УДК 003.26:004.056.55 DOI: 10.65599/III3531

ШИФРОВАНИЯ ТЕКСТОВЫХ ЭЛЕМЕНТОВ С ИСПОЛЬЗОВАНИЕМ МЕТОДА ВИЖЕНЕРА

М.Х. Гафуров

Таджикский технический университет имени академика М.С. Осими

Данная работа представляет собой о разработке нового метода шифрования текстовых объектов для повышения информационной безопасности и защиты от киберпреступников (взломщиков, хакеров). Предлагаемый подход отходит от традиционного использования нумерации отдельных символов (униграмм) в алфавите шифрования. Вместо этого в качестве основы для создания уникального (частного) алфавита используются пронумерованные лингвистические элементы, такие как корни слов, слоги, а также биграммы и триграммы. Алгоритм подробно описывает процесс разделения открытого текста на эти элементы, их произвольной нумерацию, и последующее применение закрытого ключа (на примере шифра Виженера). Демонстрируется, что использование уникальных (частных), многокомпонентных алфавитов значительно увеличивает криптографическую стойкость зашифрованного текста. Предложенный метод шифрования носит универсальный характер в отношении своего применения к различным языкам, но его конкретная реализация и эффективность тесно зависят от грамматических и лингвистических особенностей того языка, на котором написан исходный текст.

Ключевые слова: *объект, шифрование, зашифрование, расшифрование, частный алфавит, элемент, символ, ключ, вариант, стабильность, киберпреступность.*

БАДАЛСОЗИИ УНСУРҲОИ МАТН БО ИСТИФОДАИ УСУЛИ ВИЖЕНЕР

М.Ҳ. Ғафуров

Мақолаи мазкур дар бораи таҳияи усули нави бадалсозии объектҳои матнӣ ҷиҳати беҳтар кардани амнияти иттилоотӣ ва муҳофизат аз киберҷиноятқорон (қулфшиканон, хакерон) мебошад. Ҷараёни бадалсозии пешниҳодшуда аз истифодаи анъанавии рақамгузории символҳои алоҳида (унограммаҳо) дар алифбои бадалсозӣ сарфи назар мекунад. Ба ҷои он, унсурҳои забоншиносии (лингвистикии) рақамгузоришуда, аз қабили калимаҳои решагӣ, пешоянду пасояндҳо, ҳиҷою аъзоёни пайрави ҷумла ва биграммаю триграммаҳо ҳамчун асос барои таҳияи алифбои хусусии (беназири) бадалсозӣ истифода мешаванд. Дар алгоритм раванди ба унсурҳо тақсим кардани матни кушода, таҳияи алифбои хусусӣ, рақамгузории унсурҳои он ва истифодаи минбаъдаи калиди ихтиёрии бадалсозӣ (бо истифода аз усули бадалсозии Виженер дар мисоли объекти матнии кушода) муфассал тасвир шудааст. Нишон дода шудааст, ки истифодаи алифбои хусусии таҳияшудаи бисёркомпонентаи унсурҳояш ихтиёрӣ рақамгузоришуда устувории криптографии матни бадалшударо хеле баланд мекунад. Усули бадалсозии пешниҳодшуда нисбати тадбиқи забонҳои гуногун характери унверсалӣ дошта, истифодаи мушаххас ва самаранокии он ба хусусиятҳо хоси граматика ва лингивистикаи он забоне, ки матни додашуда иншо гардидааст алоқаи зич дорад.

Калидвожаҳо: *объект, бадалсозӣ, бадалкунӣ, аксбадалкунӣ, алифбои хусусӣ, унсур, аломат, калид, вариант, устуворӣ, киберҷиноят.*

ENCRYPTION OF TEXTUAL ELEMENTS USING THE VIGENÈRE METHOD

M.H. Gafurov

This work presents the development of a novel method for encrypting text objects to enhance information security and provide protection against cybercriminals (crackers, hackers). The proposed approach moves away from the traditional use of single-character (unigram) numbering in the encryption alphabet. Instead, it utilizes numbered linguistic elements-such as word roots, syllables, bigrams, and trigrams-as the foundation for creating a unique (private) alphabet. The algorithm details the process of decomposing plaintext into these elements, their arbitrary numbering, and the subsequent application of a private key (using the Vigenère cipher as an example). It is demonstrated that the use of unique, multi-component alphabets significantly increases the cryptographic strength of the ciphertext. While the proposed encryption method is universal in its applicability to various languages, its specific implementation and effectiveness are closely dependent on the grammatical and linguistic features of the source language.

Keywords: *object, encryption, enciphering, deciphering, private alphabet, element, symbol, key, variant, stability, cybercrime.*

Введение

В большинстве методов шифрования текстовых объектов используется последовательная нумерация символов стандартного алфавита языка. Этот процесс, начиная с древних времен (например, шифр Цезаря) и заканчивая некоторыми современными методами, применяет способ замены одного символа другим (**униграммное шифрование**). В таких методах каждый символ стандартного алфавита нумеруется, после чего на основе определенного правила создается

секретный ключ и осуществляется процесс шифрования. Данные подходы подробно рассмотрены в работах [1-14].

Использование мощных электронных устройств (суперкомпьютеров, квантовых компьютеров) позволяет киберпреступникам и заинтересованным лицам проводить **частотный анализ** символов текста и восстанавливать исходный (открытый) текст из зашифрованного.

Для решения вышеуказанной проблемы предлагается метод создания **индивидуального алфавита**, элементы которого извлекаются непосредственно из самого текста. Такой алфавит формируется уникально для каждого отдельного документа.

В научной источниках этот вопрос рассматривался следующим образом. В работах [15-18] изучена подготовка базы предлогов, послеслогов, морфологический анализ слов таджикского литературного языка и оценка использования биграмм. В работе [19] приведено использование униграмм, способ создания частного алфавита на их основе и разработка варианта произвольного закрытого ключа. В работе [20] рассмотрен способ создания уникального алфавита шифрования. В работе [21] описан метод шифрования текстового объекта с использованием элементов текста, где закрытый ключ создается **стохастическим способом**. В работах [22-24] рассмотрен метод шифрования текста с применением **шифра Полибия**, использующий элементы текста (корневые слова, предлоги, послеслоги), а также левосторонние биграммы и триграммы. В работах [25, 26] представлены матричные и операторно-матричные методы шифрования. В работе [27] описано создание ключей на основе элементов текста и многозначных чисел. В работе [28] рассматривается усовершенствование **метода Виженера**. В работе [29] изучено шифрования текстовых объектов с использованием стохастической нумерации их элементов.

Алгоритм решения задачи

На основе вышеупомянутых работ рассмотрим методы шифрования заданного открытого текстового объекта, элементы которого нумеруются стохастическим (произвольным) методом.

Алгоритм решения рассматриваемой задачи заключается в следующем.

1. Задается, создается или выбирается открытый текстовый объект на произвольном языке. Предположим, он выглядит следующим образом:

$$\mathcal{B} = \{ b_i, i = \overline{1, n};\ b_i \in \mathcal{B} \} \quad (1)$$

2. Открытый текст $\mathcal{B}$ разделяется на следующие группы в соответствии с морфологическим анализом языка текста:

а) элементы, состоящие из корневых слов, предлогов, послеслогов, аффиксов и остальных членов предложения;

б) элементы, состоящие из слогов и остальных членов предложения;

в) элементы, состоящие из *N*-грамм (биграмм, триграмм и т.д.).

3.Для обеспечения корректного восстановления структуры объекта при расшфровании (обратном преобразовании), перед началом процесса шифрования создается вспомогательный ключ K_1. В нем все небуквенные знаки (символ абзаца, пробел, цифры и т.д.) заменяются символами из кодировок ASCII или Unicode, которые отсутствуют в исходном текстовом объекте. Для разграничения элементов зашифрованного объекта в пунктах **а)** и **б)** создается вспомогательный ключ K_2, который сохраняет структуру разделенных элементов. Использование вспомогательных ключей является одним из уровней защиты текстового объекта.

4.Выбрав одну из групп (**а** или **б**) из пункта 2, разделяем исходный открытый текст на элементы согласно требованиям морфологического анализа. Затем, применяя первый вспомогательный ключ K_1, представляем его в виде последовательности символов $\mathcal{B}1 = F(\mathcal{B}, K_1)$, что считается первым уровнем защиты текстового объекта.

В случае разделения текстового объекта на N-граммы (биграммы, триграммы и т.д.), если в конце объекта они оказываются неполными, используется пустой символ (varnothing) для их дополнения. Поскольку структура N-грамм в исходном и зашифрованном объектах не меняется, вспомогательный ключ K_2 не применяется. Также при разделении текстового объекта на левосторонние N-граммы полностью используются все имеющиеся в нем символы, знаки и цифры.

5.Используя последовательность элементов в объекте $\mathcal{B}1$, формируем частное множество, состоящее из неповторяющихся элементов. Это множество используется в качестве **частного алфавита шифрования** для данного текстового объекта и имеет следующий вид:

$$F = \{b1_i,\ i = \overline{1, n}\ ;\ \ b1_i \in \mathcal{B}1\} \tag{2}$$

6. Основной закрытый ключ преобразования K_a выбирается произвольно в зависимости от выбранного метода. В методе шифрования Виженера в качестве ключа может использоваться последовательность символов, знаков, цифр или их комбинация, также могут быть использованы слова. Пусть ключ имеет следующий вид:

$$K_a = \{\ a_j\ ,\ j = \overline{1, n1}\ \} \tag{3}$$

7. Согласно требованиям пункта 2 алгоритма, основной закрытый ключ шифрования Ка разделяется на элементы в зависимости от выбранного способа. При необходимости к нему также применяются вспомогательные ключи К1 и К2. В результате ключ шифрования для текстового объекта, разделенного на элементы, принимает следующий вид:

$$K1_a = \{\ a1_j,\ j = \overline{1, m}\ \} \tag{4}$$

8. Элементы основного закрытого ключа шифрования К1а, которые отсутствуют в частном алфавите, добавляются в него. Таким образом, мы получаем расширенный частный алфавит, который имеет следующий вид:

$$F1 = \{b1_{i+k},\ i = \overline{1, n}\ ;\ \ k = \overline{1, m}\ ;\ b1_{i+k} \in \mathcal{B}1 U K1_a\} \tag{5}$$

9. Элементы расширенного частного алфавита $F1$ нумеруются в произвольном порядке. Количество возможных вариантов нумерации алфавита $F1$ зависит от числа входящих в него элементов: чем их больше, тем выше криптостойкость зашифрованного объекта. Например, если через U обозначить варианты нумерации элементов в расширенном частном алфавите, то $U = (n+m)!$, что представляет собой ещё один уровень повышения защиты текстового объекта.

10. Используя выбранный произвольный закрытый ключ $K1_a$ и пронумерованные элементы расширенного частного алфавита $F1$, осуществляем зашифрование объекта $\mathcal{B}1$, состоящего из последовательности элементов, в соответствии с выбранным методом (метод Виженера). В результате получаем зашифрованный объект $\mathcal{B}2 = F(\mathcal{B}1,\ K1_a)$. В методе Виженера элементы зашифрованного объекта определяются по следующей формуле:

$$c_i = b1_i + a1_i(mod(n + m)), \quad i = 1, 2, 3, \ldots \tag{6}$$

где, $b1_i$ - элементы исходного объекта в соответствии с их порядковыми номерами в частном алфавите шифрования; c_i - элементы зашифрованного объекта в соответствии с их порядковыми номерами в частном алфавите; $a1_i$ - элементы произвольно выбранного ключа в соответствии с их порядковыми номерами в частном алфавите; $(n + m)$ - период алфавита шифрования (общее количество элементов в расширенном частном алфавите).

11.Для расшифровании зашифрованного объекта $\mathcal{B}2$ в открытый объект $\mathcal{B}$ (законное расшифрования), имея доступ к основному и вспомогательным ключам, расширенному частному алфавиту с пронумерованными элементами и методу шифрования, достаточно выполнить вышеуказанные пункты в обратном порядке.

В методе Виженера для легального (законного) расшифрования зашифрованного объекта используется следующая формула:

$$b1_i = c_i - a1_i(mod(n + m)), \quad i = 1, 2, 3, \ldots \tag{7}$$

Решение задачи на примерах

Применим процесс шифрования текстовых объектов для трёх случаев разделения на элементы, используя метод Виженера.

1. Шифрования текстового объекта с использованием корневых слов и других элементов.

Пусть текстовый объект задан в следующем виде (рубаи Омара Хайяма):

$$\mathscr{B} = \left\{ \begin{array}{c} \text{Гар як нафасат зи зиндагонӣ гузарад,} \\ \text{Магзор, ки ҷуз ба шодмони гузарад.} \\ \text{Зинҳор ки сармояи ин мулки вуҷуд,} \\ \text{Умр асту чунон к — аш гузарони, гузарад.} \end{array} \right\} \quad (A)$$

Теперь, основываясь на тексте объекта и согласно требованиям пункта 3 алгоритма, сформируем первый вспомогательный ключ, который примет следующий вид:

$$K1 = \{ \text{ʼ} \to z, \text{ʽ} \to g, \text{ʻ} \to s, - \to f, \hookleftarrow \to d \} \quad (8)$$

Применяя первый вспомогательный ключ ***K1*** к объекту $\mathscr{B}$, разделим текст в соответствии с правилами грамматики таджикского языка на текстовые элементы - корневые слова, предлоги и послелоги, аффиксы, второстепенные члены предложения и знаки пунктуации. Результат примет следующий вид:

$$\mathscr{B}1 = \left\{ \begin{array}{c} \text{Гар s як s нафас ат s зи s зинда гон ӣ s гузар ад z d} \\ \text{Магзор z s ки s ҷуз s ба s шод мо ни s гузар ад g d} \\ \text{Зинҳор s ки s сар моя и s ин s мулк и s вуҷуд z d} \\ \text{Умр s аст у s чун он s к f аш s гузар он и z s гузар ад g} \end{array} \right\} \quad (9)$$

Как видно, общее количество элементов в текстовом объекте $\mathscr{B}1$, представленном в виде последовательности, равно 70, при этом некоторые элементы повторяются несколько раз.

Используя текстовый объект $\mathscr{B}1$ и следуя требованиям пункта 4 алгоритма, сформируем один из вариантов **частного алфавита шифровании**. Элементы данного алфавита пронумеруем в произвольном порядке, как показано в Таблице 1.

Таблица 1 – Частный алфавит шифровании, состоящий из корневых слов и дополнительных элементов

1	2	3	4	5	6	7	8	9	10	11	12	13
Гар	s	як	у	мулк	зи	зинда	гон	ӣ	гузар	ад	чун	d

14	15	16	17	18	19	20	21	22	23	24	25
Магзор	ат	ҷуз	ба	шод	Умр	ни	g	он	нафас	к	и

26	27	28	29	30	31	32	33	34	35	36
ин	ки	вуҷуд	мо	сар	z	Зинҳор	моя	f	аш	аст

Теперь для применения метода Виженера представим таблицу, состоящую из трёх строк, количество столбцов которой равно количеству элементов в объекте. В первую строку последовательно вносятся элементы исходного объекта. Во вторую строку циклически, до последней ячейки, вносятся элементы произвольно выбранного основного ключа.

Примечание. В качестве основного ключа может использоваться последовательность символов, знаков, цифр или их комбинация, а также отдельные слова. При разделении произвольного основного ключа на элементы те из них, которые отсутствуют в частном множестве (алфавите), добавляются в него и нумеруются. В этом случае созданный алфавит называется **расширенным частным алфавитом шифровании**.

В третьей строке размещаются элементы зашифрованного объекта, которые определяются согласно формуле (6).

Пусть произвольно выбранный основной ключ шифрования будет следующим:

$$K_a = \{ \text{ҳаёти осуда} \} \quad (10)$$

Затем разделим основной ключ на элементы - «ҳаёт, и, осуда». После применения первого вспомогательного ключа ***K1*** он примет следующий вид:

$$K_a = \{ \text{ҳаёт-и-s-осуда} \} \quad (10.1)$$

Из разделенных элементов ключа следует, что слова «ҳаёт» и «осуда» отсутствуют в частном множестве шифровании. Мы добавляем их в произвольные позиции частного алфавита шифровании, нумеруем их и получаем один из вариантов **пронумерованного расширенного частного алфавита шифровании** следующего вида (Таблица 2):

Таблица 2 – Расширенный частный алфавит шифровании, состоящий из корневых слов и дополнительных элементов

1	2	3	4	5	6	7	8	9	10	11	12	13	14
Гар	s	як	у	мулк	зи	зинда	гон	ӣ	гузар	ад	чун	d	ат

15	16	17	18	19	20	21	22	23	24	25	26	27
f	ҳаёт	ҷуз	ба	шод	Умр	ни	g	он	нафас	к	и	осуда

28	29	30	31	32	33	34	35	36	37	38
ин	ки	вуҷуд	мо	сар	z	Зинҳор	моя	Магзор	аш	аст

Из представленного варианта расширенного частного алфавита шифровании следует, что он состоит из 38 элементов. Это означает, что количество возможных вариантов нумерации элементов расширенного алфавита составляет $\mathbf{38! \approx 5.23*10^{44}}$, что существенно повышает криптостойкость системы.

Теперь, построив трехстрочную таблицу, приступим к зашифрованию заданного текстового объекта. Для зашифрования исходного текста будем использовать произвольно выбранный ключ (10.1), числовые значения элементов из Таблицы 2 и формулу (6). В результате имеем (Таблица 3):

Таблица 3 – Процесс зашифрования текстового объекта с использованием корневых слов

Гар	s	як	s	нафас	ат	s	зи	s	зинда	гон	ӣ	s	гузар
ҳаёт	и	s	осуда	ҳаёт	и	s	осуда	ҳаёт	и	s	осуда	ҳаёт	и
ҷуз	ин	мулк	ки	s	s	у	z	ба	z	гузар	Магзор	ба	Магзор

ад	z	d	Магзор	z	s	ки	s	ҷуз	s	ба	s	шод	мо
s	осуда	ҳаёт	и	s	осуда	ҳаёт	и	s	осуда	ҳаёт	и	s	осуда
d	g	ки	нафас	моя	ки	зинда	ин	шод	ки	Зинҳор	ин	ни	Умр

ни	s	гузар	ад	g	d	Зинҳор	s	ки	s	сар	моя	и
ҳаёт	и	s	осуда	ҳаёт	и	s	осуда	ҳаёт	и	s	осуда	ҳаёт
аш	ин	чун	аст	аст	Гар	Магзор	ки	зинда	ин	Зинҳор	нафас	у

s	ин	s	мулк	и	s	вуҷуд	z	d	Умр	s	аст	у	s	чун
и	s	осуда	ҳаёт	и	s	осуда	ҳаёт	и	s	осуда	ҳаёт	и	s	осуда
ин	вуҷуд	ки	ни	ат	у	шод	ад	Гар	g	ки	ҳаёт	вуҷуд	у	Гар

он	s	к	f	аш	s	гузар	он	и	z	s	гузар	ад	g
ҳаёт	и	s	осуда	ҳаёт	и	s	осуда	ҳаёт	и	s	осуда	ҳаёт	и
Гар	ин	осуда	у	f	ин	чун	чун	у	ни	у	аш	осуда	гузар

В процессе извлечения элементов третьей строки Таблицы 3, которые представляют собой текстовые элементы зашифрованного объекта, выполняются следующие действия: создается второй вспомогательный ключ ***K2***, который необходим для сохранения структуры каждого зашифрованного

элемента в процессе расшифрования и при извлечении элементов третьей строки между ними вставляются символы из ***К2***, которые не входят в состав частного алфавита.

Примечание. После завершения процесса расшифрования и извлечения элементов исходного объекта из третьей строки, символы второго вспомогательного ключа исключаются.

Пусть второй вспомогательный ключ имеет следующий вид:

$$К2 = \{*, ?, /, !, \$, €\} \quad (11)$$

С учетом второго вспомогательного ключа, последовательно извлекая полученные элементы из ячеек третьей строки Таблицы 3, формируем текст зашифрованного объекта $В2 = F(В1; К1, К_а, К2)$, который принимает следующий вид:

$$В2 = \left\{ \begin{array}{c} ҷуз * ин? мулк/ки! s$s€y? z! ба€z * гузар/Магзор$ба \\ * Магзор? d/g! ки€нафас? моя$ки$зинда/ин! шод€ки? \\ Зинҳор$ин? ни€Умр! аш$ин? чун$аст€аст$Гар! Магзор \\ ? ки$зинда€ин/Зинҳор! нафас? у$ин/вуҷуд! ки€ни/ат \\ $у? шод€ад? Гар! g$ки€ҳаёт$вуҷуд€у? Гар! Гар€ин$ \\ осуда? у/f€ин/чун! чун/у$ни€у$аш! осуда$гузар \end{array} \right\} \quad (12)$$

2. Шифрование текстового объекта с использованием слогов и других элементов

Рассмотрим второй случай зашифрования заданного текстового объекта. В данном случае основными элементами являются **слоги**, а дополнительными - второстепенные члены предложения, знаки пунктуации и специальные символы.

В этом случае также применяется первый вспомогательный ключ (8) к заданному текстовому объекту, после чего слова, содержащиеся в нем, разделяются на слоги, т.е. имеем:

$$В3 = \left\{ \begin{array}{c} Гар s як s на фа сат s зи s зин да го нӣ s гу за рад z d \\ Маг зор z s ки s ҷуз s ба s шод мо ни s гу за рад g d \\ Зин ҳор s ки s сар мо я и s ин s мул ки s ву ҷуд z d \\ Умр s ас ту s чу нон s к f аш s гу за ро ни z s гу за рад g \end{array} \right\} \quad (13)$$

Общее количество элементов в текстовом объекте ***В3***, представленном в виде последовательности, равно 80, при этом некоторые из них повторяются несколько раз.

На основе текстового объекта ***В3*** сформируем частный алфавит шифровании. Элементы данного алфавита пронумеруем в произвольном порядке, как показано в Таблице 4.

Таблица 4 – Частный алфавит шифровании на основе слогов и дополнительных элементов

1	2	3	4	5	6	7	8	9	10	11	12	13	14
рад	ин	сат	як	ки	на	Зин	нӣ	за	го	ҳор	Маг	ни	ҷуз

15	16	17	18	19	20	21	22	23	24	25	26	27	28
s	гу	чу	и	шод	ба	сар	мул	ҷуд	z	зор	аш	Умр	d

29	30	31	32	33	34	35	36	37	38	39	40	41	42
ро	Гар	мо	фа	ас	ту	зи	я	f	ву	нон	к	да	g

Теперь выберем произвольный основной ключ шифрования. Пусть он, как и прежде, будет $К_а$ **= { ҳаёти осуда }**. Разделим его на слоги: $К_а$ **= { ҳа-ё-ти ‿ о-су-да }**. После применения первого вспомогательного ключа ***К1*** основной ключ примет следующий вид:

$$К_а = \{ ҳа\text{-}ё\text{-}ти\text{-}s\text{-}о\text{-}су\text{-}да \} \quad (14)$$

Из разделенных элементов ключа следует, что слоги «ҳа, ё, ти, о, су» отсутствуют в частном множестве шифровании. Добавляем их в произвольные позиции частного алфавита шифровании, нумеруем их и получаем один из вариантов **пронумерованного расширенного частного алфавита шифровании**, который представлен в Таблице 5.

Таблица 5 – Расширенный частный алфавит шифровании на основе слогов и дополнительных элементов

1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
рад	ин	сат	як	ки	на	Зин	ҳа	нӣ	за	го	ҳор	Маг	ни	ҷуз	s

17	18	19	20	21	22	23	24	25	26	27	28	29	30	31
ти	гу	чу	и	шод	ба	сар	мул	ё	ҷуд	z	зор	аш	Умр	d

32	33	34	35	36	37	38	39	40	41	42	43	44	45	46	47
ро	о	Гар	мо	фа	ас	су	ту	зи	я	f	ву	нон	к	да	g

Из представленного варианта расширенного частного алфавита шифровании следует, что он состоит из 47 элементов. Таким образом, количество возможных вариантов нумерации этого алфавита составляет $\mathbf{47! \approx 2.59*10^{59}}$.

Для зашифрования заданного объекта, разделенного на элементы, используем произвольно выбранный основной ключ (14), числовые значения элементов из Таблицы 5 и формулу (6). Результаты вычислений приведены в Таблице 6.

Таблица 6 – Процесс зашифрования текстового объекта с использованием слогов

Гар	s	як	s	на	фа	сат	s	зи	s	зин	да	го	нӣ	s	гу
ҳа	ё	ти	s	о	су	да	ҳа	ё	ти	s	о	су	да	ҳа	ё
f	я	шод	ро	ту	мул	ин	мул	гу	о	сар	ро	ин	ҳа	мул	ву

за	рад	z	d	Маг	зор	z	s	ки	s	ҷуз	s	ба	s	шод	мо
ти	s	о	су	да	ҳа	ё	ти	s	о	су	да	ҳа	ё	ти	s
z	ти	Маг	ба	и	фа	ки	о	шод	ин	на	ҷуз	нӣ	я	су	як

ни	s	гу	за	рад	g	d	Зин	ҳор	s	ки	s	сар	мо	я	и	s
о	су	да	ҳа	ё	ти	s	о	су	да	ҳа	ё	ти	s	о	су	да
g	Зин	ти	гу	ҷуд	ти	g	зи	сат	ҷуз	Маг	я	зи	як	z	ти	ҷуз

ин	s	мул	ки	s	ву	ҷуд	z	d	Умр	s	ас	ту	s	чу
ҳа	ё	ти	s	о	су	да	ҳа	ё	ти	s	о	су	да	ҳа
за	я	я	шод	ин	Гар	ё	мо	нӣ	g	ро	сар	Умр	ҷуз	z

нон	s	к	f	аш	s	гу	за	ро	ни	z	s	гу	за	рад	g
ё	ти	s	о	су	да	ҳа	ё	ти	s	о	су	да	ҳа	ё	ти
ба	о	ни	зор	и	ҷуз	ҷуд	мо	ин	Умр	Маг	Зин	ти	гу	ҷуд	ти

С учетом второго вспомогательного ключа ***К2***, последовательно извлекая полученные элементы из ячеек третьей строки Таблицы 6, формируем текст зашифрованного объекта ***В4 = F(В3; К1, К$_а$, К2)***, который имеет следующий вид:

$$\mathbf{В4} = \left\{ \begin{array}{c} \text{f\$я} * \text{шод! ро/ту? мул} * \text{ин/мул\$гу! о? сар} * \text{ро! ин/} \\ \text{ҳа? мул} * \text{ву/z? ти\$Маг/ба! и} * \text{фа/ки! о? шод\$ин! на} \\ * \text{ҷуз/нӣ? я! су/як\$g! Зин} * \text{ти/гу? ҷуд! ти\$g€зи/сат} \\ * \text{ҷуз? Маг! я€зи\$як/z} * \text{ти! ҷуз? за€я\$я/шод} * \text{ин? Гар} \\ \text{! ё\$мо/нӣ} * \text{g! ро? сар\$ Умр/ҷуз} * \text{z€ба? о\$ни/зор! и} * \\ \text{ҷуз/ҷуд? мо/ин\$Умр} * \text{Маг? Зин\$ти/гу? ҷуд} * \text{ти} \end{array} \right\} \quad (15)$$

3. Шифрование текстового объекта с использованием биграмм

Рассмотрим процесс шифрования текстового объекта с использованием биграмм. В данном случае также применяется первый вспомогательный ключ (8) к заданному текстовому объекту, после чего объект разделяется на левосторонние биграммы (последовательности из двух символов). В результате получаем объект $\mathcal{B}5$ следующего вида:

$$\mathcal{B}5 = \begin{Bmatrix} \text{Га ps як sн аф ас ат sз иs зи нд аг он ӣs гу за ра дz} \\ \text{dM аг зо pz sк иs ҷу зs ба sш од мо ни sг уз ар ад gd} \\ \text{Зи нҳ ор sк иs са рм оя иs ин sм ул ки sв уҷ уд zd Ум} \\ \text{ps ас ту sч ун он sк fa шs гу за ро ни zs гу за ра дg} \end{Bmatrix} \quad (16)$$

Общее количество элементов в текстовом объекте $\mathcal{B}5$, представленном в виде последовательности биграмм, равно 72, при этом некоторые из элементов повторяются несколько раз.

На основе текстового объекта $\mathcal{B}5$ сформируем частный алфавит шифровании. Элементы данного алфавита пронумеруем в произвольном порядке, как показано в Таблице 7.

Таблица 7 – Частный алфавит шифровании на основе биграмм

1	2	3	4	5	6	7	8	9	10	11	12	13	14
нд	рм	fa	за	ps	од	sв	ту	аф	зо	sм	ун	он	dM

15	16	17	18	19	20	21	22	23	24	25	26	27	28
sк	zs	иs	ин	гу	ат	sш	ор	ад	шs	ас	ро	як	Га

29	30	31	32	33	34	35	36	37	38	39	40	41	42
нҳ	ул	Зи	sн	pz	ни	оя	ра	аг	ҷу	Ум	sч	зs	мо

43	44	45	46	47	48	49	50	51	52	53	54	55	56
sз	ба	са	ӣs	sг	ки	уд	ар	gd	уҷ	дz	уз	zd	дg

Теперь выберем произвольный основной ключ преобразования. Пусть он, как и прежде, будет K_a = { *ҳаёти осуда* }. Применим первый вспомогательный ключ (8) к основному ключу K_a и разделим его на левосторонние биграммы, в результате чего он примет следующий вид:

$$K_a = \{ \text{ҳа-ёт-иs-ос-уд-а}\varnothing \} \quad (17)$$

В последнем элементе основного ключа K_a для завершения формирования левосторонней биграммы добавлен символ «∅» - пустая позиция.

Из выделенных элементов ключа K_a следует, что биграммы «ҳа, ёт, ос, а∅» не были включены в предварительный частный алфавит шифровании. Мы добавляем их в произвольные позиции частного алфавита шифрования, нумеруем их и получаем один из вариантов **расширенного частного алфавита шифровании на основе биграмм**, представленный в Таблице 8.

Таблица 8 – Расширенный частный алфавит шифровании на основе биграмм

1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
нд	рм	fa	за	ps	од	sв	ёт	ту	аф	зо	sм	ун	он	dM

16	17	18	19	20	21	22	23	24	25	26	27	28	29	30
sк	zs	иs	ин	а∅	гу	ат	sш	ор	ад	шs	ас	ро	як	Га

31	32	33	34	35	36	37	38	39	40	41	42	43	44	45
нҳ	ул	Зи	sн	pz	ни	оя	ра	аг	ҳа	ҷу	Ум	sч	зs	мо

46	47	48	49	50	51	52	53	54	55	56	57	58	59	60
sз	ба	ос	са	ӣs	sг	ки	уд	ар	gd	уҷ	дz	уз	zd	дg

Общее количество элементов расширенного частного алфавита шифрования равно 60, а количество вариантов их нумерации составляет $\mathbf{60! \approx 8.32*10^{81}}$.

Для зашифрования данного объекта, разделенного на биграммы, используя произвольно выбранный основной ключ (17), номера элементов из Таблицы 8 и формулу (6), получаем результат, представленный в Таблице 9.

Таблица 9 – Процесс шифрования текстового объекта с использованием биграмм

Га	ps	як	sн	аф	ас	ат	sз	ив	зи	нд	аг	он	йs	гу	за	ра	дz
ҳа	ёт	ив	ос	уд	а⌀	ҳа	ёт	ив	ос	уд	а⌀	ҳа	ёт	ив	ос	уд	а⌀
аф	ун	ба	ат	fa	ба	рм	ар	ни	гу	ар	zd	ар	уз	аг	ки	нҳ	zs

dM	аг	зо	pz	sк	ив	ҷу	зs	ба	sш	од	мо	ни	sг	уз	ар	ад	gd
ҳа	ёт	ив	ос	уд	а⌀	ҳа	ёт	ив	ос	уд	а⌀	ҳа	ёт	ив	ос	уд	а⌀
gd	ба	як	sш	ту	ра	гу	ки	ps	зо	zd	ps	sк	zd	sк	Ум	ив	dM

Зи	нҳ	ор	sк	ив	са	рм	оя	ив	ин	sм	ул	ки	sв	уҷ	уд	zd	Ум
ҳа	ёт	ив	ос	уд	а⌀	ҳа	ёт	ив	ос	уд	а⌀	ҳа	ёт	ив	ос	уд	а⌀
ун	аг	Ум	за	зо	ту	Ум	мо	ни	sв	ps	ки	ул	dM	он	ҷу	ки	рм

ps	ас	ту	sч	ун	он	sк	fa	шs	гу	за	ро	ни	zs	гу	за	ра	дg
ҳа	ёт	ив	ос	уд	а⌀	ҳа	ёт	ив	ос	уд	а⌀	ҳа	ёт	ив	ос	уд	а⌀
мо	pz	ас	нҳ	од	sн	уҷ	зо	зs	ту	дz	ос	sк	ад	аг	ки	нҳ	а⌀

Последовательно извлекая полученные элементы из третьей строки Таблицы 9, получаем зашифрованный объект $\mathcal{B}6 = F(\mathcal{B}5; K1, K_a)$, который принимает следующий вид:

$$\mathcal{B}6 = \begin{Bmatrix} \text{афунбаатfабармарнигуарzdарузагкинҳzsg} \\ \text{dбаякштурагукирsзozdpsskzdskУмisdМу} \\ \text{нагУмзазотуУммонисврскиулdМонҷукир} \\ \text{мморzаснҳодsнуҷзозstудzосскадагкинҳа⌀} \end{Bmatrix} \quad (18)$$

Из применения биграмм при зашифроавние данного текстового объекта следует, что второй вспомогательный ключ ***K2*** для разделения элементов зашифрованного объекта не использовался. Действительно, при разделении зашифрованного объекта на левосторонние биграммы и их использовании в процессе расшифрования необходимость в разделительных символах отпадает, так как во всех случаях структура биграмм (***N***-грамм) остается неизменной, и зашифрованный объект легко и безошибочно разделяется на пары символов (левосторонние биграммы).

Для восстановления зашифрованных объектов к исходному (процесс законного расшифрования) во всех трёх случаях пункта 2, согласно пункту 11 алгоритма, достаточно получить основные и вспомогательные ключи (для подпунктов а) и б) - основные ключи K_a и вспомогательные ***K1, K2***; для подпункта в) - основные ключи K_a и вспомогательный ***K1***), нумерованный расширенный частный алфавит и метод шифрованиия, а затем выполнить вышеуказанные пункты в обратном порядке (от конца к началу).

Заключение

1. Предложенный метод шифрования имеет универсальный характер с точки зрения применимости к различным языкам, однако его конкретная реализация и эффективность тесно связаны с грамматическими и лингвистическими особенностями языка, на котором написан исходный текст.

Универсальность метода проявляется в следующем:

Произвольный язык текста. Алгоритм шифрования начинается с предоставленного открытого текстового объекта $\mathcal{B}$, который может быть представлен на любом языке.

Использование адаптируемого алфавита. В предложенном методе используется частный алфавит шифрования, который включает в себя все символы и элементы, присутствующие в языке данного открытого текстового объекта.

Лингвистические элементы как основа. Вместо традиционного использования только одиночных символов (униграмм), метод основан на использовании языковых элементов. Эти элементы (корни слов, префиксы и суффиксы, аффиксы, слоги, биграммы, триграммы и

второстепенные члены предложения) являются общими лингвистическими категориями, существующими во многих языках. Таким образом, концептуально метод универсален, так как использует гибкий и адаптируемый алфавит, состоящий из элементов языка, применимых независимо от того, о каком языке идет речь.

2. Зависимость от грамматических особенностей. Хотя метод можно применить к любому языку, его реализация напрямую зависит от грамматических характеристик конкретного языка.

Разделение на языковые элементы. Процесс шифрования требует, чтобы текстовый объект $\mathcal{B}$ был разделен на группы элементов. Эти группы включают: - корни слов и дополнительные элементы (префиксы, суффиксы, аффиксы, второстепенные члены предложения); - слоги и дополнительные элементы; - биграммы или триграммы (***N***-граммы).

Для успешной шифрования необходимо правильно выделить (идентифицировать) эти лингвистические элементы.

3. На примере таджикского языка показано, как таджикский текст (рубаи Омара Хайяма) разделяется на конкретные морфологические компоненты.

Поскольку метод шифрования основан не на простых буквах, а на нумерации и замене именно лингвистических элементов, для его использования в конкретных случаях (например, для русского, английского или таджикского языков и др.) необходимо наличие инструментов (например, компьютерных программ, упомянутых в источниках). Эти инструменты должны позволять правильно определять и группировать такие элементы, как аффиксы, корни и слоги, в соответствии с правилами данного языка.

Таким образом, метод универсален по своей структуре, но его практическая реализация требует специальной лингвистической обработки исходного текста, основанной на грамматических и морфологических особенностях его языка.

4. Преимуществом предложенного метода является разработка частного алфавита, элементы которого нумеруются произвольно и извлекаются непосредственно из исходного текста. Процесс создания такого частного алфавита и способ его произвольной нумерации способствуют повышению стойкости шифрования текста.

Низкая вероятность несанкционированного расшифрования (дешифрования) при использовании языковых элементов в качестве основы для замены приводит к созданию зашифрованного объекта, стойкость которого можно выразить через крайне низкую вероятность незаконного доступа. Анализ показателей математической стойкости для шифрования текстового объекта $\mathcal{B}$ представлен в следующей таблице (Таблица 10):

Таблица 10 – Анализ сравнительных математических показателей стойкости

Тип лингвистической единицы	Количество элементов алфавита (n)	Количество вариантов нумерации (n!)	Количество элементов объекта	Вероятность незаконного дешифрования
Корни слов и дополнительные элементы	38	$\approx 5.23*10^{44}$	70	$\approx 1.2*10^{-100}$
Слоги и дополнительные элементы	47	$\approx 2.59*10^{59}$	80	$\approx 7.16*10^{-118}$
Биграммы	60	$\approx 8.32*10^{81}$	72	$\approx 6.12*10^{-103}$

Как видно из таблицы, переход к более крупным единицам (от корней к биграммам) увеличивает размер алфавита с 38 до 60, что приводит к колоссальному росту числа вариантов нумерации. Вероятность дешифрования (например, 10^{-118}) настолько мала, что в криптографии это считается «абсолютной безопасностью» по отношению к прямым атакам (полному перебору или brute-force). Основное преимущество заключается в том, что стойкость обеспечивается не только за счет длины ключа, но и за счет лингвистической структуры исходного текста.

5. Одной из важных задач является передача минимально необходимых данных в адрес расшифрователя с использованием закрытых (секретных) и открытых каналов электронной связи.

Известно, что зашифрованный объект передается по открытым каналам связи. По дорогостоящей закрытой связи должны передаваться: частный алфавит с пронумерованными элементами, основные и вспомогательные ключи, а также метод шифрования.

Из данных, приведенных в Таблице 10, следует, что количество элементов частного алфавита вместе с элементами основных и вспомогательных ключей всегда меньше количества элементов открытого объекта. В случае большого объема открытого объекта этот показатель увеличивается в несколько раз. В связи с этим, с экономической точки зрения, использование данного метода считается выгодным.

Рецензент: Мирзоев С.Х. – д.т.н., профессор кафедры информатики Таджикского национального университета.

Литература

1. Басалова Г.В. Основы криптографии. Тула: Изд-во Тульского государственного университета, 2009.
2. Бернет С., Пэйн С. Криптография. Официальное руководство RSA Security. М.: Бином, 2007.
3. Ященко В.В. Введение в криптографию. М.: МЦНМО: «ЧеРо», 2000.
4. Зубов А.Ю. Криптографические методы защиты информации. Совершенные шифры. М.: Гелиос АРВ, 2005.
5. Левин М. Криптография без секретов: Руководство пользователя. М.: Новый издательский дом, 2005.
6. Мао В. Современная криптография. Теория и практика. М.: Вильямс, 2005. 763 с.
7. Романец Ю.В., Тимофеев П.А., Шаньгин В.Ф. Защита информации в компьютерных системах и сетях. М.: Радио и связь, 2001.
8. Рябко Б.Я., Фионов А.Н. Криптографические методы защиты информации. М.: Горячая линия-Телеком, 2005.
9. Шнайер Б., Фергюсон Н. Практическая криптография. М.: Диалектика, 2005.
10. Шеннон К. Теория связи в секретных системах. Сборник «Работы по теории информации и кибернетике». М., ИЛ, 1963. С. 333-369.
11. Сингх С. Книга шифров. Тайная история шифров и их расшифровки. М.: Аст, Астрель, 2006. 447 с.
12. Бауэр Ф. Расшифрованные секреты. Методы и принципы криптологии. М.: Мир, 2007. 550 с.
13. Шнайер Б. Прикладная криптография. Протоколы, алгоритмы, исходные тексты на языке Си. М.: Триумф, 2003. 806 с.
14. Алферов А.П., Зубов А.Ю., Кузьмин А.С., Черемушкин А. В. Основы криптографии. М.: Гелиос АРВ, 2001. 479 с.
15. Усманов З.Д. О формировании базы префиксов таджикского литературного языка. / З.Д. Усманов, Г.М. Довудов // Доклады АН Республики Таджикистан. 2009, т. 52, № 6. С.431-436.
16. Усманов З.Д. О множестве постфиксов таджикского литературного языка. / З.Д. Усманов, О.М. Солиев, Г.М. Довудов // Доклады АН Республики Таджикистан. 2009, т. 53, № 2. С.99-103.
17. Усманов З.Д. Морфологический анализ словоформ таджикского языка: монография. / З.Д. Усманов, Г.М. Довудов / Душанбе: «Дониш», 2015. – 130 с.
18. Косимов А.А. Оценка эффективности использования триграмм при идентификации текста. / А.А. Косимов // Известия Академии наук Республики Таджикистан. Отделение физико-математических, химических, геологических и технических наук. 2017. № 1 (166). С. 51-57.
19. Ғафуров, М. Ҳ. Бадалсозии объекти матнӣ бо истифодаи символҳои забон / М. Ҳ. Ғафуров // Политехнический вестник. Серия: Интеллект. Инновации. Инвестиции. 2020. No. 4(52). P. 31-35.
20. Гафуров, М. Х. Об одном способе разработки уникальных вариантов алфавита шифрования / М. Х. Гафуров, А. А. Косимов, А. Абдукарим // Политехнический вестник. Серия: Интеллект. Инновации. Инвестиции. 2022. № 1 (57). С. 47-50.
21. Гафуров, М. Х. Об одном способе шифрования объекта с использованием элементов языка / М. Х. Гафуров // Политехнический вестник. Серия: Интеллект. Инновации. Инвестиции. 2023. № 2 (62). С. 22-29.
22. Гафуров, М. Х. Операторное применение шифрования элементов языка с квадратом Полибея / М. Х. Гафуров // Вестник Технологического университета Таджикистана. 2024. № 1 (56). С. 159-164.
23. Гафуров, М. Х. Применение биграмм и триграмм при шифровании объекта с использованием квадрата Полибея / М. Х. Гафуров // Политехнический вестник. Серия: Интеллект. Инновации. Инвестиции. 2024. № 1 (65). С. 72-75.
24. Гафуров, М. Х. Применение биграмм в шифровании текстовых объектов с использованием квадрата Полибея и двойного ключа / М. Х. Гафуров, Р. Б. Гиесов // Известия Национальной академии наук Таджикистана. Отделение физико-математических, химических, геологических и технических наук. 2025. № 3(200). С. 59-66.
25. Гафуров, М. Х. Шифрование элементов текста матричным и оператор - матричным методами / М. Х. Гафуров // Политехнический вестник. Серия: Интеллект. Инновации. Инвестиции. 2024. № 4 (68). С. 30-35.

26. Гафуров, М. Х. Применение матричных и оператор-матричных методов при шифровании текстового объекта с использованием биграммы и триграммы / М. Х. Гафуров // Вестник Таджикского национального университета. Серия естественных наук. 2025. № 2. С. 23-32.

27. Гафуров, М. Х. Применение биграмм, триграмм и многозначных чисел при разработке ключа шифрования текстового объекта / М. Х. Гафуров, М. А. Тоирова // Доклады Национальной академии наук Таджикистана. 2025. Т. 68, № 5. С. 445-451.

28. Ғафуров, М. Ҳ. Такмили усули Виженер дар бадалсозии объектҳои матнӣ / М. Ҳ. Ғафуров // Паёми политехникӣ. Бахши: Интеллект, Инноватсия, инвеститсия. 2025. No. 2(70). P. 67-71.

29. Гафуров, М. Х. Об одном способе шифрования текстовых объектов с использованием нумерации его элементов / М. Х. Гафуров // Доклады Национальной академии наук Таджикистана. 2025. Т. 68, № 7. С. 655-664.

СВЕДЕНИЯ ОБ АВТОРЕ - МАЪЛУМОТ ДАР БОРАИ МУАЛЛИФ -INFORMATION ABOUT THE AUTHOR

TJ	RU	EN
Ғафуров Миршафи Ҳамитович	Гафуров Миршафи Хамитович	Gafurov Mirshafi Hamitovich
Номзади илмҳои техникӣ, дотсент	Кандидат технических наук, доцент	Candidate of technical sciences, associate professor
Донишгоҳи техникии Тоҷикистон ба номи академик М.С. Осимӣ	Таджикский технический университет имени академика М.С. Осими	Tajik technical university named after academician M.S. Osimi
E-mail: mirugaf56@gmail.com		