_t = 0\} + \delta P\{m_{\tau_t} = 1, \gamma_t = 1\} = (1-\delta)p + \delta\theta;$
Тогда:
$P\{y_t = 0\} = 1 - P\{y_t = 1\} = (1-\delta)(1-p) + \delta(1-\theta);$ □

$$h = \frac{H(y_1,...,y_t)}{T} = \frac{TH(y_1)}{T} = H(y_1); \tag{11}$$

Воспользуемся леммой 4.1:
$h = -P\{y_t = 1\}\log_2 P(y_t = 1) - P\{y_t = 0\}\log_2 P(y_t = 0) = -((1-\delta)p + \delta\theta)\log_2((1-\delta)p + \delta\theta) - ((1-\delta)(1-p) + \delta(1-\theta))\log_2((1-\delta)(1-p) + \delta(1-\theta)).$

5 Математическая модель вкраплений в двоичную стационарную марковскую последовательность 1-го порядка и ее свойства

Рассмотрим модель (5)-(8).

Пусть контейнер (5) пердставляет собой цепь Маркова 1-го порядка с вектором распределения вероятностей $\pi = (\frac{1}{2}, \frac{1}{2})$ и матрицей вероятностей одношаговых переходов

$$P(\varepsilon) = \frac{1}{2}\begin{pmatrix}1+\varepsilon, 1-\varepsilon\\ 1-\varepsilon, 1+\varepsilon\end{pmatrix}, |\varepsilon| < 1, \varepsilon \neq 0. \tag{12}$$

Лемма 5.1. *Для модели (5)-(8) с условием (12):*

$$P\{y_{t-1} = 1, y_t = 1\} = \frac{1}{4}(1+\varepsilon)(1-\delta)^2 + \theta\delta(1-\delta) + \theta^2\delta^2; \tag{13}$$

$$P\{y_{t-1} = 1, y_t = 0\} = \frac{1}{4}(1-\varepsilon)(1-\delta)^2 + \frac{1}{2}\delta(1-\delta) + \theta(1-\theta)\delta^2; \tag{14}$$

$$P\{y_{t-1} = 0, y_t = 1\} = \frac{1}{4}(1-\varepsilon)(1-\delta)^2 + \frac{1}{2}\delta(1-\delta) + \theta(1-\theta)\delta^2; \tag{15}$$

$$P\{y_{t-1} = 0, y_t = 0\} = \frac{1}{4}(1+\varepsilon)(1-\delta)^2 + \delta(1-\theta)(1-\delta) + \delta^2(1-\theta)^2. \tag{16}$$

Доказательство. Рассмотрим биграмм: $\{y_{t-1}, y_t\}$
$(a_1, a_2) \in \{0,1\}, P\{y_{t-1} = a_1, y_t = a_2\} = \sum_{(b_1,b_2)\in\{0,1\}^2} P\{y_{t-1} = b_1, y_t = b_2, \gamma_{t-1} = a_1, \gamma_t = a_2\} = \sum_{(b_1,b_2)\in\{0,1\}^2} P\{y_{t-1} = b_1, y_t = b_2 | \gamma_{t-1} = a_1, \gamma_t = a_2\} P\{\gamma_{t-1} = a_1, \gamma_t = a_2\}$.
Для (13):
$\sum_{(b_1,b_2)\in\{0,1\}^2} P\{y_{t-1} = b_1, y_t = b_2 | \gamma_{t-1} = a_1, \gamma_t = a_2\} P\{\gamma_{t-1} = a_1, \gamma_t = a_2\} = \frac{1}{2}\cdot\frac{1}{2}(1+\varepsilon)(1-\delta)^2 + \theta\delta(1-\theta) + \theta^2\delta^2$.
Для случаев (14)-(16) доказывается аналогично. □

Для формул (13)-(16) справедливо условие нормировки:
$\sum_{(a_1,a_2)\in\{0,1\}^2} P\{y_{t-1} = a_1, y_t = a_2\} = 1$.
Далее полагаем, что $\theta = \frac{1}{2}$.
Тогда:

$$P\{y_{t-1} = 1, y_t = 1\} = P\{y_{t-1} = 0, y_t = 0\} = \delta^2\frac{\varepsilon}{4} - \delta\frac{\varepsilon}{2} + \frac{1+\varepsilon}{4}; \tag{17}$$

$$P\{y_{t-1} = 1, y_t = 0\} = P\{y_{t-1} = 0, y_t = 1\} = -\delta^2\frac{\varepsilon}{4} + \delta\frac{\varepsilon}{2} + \frac{1-\varepsilon}{4}. \tag{18}$$

Определение 5.1. *[1] Дискретный стационарный источник называется марковским источником порядка m, если для любого l(l>m) и любой последовательности $c_l = (a_{i_1}, ..., a_{i_l})$ выполняется: $P\{a_{i_l} | a_{i_{l-1}}, ..., a_{i_1}\} = P\{a_{i_l} | a_{i_{l-1}}, ..., a_{i_{l-m+1}}\}$.*

Определение 5.2. *[1] Величина: $H^{(k)} = \sum_{\{c_k\}} P\{a_{i_1}, ..., a_{i_k}\} \log P\{a_{i_k} | a_{i_{k-1}} ..., a_{i_1}\}$ называется шаговой энтропией марковского источника порядка k.*

Введем понятие энтропии на знак для l-граммы:

$$H_l(\delta) = -\frac{1}{l}\sum_{(a_1,...,a_l)\in\{0,1\}} P\{y_{t-l} = a_1, ..., y_{t-1} = a_l\} \log P\{y_{t-l} = a_1, ..., y_{t-1} = a_l\}. \tag{19}$$

При $\delta = 0$ стегоконтейнер y совпадает с контейнером x, тогда:

$$H_l(0) = -\frac{1}{l}(H\{x_1\} + (l-1)H\{x_2|x_1\}); \tag{20}$$

$$\lim_{l\to\infty} H_l(0) = \lim_{l\to\infty} -\frac{1}{l}(H\{x_1\} + (l-1)H\{x_2|x_1\}) = H\{x_2|x_1\}. \tag{21}$$

Рассмотрим случайную величину $\xi \in B = \{b_1, ..., b_m\}$, заданную на вероятностном пространстве $(\Omega, \mathcal{F}, \mathcal{P})$, $P\{\xi = b_i\} = p_i$.

Определение 5.3. *[1] Величина*

$$I\{b_i\} = -\log p_i \tag{22}$$

называется собственной информацией, содержащейся в исходе $b_i \in B$.

Велиина $I\{b_i\}$ изменяется от нуля в случае реализации достоверного исхода до бесконечности, когда $p(b_i) = p_i \to 0$. Величину $I\{b_i\}$ можно интерпретировать как априорную неопределённость события $\{\xi = b_i\}$.

Случайная величина $I\{\xi\}$ имеет математическое ожидание

$$EI\{\xi\} = -\sum_{b_i\in B} p_i \log p_i. \tag{23}$$

Определение 5.4. *Величина $EI\{\xi\}$ называется средней собственной информацией.*

Средняя собственная информация равна энтропии: $EI\{\xi\} = H\{\xi\}$.

Для представления функции логарифма воспользуемся формулой Маклорена первого порядка:

$$f(\delta) = f(\delta_0) + (\delta - \delta_0)f'(\delta_0) + o((\delta - \delta_0)^2). \tag{24}$$

Для краткости, в обозначении функции log будем использовать основание b.

Согласно (24) в точке $\delta = 0$ справедливо асимптотическое разложение 1-го порядка:

$$\log_b(a_0(\varepsilon) + \delta a_1(\varepsilon) + \delta^2 a_2(\varepsilon) + o(\delta^2)) =$$
$$= \log a_0(\varepsilon) + \delta\left(\log(a_0(\varepsilon) + \delta a_1(\varepsilon) + \delta^2 a_2(\varepsilon) + o(\delta^2))\right)'|_{\delta=0} + o(\delta) =$$
$$= \log a_0(\varepsilon) + \delta\frac{1}{\ln b}\frac{a_1(\varepsilon)}{a_0(\varepsilon)} + o(\delta).$$

Учитывая (24), найдем асимптотические выражения при $\delta \to 0$ для собственной информации $I\{y_{t-1} = i_1, y_t = i_2\}, i_1, i_2 \in \{0, 1\}$:

$$I\{y_{t-1} = 0, y_t = 0\} = I\{y_{t-1} = 1, y_t = 1\} = -\log(\frac{1}{4}(1+\varepsilon)(1-\delta)^2 + \frac{1}{2}\delta(1-\delta) + \frac{1}{4}\delta^2) =$$
$$= -\log\frac{1+\varepsilon}{4} + \delta\frac{1}{\ln b}\frac{2\varepsilon}{1+\varepsilon} + o(\delta);$$
$$I\{y_{t-1} = 0, y_t = 1\} = I\{y_{t-1} = 1, y_t = 0\} = -\log(\frac{1}{4}(1-\varepsilon)(1-\delta)^2 + \frac{1}{2}\delta(1-\delta) + \frac{1}{4})\delta^2) =$$
$$= -\log\frac{1-\varepsilon}{4} - \delta\frac{1}{\ln b}\frac{2\varepsilon}{1-\varepsilon} + o(\delta).$$

Лемма 5.2. *Если имеет место монобитная модель вкраплений (5)-(8), то для энтропии при $l = 2$ справедливо асимптотическое разложение 1-го порядка*

$$H_2(\delta) = H_2(0) + 2\delta\varepsilon\log\frac{1+\varepsilon}{1-\varepsilon} + O(\delta^2). \tag{25}$$

Доказательство. $H_2(\delta) = -(P\{y_{t-1}=0, y_t=0\}\log(P\{y_{t-1}=0, y_t=0\}) + P\{y_{t-1}=0, y_t=1\}\log(P\{y_{t-1}=0, y_t=1\}) + P\{y_{t-1}=1, y_t=0\}\log(P\{y_{t-1}=1, y_t=0\}) + P\{y_{t-1}=1, y_t=1\}\log(P\{y_{t-1}=1, y_t=1\})) = -2(P\{y_{t-1}=0, y_t=0\}\log(P\{y_{t-1}=0, y_t=0\}) + P\{y_{t-1}=0, y_t=0\}\log(P\{y_{t-1}=0, y_t=1\})) = -2((\delta^2\frac{\varepsilon}{4} - \delta\frac{\varepsilon}{2} + \frac{1+\varepsilon}{4})(-\log(\frac{1+\varepsilon}{4}) + \delta\frac{1}{\ln b}\cdot\frac{2\varepsilon}{1+\varepsilon} + o(\delta^2)) + (-\delta^2\frac{\varepsilon}{4} + \delta\frac{\varepsilon}{2} + \frac{1-\varepsilon}{4})(-\log(\frac{1-\varepsilon}{4}) - \delta\frac{1}{\ln b}\cdot\frac{2\varepsilon}{1-\varepsilon} + o(\delta^2))) = \frac{1}{2}(-(1+\varepsilon)\log(\frac{1+\varepsilon}{4}) - (1-\varepsilon)\log(\frac{1-\varepsilon}{4}) + 2\delta\varepsilon\log(\frac{1+\varepsilon}{1-\varepsilon}))) + O(\delta^2) = H_2(0) + 2\delta\varepsilon\log(\frac{1+\varepsilon}{1-\varepsilon}))) + O(\delta^2)$. □

Определение 5.5. *[1] Величина* $\lim_{k\to\infty} H^{(k)} = \lim_{k\to\infty} H_k = H_\infty \geqslant 0$ *называется энтропий марковского источника, где* H_k *- энтропия на знак.*

Рассмотрим условные вероятности появления трехграммы $(0,0,0)$ при условии всевозможных стегоключей.

$P\{y_{i-1}=0, y_i=0, y_{i+1}=0\} = (1-\delta)^3 P\{y_{i-1}=0, y_i=0, y_{i+1}=0|\gamma_{i-1}=0, \gamma_i=0, \gamma_{i+1}=0\} +$
$+ \delta(1-\delta)^2(P\{y_{i-1}=0, y_i=0, y_{i+1}=0|\gamma_{i-1}=1, \gamma_i=0, \gamma_{i+1}=0\} + P\{y_{i-1}=0, y_i=0, y_{i+1}=0|\gamma_{i-1}=0, \gamma_i=1, \gamma_{i+1}=0\} + P\{y_{i-1}=0, y_i=0, y_{i+1}=0|\gamma_{i-1}=0, \gamma_i=0, \gamma_{i+1}=1\}) +$
$+ \delta^2(1-\delta)(P\{y_{i-1}=0, y_i=0, y_{i+1}=0|\gamma_{i-1}=1, \gamma_i=1, \gamma_{i+1}=0\} + P\{y_{i-1}=0, y_i=0, y_{i+1}=0|\gamma_{i-1}=0, \gamma_i=1, \gamma_{i+1}=1\} + P\{y_{i-1}=0, y_i=0, y_{i+1}=0|\gamma_{i-1}=1, \gamma_i=0, \gamma_{i+1}=1\}) +$
$+ \delta^3(P\{y_{i-1}=0, y_i=0, y_{i+1}=0|\gamma_{i-1}=1, \gamma_i=1, \gamma_{i+1}=1\})$.

$P\{y_{i-1}=0, y_i=0, y_{i+1}=0|\gamma_1=0, \gamma_2=0, \gamma_3=0\} = P\{x_{i-1}=0, x_i=0, x_{i+1}=0\} = P\{x_{i-1}=0\}P\{x_i=0, x_{i+1}=0|x_{i-1}=0\} = P\{x_{i-1}=0\}P\{x_i=0\}P\{x_{i+1}=0|x_i=0\} = \frac{1}{2}\cdot\frac{1}{2}(1+\varepsilon)\cdot\frac{1}{2}(1+\varepsilon) = \frac{1}{8}(1+\varepsilon)^2$;

$P\{y_{i-1}=0, y_i=0, y_{i+1}=0|\gamma_1=1, \gamma_2=0, \gamma_3=0\} = P\{\xi=0, x_i=0, x_{i+1}=0\} = P\{\xi=0\}P\{x_i=0, x_{i+1}=0\} = P\{\xi=0\}P\{x_i=0\}P\{x_{i+1}=0|x_i=0\} = \frac{1}{2}\cdot\frac{1}{2}(1+\varepsilon)\cdot\frac{1}{2} = \frac{1}{8}(1+\varepsilon)$;

$P\{y_{i-1}=0, y_i=0, y_{i+1}=0|\gamma_1=0, \gamma_2=1, \gamma_3=0\} = P\{x_{i-1}=0, \xi=0, x_{i+1}=0\} = P\{\xi=0\}P\{x_{i-1}=0, x_{i+1}=0\} = \frac{1}{2}\cdot\frac{1}{2}\cdot\frac{1}{2}(1+\varepsilon^2) = \frac{1}{8}(1+\varepsilon^2)$;

$P\{y_{i-1}=0, y_i=0, y_{i+1}=0|\gamma_1=0, \gamma_2=0, \gamma_3=1\} = P\{x_{i-1}=0, x_i=0, \xi=0\} = P\{\xi=0\}P\{x_{i-1}=0, x_i=0\} = \frac{1}{2}\cdot\frac{1}{2}\cdot\frac{1}{2}(1+\varepsilon) = \frac{1}{8}(1+\varepsilon)$;

$P\{y_{i-1}=0, y_i=0, y_{i+1}=0|\gamma_1=0, \gamma_2=1, \gamma_3=1\} = P\{y_{i-1}=0, y_i=0, y_{i+1}=0|\gamma_1=1, \gamma_2=0, \gamma_3=1\} = P\{y_{i-1}=0, y_i=0, y_{i+1}=0|\gamma_1=1, \gamma_2=1, \gamma_3=0\} = P\{y_{i-1}=0, y_i=0, y_{i+1}=0|\gamma_1=1, \gamma_2=1, \gamma_3=1\} = P\{\xi=0, \xi=0, \xi=0\} = P\{\xi=1\}P\{x_{i-1}=1\}P\{x_i=0\} = \frac{1}{2}\cdot\frac{1}{2}\cdot\frac{1}{2} = \frac{1}{8}$;

$P\{y_{i-1}=0, y_i=0, y_{i+1}=0\} = \frac{1}{8}\big(\varepsilon(\varepsilon+2)\delta^2 - 2\varepsilon(\varepsilon+1)\delta + (1+\varepsilon)^2\big)$.

Найдем вероятности появления всевозможных трехграмм в стегоконтейнере $\{y_t\}$:

$$P\{y_{i-1}=1, y_i=1, y_{i+1}=1\} = P\{y_{i-1}=0, y_i=0, y_{i+1}=0\} = \frac{1}{8}\big(\varepsilon(\varepsilon+2)\delta^2 - 2\varepsilon(\varepsilon+2)\delta + (1+\varepsilon)^2\big),$$
$$P\{y_{i-1}=1, y_i=1, y_{i+1}=0\} = P\{y_{i-1}=0, y_i=0, y_{i+1}=1\} = P\{y_{i-1}=0, y_i=1, y_{i+1}=1\} = P\{y_{i-1}=1, y_i=0, y_{i+1}=0\} = \frac{1}{8}\big(-\varepsilon^2\delta^2 + 2\varepsilon^2\delta - \varepsilon^2 + 1\big),$$
$$P\{y_{i-1}=1, y_i=0, y_{i+1}=1\} = P\{y_{i-1}=0, y_i=1, y_{i+1}=0\} = \frac{1}{8}\big(\varepsilon(\varepsilon-2)\delta^2 - 2\varepsilon(\varepsilon-2)\delta + (1-\varepsilon)^2\big).$$

Теорема 5.1. *Если имеет место монобитная модель вкраплений (5)-(8) , то для энтропии при $l = 3$ справедливо асимптотическое разложение 1-го порядка:*

$$H_3(\delta) = H_3(0) + 2\varepsilon\delta \log \frac{1+\varepsilon}{1-\varepsilon} + O(\delta^2); \tag{26}$$

собственная информация имеет вид:

$$I\{y_{i-1} = 0, y_i = 0, y_{i+1} = 0\} = -\log \frac{(1+\varepsilon)^2}{8} + \delta \frac{1}{\ln b} \cdot \frac{2\varepsilon^2 + 4\varepsilon}{(1+\varepsilon)^2} + O(\delta^2),$$

$$I\{y_{i-1} = 1, y_i = 0, y_{i+1} = 0\} = I\{y_{i-1} = 0, y_i = 0, y_{i+1} = 1\} =$$

$$= -\log \frac{1-\varepsilon^2}{8} - \delta \frac{1}{\ln b} \cdot \frac{2\varepsilon^2}{1-\varepsilon^2} + O(\delta^2),$$

$$I\{y_{i-1} = 0, y_i = 1, y_{i+1} = 0\} = -\log \frac{(1-\varepsilon)^2}{8} + \delta \frac{1}{\ln b} \cdot \frac{2\varepsilon^2 - 4\varepsilon}{(1-\varepsilon)^2} + O(\delta^2);$$

$I\{y_{i-1} = j_1, y_i = j_2, y_{i+1} = j_3\} = I\{y_{i-1} = 1 - j_1, y_i = 1 - j_2, y_{i+1} = 1 - j_3\},\ j_1, j_2, j_3 \in \{0, 1\}.$

Доказательство. Подставляя в (24) найденные выражения для вероятностей трехграмм, получим асимптотические выражения для собственной информации.

Используя выражения для собственной информации, получим:
$H_3(\delta) = -2(\frac{1}{8}\big(\varepsilon(\varepsilon+2)\delta^2 - 2\varepsilon(\varepsilon+2)\delta + (1+\varepsilon)^2\big)(\log(\frac{(1+\varepsilon)^2}{8}) + \delta\frac{1}{\ln b} \cdot \frac{-2\varepsilon^2-4\varepsilon}{(1+\varepsilon)^2}) + 2\frac{1}{8}\big(-\varepsilon^2\delta^2 + 2\varepsilon^2\delta - \varepsilon^2 + 1\big)(\log(\frac{1-\varepsilon^2}{8}) + \delta\frac{1}{\ln b} \cdot \frac{2\varepsilon^2}{1-\varepsilon^2}) + \frac{1}{8}\big(\varepsilon(\varepsilon-2)\delta^2 - 2\varepsilon(2+\varepsilon)\delta + (1-\varepsilon)^2\big)(\log(\frac{(1-\varepsilon)^2}{8}) + \delta\frac{1}{\ln b} \cdot \frac{-2\varepsilon^2+4\varepsilon}{(1-\varepsilon)^2})) + O(\delta^2) = -((1-\varepsilon)\log(1-\varepsilon) + (1+\varepsilon)\log(1+\varepsilon) + \log(\frac{1}{8}) + 2\varepsilon\delta\log\frac{1-\varepsilon}{1+\varepsilon}) + O(\delta^2) = H_3(0) - 2\varepsilon\delta\log\frac{1-\varepsilon}{1+\varepsilon} + O(\delta^2).$ □

Рассмотрим условные вероятности появления четырехграммы $(0,0,0,0)$ при условии всевозможных стегоключей.

$P\{y_{i-1} = 0, y_i = 0, y_{i+1} = 0, y_{i+2} = 0\} = (1-\delta)^4 P\{y_{i-1} = 0, y_i = 0, y_{i+1} = 0, y_{i+2} = 0|\gamma_{i-1} = 0, \gamma_i = 0, \gamma_{i+1} = 0, \gamma_{i+2} = 0\} + \delta(1-\delta)^3\Big(P\{y_{i-1} = 0, y_i = 0, y_{i+1} = 0, y_{i+2} = 0|\gamma_{i-1} = 0, \gamma_i = 0, \gamma_{i+1} = 0, \gamma_{i+2} = 1\} + P\{y_{i-1} = 0, y_i = 0, y_{i+1} = 0, y_{i+2} = 0|\gamma_{i-1} = 0, \gamma_i = 0, \gamma_{i+1} = 1, \gamma_{i+2} = 0\} + P\{y_{i-1} = 0, y_i = 0, y_{i+1} = 0, y_{i+2} = 0|\gamma_{i-1} = 0, \gamma_i = 1, \gamma_{i+1} = 0, \gamma_{i+2} = 0\} + P\{y_{i-1} = 0, y_i = 0, y_{i+1} = 0, y_{i+2} = 0|\gamma_{i-1} = 1, \gamma_i = 0, \gamma_{i+1} = 0, \gamma_{i+2} = 0\}\Big) + \delta^2(1-\delta)^2\Big(P\{y_{i-1} = 0, y_i = 0, y_{i+1} = 0, y_{i+2} = 0|\gamma_{i-1} = 0, \gamma_i = 0, \gamma_{i+1} = 1, \gamma_{i+2} = 1\} + P\{y_{i-1} = 0, y_i = 0, y_{i+1} = 0, y_{i+2} = 0|\gamma_{i-1} = 0, \gamma_i = 1, \gamma_{i+1} = 1, \gamma_{i+2} = 0\} + P\{y_{i-1} = 0, y_i = 0, y_{i+1} = 0, y_{i+2} = 0|\gamma_{i-1} = 1, \gamma_i = 1, \gamma_{i+1} = 0, \gamma_{i+2} = 0\} + P\{y_{i-1} = 0, y_i = 0, y_{i+1} = 0, y_{i+2} = 0|\gamma_{i-1} = 1, \gamma_i = 0, \gamma_{i+1} = 1, \gamma_{i+2} = 0\} + P\{y_{i-1} = 0, y_i = 0, y_{i+1} = 0, y_{i+2} = 0|\gamma_{i-1} = 0, \gamma_i = 1, \gamma_{i+1} = 0, \gamma_{i+2} = 1\} + P\{y_{i-1} = 0, y_i = 0, y_{i+1} = 0, y_{i+2} = 0|\gamma_{i-1} = 1, \gamma_i = 0, \gamma_{i+1} = 0, \gamma_{i+2} = 1\}\Big) + \delta^3(1-\delta)\Big(P\{y_{i-1} = 0, y_i = 0, y_{i+1} = 0, y_{i+2} = 0|\gamma_{i-1} = 0, \gamma_i = 1, \gamma_{i+1} = 1, \gamma_{i+2} = 1\} + P\{y_{i-1} = 0, y_i = 0, y_{i+1} = 0, y_{i+2} = 0|\gamma_{i-1} = 1, \gamma_i = 0, \gamma_{i+1} = 1, \gamma_{i+2} = 1\} + P\{y_{i-1} = 0, y_i = 0, y_{i+1} = 0, y_{i+2} = 0|\gamma_{i-1} = 1, \gamma_i = 1, \gamma_{i+1} = 0, \gamma_{i+2} = 1\} + P\{y_{i-1} = 0, y_i = 0, y_{i+1} = 0, y_{i+2} = 0|\gamma_{i-1} = 1, \gamma_i = 1, \gamma_{i+1} = 1, \gamma_{i+2} = 0\}\Big) + \delta^4(P\{y_{i-1} = 0, y_i = 0, y_{i+1} = 0, y_{i+2} = 0|\gamma_{i-1} = 1, \gamma_i = 1, \gamma_{i+1} = 1, \gamma_{i+2} = 1\}).$

$P\{y_{i-1} = 0, y_i = 0, y_{i+1} = 0, y_{i+2} = 0|\gamma_{i-1} = 0, \gamma_i = 0, \gamma_{i+1} = 0, \gamma_{i+2} = 0\} = \frac{(1+\varepsilon)^3}{16};$

$P\{y_{i-1} = 0, y_i = 0, y_{i+1} = 0, y_{i+2} = 0|\gamma_{i-1} = 1, \gamma_i = 0, \gamma_{i+1} = 0, \gamma_{i+2} = 0\} = P\{y_{i-1} = 0, y_i = 0, y_{i+1} = 0, y_{i+2} = 0|\gamma_{i-1} = 0, \gamma_i = 0, \gamma_{i+1} = 0, \gamma_{i+2} = 1\} = \frac{(1+\varepsilon)^2}{16};$

$P\{y_{i-1}=0,y_i=0,y_{i+1}=0,y_{i+2}=0|\gamma_{i-1}=0,\gamma_i=1,\gamma_{i+1}=0,\gamma_{i+2}=0\}=P\{y_{i-1}=0,y_i=0,y_{i+1}=0,y_{i+2}=0|\gamma_{i-1}=0,\gamma_i=0,\gamma_{i+1}=1,\gamma_{i+2}=0\}=\frac{(1+\varepsilon)(1+\varepsilon^2)}{16}$;

$P\{y_{i-1}=0,y_i=0,y_{i+1}=0,y_{i+2}=0|\gamma_{i-1}=1,\gamma_i=1,\gamma_{i+1}=0,\gamma_{i+2}=0\}=P\{y_{i-1}=0,y_i=0,y_{i+1}=0,y_{i+2}=0|\gamma_{i-1}=0,\gamma_i=0,\gamma_{i+1}=0,\gamma_{i+2}=1\}=P\{y_{i-1}=0,y_i=0,y_{i+1}=0,y_{i+2}=0|\gamma_{i-1}=1,\gamma_i=0,\gamma_{i+1}=0,\gamma_{i+2}=1\}=P\{y_{i-1}=0,y_i=0,y_{i+1}=0,y_{i+2}=0|\gamma_{i-1}=0,\gamma_i=0,\gamma_{i+1}=1,\gamma_{i+2}=1\}=\frac{1+\varepsilon}{16}$;

$P\{y_{i-1}=0,y_i=0,y_{i+1}=0,y_{i+2}=0|\gamma_{i-1}=0,\gamma_i=1,\gamma_{i+1}=0,\gamma_{i+2}=1\}=P\{y_{i-1}=0,y_i=0,y_{i+1}=0,y_{i+2}=0|\gamma_{i-1}=1,\gamma_i=0,\gamma_{i+1}=1,\gamma_{i+2}=0\}=\frac{1+\varepsilon^2}{16}$;

$P\{y_{i-1}=0,y_i=0,y_{i+1}=0,y_{i+2}=0|\gamma_{i-1}=1,\gamma_i=1,\gamma_{i+1}=1,\gamma_{i+2}=0\}=P\{y_{i-1}=0,y_i=0,y_{i+1}=0,y_{i+2}=0|\gamma_{i-1}=1,\gamma_i=1,\gamma_{i+1}=0,\gamma_{i+2}=1\}=P\{y_{i-1}=0,y_i=0,y_{i+1}=0,y_{i+2}=0|\gamma_{i-1}=1,\gamma_i=0,\gamma_{i+1}=1,\gamma_{i+2}=1\}=P\{y_{i-1}=0,y_i=0,y_{i+1}=0,y_{i+2}=0|\gamma_{i-1}=0,\gamma_i=1,\gamma_{i+1}=1,\gamma_{i+2}=1\}=\frac{1}{16}$;

$P\{y_{i-1}=0,y_i=0,y_{i+1}=0,y_{i+2}=0\}=P\{y_{i-1}=1,y_i=1,y_{i+1}=1,y_{i+2}=1\}=\frac{1}{16}\big(\delta^4\varepsilon^2+\delta^3(-4\varepsilon^2)+\delta^2(\varepsilon^3+8\varepsilon^2+3\varepsilon)+\delta(-2\varepsilon^3-8\varepsilon^2-6\varepsilon)+\varepsilon^3+3\varepsilon^2+3\varepsilon+1\big)$.
Найдем вероятности появления всевозможных четырехграмм в стегоконтейнере $\{y_t\}$:

$$
\begin{aligned}
&P\{y_{i-1}=0,y_i=0,y_{i+1}=0,y_{i+2}=1\}=P\{y_{i-1}=1,y_i=1,y_{i+1}=1,y_{i+2}=0\}=\\
&=P\{y_{i-1}=1,y_i=0,y_{i+1}=0,y_{i+2}=0\}=P\{y_{i-1}=0,y_i=1,y_{i+1}=1,y_{i+2}=1\}=\\
&=\tfrac{1}{16}\big(\delta^4(-\varepsilon^2)+\delta^3\cdot4\varepsilon^2+\delta^2(-\varepsilon^3-6\varepsilon^2+\varepsilon)+\delta(2\varepsilon^3+4\varepsilon^2-2\varepsilon)-\varepsilon^3-\varepsilon^2+\varepsilon+1\big);\\
&P\{y_{i-1}=0,y_i=1,y_{i+1}=0,y_{i+2}=0\}=P\{y_{i-1}=1,y_i=0,y_{i+1}=1,y_{i+2}=1\}=\\
&=P\{y_{i-1}=0,y_i=0,y_{i+1}=1,y_{i+2}=0\}=P\{y_{i-1}=1,y_i=1,y_{i+1}=0,y_{i+2}=1\}=\\
&=\tfrac{1}{16}\big(\delta^4(-\varepsilon^2)+\delta^3\cdot4\varepsilon^2+\delta^2(\varepsilon^3-6\varepsilon^2-\varepsilon)+\delta(-2\varepsilon^3+4\varepsilon^2+2\varepsilon)+\varepsilon^3-\varepsilon^2-\varepsilon+1\big);\\
&P\{y_{i-1}=0,y_i=0,y_{i+1}=1,y_{i+2}=1\}=P\{y_{i-1}=1,y_i=1,y_{i+1}=0,y_{i+2}=0\}=\\
&\quad=\tfrac{1}{16}\big(\delta^4\varepsilon^2+\delta^3(-4\varepsilon^2)+\delta^2(-\varepsilon^3+4\varepsilon^2+\varepsilon)+\delta(2\varepsilon^3-2\varepsilon)-\varepsilon^3-\varepsilon^2+\varepsilon+1\big);\\
&P\{y_{i-1}=0,y_i=1,y_{i+1}=1,y_{i+2}=0\}=P\{y_{i-1}=1,y_i=0,y_{i+1}=0,y_{i+2}=1\}=\\
&\quad=\tfrac{1}{16}\big(\delta^4\varepsilon^2+\delta^3(-4\varepsilon^2)+\delta^2(\varepsilon^3+4\varepsilon^2-\varepsilon)+\delta(-2\varepsilon^3+2\varepsilon)+\varepsilon^3-\varepsilon^2-\varepsilon+1\big);\\
&P\{y_{i-1}=1,y_i=0,y_{i+1}=1,y_{i+2}=0\}=P\{y_{i-1}=0,y_i=1,y_{i+1}=0,y_{i+2}=1\}=\\
&=\tfrac{1}{16}\big(\delta^4\varepsilon^2+\delta^3(-4\varepsilon^2)+\delta^2(-\varepsilon^3+8\varepsilon^2-3\varepsilon)+\delta(2\varepsilon^3-8\varepsilon^2+6\varepsilon)-\varepsilon^3+3\varepsilon^2-3\varepsilon+1\big).
\end{aligned}
$$

Теорема 5.2. *Если имеет место монобитная модель вкраплений (5)-(8) , то для энтропии при $l=4$ справедливо асимптотическое разложение 1-го порядка:*

$$H_4(\delta)=H_4(0)+\frac{24\varepsilon\delta}{16}\log\frac{1+\varepsilon}{1-\varepsilon}+O(\delta^2);\tag{27}$$

собственная информация имеет вид:

$$I\{y_{i-1}=0,y_i=0,y_{i+1}=0,y_{i+2}=0\}=I\{y_{i-1}=1,y_i=1,y_{i+1}=1,y_{i+2}=1\}=$$
$$=-\left(\log\frac{(1+\varepsilon)^3}{16}+\delta\frac{-2\varepsilon^3-8\varepsilon^2-6\varepsilon}{(1+\varepsilon)^3\ln b}\right)+O(\delta^2);$$
$$I\{y_{i-1}=0,y_i=0,y_{i+1}=0,y_{i+2}=1\}=I\{y_{i-1}=1,y_i=1,y_{i+1}=1,y_{i+2}=0\}=$$
$$=I\{y_{i-1}=1,y_i=0,y_{i+1}=0,y_{i+2}=0\}=I\{y_{i-1}=0,y_i=1,y_{i+1}=1,y_{i+2}=1\}=$$
$$=-\left(\log\frac{(1-\varepsilon)(1+\varepsilon)^2}{16}+\delta\frac{2\varepsilon^3+4\varepsilon^2-2\varepsilon}{(1-\varepsilon)(1+\varepsilon)^2\ln b}\right)+O(\delta^2);$$
$$I\{y_{i-1}=0,y_i=0,y_{i+1}=1,y_{i+2}=0\}=I\{y_{i-1}=1,y_i=1,y_{i+1}=0,y_{i+2}=1\}=$$
$$=I\{y_{i-1}=0,y_i=1,y_{i+1}=0,y_{i+2}=0\}=I\{y_{i-1}=1,y_i=0,y_{i+1}=1,y_{i+2}=1\}=$$
$$=-\left(\log\frac{(1-\varepsilon)^2(1+\varepsilon)}{16}+\delta\frac{-2\varepsilon^3+4\varepsilon^2+2\varepsilon}{(1-\varepsilon)^2(1+\varepsilon)\ln b}\right)+O(\delta^2);$$
$$I\{y_{i-1}=0,y_i=0,y_{i+1}=1,y_{i+2}=1\}=I\{y_{i-1}=1,y_i=1,y_{i+1}=0,y_{i+2}=0\}=$$
$$=-\left(\log\frac{(1-\varepsilon)(1+\varepsilon)^2}{16}+\delta\frac{2\varepsilon^3-2\varepsilon}{(1-\varepsilon)(1+\varepsilon)^2\ln b}\right)+O(\delta^2);$$
$$I\{y_{i-1}=0,y_i=1,y_{i+1}=1,y_{i+2}=0\}=I\{y_{i-1}=1,y_i=0,y_{i+1}=0,y_{i+2}=1\}=$$
$$=-\left(\log\frac{(1-\varepsilon)^2(1+\varepsilon)}{16}+\delta\frac{-2\varepsilon^3+2\varepsilon}{(1-\varepsilon)^2(1+\varepsilon)\ln b}\right)+O(\delta^2);$$
$$I\{y_{i-1}=1,y_i=0,y_{i+1}=1,y_{i+2}=0\}=I\{y_{i-1}=0,y_i=1,y_{i+1}=0,y_{i+2}=1\}=$$
$$=-\left(\log\frac{(1-\varepsilon)^3}{16}+\delta\frac{2\varepsilon^3-8\varepsilon^2+6\varepsilon}{(1-\varepsilon)^3\ln b}\right)+O(\delta^2).$$

Доказательство. Подставляя в (24) найденные выражения для вероятностей четырехграмм, получим асимптотические выражения для собственной информации.
Используя выражения для собственной информации, получим:

$$H_4(\delta)=-2\Big(\tfrac{1}{16}\big(\delta^4\varepsilon^2+\delta^3(-4\varepsilon^2)+\delta^2(\varepsilon^3+8\varepsilon^2+3\varepsilon)+\delta(-2\varepsilon^3-8\varepsilon^2-6\varepsilon)+\varepsilon^3+3\varepsilon^2+3\varepsilon+1\big)\big(\log\tfrac{(1+\varepsilon)^3}{16}+\delta\tfrac{-2\varepsilon^3-8\varepsilon^2-6\varepsilon}{(1+\varepsilon)^3\ln b}+O(\delta^2)\big)+2\tfrac{1}{16}\big(\delta^4(-\varepsilon^2)+\delta^3\cdot4\varepsilon^2+\delta^2(-\varepsilon^3-6\varepsilon^2+\varepsilon)+\delta(2\varepsilon^3+4\varepsilon^2-2\varepsilon)-\varepsilon^3-\varepsilon^2+\varepsilon+1\big)\big(\log\tfrac{(1-\varepsilon)(1+\varepsilon)^2}{16}+\delta\tfrac{2\varepsilon^3+4\varepsilon^2-2\varepsilon}{(1-\varepsilon)(1+\varepsilon)^2\ln b}+O(\delta^2)\big)+2\tfrac{1}{16}\big(\delta^4(-\varepsilon^2)+\delta^3\cdot4\varepsilon^2+\delta^2(\varepsilon^3-6\varepsilon^2-\varepsilon)+\delta(-2\varepsilon^3+4\varepsilon^2+2\varepsilon)+\varepsilon^3-\varepsilon^2-\varepsilon+1\big)\big(\log\tfrac{(1-\varepsilon)^2(1+\varepsilon)}{16}+\delta\tfrac{-2\varepsilon^3+4\varepsilon^2+2\varepsilon}{(1-\varepsilon)^2(1+\varepsilon)\ln b}+O(\delta^2)\big)+\tfrac{1}{16}\big(\delta^4\varepsilon^2+\delta^3(-4\varepsilon^2)+\delta^2(-\varepsilon^3+4\varepsilon^2+\varepsilon)+\delta(2\varepsilon^3-2\varepsilon)-\varepsilon^3-\varepsilon^2+\varepsilon+1\big)\big(log\tfrac{(1-\varepsilon)(1+\varepsilon)^2}{16}+\delta\tfrac{2\varepsilon^3-2\varepsilon}{(1-\varepsilon)(1+\varepsilon)^2\ln b}+O(\delta^2)\big)+\tfrac{1}{16}\big(\delta^4\varepsilon^2+\delta^3(-4\varepsilon^2)+\delta^2(\varepsilon^3+4\varepsilon^2-\varepsilon)+\delta(-2\varepsilon^3+2\varepsilon)+\varepsilon^3-\varepsilon^2-\varepsilon+1\big)\big(\log\tfrac{(1-\varepsilon)^2(1+\varepsilon)}{16}+\delta\tfrac{-2\varepsilon^3+2\varepsilon}{(1-\varepsilon)^2(1+\varepsilon)\ln b}+O(\delta^2)\big)+\tfrac{1}{16}\big(\delta^4\varepsilon^2+\delta^3(-4\varepsilon^2)+\delta^2(-\varepsilon^3+8\varepsilon^2-3\varepsilon)+\delta(2\varepsilon^3-8\varepsilon^2+6\varepsilon)-\varepsilon^3+3\varepsilon^2-3\varepsilon+1\big)\big(\log\tfrac{(1-\varepsilon)^3}{16}+\delta\tfrac{2\varepsilon^3-8\varepsilon^2+6\varepsilon}{(1-\varepsilon)^3\ln b}+O(\delta^2)\big)\Big)=\tfrac{(1+\varepsilon)^3}{16}\log\tfrac{(1+\varepsilon)^3}{16}+3\tfrac{(1+\varepsilon)^2(1-\varepsilon)}{16}\log\tfrac{(1+\varepsilon)^2(1-\varepsilon)}{16}+3\tfrac{(1+\varepsilon)(1-\varepsilon)^2}{16}\log\tfrac{(1+\varepsilon)(1-\varepsilon)^2}{16}+\tfrac{(1-\varepsilon)^3}{16}\log\tfrac{(1-\varepsilon)^3}{16}+\tfrac{24\varepsilon\delta}{16}\log\tfrac{1+\varepsilon}{1-\varepsilon}+O(\delta^2)=H_4(0)+\tfrac{24\varepsilon\delta}{16}\log\tfrac{1+\varepsilon}{1-\varepsilon}+O(\delta^2).\qquad\square$$

Оценим остаточный член для асимптотического выражения энтропии биграммы:

$$r_n(\delta)=\frac{f^{(n+1)}(\overline{\delta})}{(n+1)!}(\delta-\delta_0),\overline{\delta}\in[\delta_0,\delta].\tag{28}$$

Для асимптотического разложения 1-го порядка при $\delta_0=0$ остаточный член имеет вид:

$$r_n(\delta)=\frac{f''(\overline{\delta})}{2}\delta,\overline{\delta}\in[0,\delta].\tag{29}$$

$$(\log(P_{00}))'' = \frac{-2\varepsilon(\delta^2\varepsilon-2\delta\varepsilon+\varepsilon-1)}{\ln b(\delta^2\varepsilon-2\delta\varepsilon+\varepsilon+1)^2};$$
$$(\log(P_{01}))'' = \frac{-2\varepsilon(\delta^2\varepsilon-2\delta\varepsilon+\varepsilon+1)}{\ln b(\delta^2\varepsilon-2\delta\varepsilon+\varepsilon-1)^2};$$

Тогда остаточный член для $\log(P_{00}) = \log(P_{11})$ равен:

$$r_{n_{00}}(\delta) = \frac{\delta}{2} \cdot \frac{-2\varepsilon(\delta^2\varepsilon - 2\delta\varepsilon + \varepsilon - 1)}{\ln b(\delta^2\varepsilon - 2\delta\varepsilon + \varepsilon + 1)^2}|_{\delta=\overline{\delta}} \tag{30}$$

Остаточный член для $\log(P_{10}) = \log(P_{01})$ равен:

$$r_{n_{10}}(\delta) = \frac{\delta}{2} \cdot \frac{-2\varepsilon(\delta^2\varepsilon - 2\delta\varepsilon + \varepsilon + 1)}{\ln b(\delta^2\varepsilon - 2\delta\varepsilon + \varepsilon - 1)^2}|_{\delta=\overline{\delta}} \tag{31}$$

Тогда на основании (30) и (31) остаточный член для энтропиии будет равен:

$$R_n(\delta) = -2(P_{00}r_{n_{00}}(\delta) + P_{10}r_{n_{01}}(\delta)) = -2\big((\delta^2\frac{\varepsilon}{4} - \delta\frac{\varepsilon}{2} + \frac{1+\varepsilon}{4})(\frac{\delta}{2} \cdot \frac{-2\varepsilon(\delta^2\varepsilon - 2\delta\varepsilon + \varepsilon - 1)}{\ln b(\delta^2\varepsilon - 2\delta\varepsilon + \varepsilon + 1)^2}|_{\delta=\overline{\delta}}) +$$
$$(-\delta^2\frac{\varepsilon}{4} + \delta\frac{\varepsilon}{2} + \frac{1-\varepsilon}{4})(\frac{\delta}{2} \cdot \frac{-2\varepsilon(\delta^2\varepsilon - 2\delta\varepsilon + \varepsilon + 1)}{\ln b(\delta^2\varepsilon - 2\delta\varepsilon + \varepsilon - 1)^2}|_{\delta=\overline{\delta}})\big)$$

6 Линейный дискриминантный анализ

Линейный дискриминантный анализ (ЛДА), а также связанный с ним линейный дискриминант Фишера — методы статистики и машинного обучения, применяемые для нахождения линейных комбинаций признаков, наилучшим образом разделяющих два или более класса объектов или событий. Полученная комбинация может быть использована в качестве линейного классификатора или для сокращения размерности пространства признаков перед последующей классификацией. ЛДА тесно связан с дисперсионным анализом и регрессионным анализом, также пытающимися выразить какую-либо зависимую переменную через линейную комбинацию других признаков или измерений. В этих двух методах зависимая переменная — численная величина, а в ЛДА она является величиной номинальной (меткой класса). Помимо того, ЛДА имеет схожие черты с методом главных компонент и факторным анализом, которые ищут линейные комбинации величин, наилучшим образом описывающие данные. Для использования ЛДА признаки должны быть непрерывными величинами, иначе следует использовать анализ соответствий (англ. Discriminant Correspondece Analysis).

6.1 Линейный дискриминантный анализ для случая двух классов

Для каждого образца объекта или события с известным классом y рассматривается набор наблюдений x (называемых ещё признаками, переменными или измерениями). Набор таких образцов называется обучающей выборкой (или набором обучения, обучением). Задачи классификации состоит в том, чтобы построить хороший прогноз класса y для всякого так же распределённого объекта (не обязательно содержащегося в обучающей выборке), имея только наблюдения x.
При ЛДА предполагается, что функции совместной плотности распределения вероятностей $p(\vec{x}|y=1)$ и $p(\vec{x}|y=0)$ - нормальны. В этих предположениях оптимальное байесовское решение - относить точки ко второму классу если отношение правдоподобия ниже некоторого порогового значения T:

$$(\vec{x}-\vec{\mu}_0)^T\Sigma_{y=0}^{-1}(\vec{x}-\vec{\mu}_0)+\ln|\Sigma_{y=0}|-(\vec{x}-\vec{\mu}_1)^T\Sigma_{y=1}^{-1}(\vec{x}-\vec{\mu}_1)-\ln|\Sigma_{y=0}|<T$$

Если не делается никаких дальнейших предположений, полученную задачу классификации называют квадратичным дискриминантным анализом (англ. quadratic discriminant analysis, QDA). В ЛДА делается дополнительное предположение о гомоскедастичности (т.е. предполагается, что ковариационные матрицы равны, $\Sigma_{y=0}=\Sigma_{y=1}=\Sigma$) и считается, что ковариационные матрицы имеют полный ранг. При этих предположениях задача упрощается и сводится к сравнению скалярного произведения с пороговым значением

$$\vec{\omega}\cdot\vec{x}<c$$

для некоторой константы c, где

$$\vec{\omega}=\Sigma^{-1}(\vec{\mu_1}-\vec{\mu_0}).$$

Это означает, что вероятность принадлежности нового наблюдения x к классу y зависит исключительно от линейной комбинации известных наблюдений.

6.2 Результаты линейного дискриминантного анализа

Линейный дискриминантный анализ применен для классификации последовательностей с вкраплениями и без вкраплений при фиксированном параметре ε.
Пусть имеется последовательность $Y=\{y_1,...,y_T\}$, на основании Y вычисляем $(H_3(\delta),H_4(\delta))$

при фиксированном $\varepsilon = 0.55$, тогда:
H_0: последовательность Y имеет вкрапления
H_1: последовательность Y не имеет вкраплений
тогда для $n = n_0 + n_1$, (где n_0 - количество заведомо пустых последовательностей, n_1 - количество последовательностей с вкрапениями) последовательностей можно провести дискриминантный анализ и оценить вероятновсть правильной классификации и мощность критерия.
Тогда вероятности ошибок первого и второго рода:

$$\alpha = \frac{n_0 - \nu_0}{n_0} - ; \tag{32}$$

$$\beta = \frac{n_1 - \nu_1}{n_1} - ; \tag{33}$$

ν_0 - количество верно отпределенных пустых последоваательностей, ν_1 - количество верно отпределенных последоваательностей с вкраплениями.
Мощность критерия:

$$w = \frac{\nu_1}{n_1} \tag{34}$$

δ	α	β	Мощность критерия
0.07	21%	14%	0.86
0.08	10%	13%	0.87
0.09	14%	8%	0.92
0.1	13%	5%	0.95

Таблица 1: Результаты дискриминантного анализа.

6.3 Вывод

Методом линейного дискриминантного анализа на основании энтропийных характеристик $(H_3(\delta), H_4(\delta))$ можно определить наличие вкраплений в последовательность с вероятностью ошибки второго рода 5% при доле вкраплений $\delta \geq 0.1$.

7 Исследование реальных данных на основании изображений в формате JPEG

Формат файла JPEG (Joint Photographic Experts Group - Объединенная экспертная группа по фотографии, произносится "джейпег) был разработан компанией C-Cube Microsystems как эффективный метод хранения изображений с большой глубиной цвета, например, получаемых при сканировании фотографий с многочисленными едва уловимыми (а иногда и неуловимыми) оттенками цвета. Самое большое отличие формата JPEG от других рассмотренных здесь форматов состоит в том, что в JPEG используется алгоритм сжатия с потерями (а не алгоритм без потерь) информации. Алгоритм сжатия без потерь так сохраняет информацию об изображении, что распакованное изображение в точности соответствует оригиналу. При сжатии с потерями приносится в жертву часть информации об изображении, чтобы достичь большего коэффициента сжатия. Распакованное изображение JPEG редко соответствует оригиналу абсолютно точно, но очень часто эти различия столь незначительны, что их едва можно (если вообще можно) обнаружить.

Процесс сжатия изображения JPEG достаточно сложен и часто для достижения приемлемой производительности требует специальной аппаратуры. Вначале изображение разбивается на квадратные блоки со стороной размером 8 пиксел. Затем производится сжатие каждого блока отдельно за три шага. На первом шаге с помощью формулы дискретного косинусоидального преобразования (DCT) производится преобразование блока 8x8 с информацией о пикселах в матрицу 8x8 амплитудных значений, отражающих различные частоты (скорости изменения цвета) в изображении. На втором шаге значения матрицы амплитуд делятся на значения матрицы квантования, которая смещена так, чтобы отфильтровать амплитуды, незначительно влияющие на общий вид изображения. На третьем и последнем шаге квантованная матрица амплитуд сжимается с использованием алгоритма сжатия без потерь.

Поскольку в квантованной матрице отсутствует значительная доля высокочастотной информации, имеющейся в исходной матрице, первая часто сжимается до половины своего первоначального размера или даже еще больше. Реальные фотографические изображения часто совсем невозможно сжать с помощью методов сжатия без потерь, поэтому 50%-ное сжатие следует признать достаточно хорошим. С другой стороны, применяя методы сжатия без потерь, можно сжимать некоторые изображения на 90%. Такие изображения плохо подходят для сжатия методом JPEG.

При сжатии методом JPEG потери информации происходят на втором шаге процесса. Чем больше значения в матрице квантования, тем больше отбрасывается информации из изображения и тем более плотно сжимается изображение. Компромисс состоит в том, что более высокие значения квантования приводят к худшему качеству изображения. При формировании изображения JPEG пользователь устанавливает показатель качества, величине которого "управляет"значениями матрицы квантования. Оптимальные показатели качества, обеспечивающие лучший баланс между коэффициентом сжатия и качеством изображения, различны для разных изображений и обычно могут быть найдены только методом проб и ошибок.

В качестве контейнера рассматривалась последовательность младших бит ДКП коэффициентов jpeg изображения.
Пусть $X = x_1, ..., x_n$ последовательность младших бит коэффициентов ДКП изображения.

$$\hat{p}\{x = 1\} = \frac{\sum_{i=1}^{n} x_i}{n} \tag{35}$$

На основании экспериментального исследования некоторой базы изображений и формате jpeg можно сделать вывод, что экпериментальная оценка вероятности появления единич-

ного бита находится в интервале [0.350705572829466, 0.480247228960331] для исследуемой базы изображений. Следовательно младшие биты ДКП коэффициентов изображений формата jpeg не являются равномерно распределенными. Поэтому изображения в формате jpeg не являются подходящим контейнером для исследования изложеной теории обнаружения вкраплений.

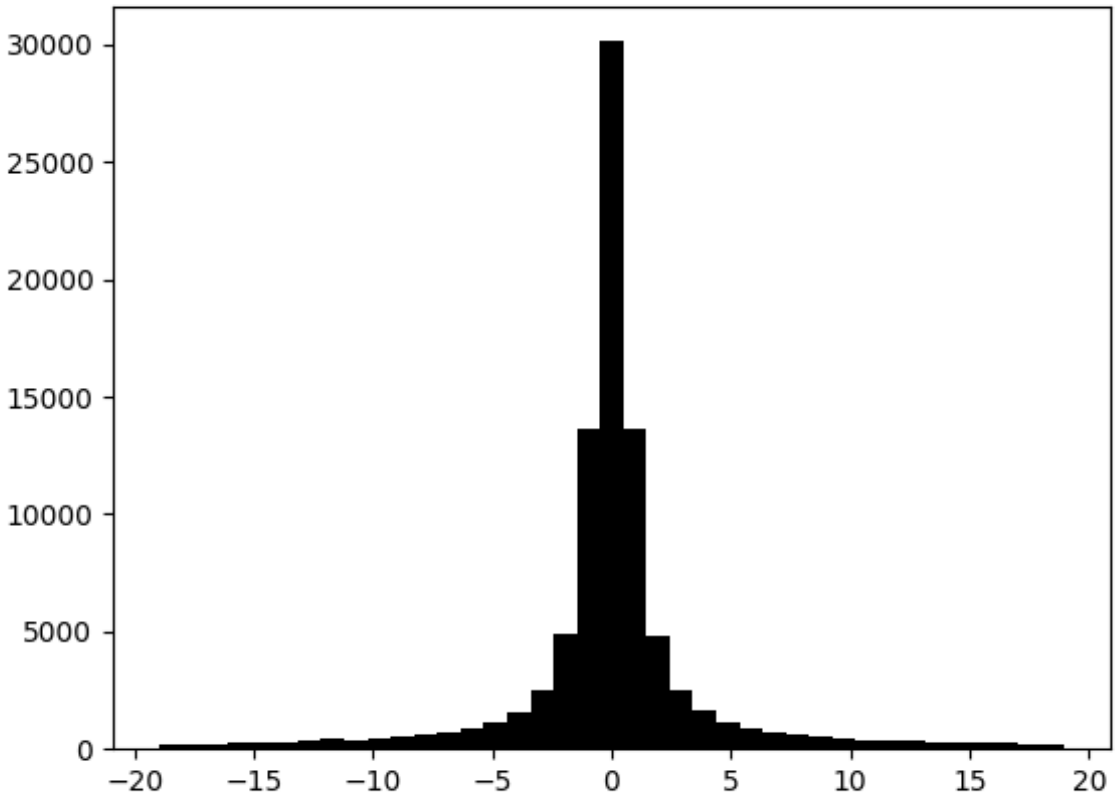

Рис. 1: Гистограмма ДКП коэффициентов jpeg изображения

8 Компьютерные эксперименты

Экспериментально построим $H_2(\delta)$ и $H_2(0)$

Рис. 2: График зависимости энтропии $H_2(\delta)$ от длины последовательности при $\delta = 0.1$

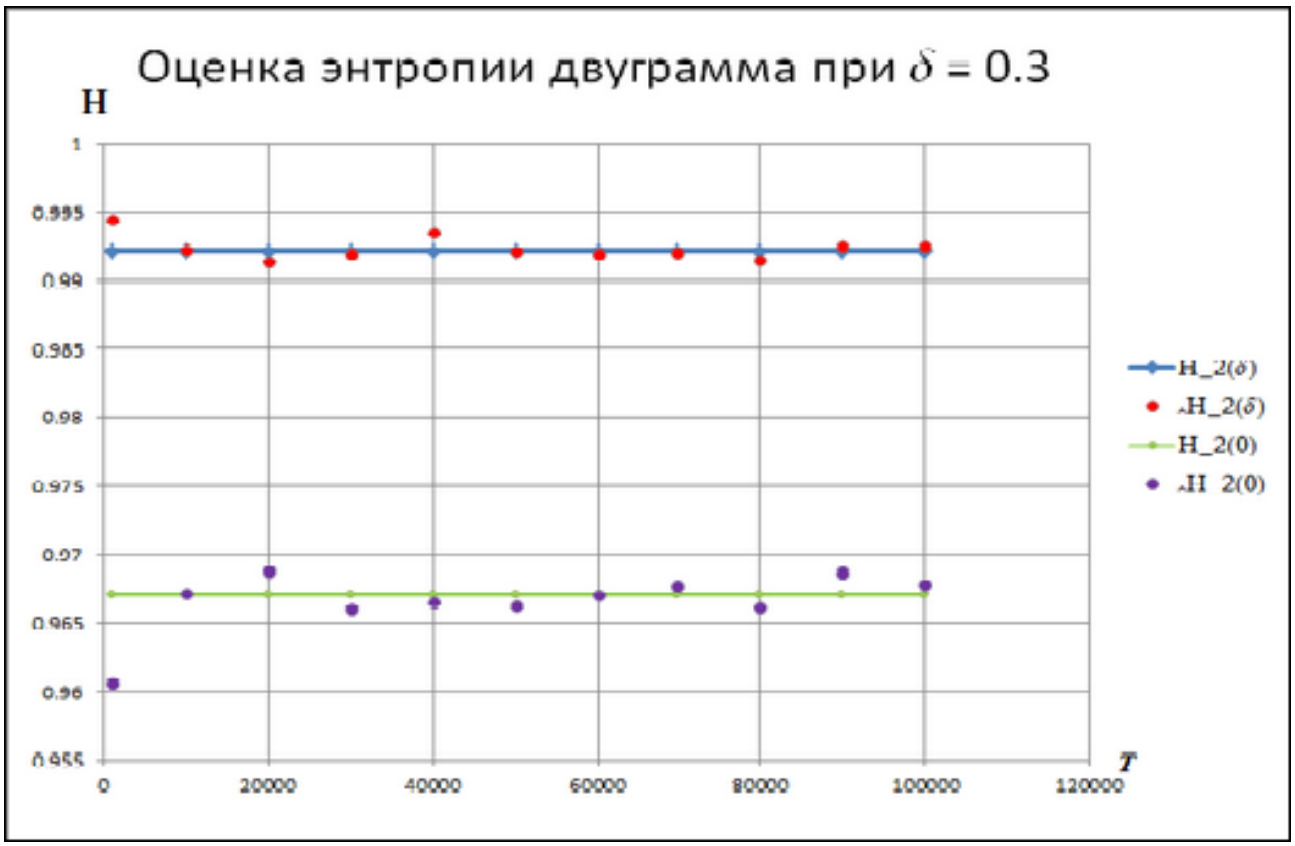

Рис. 3: График зависимости энтропии $H_2(\delta)$ от длины последовательности при $\delta = 0.3$

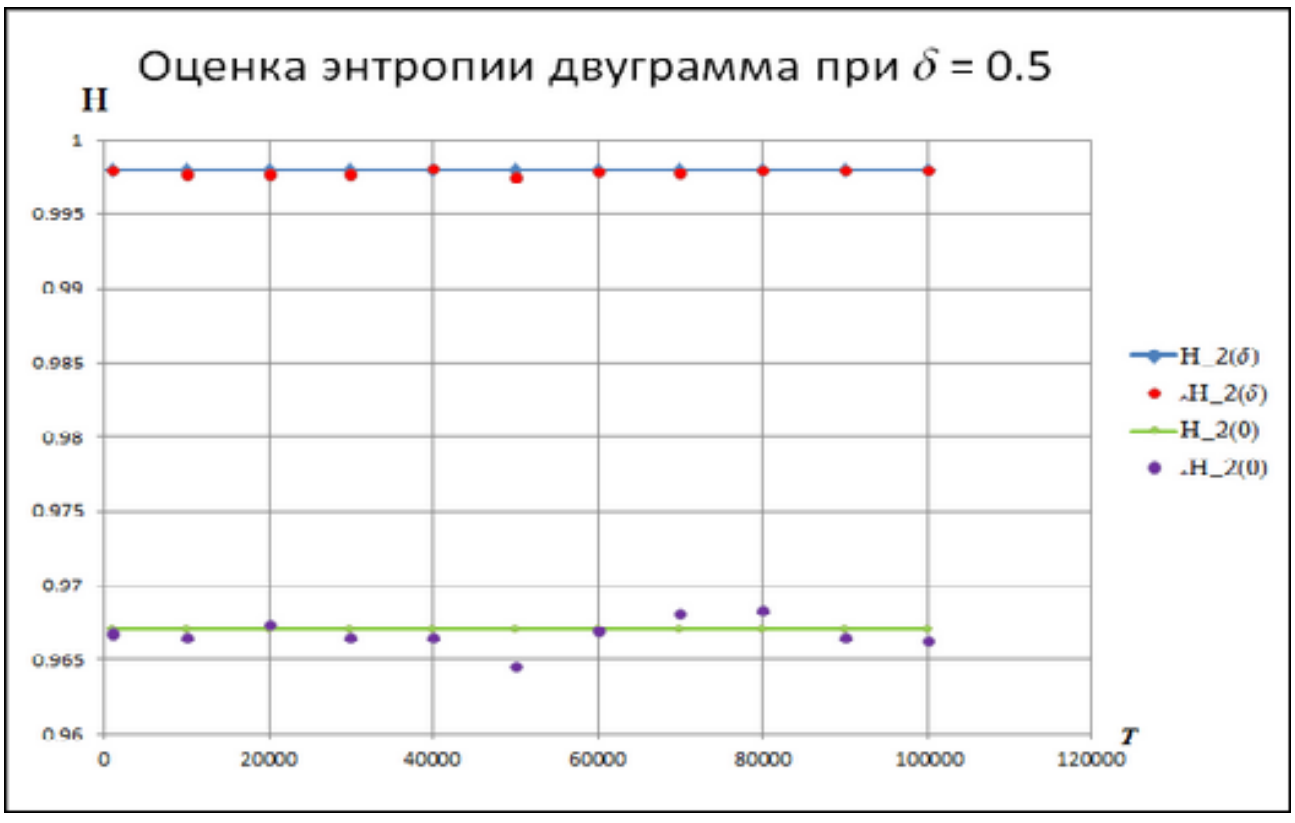

Рис. 4: График зависимости энтропии $H_2(\delta)$ от длины последовательности при $\delta = 0.5$

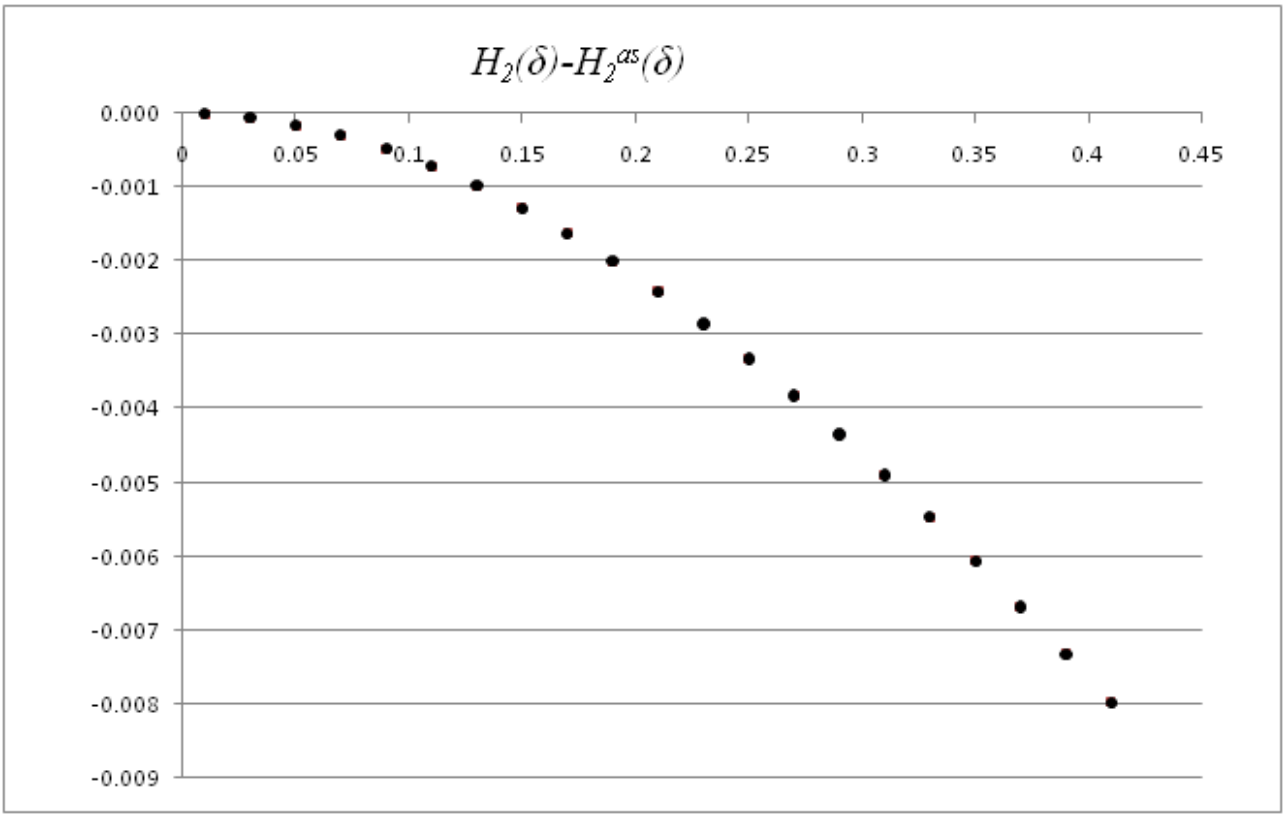

Рис. 5: График зависимости разности ассимптотического и точного значений энтропии биграммы от доли вкрапления

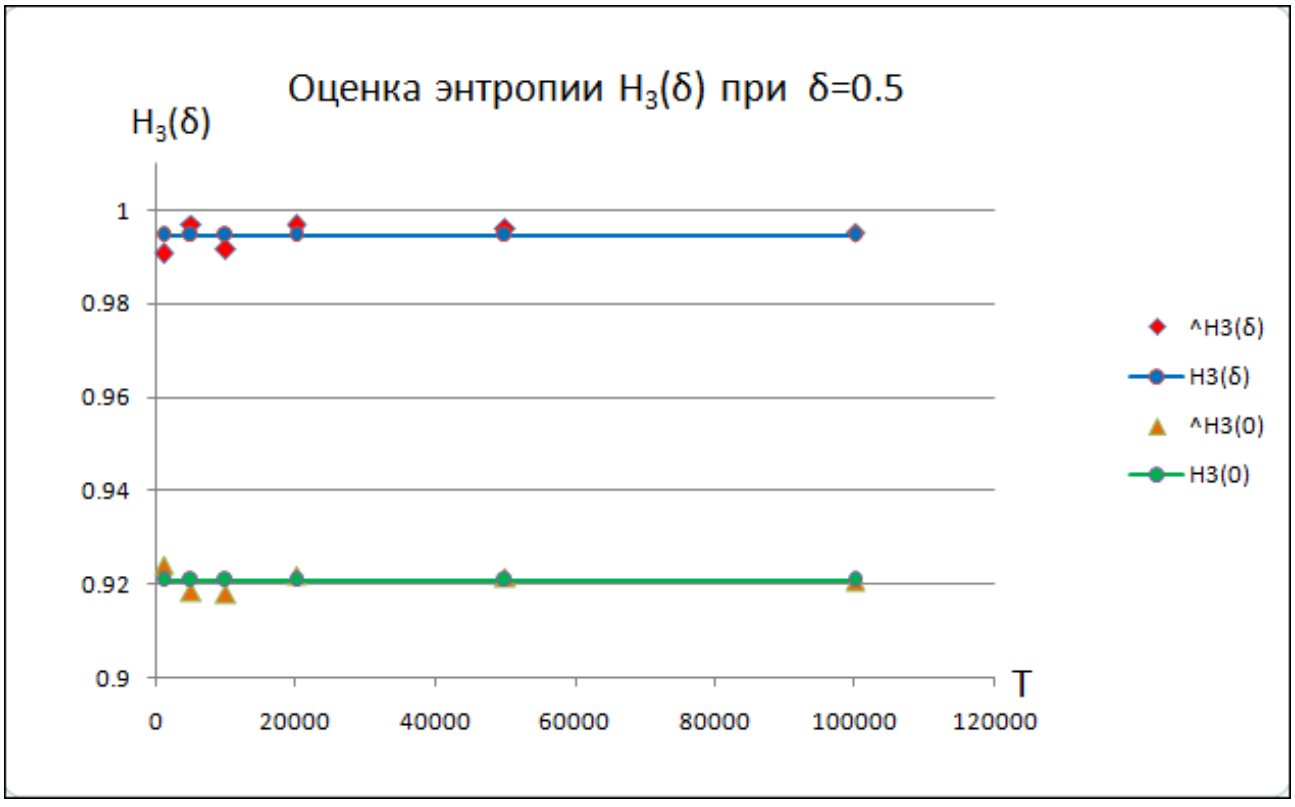

Рис. 6: График зависимости энтропии $H_3(\delta)$ от длины последовательности при $\delta = 0.5$

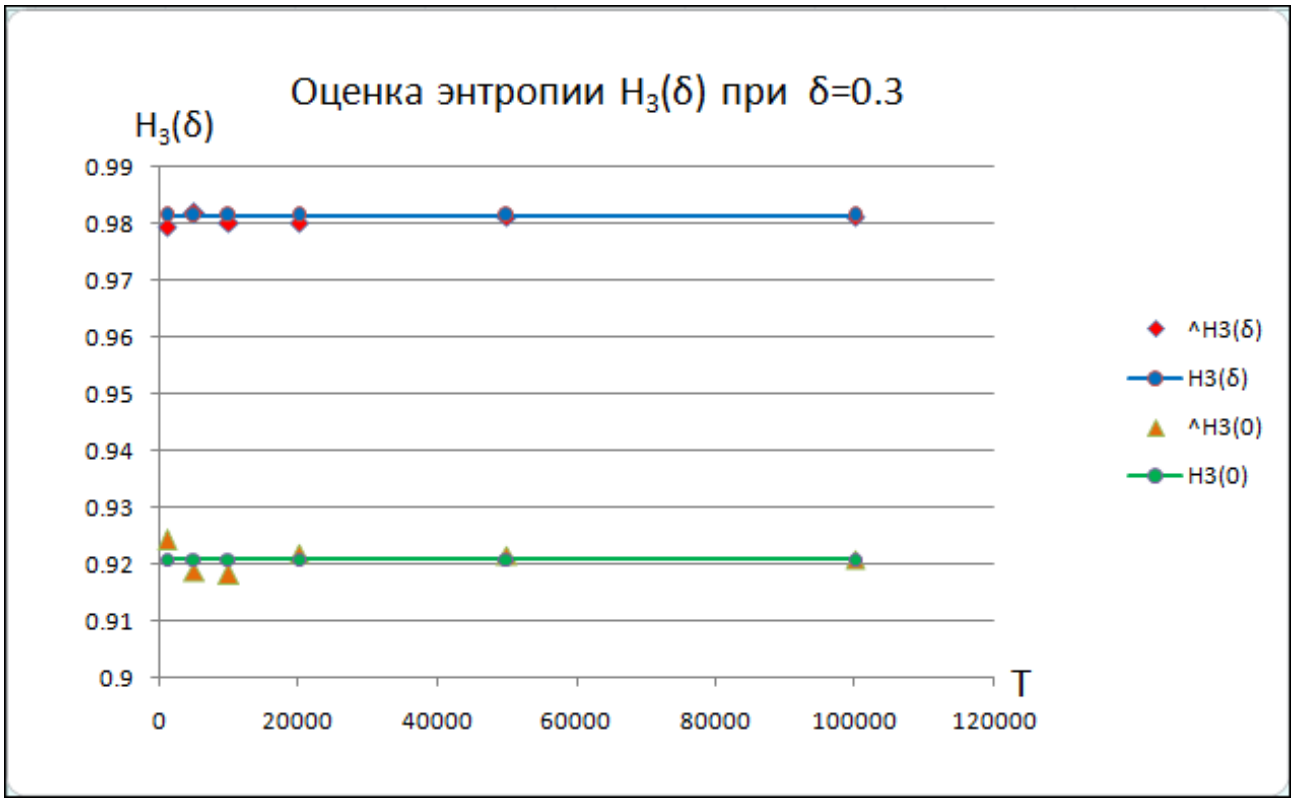

Рис. 7: График зависимости энтропии $H_3(\delta)$ от длины последовательности при $\delta = 0.3$

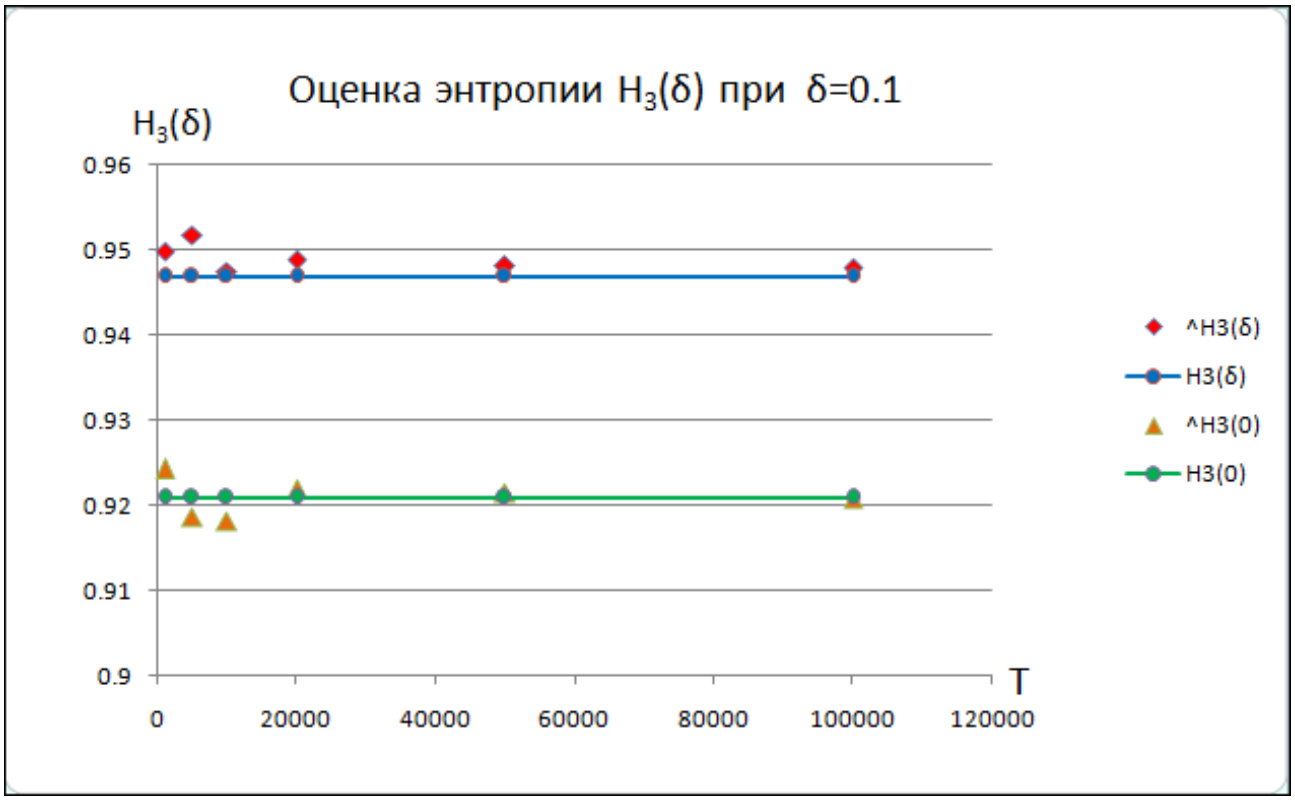

Рис. 8: График зависимости энтропии $H_3(\delta)$ от длины последовательности при $\delta = 0.1$

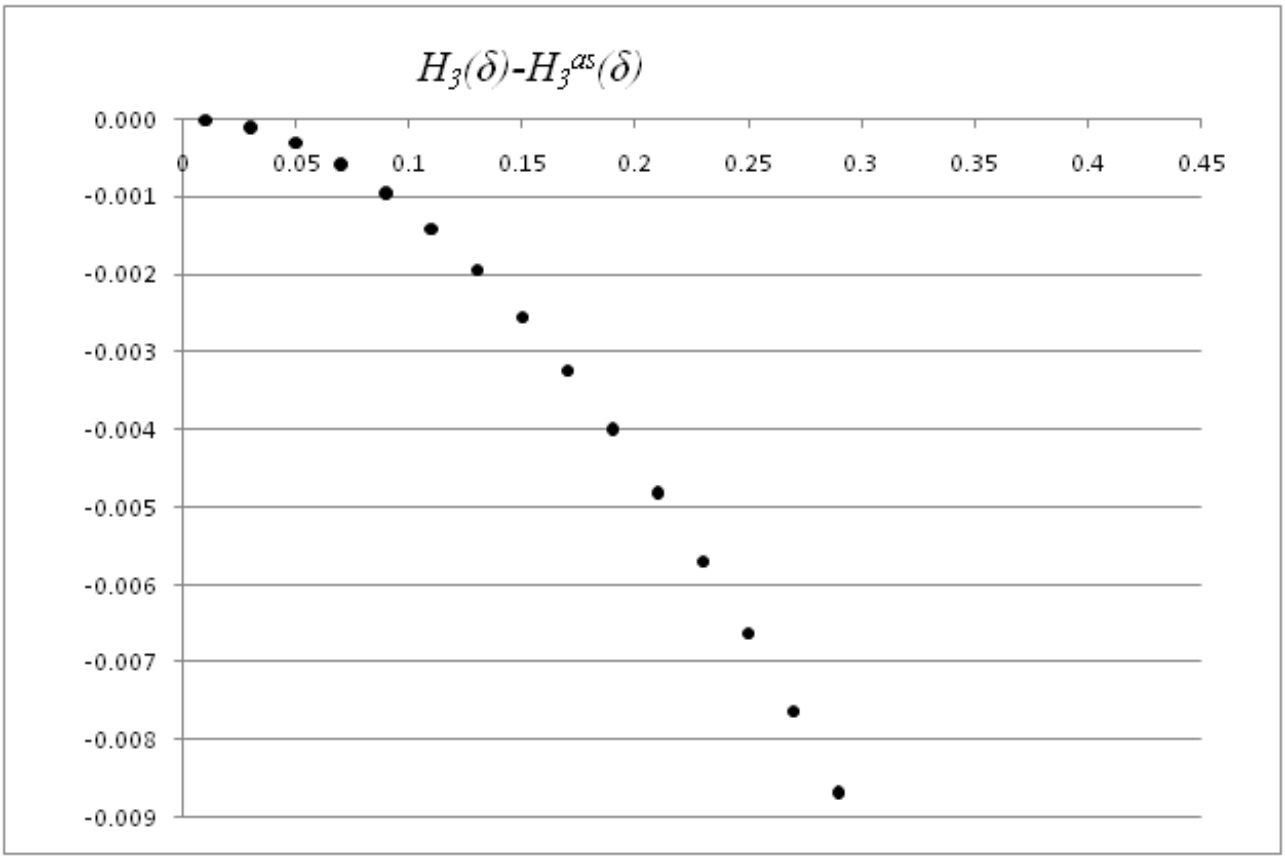

Рис. 9: График зависимости разности ассимптотического и точного значений энтропии 3-граммы от доли вкрапления

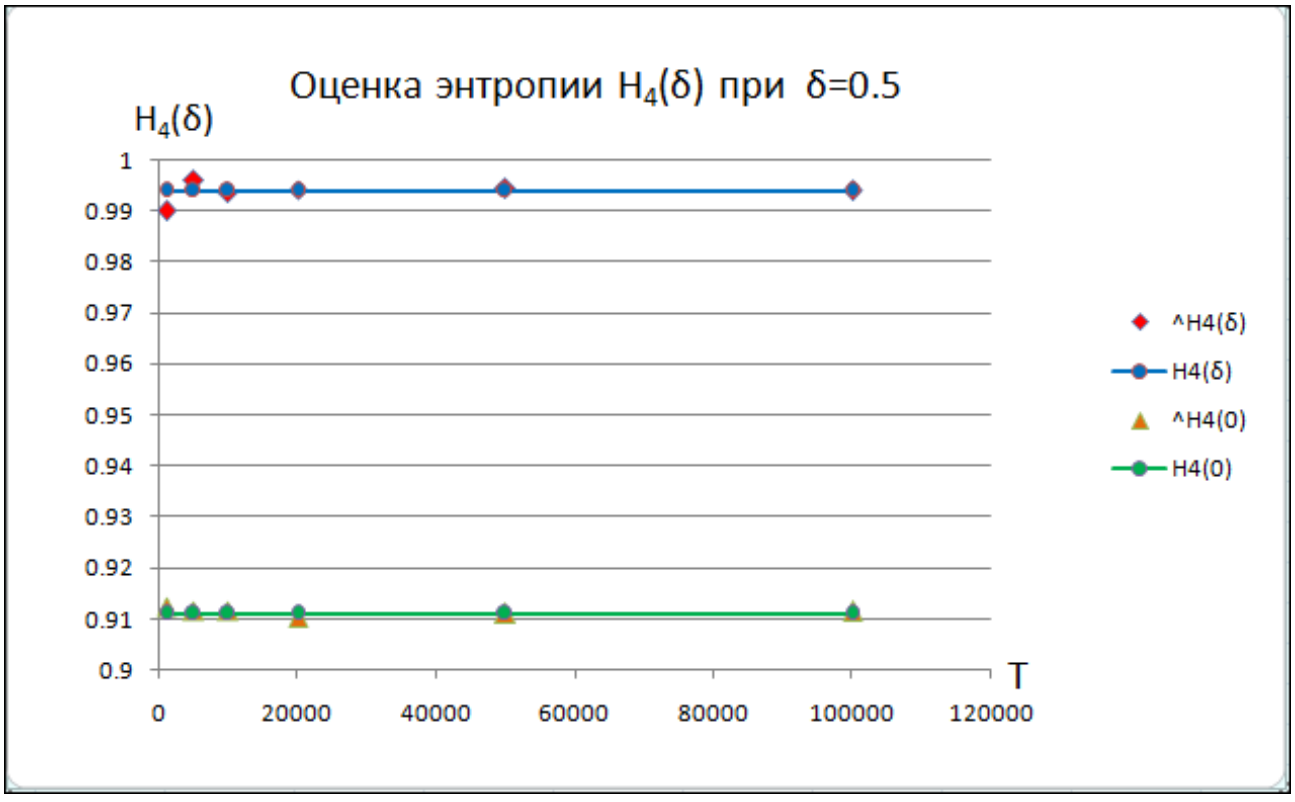

Рис. 10: График зависимости энтропии $H_4(\delta)$ от длины последовательности при $\delta = 0.5$

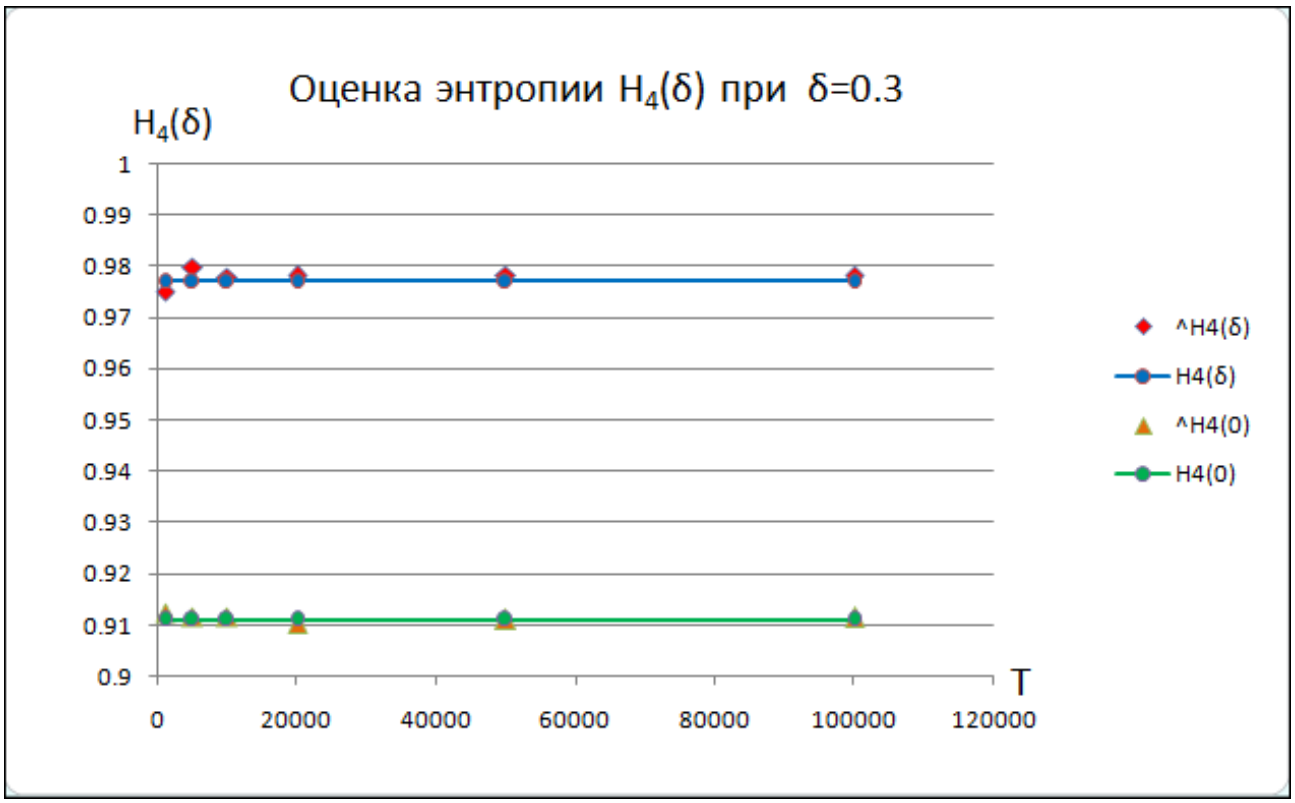

Рис. 11: График зависимости энтропии $H_4(\delta)$ от длины последовательности при $\delta = 0.3$

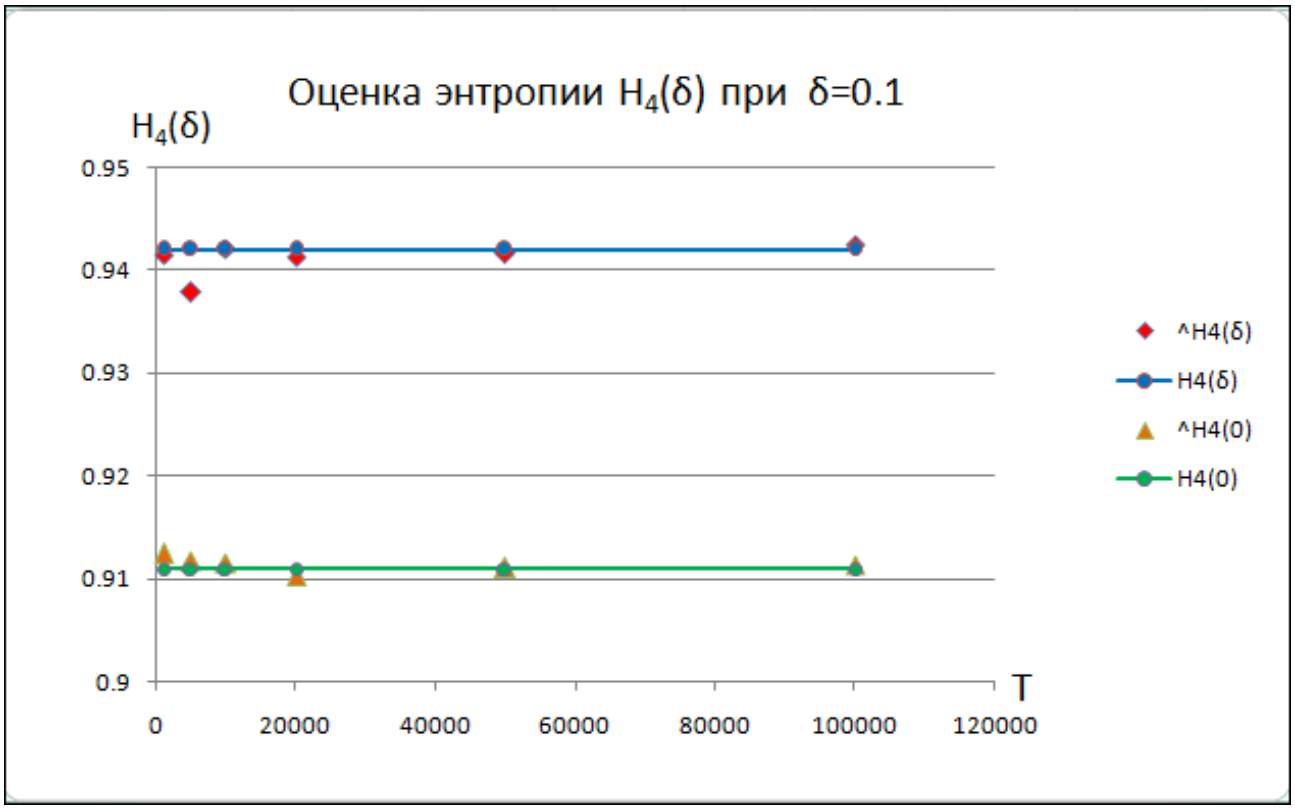

Рис. 12: График зависимости энтропии $H_4(\delta)$ от длины последовательности при $\delta = 0.1$

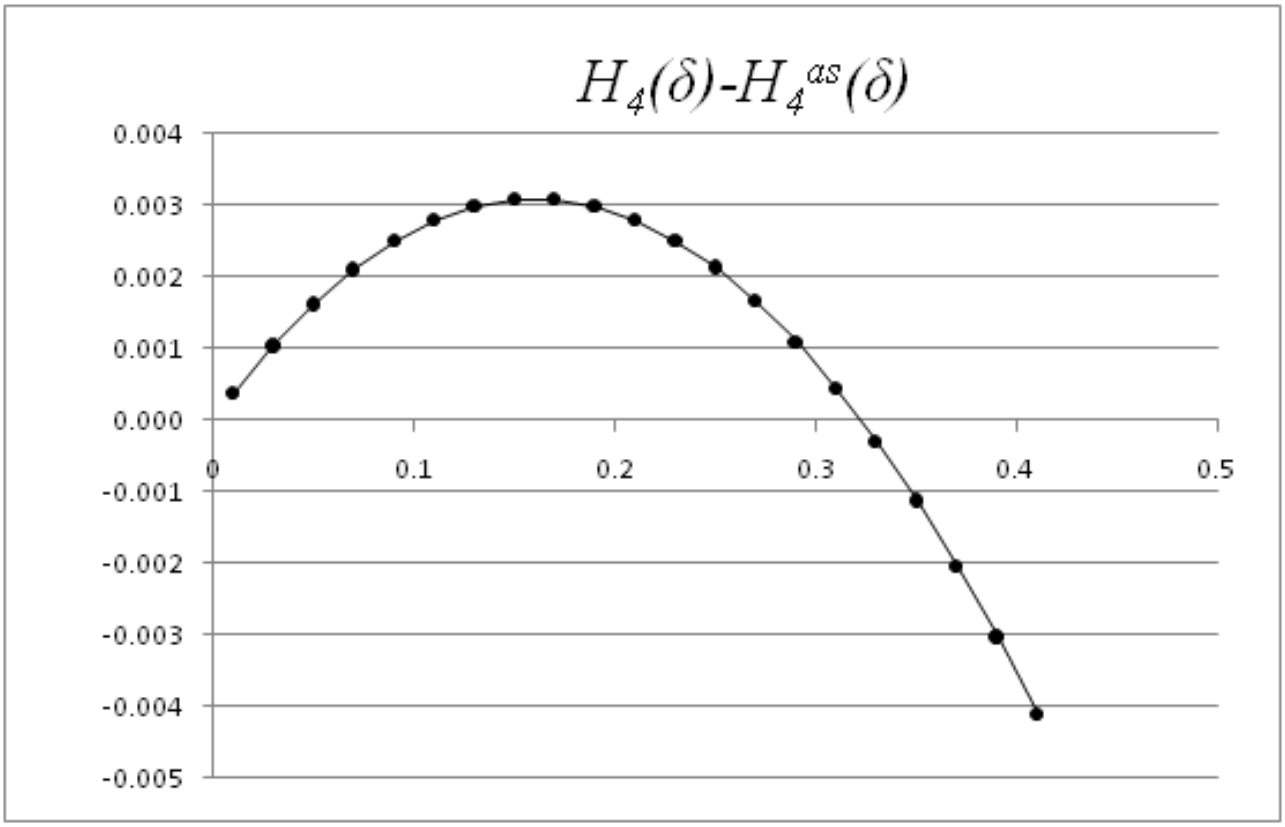

Рис. 13: График зависимости разности ассимптотического и точного значений энтропии биграммы от доли вкрапления

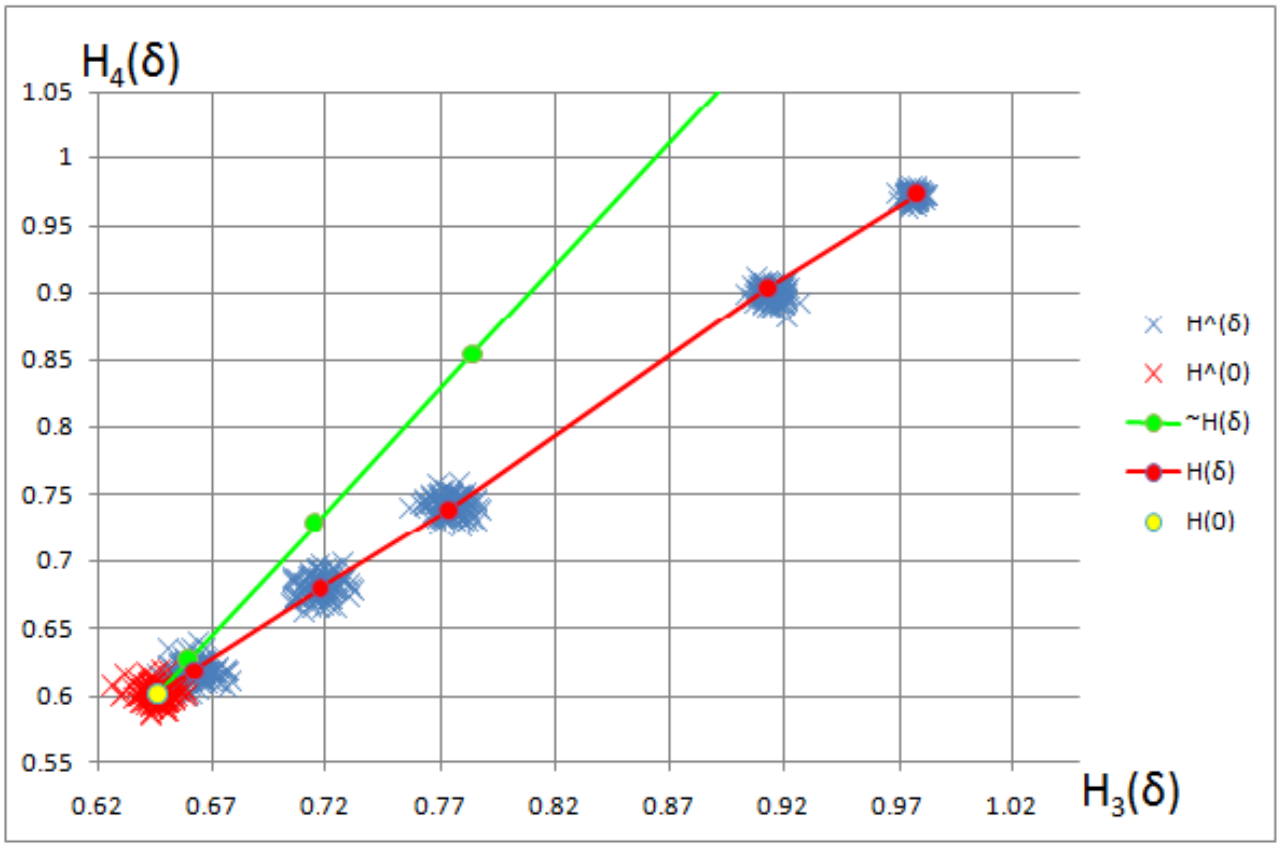

Рис. 14: График зависимости энтропии $H_4(\delta)$ от $H_3(\delta)$ при различных долях вкраплений

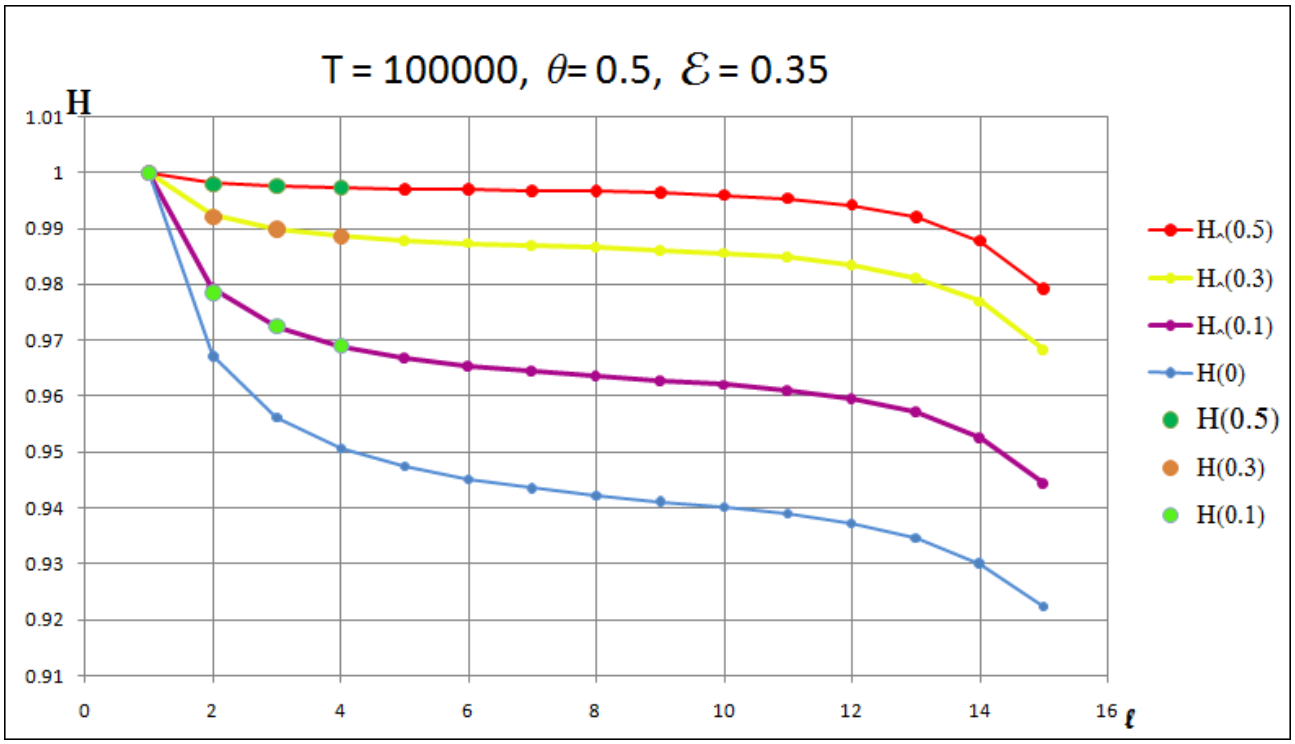

Рис. 15: Семейство графиков зависимости энтропии $H_l(\delta)$ от L при различных δ

Список литературы

[1] А. А. Духин: Теория информации - М.: "Гелиос АРВ 2007.

[2] А.В. Аграновский, А. В. Балакин: Стеганография, цифровые водяные знаки о стегоанализ - М.: Вузовская книга, 2009.

[3] В. Г. Грибунин, И. Н. Оков, И. В. Туринцев: Цифровая стеганография - М.: Солон-Прессб, 2002.

[4] Н. П. Варновский, Е. А. Голубев, О. А. Логачев: Современные направления стеганографии. Математика и безопасность информационных технологий. Материалы конференции в МГУ 28-29 октября 2004 г., МЦМНО, М., 2005, с. 32-64.

[5] Ю. С. Харин [и др.]: Криптология - Минск: БГУ, 2013.

[6] Ю. С. Харин, Е. В. Вечерко "Статистическое оценивание параметров модели вкраплений в двоичную цепь Маркова Дискрет. матем., 25:2 (2013), 135-148.