О ГРАНИЦАХ ЗАШУМЛЕНИЯ ТЕКСТОВ ПРИ СОХРАНЕНИИ ИХ СОДЕРЖАНИЯ. ПРИЛОЖЕНИЯ К КРИПТОГРАФИИ

Бабаш А.В.[1], Баранова Е.К.[2], Лютина А.А.[3], Мурзакова А.А.[4], Мурзакова Е.А.[5], Рябова Д.М.[6], Семис-оол Е.С.[7]

Цель статьи: *ввести математическую модель искаженного содержательного текста и меры его искажения, дать численную классификацию искажения содержательных текстов, привести приложения модели в криптографии.*

Метод: *дешифрование усложненного шифра Виженера, в котором используется почти периодический ключ (зашумленный), осуществляется как дешифрование зашумленного открытого текста на периодическом ключе (известными методами дешифрования Виженера), но с другим распределением вероятности встречаемости символов открытого текста, с дальнейшим сведением задачи к определению допустимого уровня шума в открытом тексте для понимания его содержания.*

Полученный результат: *представлены способы определения содержания текста (дешифрование) шифра гаммирования при использовании слабых ключей. Получена новая трудоемкость и повышена надежность метода за счет того, что k-зашумленных слабых (почти периодических) ключей больше, чем периодических. Получена формула расчета вероятности встречаемости символов после k-го зашумления. Введены искусственные языки для удобства расчетов и рассмотрены практические примеры зашумления текста (необходимые расчеты производились при помощи написанной программы на языке программирования python). Качество содержания искаженного открытого текста было оценено при помощи выделения двух границ понимания опытным путем.*

Ключевые слова*: содержательный текст, искаженный текст, мера зашумления текста, дешифрование шифра.*

DOI:10.21681/2311-3456-2020-1-74-86

1. Введение

Существуют шифры и методы их дешифрования, в результате которых дешифровальщик получает искаженный открытый текст. В данной работе рассмотрен один из методов дешифрования шифра случайного гаммирования с расчетом параметров его сложности на k-зашумленном слабом ключе. Полагаем, что дешифрование открытого текста удалось, если можно понять содержание полученного текста. В связи с этим, в работе проводится формализация понятия содержания получаемого в результате дешифрования искаженного открытого текста и дается количественная классификация качества этого содержания. Для определения границ понимания содержания текста при зашумлении рассматриваются открытые тексты на искусственных языках («Детский язык»), состоящем из ограниченного набора слов и фраз детской речи (язык радости, плача и требования (просьбы)).

2. Модель содержательного текста

В вероятностных моделях источников сообщений источник открытого текста рассматривается как источник случайных последовательностей. Считается, что источник генерирует конечную или бесконечную последовательность случайных символов $\mathrm{x}_1, \mathrm{x}_2, \ldots, \mathrm{x}_n$ из алфавита I. Вероятность случайного сообщения «$i_1, i_2, \ldots, i_n$» определяется как вероятность совместного события

$$\mathrm{P}(i_1, i_2, \ldots, i_n) = \mathrm{P}(\mathrm{x}_1 = i_1, \mathrm{x}_2 = i_2, \ldots, \mathrm{x}_n = i_n).$$

При этом, естественно, требуют выполнения условий:

1. для любого случайного сообщения «$i_1, i_2, \ldots, i_n$»

$$\mathrm{P}(i_1, i_2, \ldots, i_n) \geq 0$$

2. $\sum_{i_1, i_2, \ldots, i_n} \mathrm{P}(i_1, i_2, \ldots, i_n) = 1;$

для любого случайного сообщения «$i_1, i_2, \ldots, i_n$»

$$\mathrm{P}(i_1, i_2, \ldots, i_n) =$$

1 Бабаш Александр Владимирович, доктор физико-математических наук, профессор НИУ ВШЭ, г. Москва, Россия. E-mail: ababash@hse.ru

2 Баранова Елена Константиновна, доцент НИУ ВШЭ, г. Москва, Россия. E-mail: ekbaranova@hse.ru

3 Лютина Анна Андреевна, студент магистратуры НИУ ВШЭ, г. Москва, Россия. E-mail: aalyutina@edu.hse.ru

4 Мурзакова Александра Андреевна, студент магистратуры НИУ ВШЭ, г. Москва, Россия. E-mail: aamurzakova@edu.hse.ru

5 Мурзакова Екатерина Андреевна, студент магистратуры НИУ ВШЭ, г. Москва, Россия. E-mail: eamurzakova@edu.hse.ru

6 Рябова Дарья Михайловна, студент магистратуры НИУ ВШЭ, г. Москва, Россия. E-mail: dmryabova@edu.hse.ru

7 Семис-оол Елена Сергеевна, студент магистратуры НИУ ВШЭ, г. Москва, Россия. E-mail: essemisool@edu.hse.ru

$$= \sum_{i_1,i_2,\dots,i_s} \mathrm{P}(i_1, i_2, \dots, i_s), s \geq n+1$$

Смысл последнего условия состоит в том, что вероятность всякого случайного сообщения длины n есть сумма вероятностей всех «продолжений» этого сообщения до длины $s > n$ (некоторый вариант аксиомы Колмогорова). Текст, порождаемый таким источником, является вероятностным аналогом языка. Он обладает одинаковыми с языком частотными характеристиками k-грамм. Задавая конкретное вероятностное распределение на множестве открытых текстов, мы задаем соответствующую модель источника сообщений.

В данной работе мы будем опираться на частный случай этой общей модели.

Стационарный источник независимых символов алфавита. В этой модели предполагается, что вероятности сообщений полностью определяются вероятностями отдельных символов алфавита $I : P = \left(p_i, i \in I\right)$.

Определение 1. **Под открытым текстом понимается реализация последовательности независимых испытаний в полиномиальной вероятностной схеме** $P = \left(p_i, i \in I\right)$.

Исходу взаимно однозначно соответствует символ алфавита I. Эта модель позволяет разделить буквы алфавита на классы высокой, средней и низкой частот использования. В связи с чем

$$\mathrm{P}(i_1, i_2, \dots, i_L) = \prod_{j=1}^{L} P(x_j = i_j)$$

и

$$\mathrm{P}\left(\mathrm{x}_j = i\right) > 0, \sum_{i \in I} \mathrm{P}\left(\mathrm{x}_j = i\right) = 1.$$

Зашумление открытых текстов

Под одиночным случайным искажением (ОСИ) открытого текста $x = i_1 i_2 \dots i_L$ мы будем понимать случайный и равновероятный выбор элемента $j \in \{1, 2, \dots, L\}$ и случайную и равновероятную замену буквы i_j на букву $i \in I, i \neq i_j$. Искаженный открытый текст в результате ОСИ будет иметь вид $x^1 = i_1 i_2 \dots i_{j-1} i_j i_{j+1} \dots i_L$. При последовательном искажении текста $x = i_1 i_2 \dots i_L$ одиночным искажением n раз (n-ОСИ) мы получим искаженный текст $x^n = i_1^{'} i_2^{'} \dots i_L^{'}$ [1-5].

Определение 2. **Слабым ключом шифра назовем ключ, при котором по шифрованному тексту имеется метод его нахождения или его части, или имеется способ определить отдельные части открытого текста.** При этом ключ составляется из случайной равновероятной выборки.

Рассмотрим ключ в шифре гаммирования. Предположим, он состоит из одинаковых последовательностей длины d, как лозунговый ключ в шифре Виженера. Тогда можно определить длину ключевого слова с помощью известных методов, например, теста Казиски и теста Фридмана. Обозначим через I_d^N – множество всех локально-периодических последовательностей периода d, тогда длину ключевой последовательности можно представить, как $N = kd + r$ [1, 6-10].

Рассмотрим, в чем заключаются методы нахождения периода d.

Тест Казиски анализирует повторения в шифртексте. Метод основывается на том, что если гамма локально периодическая, то две одинаковые m-граммы открытого текста, отстоящие друг от друга на расстоянии, которое кратно периоду m-граммы, будут одинаково зашифрованы в некоторые одинаковые m-граммы, находящиеся на том же расстоянии друг от друга. Появление же одинаковых m-грамм в шифрованном тексте по другим причинам маловероятно (при некоторых разумных ограничениях на величину m и на длину шифртекста N). Из этого следует, что большинство расстояний между одинаковыми m-граммами шифртекста делится на минимальный период. По этой причине на практике в качестве предполагаемого периода гаммы рассматривают наибольший общий делитель длин большинства расстояний между повторениями m-грамм. Эксперименты показали хорошую надежность данного метода, при условии, что в шифртексте имеются повторения триграмм и m-грамм при m, большем трёх.

Метод Фридмана основан на введенном им понятии «индекса совпадения».

Первый метод Фридмана

Первый метод Уильяма Фридмана состоит в том, что для данного шифртекста B вычисляют индекс совпадения $IC\left(B\right)$ и сравнивают его с величинами

$$\frac{N-d}{d\left(N-1\right)} \sum_{i \in I} P_i^2 + \frac{N\left(d-1\right)}{\left(N-1\right)d} * \frac{1}{|I|}, d = 1, 2, 3, \dots$$

При достаточной близости $IC\left(B\right)$ к одной из этих величин при некотором d, предполагают, что период равен этому d. Данный метод удовлетворительно работает при использовании для зашифрования локально-периодических последовательностей периода не более 5. С точки зрения обоснованности применения данного метода лучше сравнивать IC(B) с более точным выражением для $N = kd + r$

$$\frac{(k+1)kr + k(k-1)(d-r)}{N(N-1)} \sum_i P_i^2 +$$

$$+ \left(1 - \frac{(k+1)kr + k(k-1)(d-r)}{N(N-1)}\right) \frac{1}{|I|}.$$

Второй метод Уильяма Фридмана

Второй метод Уильяма Фридмана также основан на вычислении индекса совпадения. Он состоит в опробо-

вании возможных периодов по следующей схеме. Для предполагаемого периода d выписываются d подпоследовательностей

$$b_1, b_{1+d}, b_{1+2d}, \ldots$$
$$b_2, b_{2+d}, b_{2+2d}, \ldots$$
$$\ldots\ldots\ldots\ldots\ldots\ldots$$
$$b_d, b_{d+d}, b_{d+2d}, \ldots$$

Для каждой подпоследовательности подсчитывается ее индекс совпадения. Если все индексы совпадения в среднем близки к значению

$$\frac{1}{d}\sum_{i\in I} P_i^2,$$

то есть к среднему значению индекса совпадения случайных шифртекстов, полученных с помощью гамм периода 1, то принимают величину d за истинный период, в противном случае опробуют следующую величину периода. Приведенный способ нахождения периода гаммы по шифрованному тексту удовлетворительно работает для периодов, не превышающих 30.

Метод БШ

Предлагаемый метод определения периода гаммы в шифре гаммирования по известному шифртексту $B = b_1, b_2, \ldots, b_N$ состоит в следующем. Выписываются все пары номеров j, j', для которых $b_j = b_{j'}$. Пусть $\Pi(B, =)$ – множество таких пар. Очевидно,

$$|\Pi(B, =)| = \sum_{i\in I} F_i(F_i - 1),$$

где F_i - частота встречаемости буквы i в шифртексте В.

Каждой паре (j, j') из $\Pi(B, =)$ ставится в соответствие расстояние $p(j, j')$, равное абсолютной величине разности между j и j'. Ищется максимальное по мощности подмножество $\Pi(B, d, =)$ пар в $\Pi(B, =)$ такое, что их расстояния $p(j, j')$ имеют некоторый общий наибольший делитель d, отличный от 1. Подсчитывается величина

$$\text{ИБШ}(B, d) = \frac{|\Pi(B, d, =)|}{N(N-1)}$$

и сравнивается с величиной

$$E_U\big(\text{ИБШ}(B, d)\big) =$$
$$= \frac{(k+1)kr + k(k-1)(d-r)}{N(N-1)}\sum_i P_i^2,$$

где k, r определены равенством $N = kd + r$. Если эти величины близки, принимается гипотеза о том, что шифрование проводилось гаммой периода d. В противном случае эта гипотеза отвергается [1].

Такой периодический ключ в шифре гаммирования является слабым ключом (выше были рассмотрены методы определения периода такого ключа). Введем обозначение для такого ключа – d -слабый ключ.

Генератор псевдослучайных последовательностей может выработать d -слабый ключ с вероятностью $\frac{|I|^d}{|I|^L}$.

Если возьмем k -зашумленный слабый ключ (произведем k раз зашумление ключевой последовательности), то таких ключей будет больше, чем периодических последовательностей длины d (d -слабых ключей), следовательно, с большей вероятностью сможем дешифровать текст, зашифрованный на таком k -зашумленном слабом ключе. Можем назвать такие ключи d-слабыми ключами с неким % зашумления.

Возьмем ключ в шифре гаммирования как последовательность локального периода (период d).

$x = i_1, i_2 \ldots i_j \ldots i_N$ - содержательный текст;

$\gamma = \gamma_1, \gamma_2 \ldots \gamma_j \ldots \gamma_N$ – ключ, который выбирается из множества равновероятных последовательностей $\gamma \in |I|^N$;

$y = y_1, y_2 \ldots y_j \ldots y_N$ – зашифрованный текст.

Произведем k раз зашумление ключевой последовательности, то есть получим зашумленный ключ $\gamma' = \gamma_1', \gamma_2' \ldots \gamma_j' \ldots \gamma_N'$. Далее произведем зашифрование на измененном ключе:

$$\frac{i_1, i_2 \ldots i_d, i_{d+1} \ldots i_j \ldots i_N}{\gamma_1', \gamma_2' \ldots \gamma_d' \gamma_{d+1}' \ldots \gamma_j' \ldots \gamma_N'}$$
$$y_1', y_2' \ldots y_d' y_{d+1}' \ldots y_j' \ldots y_N'$$

Но можем в силу ассоциативности сначала наложить шум на открытый текст, а потом зашифровать его на ключе с периодом d. Распределение букв в таком открытом тексте будет уже другим, нежели в не зашумлённом тексте.

$$i_1, i_2 \ldots i_d, i_{d+1} \ldots i_j \ldots i_N$$
$$\oplus\ \xi_1, \xi_2 \ldots \xi_d, \xi_{d+1} \ldots \xi_j \ldots \xi_N$$
$$\frac{\gamma_1, \gamma_2 \ldots \gamma_d, \gamma_{d+1} \ldots \gamma_j \ldots \gamma_N}{y_1', y_2' \ldots y_d' y_{d+1}' \ldots y_j' \ldots y_N'}$$

$x = i_1, i_2 \ldots i_j \ldots i_N$ – открытый текст,

$\xi = \xi_1, \xi_2 \ldots \xi_j \ldots \xi_N$ – шум,

$\gamma = \gamma_1, \gamma_2 \ldots \gamma_j \ldots \gamma_N$ – ключ,

$y' = y'_1, y'_2 \dots y'_j \dots y'_N$ – зашифрованный текст,
$\oplus$ – сложение по модулю $|I|$.

$$\begin{array}{c} \oplus \begin{array}{c} i'_1, i'_2 \dots i'_d i'_{d+1} \dots i'_j \dots i'_N \\ \gamma_1, \gamma_2 \dots \gamma_d, \gamma_{d+1} \dots \gamma_j \dots \gamma_N \end{array} \\ \hline y'_1, y'_2 \dots y'_d y'_{d+1} \dots y'_j \dots y'_N \end{array}$$

$$i_j \oplus \xi_j \oplus \gamma_j = i'_j \oplus \gamma_j = y'_j$$

Получим зашифрованный текст на d-слабом ключе, который можно дешифровать как шифр Виженера, о чем говорилось выше. Так как d-слабых ключей с неким % зашумления больше, чем просто d-слабых ключей, то получаем новую трудоемкость и повышаем надежность метода [11-12].

Исходя из всего вышеописанного, встала задача описания подхода к классификации «на хорошее или плохое содержание» искаженного открытого текста. Чтобы понять, сможем ли мы прочесть открытый текст, нам нужно знать, какой уровень шума допустим.

Работа с искаженными (зашумленными) данными актуальна в разных сферах. Такие сферы можно поделить на два смысловых блока.

Блок 1. Задача «прочтения» искаженного текста, то есть поиск методов и способов восстановления искаженных данных, приведения полученных зашумленных сигналов к воспринимаемому человеком виду, чтобы можно было разобрать, о чем было переданное сообщение. Так, в работах [13-15] представлены способы восстановления зашумленных изображений. В работах [16-20] приводятся методы компенсации шума в канале при распознавании речевых и других сигналов на фоне помех; рассматривается отношение сигнал/шум и границы ошибок восприятия переданного сообщения по каналу связи.

Блок 2. Использование зашумления исходных данных как мера защиты конфиденциальной информации. Так, в [21-27] говорится о применении пространственного зашумления (шумогенераторы) и линейного зашумления (маскировка сигналов в проводах, кабелях и т.д.) для защиты от утечки конфиденциальных данных как внутри контролируемой зоны, так и за ее пределами. А в [28] приводится материал об обеспечении конфиденциальности местоположения лица (запутывание данных о местоположении с добавлением случайного шума).

В данной работе рассматриваются искаженные (зашумленные) данные применимо к криптографии.

3. Подход к классификации «на хорошее или плохое содержание» искаженного открытого текста

Теорема 1. Пусть $p_i, i \in I$ – вероятностное распределение букв алфавита I в открытых текстах, а $x = i_1, i_2 \dots i_L$ – выборка из указанного распределения. Тогда искаженный открытый текст $x' = i'_1, i'_2 \dots i'_L$ (n-ОСИ)-искажением является выборкой из распределения $p'_i, i \in I$:

$$P'_i = P_i - \frac{P_i}{N} + \frac{(1-P_i)}{N*(n-1)} \quad (1)$$

где:
P_i – вероятность встречаемости i-ой буквы в исходном тексте;
N – длина текста;
n – количество уникальных символов в алфавите.
Для k зашумление формула (1) примет вид:

$$P'_{ik} = \frac{P_i * \left(N*(n-1)-n\right)^k}{\left(N*(n-1)\right)^k} + \sum_{m=0}^{k-1} \frac{\left(N*(n-1)-n\right)^m}{\left(N*(n-1)\right)^{m+1}} \quad (2)$$

Видно, что качество содержания искаженного открытого текста $x' = i'_1, i'_2 \dots i'_L$ (n-ОСИ)-искажением вполне характеризуется распределением (1). Но как классифицировать «на хорошее или плохое» содержание текста, полученного выборкой из заданного дискретного распределения. Наше первое предложение состоит в сравнении энтропий распределений и поиска значений энтропии распределения (1), при котором практические примеры (n-ОСИ) искажений указывают на различные уровни содержания.

Доказательство формулы (2)

Для внесения 1 зашумления формула (2) примет вид:

$$\nu_1 = \frac{\nu_0 \left(Nn - N - n\right) + N}{N\left(n-1\right)}$$

Отметим, что $P_i = \frac{\nu_i}{N}$, тогда $\nu_i = P_i * N$, где ν_i – количество i-ых символов в тексте.

Для доказательства применяется метод математической индукции.

$$\nu_2 = \left(\nu_0 * \left(1 - \frac{n}{N(n-1)}\right) + \frac{1}{n-1}\right) * \left(1 - \frac{n}{N(n-1)}\right) + \frac{1}{n-1}$$

Выполним замену: $N\left(n-1\right) = k$,

$$\nu_1 = \frac{\nu_0 \left(k-n\right) + N}{k}$$

$$\nu_2 = \frac{\nu_0 * \left(k-n\right)^2}{k^2} + \frac{N*\left(k-n\right)}{k^2} + \frac{N}{k}$$

$$\nu_3 = \frac{\nu_0 \left(k-n\right)^3}{k^3} + \frac{N\left(k-n\right)^2}{k^3} + \frac{N\left(k-n\right)}{k^2} + \frac{N}{k}$$

$$\nu_4 = \frac{\nu_0 (k-n)^4}{k^4} + \frac{N(k-n)^3}{k^4} + \cdot$$

$$+ \frac{N(k-n)^2}{k^3} + \frac{N(k-n)}{k^2} + \frac{N}{k}$$

Заметна закономерность, тогда для m -ых зашумлений ν_m примет вид:

$$\nu_m = \frac{\nu_0 (k-n)^m}{k^m} + \sum_{i=0}^{m-1} \frac{N(k-N)^i}{k^{i+1}}$$

Выполним замену $k = N(n-1)$, тогда ν_m примет вид:

$$\nu_m = \frac{\nu_0 (N(n-1)-n)^m}{(N(n-1))^m} + \sum_{i=0}^{m-1} \frac{N(N(n-1)-n)^i}{(N(n-1))^{i+1}}$$

Данная формула верна для подсчета количества букв в тексте после внесения m изменений. Перейдем к частотам букв в тексте:

$$\nu_i = P_i * N$$

$$P_k = \frac{\nu_0 (N(n-1)-n)^k}{N^{k-1}(n-1)^k} + \sum_{i=0}^{k-1} \frac{(N(n-1)-n)^i}{(N(n-1))^{i+1}}$$

– формула доказана.

4. Практические примеры

Пример 1

Пусть задано слово ВЕТЕР. Найдем частоту встречаемости каждой буквы после одного изменения. Уникальных символов 4, то есть n = 4. Количество букв в слове 5, то есть N = 5. В таблице 1 представлены частоты встречаемости букв в слове.

Таблица 1

Частоты встречаемости букв в слове

	В	Е	Т	Р
P_i	$\frac{1}{5}$	$\frac{16}{75}$	$\frac{1}{5}$	$\frac{1}{5}$

Тогда по формуле (2):

$$P'_{\text{В}} = \frac{1}{5} - \frac{1}{5*5} + \frac{\left(1-\frac{1}{5}\right)}{5*(4-1)} = \frac{1}{5} - \frac{1}{25} + \frac{\frac{4}{5}}{15} = \frac{1}{5} - \frac{1}{25} + \frac{4}{75} = \frac{16}{75}$$

$$P'_{\text{Е}} = \frac{2}{5} - \frac{2}{5*5} + \frac{\left(1-\frac{2}{5}\right)}{5*(4-1)} = \frac{2}{5} - \frac{2}{25} + \frac{\frac{3}{5}}{15} = \frac{2}{5} - \frac{2}{25} + \frac{3}{75} = \frac{27}{75}$$

$$P'_{\text{Т}} = \frac{1}{5} - \frac{1}{5*5} + \frac{\left(1-\frac{1}{5}\right)}{5*(4-1)} = \frac{1}{5} - \frac{1}{25} + \frac{\frac{4}{5}}{15} = \frac{1}{5} - \frac{1}{25} + \frac{4}{75} = \frac{16}{75}$$

$$P'_{\text{Р}} = \frac{1}{5} - \frac{1}{5*5} + \frac{\left(1-\frac{1}{5}\right)}{5*(4-1)} = \frac{1}{5} - \frac{1}{25} + \frac{\frac{4}{5}}{15} = \frac{1}{5} - \frac{1}{25} + \frac{4}{75} = \frac{16}{75}$$

В результате новые частоты букв распределяются следующим образом (см. Таблица 2):

Таблица 2

Получившиеся частоты букв (Пример 1)

	В	Е	Т	Р
P'_i	$\frac{16}{75}$	$\frac{27}{75}$	$\frac{16}{75}$	$\frac{16}{75}$

Пример 2

Рассчитаем, какая будет частота встречаемости букв после 5 изменения в слове ВЕТЕР. Исходные данные остаются такими же. Количество внесенных шумов будет 5, k = 5; длина текста остается 5, N =5; количество уникальных символов 4, n =4.

Используя формулу (2), получим:

Вероятность встречи буквы В после 5-ого шума в слове будет равна:

$$P'_{\text{В5}} = \frac{\frac{1}{5} * (5*(4-1)-4)^5}{(5*(4-1))^5} +$$

$$+ \sum_{m=0}^{4} \frac{(5*(4-1) - 4)^m}{(5*(4-1))^{m+1}} = 0{,}2394$$

Вероятность встречи буквы Е после 5-ого шума в слове будет равна:

$$P'_{\text{Е5}} = \frac{\frac{2}{5} * (5*(4-1)-4)^5}{(5*(4-1))^5} +$$

$$+ \sum_{m=0}^{4} \frac{(5*(4-1) - 4)^m}{(5*(4-1))^{m+1}} = 0{,}2818$$

Поскольку исходные частоты у букв В, Р, Т равны, то их вероятностные частоты после изменений также будут равны:

$$P'_{\text{В}} = P'_{\text{Т}} = P'_{\text{Р}} = 0{,}2394$$

В результате новые частоты букв распределяются следующим образом (см. Таблица 3):

Таблица 3

Получившиеся частоты букв (Пример 2)

	В	Е	Т	Р
P'_{i5}	$0,2394$	$0,2818$	$0,2394$	$0,2394$

Обратим внимание, что чем больше замен букв совершаем, тем более равновероятна встречаемость букв в слове.

Пример 3

Для дальнейшего анализа текстов была разработана программа на языке python, которая производит замены букв в тексте и рассчитывает наблюдаемую частоту встречаемости буквы и теоретическую частоту каждой буквы по выведенной выше формуле. Были собраны фразы для трех категорий языка ребенка: радость, плач и требование (просьба). Программа предлагает пользователю выбрать одну из категорий языка ребенка, количество файлов для формирования вариантов зашумления и количество фраз из подготовленных наборов придуманного языка, которые будут формировать исходный текст. Общее количество зашумлений совпадает с количеством символов в тексте. Итогом работы программы являются автоматически сформированные файлы (по два на каждый вариант зашумления). Первый файл содержит информацию о количестве символов в тексте, количестве букв в алфавите для данной фразы, среднюю частоту встречаемости каждой буквы, общее число зашумлений и сами зашумленные фразы на каждом этапе. Второй файл содержит частоты букв: для каждой строки с <<Eqn0164.wmf>>-ым изменением рассчитывается частота встречаемости буквы во фразе, далее рассчитывается частота встречи данной буквы по выведенной формуле.

Рассмотрим предложение на языке ребенка в ситуации «требование (просьба)» для двух фраз: «хочу пить дай мне игрушку». С помощью программного приложения были произведены зашумления (за каждый ход зашумляется одна буква в тексте, то есть заменяется на другую из этого же алфавита). Полученные шаги зашумления приведены в Таблице 4. Исходя из представленных данных в таблице были выделены две границы зашумления текста. Первая граница – номер шага зашумления, до которого совершенно ясно, к какой категории отнести текст, после этой границы текст сложнее отнести к конкретной категории. Можно понять какие-то отдельные слова, но смысл или эмоция ребенка может быть определена не точно. Вторая граница – номер шага зашумления, после которой нельзя определить, к какой категории относится текст. После анализа полученных данных был сделан вывод, что первая граница находится на 4-ом шаге, а вторая граница – на 22 шаге.

Сравним полученные частоты букв на каждом шаге и вычисленные частоты по формуле (2). В теории текст зашифрован, если вероятности встречаемости букв в тексте равны. То есть после k-ого шага, если текст уже не читаемый, вероятность встречаемости каждой буквы в тексте должна быть равна $\frac{1}{n} = \frac{1}{19} = 0,053$, где n – количество уникальных символов, или находиться в пределах этого значения с учетом возможной погрешности. Для этого обратимся к автоматически сформированным программой данным, которые для удобства представлены в виде таблиц на Рис. 1 и 2.

Таблица 4

Пример зашумления текста в категории «требование (просьба)»

Длина текста	Букв в алфавите	Средняя частота буквы	№ изменения	Предложение
26	19	0.053	0	хочу пить дай мне игрушку
26	19	0.053	1	хочу пидь дай мне игрушку
26	19	0.053	2	хочу пидь даймме игрушку
26	19	0.053	3	хочу пидь дкймме игрушку
26	19	0.053	4	хочу пидь дкймчигрушку
26	19	0.053	5	хочуепидь дкймчигрушку
26	19	0.053	6	хочуепидь дкйчигрушку
26	19	0.053	7	хочуепидь дкчйчигрушку
26	19	0.053	8	хочуепидь дкчймнеригрушку
26	19	0.053	9	хочуепимь дкчймнеригрушку
26	19	0.053	10	хочуепимь дкчймнеригруаку
26	19	0.053	11	хочуепимььдкчймнеригруаку
26	19	0.053	12	хочуепимььдйчймнеригруаку

Длина текста	Букв в алфавите	Средняя частота буквы	№ изменения	Предложение
26	19	0.053	13	хочуепимдьдйчймнеригруаку
26	19	0.053	14	хочуепимдьдйчймнхригруаку
26	19	0.053	15	хочуепимдьдйчймнхригриаку
26	19	0.053	16	хочуепимдьчйчймнхригриаку
26	19	0.053	17	хочуепимдьчйиймнхригриаку
26	19	0.053	18	хочуепимдьчйиймноригриаку
26	19	0.053	19	хочуепимдьчйиьмноригриаку
26	19	0.053	20	хочуепимдьчйипмноригриаку
26	19	0.053	21	хочуепимдьчйтпмноригриаку
26	19	0.053	22	хочуепимдьчйтумноригриаку
26	19	0.053	23	х чуепимдьчйтумноригриаку
26	19	0.053	24	х чуапимдьчйтумноригриаку
26	19	0.053	25	х чуапимдькйтумноригриаку
26	19	0.053	26	х чуапимдькйтумноринриаку

На Рис. 1 представлены наблюдаемые (эмпирические) частоты встречаемости букв в тексте после каждого изменения буквы. На Рис. 2 представлены частоты, вычисленные по формуле. Если посмотреть на 27 шаг зашумления, то наблюдаемые частоты встречаемости букв далеки от средней частоты 0,053. Частоты разбросаны от 0 до 0,115. При этом из 19 букв алфавита 6 имеют частоту 0,077, 8 букв с частотой 0,038 и 3 буквы не встречаются в тексте совсем. Частоты не равны между собой, однако сформировалось 3 группы с одинаковыми частотами, что может свидетельствовать о приближении текста к зашифрованному.

Другую картину можно наблюдать в теоретических частотах. Частоты букв находятся в пределах от 0,035 до 0,1, что гораздо ближе к среднему. При этом 14 из 19 букв алфавита имеют одинаковую вероятность 0,48, что может подтверждать то, что текст зашифрован.

№	А	Г	Д	Е	И	Й	К	М	Н	О	П	Р	Т	У	Х	Ч	Ш	Ь	ср. знач.
1	0.038	0.038	0.038	0.038	0.077	0.038	0.038	0.038	0.038	0.038	0.038	0.038	0.038	0.115	0.038	0.038	0.038	0.038	0.192
2	0.038	0.038	0.077	0.038	0.077	0.038	0.038	0.038	0.038	0.038	0.038	0.038	0.0	0.115	0.038	0.038	0.038	0.038	0.192
3	0.038	0.038	0.077	0.038	0.077	0.038	0.038	0.077	0.038	0.038	0.038	0.038	0.0	0.115	0.038	0.038	0.038	0.038	0.154
4	0.0	0.038	0.077	0.038	0.077	0.038	0.077	0.077	0.038	0.038	0.038	0.038	0.0	0.115	0.038	0.038	0.038	0.038	0.154
5	0.0	0.038	0.077	0.038	0.077	0.038	0.077	0.077	0.038	0.038	0.038	0.038	0.0	0.115	0.038	0.077	0.038	0.038	0.115
...																			
22	0.038	0.038	0.038	0.038	0.115	0.038	0.038	0.077	0.038	0.077	0.077	0.077	0.038	0.077	0.038	0.077	0.0	0.038	0.038
...																			
25	0.077	0.038	0.038	0.0	0.115	0.038	0.038	0.077	0.038	0.038	0.038	0.077	0.038	0.115	0.038	0.077	0.0	0.038	0.077
...																			
27	0.077	0.0	0.038	0.0	0.115	0.038	0.077	0.077	0.077	0.038	0.038	0.077	0.038	0.115	0.038	0.038	0.0	0.038	0.077

№	А	Г	Д	Е	И	Й	К	М	Н	О	П	Р	Т	У	Х	Ч	Ш	Ь	ср. знач.
1	0.038	0.038	0.038	0.038	0.077	0.038	0.038	0.038	0.038	0.038	0.038	0.038	0.038	0.115	0.038	0.038	0.038	0.038	0.192
2	0.039	0.039	0.076	0.039	0.076	0.039	0.039	0.039	0.039	0.039	0.039	0.039	0.002	0.113	0.039	0.039	0.039	0.039	0.187
3	0.04	0.04	0.075	0.04	0.075	0.04	0.04	0.04	0.04	0.04	0.04	0.04	0.004	0.11	0.04	0.04	0.04	0.04	0.181
4	0.04	0.04	0.074	0.04	0.074	0.04	0.04	0.04	0.04	0.04	0.04	0.04	0.006	0.108	0.04	0.04	0.04	0.04	0.176
5	0.041	0.041	0.073	0.041	0.073	0.041	0.041	0.041	0.041	0.041	0.041	0.041	0.008	0.106	0.041	0.041	0.041	0.041	0.171
...																			
22	0.047	0.047	0.063	0.047	0.063	0.047	0.047	0.047	0.047	0.047	0.047	0.047	0.031	0.079	0.047	0.047	0.047	0.047	0.111
...																			
25	0.047	0.047	0.062	0.047	0.062	0.047	0.047	0.047	0.047	0.047	0.047	0.047	0.033	0.076	0.047	0.047	0.047	0.047	0.104
...																			
27	0.048	0.048	0.061	0.048	0.061	0.048	0.048	0.048	0.048	0.048	0.048	0.048	0.035	0.074	0.048	0.048	0.048	0.048	0.1

Пример 4

Проанализируем язык ребенка для трех различных категорий: радость, плач и требование (просьба) с целью численной классификации «на хорошее или плохое содержание» искаженного открытого текста (выделения границ понимания зашумленного текста). Язык ребенка был выбран для упрощения задачи и сокращения числа букв в алфавите. Для каждой категории были проанализированы 25 фраз разной длины по 10 вариантов зашумлений для каждой длины (всего проанализировано 750 различных вариантов фраз и зашумлений). Для каждого из полученных вариантов фраз были выделены две границы зашумления:

- **Граница 1.** До какой замены совершенно ясно, к какой категории относится текст, после нее текст сложнее отнести к конкретной категории;
- **Граница 2.** С какой замены совершенно непонятно, к какой категории относится текст.

Задача состояла в определении границ и соотнесении полученных данных с теоретической частотой по формуле для выявления возможных закономерностей. Для удобства восприятия полученные результаты представлены в виде графиков. По горизонтальной оси указана длина текста, а по вертикальной – количество зашумлений. Точками на плоскости отмечено, при какой длине исходного текста на каком зашумлении выделена та или иная граница. Зеленым цветом показана Граница 1, а оранжевым – Граница 2.

На Рис. 3 представлена зависимость понимания текста от длины и числа зашумлений по каждой категории с выделением номера зашумления, при котором определена смысловая граница.

На Рис. 4 также представлена зависимость понимания текста от длины и числа зашумлений по каждой категории в процентном соотношении. То есть вертикальная ось показывает процент зашумления текста при выделении границы.

Как видно из графиков, с увеличением длины текста увеличивается количество зашумленных символов, при которых можно понять смысл текста. Однако графики увеличиваются не стабильно. Наблюдается то уменьшение числа зашумленных символов, то увеличение. Так для 1 границы в категории «требование (просьба)» процент зашумления текста находится в диапазоне от 15% до 77%, когда для «радости» разброс гораздо ниже – от 19% до 56%. Для «плача» диапазон для 1 границы примерно такой же, как и для «радости», составляет от 15% до 57%.

Такое нестабильное поведение графиков можно объяснить субъективной оценкой понимания текста. Для одного человека некоторые слова могут быть отнесены к нескольким категориям, а для другого только к одной.

На Рис. 5 представлены средние значения в процентах по границам зашумления для каждой из рассматриваемых категорий. Первое значение показывает, какой процент текста можно зашумить, чтобы было понятно, о чем этот текст, не терялась его смысловая нагрузка; второе – после какого процента зашумления невозможно разобрать, о чем говорится в тексте.

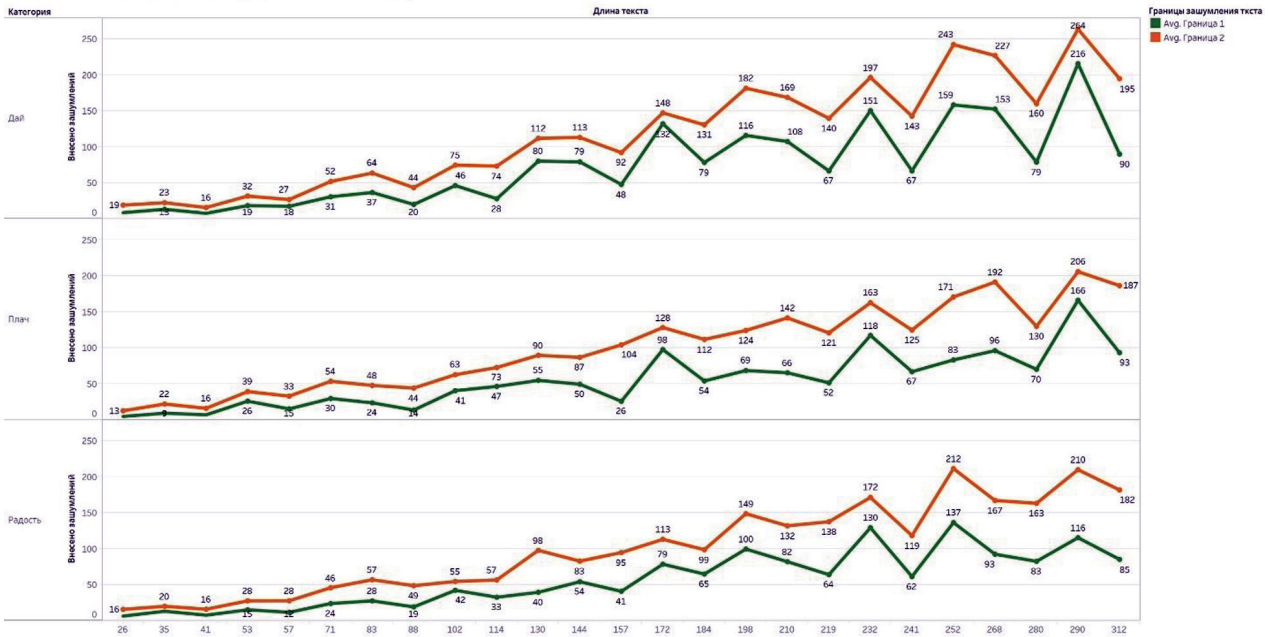
Данный рисунок подтверждает высказанное выше предположение, о том, что понимание текстов категорий «радости» и «плача» схоже. По обеим границам этих двух категорий понимание текстов ниже, чем в категории «требование (просьба)» на 10%. Это можно объяснить тем, что в категории «требование (просьба)» проще выделить слова, которые помогут понять смысл текста, а именно, что ребенок что-то просит. Категории «радость» и «плач» более эмоциональные. Из-за этого при зашумлении текстов из-за потери полных фраз в тексте трудно определить категорию. Например, слово «мама». Ребенок может радостно его выкрикивать, а может плакать, произнося это слово.

На Рис. 6 и 7 представлены объединённые результаты по всем трем категориям о зависимости понимания текста от длины и числа зашумлений. Для каждой длины текста было посчитано среднее количество зашумлений для 1 и 2 границ. На 6 рисунке представлено количество зашумлений в количественном значении, а на 7 рисунке – в процентном значении относительно длины текста.

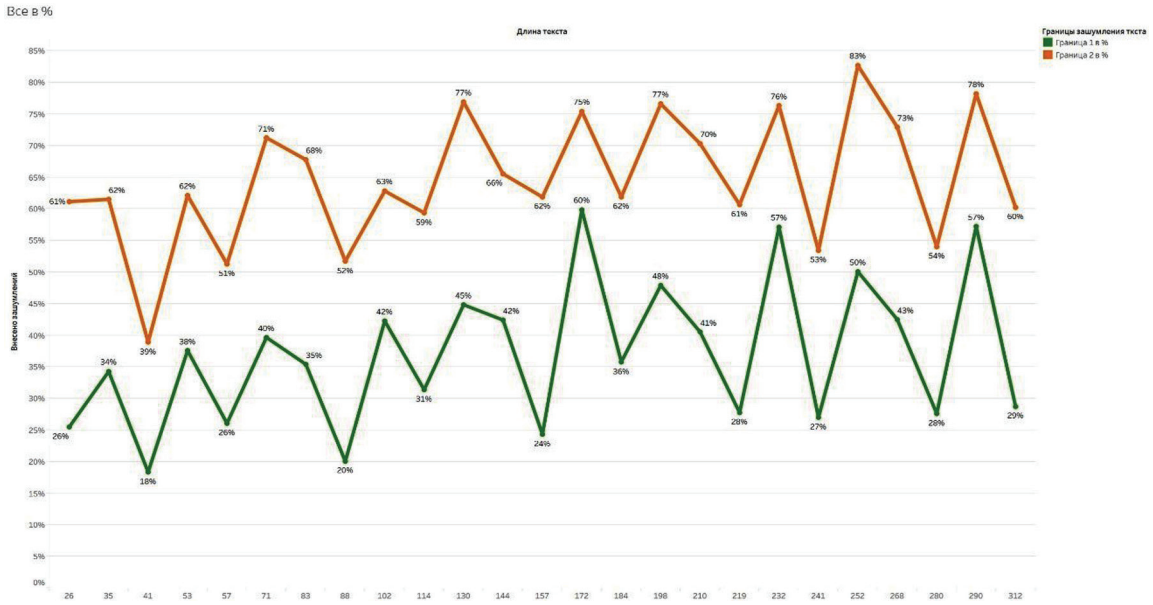
На данных графиках также не наблюдается прямой зависимости границ понимания текста от длины текста. Если посчитать средние границы понимания текстов по всем трем категориям, то для Границы 1 процент зашумления текста, до которого совершенно легко определить категорию, составляет 37%, а для Границы 2, после которой текст становится совершенно непонятен – 65%.

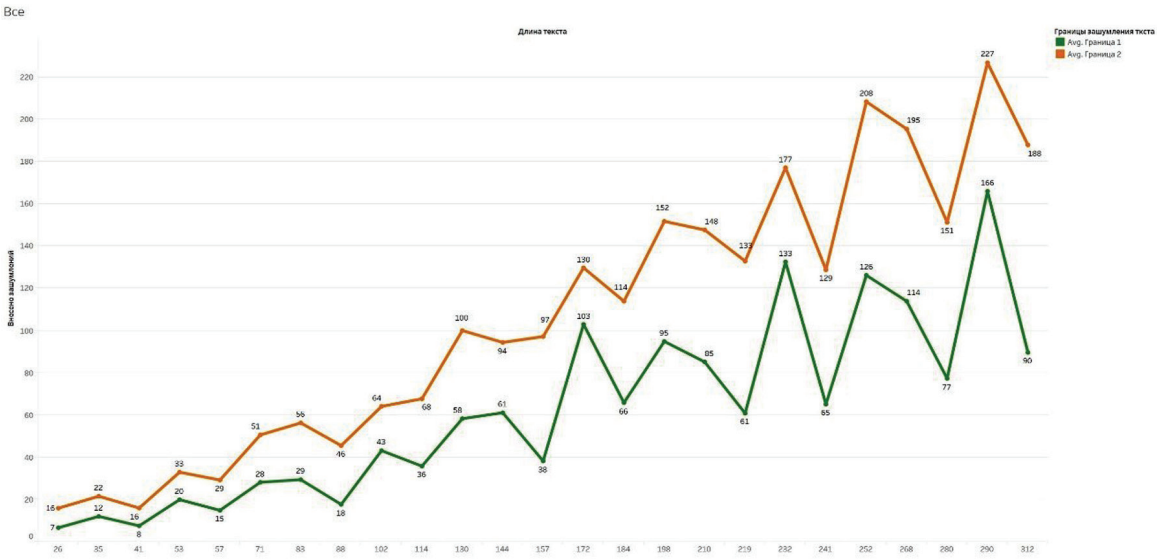
DOI:10.21681/2311-3456-2020-1-74-86

Зависимость понимания теста от длины и количества зашумлений



Зависимость в %





Средние границы понимания в % категории

Категория	
Дай	44% 71%
Плач	33% 62%
Радость	34% 61%

4. Выводы

В работе были сформулированы понятия открытого текста, зашумленного текста, слабого ключа, а также формализовано понятие содержательного текста при искажении. Были представлены способы определения содержания текста (дешифрование) при использовании слабых ключей. Показана большая надежность метода дешифрования шифра гаммирования при k-зашумленном слабом ключе. Качество содержания искаженного открытого текста было оценено при помощи специально выведенной формулы. Рассмотрены практические примеры зашумления текста и при помощи написанной программы на языке программирования python содержательного текста. На примерах получены две границы «читаемости» текста (численная классификация искажения содержательных текстов): граница первая, когда можно легко понять, к какой смысловой категории относится текст (она составила 37%), и граница вторая, когда становится невозможно распознать, о чем текст (она составила в среднем 65%).

Предполагалось, что при 30% зашумления текста еще можно разобрать, о чем он, но эти предположения строились относительно текста на русском языке, полученные же практические результаты объясняются тем, что придуманный язык из детской речи значительно меньше по алфавиту в отличие от алфавита русского языка. Частоты встречаемости букв распределены не так, как в русском языке, где существует множество смысловых категорий, а не только три. Это означает, что энтропия у искусственного языка меньше.

Практическая значимость работы, помимо криптографической составляющей, состоит в возможности применения разработанного программного приложения в образовательном процессе (для наглядного ознакомления с этапами зашумления открытого текста на искусственном языке, получения расчетов по наблюдаемым частотам встречаемости букв используемого алфавита и теоретическим частотам каждой буквы по выведенной в работе формуле, а также проведения различного рода исследований о границах понимания искаженного текста). Созданный искусственный язык может быть использован в различных исследованиях для упрощения задачи. Полученная численная классификация искажения содержательных текстов на примере текстов на языке ребенка может быть иначе интерпретирована: помимо выделенных границ понимания содержания текста, можно судить о возрасте ребенка, оценивать количество правильно произнесенных (не зашумлённых) слов и говорить об уровне развития говорящего.

Рецензент: *Чеповский Андрей Михайлович, доктор технических наук, профессор, профессор Национального исследовательского университета «Высшая школа экономики», г. Москва, Россия. E-mail: achepovsky@hse.ru*

Литература

1. Бабаш А.В. Криптография: учебное пособие / А.В. Бабаш, Г.П. Шанкин. М.:Солон-Р, 2002. 512 с.
2. Rubinstein-Salzedo S. The Vigenere Cipher / S. Rubinstein-Salzedo // Cryptography. 2018. №4. p. 41-54.
3. Venkata Subramaniam L. A survey of types of text noise and techniques to handle noisy
4. text key / L. Venkata Subramaniam [and etc.] // Analytics for Noisy Unstructured Text Data. 2009. №3. p. 115-122.
5. Колчин В.Ф. Случайные размещения: учебное пособие / В.Ф. Колчин [и др.]. М.: Наука, 1976. 224 с.
6. Бабаш А.В. Обобщенная модель шифра / А.В. Бабаш // Интеллектуальные системы в информационном противоборстве. 2015. №1. С.9-14.
7. Aized Amin Soofi. An Enhanced Vigenere Cipher For Data Security / Aized Amin Soofi [and etc.] // International journal of Scientific & Technology research. 2016. №3. p.141 – 145.
8. Kester Q. A. A cryptosystem based on Vigenere cipher with varying key / Q. A. Kester // International Journal of Advanced Research in ComputerEngineering & Technology. 2012. №10. p.108 – 113.
9. Maffre S. A Weak Key Test for Braid Based Cryptography / S. Maffre // Designs, Codes and Cryptography. 2006. №3(39). p. 347-373.
10. Berntsen M. C. Automating the cracking of simple ciphers: thesis / M. C. Berntsen.
11. Lewisburg: Bucknell University, 2005. 63 p.
12. Abiodun Esther Omolara. An Enhanced Practical Difficulty of One-Time Pad Algorithm
13. Resolving the Key Management and Distribution Problem / Abiodun Esther Omolara [and etc.] // Proceedings of the International MultiConference of Engineers and Computer Scientists 2018. 2018. №1. p.1 – 7.
14. Кормен Т. Алгоритмы: построение и анализ: учебное пособие / Т. Кормен [и др.]. М.:МЦНМО, 2002. 960 с.
15. Jean-Philippe Aumasson. Serious Cryptography. A Practical Introduction to Modern
16. Encryption: practical guide / Jean-Philippe Aumasson. San Francisco: no starch press, 2017. 312p.
17. Feng X. Reconstruction of noisy images via stochastic resonance in nematic liquid crystals / X. Feng [and etc.] // Scientific Reports. 2019. №3976. p.1 – 15.
18. Zhao H. H. Adaptive Block Compressive Sensing for Noisy Images / H. H. Zhao [and etc.] // Studies in Computational Intelligence. 2019. №2020. p.389 – 399.
19. Толстунов В.А. Нелинейный сглаживающий фильтр с показательно-степенными весами / В.А. Толстунов // Технические науки. 2015. №2(15). С.10-18.
20. Крашенинников В.Р. Зашумление эталонов в задачах обнаружения и распознавания сигналов на фоне помех / В.Р. Крашенинников, А.И. Армер // Вестник УлГТУ. 2004. №2. С.54-57.

21. Ferrand A. Using the NoiSee workflow to measure signal-to-noise ratios of confocal microscopes / A. Ferrand [and etc.] // Scientific Report. 2019. №1165. p.1-28.
22. Roy S. Fundamental noisy multiparameter quantum bounds / S. Roy // Scientific Reports. 2019. №1038. p.1 – 15.
23. Koohian A. Joint channel and phase noise estimation for mmWave full-duplex communication systems / A. Koohian [and etc.] // Eurasip Journal on Advances in Signal Processing. 2019. №18. p.1 – 12.
24. Lira de Queiroz W.J. Signal-to-noise ratio estimation for M-QAM signals in η−μand κ−μfading channels / W.J. Lira de Queiroz [and etc.] // Eurasip Journal on Advances in Signal Processing. 2019. №20. p.1 – 17.
25. Мирошниченко К.В. Организационные и технические мероприятия, направленные
26. на защиту информации ограниченного доступа / К.В. Мирошниченко, А.И. Киселев // Правоохранительная деятельность: теория и практика. 2018. №15. С.66 -70.
27. Дворников С.В. Методы предотвращения утечки информации из контролируемых помещений за счет побочных электромагнитных излучений и наводок / С.В. Дворников // Информационные технологии. 2018. №7(22). С.134-136.
28. Хорев А.А. Способы защиты объектов информатизации от утечки информации по техническим каналам: пространственное электромагнитное зашумление / А.А. Хорев // Автоматика. Вычислительная техника. 2012. №6. С.37-57.
29. Shaaban R. Visible light communication security vulnerabilities in multiuser network: power distribution and signal to noise ratio analysis / R. Shaaban [and etc.] // Lecture Notes in Networks and Systems. 2019. №2020. p.1 – 13.
30. Запечников С.В. Криптографические методы защиты информации: учебное пособие / С.В. Запечников [и др.]. М.: Юрайт, 2017. 309 с.
31. Васильева И.Н. Криптографические методы защиты информации: учебное пособие / И.Н. Васильева. М.: Юрайт, 2016. 64 с.
32. Жданов О.Н. Методы и средства криптографической защиты информации: учебное пособие / О.Н. Жданов, В.В. Золотарев. Красноярск: Сиб. ГАУ, 2007. 217 с.
33. ElSalamouny E. Optimal noise functions for location privacy on continuous regions / E.
34. ElSalamouny, S. Gambs // International Journal of Information Security. 2018. №17(6). p.613 – 630.

ABOUT TEXT NOISE BORDERS WITH THE TEXT CONTENT SAVING. APPLICATIONS TO CRYPTOGRAPHY

Babash A.V.[8], Baranova E.K.[9], Lyutina A.A.[10], Murzakova A.A.[11], Murzakova E.A.[12], Ryabova D.M.[13], Semis-ool E.S.[14]

Purpose: *to introduce mathematical model of a distorted meaningful text and a measure of its distortion, to define a numerical classification of the distortion of meaningful texts, present applications of the model in cryptography.*

Research methods: *a more complex Vigenere cipher decryption that uses an almost periodic key (noisy) is performed as decryption of noisy plaintext on a periodical key (by well-known Vigenere decryption methods), but with a different probability distribution of plaintext characters, with further lead of the task to the acceptable noise level in plaintext determining for understanding this text content.*

Results: *gamming cipher decryption ways with weak keys are presented. A new complexity was obtained and the reliability of the method was improved due to the fact that there k-noisy weak (almost periodic) keys are more than periodic ones. The formula for calculating the probability of occurrence of characters after the k-th noise was obtained. Artificial languages for ease of calculation were introduced and practical examples of text noise (the necessary*

8 Alexander Babash , Dr.Sc., Professor, National Research University Higher School of Economics, Moscow, Russia. E-mail: ababash@hse.ru

9 Elena Baranova, Associate professor, National Research University Higher School of Economics, Moscow, Russia. E-mail: ekbaranova@hse.ru

10 Anna Lyutina, student, Master's Programme National Research University Higher School of Economics, Moscow, Russia. E-mail: aalyutina@edu.hse.ru

11 Aleksandra Murzakova, student, Master's Programme National Research University Higher School of Economics, Moscow, Russia. E-mail: aamurzakova@edu.hse.ru

12 Ekaterina Murzakova, student, Master's Programme National Research University Higher School of Economics, Moscow, Russia. E-mail: eamurzakova@edu.hse.ru

13 Darya Ryabova, student, Master's Programme National Research University Higher School of Economics, Moscow, Russia. E-mail: dmryabova@edu.hse.ru

14 Elena Semis-ool, student, Master's Programme National Research University Higher School of Economics, Moscow, Russia. E-mail: essemisool@edu.hse.ru

DOI:10.21681/2311-3456-2020-1-74-86

calculations were made using a written program in the python programming language) were considered. The quality of the distorted plaintext content was assessed by highlighting of two borders of understanding.

***Keywords**: informative text; distorted text; noisy text; measure of text noise; decryption of cipher.*

References

1. Babash A.V. Cryptography: study guide / A.V. Babash, G.P. Shankin. M.: Solon-R, 2002. 512 p.
2. Rubinstein-Salzedo S. The Vigenere Cipher / S. Rubinstein-Salzedo // Cryptography. 2018. №4. p. 41-54.
3. Venkata Subramaniam L. A survey of types of text noise and techniques to handle noisy text key / L. Venkata Subramaniam [and etc.] // Analytics for Noisy Unstructured Text Data. 2009. №3. p. 115-122.
4. Kolchin V.F. Random placement: a tutorial / V.F. Kolchin [and etc.]. M.: Science, 1976. 224 p.
5. Babash A.V. Generalized cipher model / A.V. Babash // Intellectual systemsin the information confrontation. 2015. №1. p.9-14.
6. Aized Amin Soofi. An Enhanced Vigenere Cipher For Data Security / Aized Amin Soofi [and etc.] // International journal of Scientific & Technology research. 2016. №3. p.141 – 145.
7. Kester Q. A. A cryptosystem based on Vigenere cipher with varying key / Q. A. Kester //
8. International Journal of Advanced Research in Computer Engineering & Technology. 2012. №10. p.108 – 113.
9. Maffre S. A Weak Key Test for Braid Based Cryptography / S. Maffre // Designs, Codes and Cryptography. 2006. №3(39). p. 347-373.
10. Berntsen M. C. Automating the cracking of simple ciphers: thesis / M. C. Berntsen.
11. Lewisburg: Bucknell University, 2005. 63 p.
12. Abiodun Esther Omolara. An Enhanced Practical Difficulty of One-Time Pad Algorithm
13. Resolving the Key Management and Distribution Problem / Abiodun Esther Omolara [and etc.] // Proceedings of the International MultiConference of Engineers and Computer Scientists 2018. 2018. №1. p.1 – 7.
14. Kormen T. Algorithms: construction and analysis: a tutorial / T. Kormen [and etc.]. M.: MCCME, 2002. 960 p.
15. Jean-Philippe Aumasson. Serious Cryptography. A Practical Introduction to Modern
16. Encryption: practical guide / Jean-Philippe Aumasson. – San Francisco: no starch press, 2017. 312p.
17. Feng X. Reconstruction of noisy images via stochastic resonance in nematic liquid crystals / X. Feng [and etc.] // Scientific Reports. 2019. №3976. p.1 – 15.
18. Zhao H. H. Adaptive Block Compressive Sensing for Noisy Images / H. H. Zhao [and etc.] // Studies in Computational Intelligence. 2019. №2020. p.389 – 399.
19. Tolstunov V.A. Nonlinear smoothing filter with exponential power scales / V.A.
20. Tolstunov // Technical Sciences. 2015. №2 (15). p.10-18.
21. Krasheninnikov V.R. Pattern noise in the signal detection and recognition tasks with interference / V.R. Krasheninnikov, A.I. Armer // Vestnik of UlSTU. 2004. №2. p.54-57.
22. Ferrand A. Using the NoiSee workflow to measure signal-to-noise ratios of confocal microscopes / A. Ferrand [and etc.] // Scientific Report. 2019. №1165. p.1-28.
23. Roy S. Fundamental noisy multiparameter quantum bounds / S. Roy // Scientific Reports. 2019. №1038. p.1 – 15.
24. Koohian A. Joint channel and phase noise estimation for mmWave full-duplex communication systems / A. Koohian [and etc.] // Eurasip Journal on Advances in Signal Processing. 2019. №18. p.1 – 12.
25. Lira de Queiroz W.J. Signal-to-noise ratio estimation for M-QAM signals in η−μand κ−μfading channels / W.J. Lira de Queiroz [and etc.] // Eurasip Journal on Advances in Signal Processing. 2019. №20. p.1 – 17.
26. Miroshnichenko K.V. Organizational and technical activities aimedon the limited access information protection / K.V. Miroshnichenko, A.I. Kiselev // Law enforcement: theory and practice. 2018. №15. p.66 -70.
27. Dvornikov S.V. Prevention information leakage methods from controlled premises due to Transient Electromagnetic Pulse Emanation / S.V. Janitors // Information technologies. 2018. №7 (22). p.134-136.
28. Khorev A.A. information objects protection from information leakage bytechnical channels ways: spatial electromagnetic noise / A.A. Horev // Automation. Computer Engineering. 2012. №6. p.37-57.
29. Shaaban R. Visible light communication security vulnerabilities in multiuser network: power distribution and signal to noise ratio analysis / R. Shaaban [and etc.] // Lecture Notes in Networks and Systems. 2019. №2020. p.1 – 13.
30. Zapechnikov S.V. Information protecting cryptographic methods: a tutorial / S.V. Baking [and others]. M.: Yurayt, 2017. 309 p.
31. Vasilyeva I.N. Information protecting cryptographic methods: a tutorial / I.N. Vasiliev. M.: Yurayt, 2016. 64 p.
32. Zhdanov O.N. Cryptographic information protection methods and means: educational manual / O.N. Zhdanov, V.V. Zolotarev. – Krasnoyarsk: Sib. GAU, 2007. 217 p.
33. ElSalamouny E. Optimal noise functions for location privacy on continuous regions / E.
34. ElSalamouny, S. Gambs // International Journal of Information Security. 2018. №17(6). p.613 – 630.