На правах рукописи

Алхуссайн Амани

ВЕРОЯТНОСТНЫЙ АНАЛИЗ СТОЙКОСТИ ЗАЩИТЫ ИНФОРМАЦИИ МЕТОДОМ ЦЕЛОЧИСЛЕННОГО РАСЩЕПЛЕНИЯ СИМВОЛОВ

Специальность 1.2.3. Теоретическая информатика, кибернетика

Автореферат

диссертации на соискание учёной степени

кандидата физико-математических наук

Москва − 2023

Работа выполнена на кафедре информационных технологий факультета физико-математических и естественных наук федерального государственного автономного образовательного учреждения высшего образования «Российский университет дружбы народов имени Патриса Лумумбы»

Научный руководитель: доктор технических наук, профессор кафедры информационных технологий РУДН, старший научный сотрудник Института проблем передачи информации им. А.А. Харкевича Российской академии наук
Стефанюк Вадим Львович

Официальные оппоненты:

Редько Владимир Георгиевич
доктор физико-математических наук, старший научный сотрудник, Федеральное государственное учреждение «Федеральный научный центр Научно-исследовательский институт системных исследований Российской академии наук» (ФГУ ФНЦ НИИСИ РАН) главный научный сотрудник

Забежайло Михаил Иванович
доктор физико-математических наук, Федеральный исследовательский центр «информатика и управление» Российской академии наук (ФИЦ ИУ РАН), главный научный сотрудник

Дужин Василий Сергеевич
кандидат физико-математических наук, доцент кафедры алгоритмической математики, Санкт-Петербургский государственный электротехнический университет «ЛЭТИ»

Защита состоится «22» сентября 2023 г. в 15 часов 00 минут на заседании диссертационного совета ПДС 0200.006 на базе Российский университет дружбы народов имени Патриса Лумумбы, расположенного по адресу: 115419, г. Москва, ул. Орджоникидзе, д. 3, ауд. 214.

С диссертацией можно ознакомиться в библиотеке ФГАОУ ВО «Российский университет дружбы народов имени Патриса Лумумбы».

Автореферат разослан «_____»_____2023 г.

Ученый секретарь диссертационного совета ПДС 0200.006, кандидат физико-математических наук, доцент — Демидова Анастасия Вячеславовна

Общая характеристика работы

Актуальность темы исследования. В современных условиях развития информационных технологий и компьютеризации роль информации как одного из наиболее ценных и важных активов в различных сферах неуклонно растет. Защита текстовой информации при передаче по каналам связи является важной задачей для бизнес-приложений и ряда других областей жизни современного общества. Исследование проблем разработки, совершенствования и применения методов и средств защиты информации в процессе передачи и хранения информации приобрело особую важность не только в государственных, дипломатических, военных сферах, но также в банковских, коммерческих и других областях, связанных с широким кругом социально-экономических проблем.

Защита текстовой информации в России имеет свою историю. Так, первый профессиональный криптограф в России появился при Иване Грозном (1530-1584). Но в Новгороде существовала культура тайного письма с XIV в., в которой применялись в основном шифры простой замены. И всё же первым из российских государей, осознавшим всю важность криптографии для безопасности страны, стал Пётр Великий (1672−1725). Это произошло благодаря привлечению Пётром I для разработки государственного устройства России и развития образования знаменитого математика Г. В. Лейбница, задачей которого было также использование и развитие систем шифрования.

В СССР во время второй мировой войны велась разработка телефонного шифратора под руководством академика В.А. Котельникова, которому принадлежит знаменитая теорема отсчётов, лежащая в основе теории цифровой обработки сигналов.

В США К. Шеннон в 1944 г. создал основы теории секретной связи. В его работах излагается теория так называемых секретных систем, служащих, фактически, математической моделью шифров, дополняющих алгебраические и иные свойства шифров некоторыми вероятностными свойствами, что позволяет формализовать многие постановки задач синтеза и анализа шифров.

В работах К. Шеннона была доказана возможность защиты информации, обладающей свойством абсолютной стойкости[1]. Но на практике вскрылись некоторые трудности в применении этих результатов, в числе которых – необходимость использования одноразовой гаммы для каждого исходного символа и необходимость указания общей длины исходного текста. Эти ограничения приводят к сложности реализации методов защиты, тогда как нарушение этих условий облегчает задачу злоумышленнику. Поэтому возникает необходимость создания метода защиты информации, который проявлял бы стойкость при менее строгих условиях.

Согласно литературе, все известные симметричные алгоритмы, такие как DES, AES, Rijndael, гаммирование, TEA, IDEA, ГОСТ 28147-89, MARS, RC6, Serpent, Twofish, и др., а также ассиметричные алгоритмы, например: RSA, Рабина, Эль-Гамаля,

[1] Термин «абсолютная стойкость» − это принятый в литературе перевод термина perfect secrecy.

Мак-Элиса и др. не предполагается каких-то иных способов повышения степени безопасности указанных методов защиты в отношении несанкционированного доступа в канале передачи и хранения информации.

На сегодняшний день многие из известных методов защиты данных, основываются на операции деления нацело с остатком, такие как метод Цезаря, Аффинная система подстановок Цезаря, метод Хилла, метод Виженера и другие. Однако в этих методах не рассматривался вопрос многократного применения этой операции для каждого символа с целью повышения уровня безопасности и создания дополнительных затруднений для контроля передаваемых сообщений со стороны несанкционированного пользователя.

Поэтому изучение рассматриваемой в диссертации альтернативы в виде процедуры расщепления следует признать весьма актуальной задачей.

Исследование свойств предлагаемого нового метода защиты текстовой информации, чему посвящена диссертация, является актуальной задачей, возникающей как в связи с передачей по информационным сетям, так и с появлением таких новых способов хранения информации как облачная технология.

Степень разработанности темы. В отношении математических методов, используемых в диссертации для анализа свойств новой процедуры защиты, следует отметить, что предлагаемый анализ является определенным развитием известных вероятностных подходов, разработанных после классических исследований К.Шеннона и других специалистов по теории информации в таких научных учреждениях, как ИППИ РАН, ИСА РАН, ВЦ РАН, ИПУ РАН и в ряде других отечественных организаций.

Подобные методы были развиты с целью обеспечения эффективного функционирования сетей и систем связи, а также обеспечения работы сложных динамических систем, в которых вставал вопрос зашиты как от независимых внешних воздействий, так и от возможности вмешательства посторонних лиц и систем. В частности, вопросы защиты информации возникают в системах широкополосной мобильной коммуникации и в некоторых задачах эволюционного развития технических и биологических систем.

Проблеме изучения и разработки методов и систем защиты информации посвящено множество исследований как российских, так и зарубежных ученых. Среди работ авторов можно отметить труды М.И.Забежайло, В.Л.Стефанюка, А.В.Аграновского, А.В. Алексеева, А.А.Грушо, М.Венбо, А.Ю.Зубова, G.K.Alan, K.Atul, J.B. Borka, S.Bruce, J.Buchmann, R.P.Dhiren, A.J.Menezes, S.A.Vanstone и других.

Теоретическую основу исследований в области методов криптоанализа составили материалы, опубликованные следующими авторами: S.Christopher, J.Golic, S.W.Samuel, J.A.Mikhail, M.Stamp, R.M.Low, J.Zhang и многие другие.

В теории вероятностей отметим важность работ следующих авторов: Г.П.Башарин, К.Е.Самуйлов, А.А.Грушо, М.И.Забежайло, В.С.Дужин, Д.В.Смирнов, Е.Е.Тимонина, L.A.Sevastianov, B.Heinz, L.H.Lester, R.Meester, R.B.Ash, R. Durrett, P.E.Pfeiffer и других.

Среди авторов, внесших значительный вклад в область исследований генетического алгоритма и эволюции, можно отметить работы В.Г.Редько, А.Н.Аверкина, В.О.Арутюнова, J.H.Holland, S.Dutta, T.Das, Sh.Jash, D.Patra, S.Mondal, T.K.Mollah, S.Aarti, A.Suyash, A.Agarwal, A.Alecu, A.M.Salagean, M.C.Anisha, G.R.Pradyumna и многие другие.

Исследованию вопросов, связанных с использованием математических моделей в моделировании, в различные годы были посвящены труды таких ученых, как: В.О.Арутюноваа, А.Н.Аверкин, В.Г.Редько, П.А.Ляхова, М.Г.Бабенко, И.Н.Лавриненко, И.М.Белова, и др.

Существенный вклад в область математической статистики внесли Г.П.Башарин, Г.И.Ивченко, Ю.И.Медведев, Э.Леман, А.И.Кобзарь, A.Rukhin, J.Soto, J.Nechvatal, M.Smid и другие.

В семантическом анализе отметим значительность работ следующих авторов: А.Н.Аверкин, Y.Samojlik, V.Gnatyuk, V.Klimchuk, O.Shylo и другие.

В общем, теоретико-методологическую основу работы составила достаточно большая информационная база. В числе научных источников диссертации использованы такие научные сведения из журнальных статей, книг, материалов научных конференций, научных отчетов и докладов, семинаров, а также результаты собственных экспериментов автора диссертации.

Цели и задачи исследования. Целью исследования в диссертации является разработка новой теоретико-методологической и практической концепции расщепления как одной из систем защиты информации при её хранении и передаче. Достижение поставленной цели предполагает решение следующих **задач**:

1. Предложить новую процедуру – целочисленное расщепление, т.е. представление целого числа по базе другого числа в виде последовательрности из k целых чисел (расщепление k-го уровня), и доказать необходимые строгие утверждения.
2. Определить теоретическую модель защиты информации, основанную на предложенной процедуре целочисленного расщепления, и дать описание исходных положений самой модели, её параметров и свойств.
3. Разработать алгоритм защиты информации на основе целочисленного расщепления, позволяющий управлять уровнем защиты информации.
4. Доказать, что с ростом глубины расщепления вероятность несанкционированного восстановления символа экспоненциально убывает, что позволяет говорить об асимптотической стойкости расщепления.
5. Показать, что метод расщепления в значительной степени ослабляет хакеру возможность раскрытия исходного текста за счет учета его содержания, а также провести сравнение защиты методом расщепления с другими способами защиты.

Научная новизна. В процессе проведения исследований был разработан новый научный подход к защите текстовой информации и дан вероятностный анализ стойкости этого подхода:

1. Предложена новая математическая процедура – целочисленное расщепление – обобщающая известную арифметическую операцию деления с остатком.
2. С учётом доказанных теорем и утверждений, связанных с этой процедурой, построена математическая модель системы реализующей расщепление.
3. Разработан и проанализирован новый метод защиты текста, состоящий в замене каждого символа передаваемого текста на последовательность k целых чисел (расщепление k-ого уровня). Этот метод отличается от известных способов защиты, основанных на операциях модульной арифметики тем, что он обеспечивает высокую степень безопасности, поскольку взятие модуля в системе с расщеплением делается $k-1$ раз, а не только один раз, как это принято в других подходах. Важно также подчеркнуть, что параметры этого модуля изменяются на каждом шаге работы и, в частности, не совпадают с размером алфавита.

Теоретическая и практическая значимость работы. Теоретическая ценность полученных в диссертации результатов заключается в создании математического аппарата, основанного на использовании модульной арифметики в приложении к исследованию новой процедуры, названной в диссертации целочисленным расщеплением, и доказательстве ряда строгих утверждений, связанных с концепцией стойкости защиты информации с использованием предлагаемой процедуры.

Процедура расщепления, в силу которой целое число представляется уникальным образом в виде последовательности k целых чисел, может иметь ценность и для других приложений, не обязательно связанных с защитой информации. В этом отношении она напоминает китайскую теорему об остатках, отличаясь от неё кардинальным образом по математическим свойствам, отмеченным в диссертации.

Практическая значимость выполненной работы обусловлена тем, что в ней построена исчерпывающая математическая схема применения модели расщепления в процессе защиты информации при её передаче и хранении, которую можно рассматривать как основу для создания действующей программной системы.

Методология и методы исследования. В диссертационной работе применяются методология и методы модульной арифметики, методы теории вероятностей и методы теории информации, связанные со стойкостью защиты информации в различных аспектах, а также методы симметричной защиты информации на основе использования генераторов псевдослучайных чисел.

Основные положения, выносимые на защиту.

1. Предложена принципиально новая теоретическая модель защиты текстовой информации путем применения целочисленного расщепления, основанного на принципах модульной арифметики, позволяющая представить целое число по базе другого целого в виде последовательности k целых чисел.
2. Доказаны строгие утверждения, касающиеся свойств возникающего преобразования символов в процедуре расщепления.

3. Разработан и построен новый метод шифрования, позволяющий управлять уровнем защиты информации с помощью генератора псевдослучайных чисел.
4. Приведено доказательство асимптотической стойкости предлагаемого метода защиты информации.

Степень достоверности и апробация результатов. Достоверность полученных в диссертации результатов вытекает из использования строгих математических методов модульной арифметики, методов теории вероятностей, методов теории информации и методов симметричной защиты информации с использованием генераторов псевдослучайных чисел.

Основные положения диссертационной работы докладывались и обсуждались: на XIV национальной конференции по искусственному интеллекту с международным участием, Казань, 2014; на International Research Conference on Engineering, Science and Management, Dubai, 2014; на VIII Международной научной конференции «Приоритеты мировой науки: эксперимент и научная дискуссия», Южная Каролина, Северный Чарльстон, 2015; на 5й Европейской конференции по инновациям в технических и естественных науках, Австрия, Вена, 2015; на XIX Международной научно-практической конференции, Москва, 2015; на VIII Международной научно-практической конференции студентов, аспирантов и молодых учёных «Шаг в будущее: теоретические и прикладные исследования современной науки», Санкт-Петербург, 2015; на 2nd International Scientific Conference “Theoretical and Applied Sciences in the USA”, USA, New York, 2015; на First European Conference on Informational Technology and Computer Science, Austria, Vienna, 2015. на Международной научной конференции «На пути к стабильному миру: безопасность и устойчивое развитие», США, Сан-Диего, 2015; на IX International scientific conference “The latest research in modern science: experience, traditions and innovations”, Morrisville, NC, USA, 2019; на 4-й Международной научной конференции «Интеллектуальные информационные технологии в технике и на производстве», Острова Праги, Чехия, 2019; на II International Scientific Conference "MIP: Engineering-2020 - Modernization, Innovations, Progress: Advanced Technologies in Material Science, Mechanical and Automation Engineering", Krasnoyarsk, Russia, 2020; на International Scientific Conference "CAMSTECH - 2020: Advances in Material Science and Technology", Krasnoyarsk, Russia, 2020 ; на International Conference on Engineering Systems ICES 2020, Moscow, Russia, 2020; на III International Workshop on Modeling, Information Processing and Computing, Krasnoyarsk, Russia, 2021.

Работа автора была признана лучшей на III-й Международной летней школе-семинаре по искусственному интеллекту для студентов, аспирантов и молодых ученых "Интеллектуальные системы и технологии: современное состояние и перспективы", Тверь, 2015. Доклад по тематике диссертации был признан лучшим на 3rd international conference on “Engineering & Technology, Computer, Basic and Applied Sciences”, UAE, Dubai, 2016.

Другой доклад, связанный с диссертацией, был признан одним из лучших на VI Всероссийской конференции «Информационно-телекоммуникационные технологии и математическое моделирование высокотехнологичных систем», Российский

университет дружбы народов, Москва, 2016. Также работа по тематике диссертации была признана одной из лучших на конференции «Reaserch and innovation issues», Национальный исследовательский университет «Московский институт электронной техники», Москва, 2017.

Личный вклад. Представленная в диссертации модель, процедура и результаты анализа получены автором самостоятельно. Программные средства, использованные для анализа и иллюстрации работы, разработаны автором.

Публикации. Основные теоретические и практические результаты диссертации опубликованы в 25 статьях и докладах, в том числе 7 работ опубликовано в рецензируемых изданиях, рекомендованных перечнем ВАК, 6 работ в периодических научных журналах, индексируемых в системе Scopus, и получен 1 патент на изобретение.

Содержание работы

Во **введении** обоснована актуальность темы диссертации, перечислены основные направления исследований, и приведен обзор научной литературы по изучаемой проблеме. Приведено общее описание работы. Сформулирована цель работы, определены решаемые в диссертации задачи, показана научная новизна исследования, представлены теоретическая и практическая значимость представляемой работы.

В **первой главе** приведено описание объекта исследования, дан обзор проблемной области и основные задачи разработки систем защиты информации. В этой главе перечислены основные понятия и характеристики методов защиты информации, представлены классы существующих систем защиты. Перечислены некоторые известные методы замены, основанные на использовании модульной арифметики, и показаны характеризующие их недостатки. Кроме того, описаны работы по использованию метода гаммирования и метода Вернама, указана их ограниченная применимость. А также представлена важная работа К. Шеннона по теоретической стойкости шифров, в которой были исследованы требования к подобным шифрам.

Многие из перечисленных недостатков известных методов шифрования удаётся устранить с использованием разработанной в диссертации новой процедуры, получившей название целочисленного расщепления, описанию которой посвящена следующая глава.

Во **второй главе** описаны теоретические основы предлагаемой в диссертации новой процедуры защиты информации, названной целочисленным расщеплением. Определена математическая функция преобразования, отвечающего этой процедуре, и описаны ее параметры.

Основные определения и понятия процедуры целочисленного расщепления состоят в следующем:

Пусть даны два положительных целых числа r и a, для которых выполняется неравенство $r > a > 0$.

Определение 1. *Целочисленным расщеплением числа a по базе* r, называется представление a в виде последовательности целых чисел $a_1, a_2, a_3, ..., a_{k-1}, a_k$, в которой

$$a_1 = \delta^{(2)}, \text{ где } \delta^{(2)} = r \bmod a,$$

$$a_2 = \delta^{(3)}, \text{ где } \delta^{(3)} = r \bmod q^{(2)},\ q^{(2)} = \left\lfloor \frac{r}{a} \right\rfloor,$$

$$a_3 = \delta^{(4)}, \text{ где } \delta^{(4)} = r \bmod q^{(3)},\ q^{(3)} = \left\lfloor \frac{r}{q^{(2)}} \right\rfloor, \qquad (1)$$

.....

$$a_{k-1} = \delta^{(k)}, \text{ где } \delta^{(k)} = r \bmod q^{(k-1)},\ q^{(k-1)} = \left\lfloor \frac{r}{q^{(k-2)}} \right\rfloor,$$

$$a_k = q^{(k)}, \text{ где } q^{(k)} = \left\lfloor \frac{r}{q^{(k-1)}} \right\rfloor,$$

где $\delta^{(2)}$ – остаток при целочисленном делении r на a, а $q^{(i)}$ – целая часть при таком делении, $\delta^{(i)}$ – остаток при целочисленном делении r на $q^{(i-1)}$, причём символ $\lfloor\ \rfloor$ означает округление до ближайшего целого в меньшую сторону. Натуральное число k названо *уровнем расщепления*.

Целочисленное расщепление является определенным обобщением математической операции деления с остатком. Блок-схема целочисленного расщепления числа a по базе r, при уровне расщепления k, показана на рис.1.

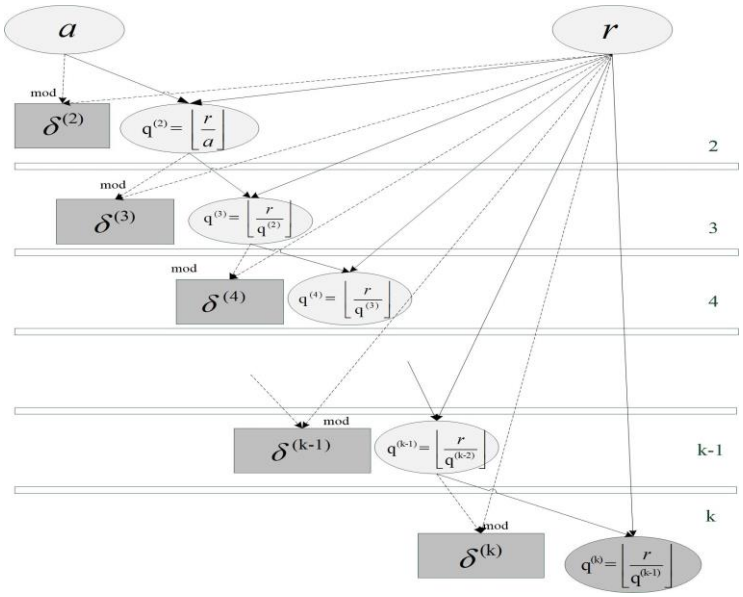


Рис.1. Блок-схема целочисленного расщепления числа a по базе r при уровне расщепления k.

Определение 2. *Функцией отображения* $\Phi_k(a,r)$ называется результат целочисленного расщепления числа a по базе r.

Согласно (1) функция отображения $\Phi_k(a,r)$ при уровне расщепления равном k задаётся следующим соотношением:

$$\Phi_k(a,r) = \left(\delta_a^{(2)}, \delta_a^{(3)}, \delta_a^{(4)}, ..., \delta_a^{(k-1)}, \delta_a^{(k)}, q_a^{(k)}\right). \qquad (2)$$

Определение 3. *Обобщённым целочисленным расщеплением числа a по векторной базе* $\vec{r} = (r_1, r_2, ..., r_l)$, $l = k-1$, называется представление числа a в виде последовательности целых чисел $a_1, a_2, a_3, ..., a_{k-1}, a_k$, в которой:

$$
\begin{aligned}
&a_1 = \delta^{(2)}, \text{ где } \delta^{(2)} = r_1 \bmod a,\ r_1 > a,\\
&a_2 = \delta^{(3)}, \text{ где } \delta^{(3)} = r_2 \bmod q^{(2)},\ q^{(2)} = \left\lfloor \frac{r_1}{a} \right\rfloor,\ r_2 > q^{(2)},\\
&a_3 = \delta^{(4)}, \text{ где } \delta^{(4)} = r_3 \bmod q^{(3)},\ q^{(3)} = \left\lfloor \frac{r_2}{q^{(2)}} \right\rfloor,\ r_3 > q^{(3)},\\
&\ldots\ldots\\
&a_{k-1} = \delta^{(k)}, \text{ где } \delta^{(k)} = r_{k-1} \bmod q^{(k-1)},\ q^{(k-1)} = \left\lfloor \frac{r_{k-2}}{q^{(k-2)}} \right\rfloor, r_{k-1} > q^{(k-1)},\\
&a_k = q^{(k)}, \text{ где } q^{(k)} = \left\lfloor \frac{r_{k-1}}{q^{(k-1)}} \right\rfloor,
\end{aligned}
\tag{3}
$$

здесь символы $\delta^{(i)}$ – обозначают остатки при целочисленном делении r_i на $q^{(i)}$, символ $\lfloor\ \rfloor$ означает округление до ближайшего целого в меньшую сторону, а натуральное число k назовём уровнем расщепления.

Целочисленное расщепление опирается на следующее известное утверждение модульной арифметики:

Утверждение 1. Пусть задана пара целых чисел r и a, где $a \neq 0$. Тогда существует единственное сочетание целых чисел q и δ, удовлетворяющее соотношению $r = a \times q + \delta$, где $0 \le \delta < |a|$.

Указанные выше определения позволяют сформулировать и доказать следующие утверждения:

Утверждение 2. Целочисленное расщепление является инъективным, т.е. выполняется следующее свойство: если $a \neq b$, то имеем $\Phi_k(a, r) \neq \Phi_k(b, r)$. (4)

Утверждение 3. Пусть задана последовательность чисел $a_1, a_2, a_3, ..., a_{k-1}, a_k$, удовлетворяющая условиям $1 \le a_k \le r$ и $0 \le a_i < r$ $, i = 1, ..., k-1$. Тогда существует такое целое число a, $0 < a < r$, расщеплением которого является эта последовательность.

Утверждение 4. Пусть даны $r > a \ge 1$ и $a_1, a_2, a_3, ..., a_{k-1}, a_k$, т.е. результат целочисленного расщепления a по базе r, тогда выполняются следующие соотношения: $0 \le a_i < r$, $i = 1, ..., k-1$ и $1 \le a_k \le r$.

Следствие 1. Последовательность чисел $a_1, a_2, a_3, ..., a_{k-1}, a_k$, возникающих при целочисленном расщеплении исходного a, определяется единственным образом.

Доказательство приведенных выше утверждений содержится в данной главе диссертации.

В принципе, целочисленное расщепление, как и обобщенное целочисленное расщепление, весьма близки к методам эволюции, используемым в генетическом алгоритме Дж. Холланда. Действительно, в процессе целочисленного расщепления можно также наблюдать построения новых поколений чисел, которые всё в большей степени отдаляются от исходного числа, что, в конечном счёте, как будет показано далее, приводит к асимптотической стойкости предлагаемой защиты.

Вопросу описания процедуры расщепления как эволюционного процесса посвящены разделы 1.5 и 2.5 диссертации, где показано, что реализация процедуры целочисленного расщепления по существу использует некую процедуру мутации, поскольку при генерации поколений применяются псевдослучайные числа.

В **третьей главе** рассмотрена математическая модель применения целочисленного расщепления для защиты текста. Определена математическая функция, позволяющая использовать эту процедуру для защиты и восстановления информации. Описан метод симметричной защиты информации и проведен вероятностный анализ стойкости защиты информации методом целочисленного расщепления.

Пусть база r превосходит максимальное значение в выбранной кодовой таблице символов. Тогда, расщеплением уровня k для символа S с кодом a в соответствии с указанной кодовой таблицей называется представление S в виде ряда соответствующих целых чисел $\delta^{(2)},\delta^{(3)},\delta^{(4)},.....,\delta^{(k-1)},\delta^{(k)},q^{(k)}$. Здесь числа $\delta^{(i)}$ и $q^{(k)}$ вычисляются в соответствии с (1) .

Назовём полученный ряд целых чисел $\delta^{(2)},\delta^{(3)},\delta^{(4)},.....,\delta^{(k-1)},\delta^{(k)},q^{(k)}$ результатом расщепления.

Из утверждений 2, 3 и 4, доказанных в главе 2, можно заключить, что если получателю известен секретный ключ, подробно описанный в главе 3 диссертации, то расщепление произвольного символа S является инъективным и обратимо, что открывает возможность однозначного восстановления этого символа на приёмном конце.

Метод расщепления применен в работе посимвольно к отдельным символам передаваемого текста.

В обсуждаемой главе отмечено, что успех защиты символов текста зависит как от предложенного метода расщепления, так и от свойств генератора псевдослучайных чисел (ГПСЧ). Защита текста и его восстановление происходят с помощью ГПСЧ, который считается известным и на приёмном и на передающем конце

От генератора псевдослучайных чисел поступает величина r_i, необходимая как при передаче с использованием расщепления, так и при восстановлении каждого символа. В результате расщепления в момент t создаются передаваемые целые числа $\delta^{(2)},\delta^{(3)},\delta^{(4)},.....,\delta^{(k-1)},\delta^{(k)},q^{(k)}$. (Предполагается, что величина $r_i > 0$ превосходит максимальное значение символов при выбранной кодовой таблице.)

В нашей модели затем поступают от генератора псевдослучайные величины r_{i+1}, r_{i+2},...., r_{i+k}, используемые для дополнительной защиты каждого компонента расщепления этого символа путём гаммирования.

Математическая модель этапа защиты символа *S* с кодом *a* [2]

Результат защиты при расщеплении

$$Y=\begin{cases} r_i \oplus a \quad при \quad k=1 \\ \delta^{(2)},\delta^{(3)},\delta^{(4)},...,\delta^{(k-1)},\delta^{(k)},q^{(k)} \quad при \quad k>1 \end{cases} \tag{5}$$

Результат защиты при гаммировании:

$$\begin{cases} \delta^{(j)} \oplus r_{i+j-1} \quad ,где \quad j=2,3,...,k \quad при\, k>1 \\ q^{(k)} \oplus r_{i+k} \end{cases} \tag{6}$$

Математическая модель этапа восстановления символа:

Результат восстановления после гаммирования

$$\begin{cases} \delta^{(j)} \oplus r_{i+j-1}, \quad где \quad j=2,3,...,k \quad при\, k>1 \\ q^{(k)} \oplus r_{i+k} \end{cases} \tag{7}$$

Результат восстановления расщеплённого символа:

$$\begin{cases} r_i \oplus Y \quad при \quad k=1 \\ \dfrac{\left(r_i-\delta^{(j)}\right)}{q^{(j)}}, \quad где \quad j=k,k-1,...,3,2 \quad при\, k>1 \end{cases} \tag{8}$$

В диссертации приведен также метод *обобщенного расщепления*, при котором на каждом шаге формирования величин, используется новая величина $r_i > q^{(i)}$.

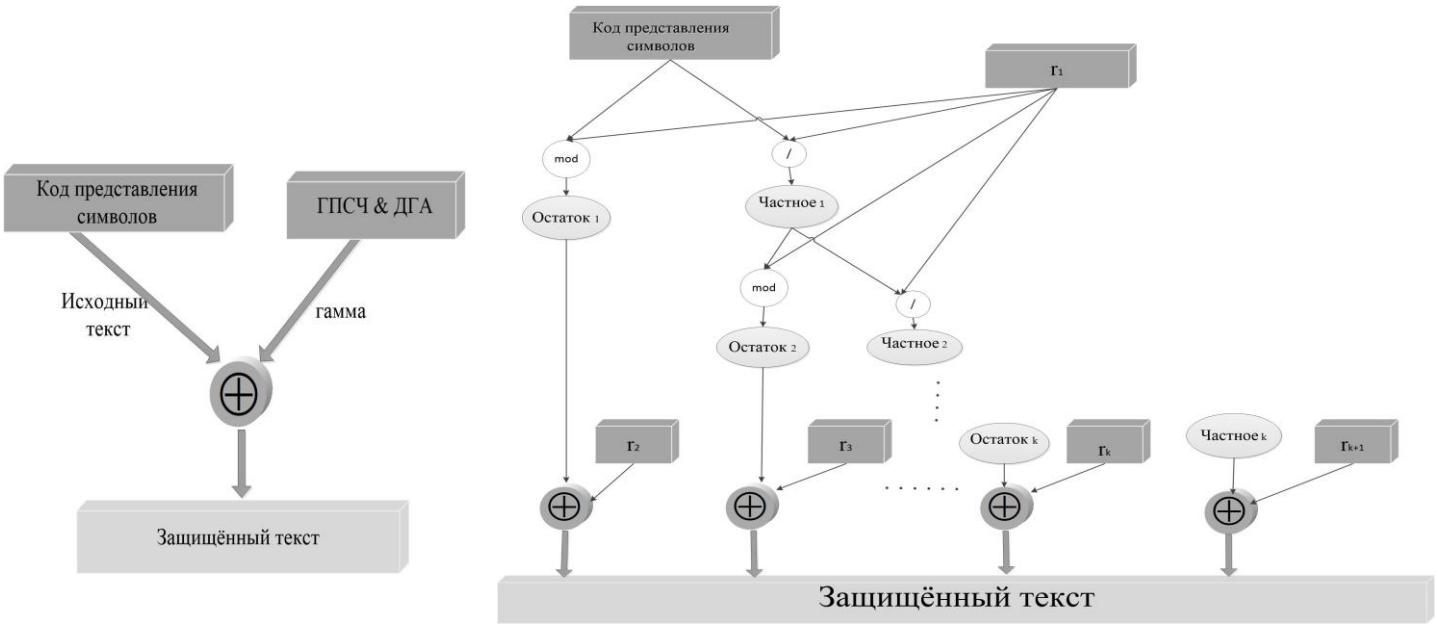
Результат защиты символа при обобщенном расщеплении:

$$\begin{cases} \delta^{(2)},\delta^{(3)},\delta^{(4)},.....,\delta^{(k-1)},\delta^{(k)},q^{(k)} \quad ,где \\ \delta^{(2)}=r_1 \bmod a, q^{(2)}=\left\lfloor \dfrac{r_1}{a} \right\rfloor \\ \delta^{(3)}=r_2 \bmod q^{(2)},\ q^{(3)}=\left\lfloor \dfrac{r_2}{q^{(2)}} \right\rfloor \\ \vdots \\ \delta^{(k)}=r_{k-1} \bmod q^{(k-1)}, q^{(k-1)}=\left\lfloor \dfrac{r_{k-2}}{q^{(k-2)}} \right\rfloor \\ q^{(k)}=\left\lfloor \dfrac{r_{k-1}}{q^{(k-1)}} \right\rfloor \end{cases} \tag{9}$$

Результат восстановления обобщенного расщеплённого символа:

$$\begin{cases} \dfrac{\left(r_i-\delta^{(j)}\right)}{q^{(j)}}, \quad где \quad j=k,k-1,...,3,2 \quad при\, k>1 \\ i=k-1,...,2,1 \quad при\, k>1 \end{cases} \tag{10}$$

[2] Символом $\oplus$ обозначена известная побитовая операция − "исключающее или" для двух целых чисел.



А) При $k = 1$ Б) При k уровнях, $k \geq 2$

Рис.2. Блок-схема метода символьного расщепления

Указанные выше определения и математические модели позволяют сформулировать и доказать следующую теорему, связанную с вероятностным анализом свойств защиты, достигаемой в результате символьного расщепления.

Обозначим через C защищённый текст, полученный в результате применения метода расщепления каждого символа исходного текста M.

Заметим, что C в силу работы генератора псевдослучайных чисел представляется хакеру как набор случайных чисел с неизвестными ему вероятностными свойствами. Последний попытается восстановить M по этим числам, применяя процедуру полного перебора.

Обозначим через $\Pr(M \mid C, k)$ вероятность успешного восстановления исходного текста M, опираясь на известный хакеру защищённый текст C и неизвестный ему уровень расщепления k.

Лемма 1. Вероятность несанкционированного восстановления исходного текста M по результату расщепления C экспоненциально убывает с ростом k, согласно приведенному в диссертации выражению:

$$\Pr(M \mid C, k) = \left(\sum_{i=2}^{k} L^{\left\lfloor \frac{N}{i} \right\rfloor} \right)^{-1}, \qquad (11)$$

где N – размер защищённого текста C, созданного для исходного текста M. Здесь L – число всех возможных событий (исходов) в ходе перебора в пространстве гамм из предполагаемых хакером чисел $\{\tilde{r}_1, \tilde{r}_2, ..., \tilde{r}_L\}$, используемых им в процессе перебора.

Следствие 2. Лемма 1 показывает, что вероятность несанкционированного восстановления исходного текста M по результату расщепления C экспоненциально убывает с ростом размера защищённого текста N.

Следствие 3. Из Леммы 1 следует, что вероятность несанкционированного восстановления исходного текста M по результату расщепления C экспоненциально убывает с ростом числа всех переборных событий L в пространстве созданной им гаммы.

На рис. 3 показан график поведения надёжности защиты расщеплением, вытекающий из леммы 1, т.е. вероятность несанкционированного восстановления исходного текста M при неизвестном ключе при различных значениях $k=2,3,\ldots,8$.

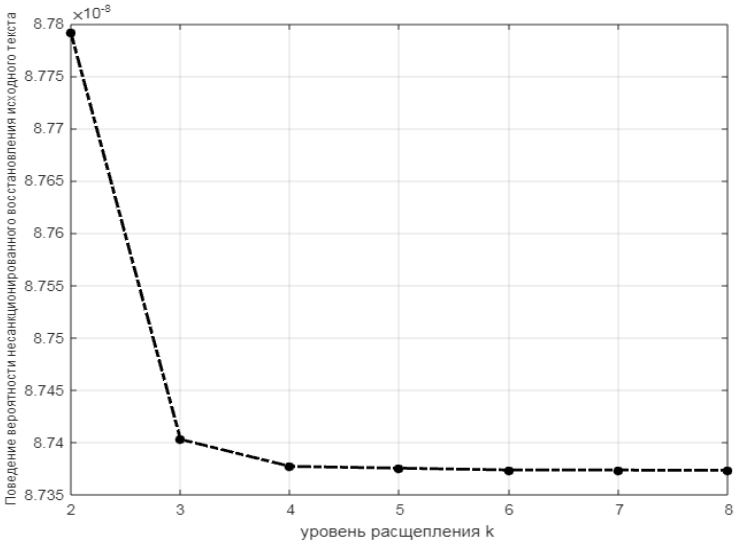
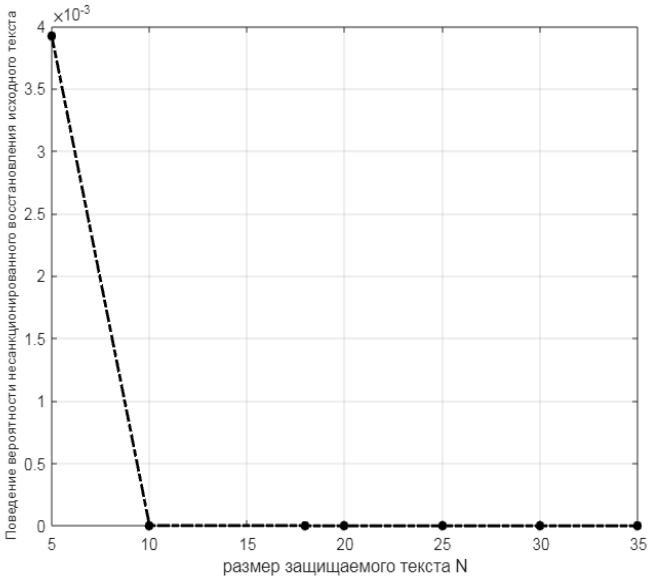
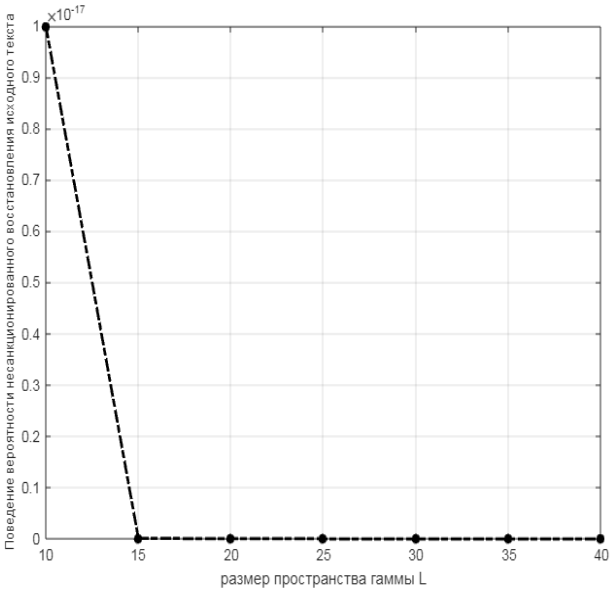


Рис. 3. Поведение вероятности несанкционированного восстановления исходного текста при различных значениях уровня расщепления k.

На рис. 4 показан график поведения вероятности несанкционированного восстановления исходного текста при различных значениях размера защищённого текста N, а также при различных значениях размера пространства гаммы L, вытекающие из Леммы 1 и Следствий 2 и 3.



А) При различных значениях размера защищаемого текста N



Б) При различных значениях размера пространства гаммы L

Рис. 4. Поведение вероятности несанкционированного восстановления исходного текста

Лемма 2. Вероятность несанкционированного восстановления исходного текста A по результату *обобщённого расщепления* C экспоненциально убывает с ростом k согласно выражению

$$\Pr(A \mid C, k) = \left(\sum_{i=2}^{k} L^{(i-1)\times\left\lfloor \frac{N}{i} \right\rfloor} \right)^{-1}, \qquad (12)$$

где $N-$ размер защищённого текста C – результат *обобщённого* расщепления исходного текста A, а L – число всех возможных событий в ходе перебора в пространстве гамм из ℓ значений случайных чисел $\{\tilde{r}_1, \tilde{r}_2, ..., \tilde{r}_L\}$, предполагаемых хакером, где $\ell = k - 1$ из определения 3.

Следствие 4. Из Леммы 2 также следует, что вероятность несанкционированного восстановления исходного текста A по результату *обобщённого расщепления* C экспоненциально убывает с ростом размера защищённого текста N.

Следствие 5. Из Леммы 2 следует, что вероятность несанкционированного восстановления исходного текста A по результату *обобщённого расщепления* C экспоненциально убывает с ростом числа всех переборных событий в пространстве гамм из случайных чисел.

На рис. 5 показан график поведения надёжности защиты при *обобщенном* расщеплении, вытекающий из леммы 2, т.е. вероятность несанкционированного восстановления исходного текста A.

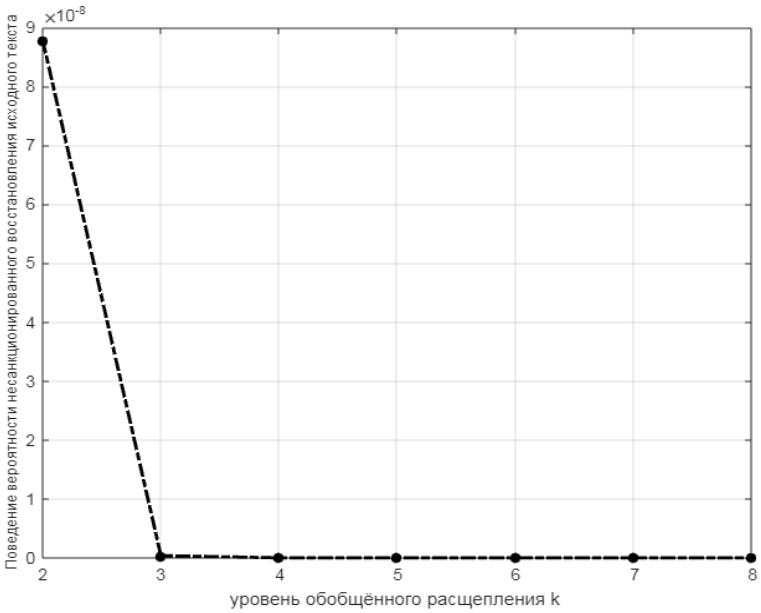
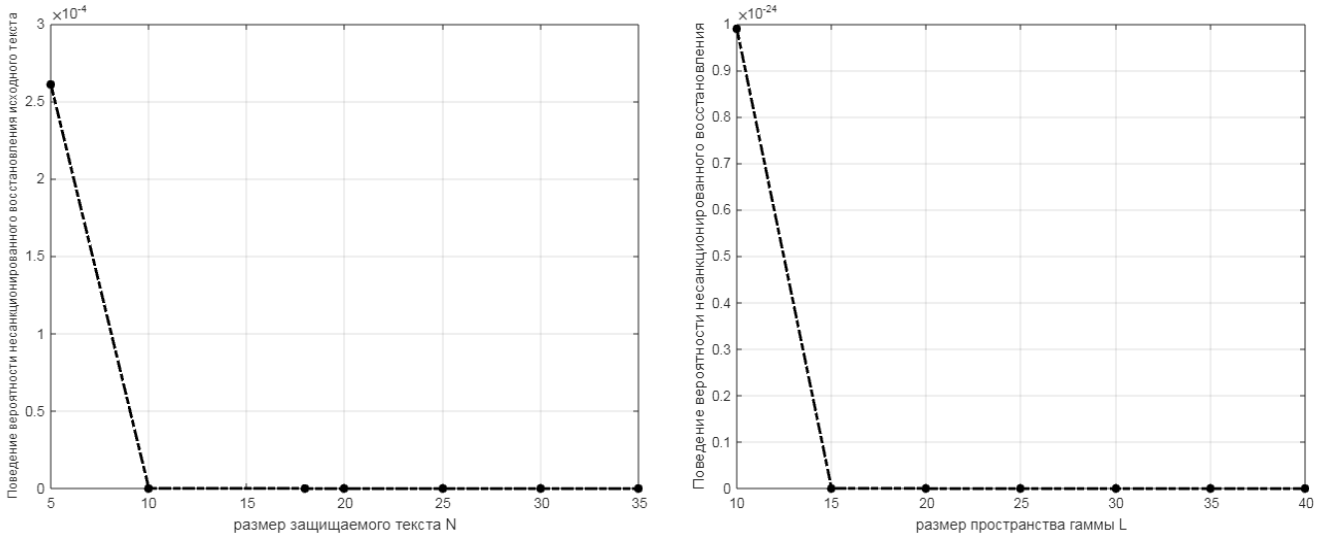


Рис. 5. Поведение вероятности несанкционированного восстановления исходного текста при различных значениях уровней *обобщённого* расщепления k.

На рис. 6 показан график поведения вероятности несанкционированного восстановления исходного текста по результату *обобщённого* расщепления при различных значениях размера защищённого текста *N*, а также при различных значениях размера пространства гаммы *L*, вытекающие из Леммы 2 и Следствий 4 и 5.



А) При различных значениях размера защищаемого текста N

Б) При различных значениях размера пространства гаммы L

Рис. 6. Поведение вероятности несанкционированного восстановления исходного текста по результату *обобщённого* расщепления

Определение 4. Будем говорить, что метод, зависящий от параметра k, обладает асимптотической стойкостью, если при $k \to \infty$ выполняется $\Pr(M \mid C,k) \to 0$. (13)

Это определение позволяет сформировать основную теорему диссертации, состоящую в следующем:

Теорема. Метод расщепления, включая обобщенный метод расщепления, обладает свойством асимптотической стойкости.

Следствие 6. Из Леммы 1 и Леммы 2 можно заключить, что, при условии $k \to \infty$, вероятность вычисления защищённого исходного текста при неизвестном ключе стремится к нулю, т.е. из Леммы 1 при $k \to \infty$ имеем $\Pr(M \mid C,k) \to 0$, и из Леммы 2 при $k \to \infty$ имеем $\Pr(A \mid C,k) \to 0$.

Замечание 1. В диссертации показано также, что метод расщепления существенно усложняет восстановление исходного текста несанкционированным пользователем опирающегося на семантическое содержание, поскольку в ходе расщепления создаётся множество посторонних символов в соответствии с кодовой таблицей, но не несущих никакой осмысленной информации.

Доказательство приведенных выше лемм и основной теоремы содержится в третьей главе диссертации.

В **четвертой главе** проведено сравнение метода символьного расщепления с известным методом защиты на основе гаммирования, с методом защиты Вернама. А также с указанными ранее традиционными методами защиты, применяющими операции модульной арифметики.

С целью выделения достоинств предлагаемого метода защиты в таблице 1 представлены отличия данного метода символьного расщепления от традиционных методов защиты с использованием гаммирования и от метода Вернама:

Таблица 1. Сравнение расщепления с некоторыми синхро-потоковыми (синхронизационными) методами защиты

	Защита методом Вернама	Защита методом гаммирования	Защита методом символьного расщепления
Период гаммы	отсутствует одноразовая гамма	Имеется	существенно увеличен
Необходимость хранения гамм	Да	Нет	Нет
Необходимость доставки получателю такой же гаммы, как у отправителя	Да	Нет	Нет
Абсолютная стойкость	Да	Нет	Асимптотическая стойкость

Вышеприведенное сравнение позволяет сформулировать следующее замечание:

Замечание 2. Метод расщепления решает проблему ненадёжности защиты, которая существует в потоковых методах защиты данных, из-за повторного использования гаммы.

В разделе 4.3 этой главы проведено сравнение расщепления с другими известными абсолютно стойкими методами защиты. Имеются два преимущества метода символьного расщепления по сравнению с известными абсолютно стойкими методами защиты информации.

Первое – метод символьного расщепления является практичным и не очень дорогим по требуемым ресурсам по сравнению с другими способами обеспечения стойкости шифров, поскольку не требуется выполнения условия об использовании каждой гаммы только один раз.

Второе – метод символьного расщепления скрывает для нарушителя информацию о длине исходного сообщения, которая определяется следующим образом:

$$c = k \times l, \tag{14}$$

где c − длина защищенного текста, k − уровень расщепления и l − длина исходного текста, причём величина k хакеру вообще неизвестна.

В разделе 4.4 отмечаются два отличия метода символьного расщепления от перечисленных известных методов защиты, применяющих операции модульной арифметики:

Первое – операция с применением модуля используется в традиционных методах только один раз для каждого символа. В предлагаемом методе символьного расщепления эта операция используется $k-1$ раз.

Второе – величины модулей, используемые во всех традиционных способах защиты информации (Цезарь, Виженер, Аффинный, Хилла и др.), основанных на операции модульной арифметики, совпадают с размером соответствующего алфавита,

тогда как в методе символьного расщепления, величина модуля изменяется на каждом шаге работы системы и не связана с общим размером алфавита.

В разделе 4.5 главы представлено подробное сравнение между расщеплением и известной китайской теоремой об остатках.

В отношении Китайской теоремы об остатках (КТО) следует отметить, что она используется в алгоритмах шифрования и в задачах разделённого секрета (sharing secret). Для вскрытия информации в таких системах требуется решение классических проблем факторизации и поиска взаимно простых чисел, на которых базируется КТО, что приводит к большим вычислительным затратам.

В методе расщепления, предлагаемом и изученном в диссертации, не возникает необходимости в использовании взаимно простых чисел.

В качестве иллюстрации приводятся примеры защиты информации с использованием метода расщепления при различных уровнях расщепления.

В случае, когда используемая кодовая таблица символов поддерживает несколько языков, то метод работает сразу для всех этих языков, как показано в диссертации.

В **заключении** к диссертации приводятся основные результаты, полученные в диссертационной работе.

Основные результаты работы

В результате исследований были решены следующие задачи:

1. Предложена новая процедура, названная в диссертации целочисленным расщеплением, позволяющая представить любые целое число по базе другого числа в виде последовательности k целых чисел, созданных по определённому правилу на основе использовании модульной арифметики, где число k названо уровнем расщепления.
2. Доказана единственность предложенной процедуры целочисленного расщепления и что соответствующее предоставление является инъективным и обратимым, что позволяет применить это представленное в системах защиты информации.
3. В диссертации предложена математическая модель защиты информации на основе целочисленного расщепления и исследованы свойства этой модели.
4. Доказано, что стойкость расщепления в отношении защиты информации растёт с ростом уровня расщепления k, что позволяет говорить об асимптотической стойкости расщепления с увеличением k.
5. Показано, что предложенный метод расщепления также существенно усложняет восстановление исходного текста несанкционированным пользователем опирающегося на семантическое содержание этого текста.
6. Проведено сравнение метода символьного расщепления и традиционных методов защиты информации и показано, что этот метод снимает ряд ограничений, характерных для защиты информации другими методов.

Публикации автора по теме диссертации

Публикации в изданиях, рекомендованных ВАК России:

1. Алхуссайн А.Х., Симметричный алгоритм шифрования с помощью генетического алгоритма и генераторов псевдослучайных чисел // Естественные и технические науки .– 2015.–Т. 85, № 7.–С. 73-79.
2. Стефанюк В.Л., Алхуссайн А.Х. Способ шифрования методом расщепления, Роспатент выдал патент на изобретение 08.12.2015.
3. Алхуссайн А.Х., Детерминированный генетический алгоритм в криптографии // Естественные и технические науки.– 2016.– Т. 93. , № 3.– С.126-129.
4. Стефанюк В.Л., Алхуссайн А.Х. Симметричное шифрование на основе метода расщепления // Естественные и технические науки.– 2016.– Т.93. , № 3.– С.130-133.
5. Stefanyuk V.L., Alhussain A.H. Symmetric Encryption on the Base of Splitting Method // Bulletin of Peoples' Friendship University of Russia, Series Mathematics. Information Sciences. Physics. – 2016.–№ 2.– P.53-61.
6. Стефанюк В.Л., Алхуссайн А.Х. Управление степенью защиты символьной информации с использованием метода целочисленного расщепления// Искусственный интеллект и принятие решений.– 2016.– № 4.– С.86-91.
7. Алхуссайн А.Х., Стефанюк В.Л. Вероятностные свойства процедуры расщепления // Искусственный интеллект и принятие решений .– 2017.– № 3.– С.49-57.
8. Алхуссайн А.Х., Некоторые результаты вероятностного анализа стойкости защиты информации методом целочисленного расщепления символов// Естественные и технические науки .– 2018.– № 12.– С. 380-381.

Другие статьи и материалы конференций:

9. Stefanuk V. L., Alhussain A. H. Symbolic Management of the Degree of Information Security by Integer Splitting // Scientific and Technical Information Processing.– 2017.–Vol. 44, Issue 6.– pp. 450-454. (Scopus)
10. Amanie Hasn Alhussain, Comparison between integer splitting cipher and traditional substitution ciphers, based on modular arithmetic // IOP Conf. Series: Materials Science and Engineering .– 2020.– Vol. 919.– 052004 .– pp. 1-6 (Scopus).
11. Amanie Hasn Alhussain, Asymptotic secrecy of the information protection by the usage of simple integer splitting method // IOP Conference Series: Materials Science and Engineering .– 2020.– Vol. 862 .– Issue 5, pp. 052032.– pp. 1-7 (Scopus).
12. Stefanuk, V.L. Alhussain, A.H., Absolute Secrecy Asymptotic for Generalized Splitting Method // Advances in Intelligent Systems and Computing .– 2020 .– 1156 AISC .– pp. 422-431 (Scopus).
13. Amanie Hasn Alhussain, The effectiveness of symbolic integer splitting method over both synchronous stream ciphers and perfectly secret ciphers// International Conference on Engineering Systems 2020.–Journal of Physics: Conference Series.– 2020.– Vol. 1687.– 012006.– pp. 01-08 (Scopus).
14. Alhussain A., Stefanuk V.L. The quantitative comparison between the integer splitting cipher and the traditional gamma cipher// CEUR Workshop Proceedings.– 2021.– Vol. 2899.– pp. 151–161(Scopus).

15. Стефанюк В.Л, Алхуссайн А.Х. Криптография с симметричным ключом с использованием генетического алгоритма // КИИ-2014, четырнадцатая национальная конференция по искусственному интеллекту с международным участием: РИЦ «Школа».−Казань, 2014.− Т. 1.− С. 267-275.
16. Alhussain A.H. A Literature Survey on the Usage of Genetic Algorithms in Recent cryptography Researches // International Journal of Scientific Research in Science, Engineering and Technology (IJSRSET).−Vol.1.−Issue 3.−India, 2015.− pp. 22-25.
17. Alhussain A.H., Stefanuk V.L. Using deterministic genetic algorithm to increase the security level of xor encryption // VIII международной научной конференции «Приоритеты мировой науки: эксперимент и научная дискуссия.−Южная Каролина, Северный Чарльстон.− США, 2015.− С.15-18.
18. Alhussain A.H. Improving the security level of cryptographic keys of gamma cipher // VIII молодёжной международной научно-практической конференции студентов, аспирантов и молодых учёных «Шаг в будущее: теоретические и прикладные исследования современной науки».− North Charleston, SC, USA, 2015.− С.12-14.
19. Alhussain A.H., Stefanuk V.L. Improvement of randomness level of pseudorandom number generators in cryptography // 2nd International Scientific Conference “Theoretical and Applied Sciences in the USA”.− New York, USA, 2015.− pp. 172-177.
20. Alhussain A.H., Stefanuk V. L. Using Genetic Algorithm to improve periodic level of pseudorandom number generators // «The First European Conference on Informational Technology and Computer Science»: East West−Vienna, Austria, 2015.− pp. 25-34.
21. Алхуссайн А.Х., Улучшение сгенерированного криптографического ключа с помощью генетического алгоритма // III-й Международной летней школы-семинара по искусственному интеллекту для студентов, аспирантов и молодых ученых "Интеллектуальные системы и технологии: современное состояние и перспективы" (ISyT’2015) .− Тверь, 2015.− С.107-118.
22. Alhussain A.H., A Literature Survey on the Usage of Genetic Algorithms in Key Generation // Труды конференции. The Strategies of Modern Science Development: Proceedings of the VIII International scientific−North Charleston, USA, 2015.− pp. 12-14.
23. Alhussain A.H., A Literature Survey on the Usage of Genetic Algorithms in Creating New Encryption Algorithm // The Strategies of Modern Science Development: Proceedings of the VIII International scientific−North Charleston, USA, 2015.− pp. 15-17.
24. Стефанюк В.Л., Алхуссайн А.Х. Криптография и кодирование как методы защиты информации// Информационно-Телекоммуникационные Технологии и Математическое моделирование высокотехнологичных систем: РУДН.− Москва, 2016 .− С.181-182.
25. Amanie Hasn Alhussain, Probabilistic analysis of the secrecy of information protection by using splitting method // The latest research in modern science: experience, traditions and innovations: Collected scientific articles of the IX International scientific conference on June 20-21, Morrisville, NC, USA, 2019.− pp. 06-11.

Алхуссайн Амани (Россия)

ВЕРОЯТНОСТНЫЙ АНАЛИЗ СТОЙКОСТИ ЗАЩИТЫ ИНФОРМАЦИИ МЕТОДОМ ЦЕЛОЧИСЛЕННОГО РАСЩЕПЛЕНИЯ СИМВОЛОВ

В диссертации предложена новая процедура, названная целочисленным расщеплением, как один из способов применения модульной арифметики в области защиты информации, и приведены основные определения и понятия этой процедуры. Описаны математические функции возникающих преобразований, исследованы их свойства и доказаны основные утверждения, оправдывающие применимость процедуры расщепления в задачах обеспечения стойкости защиты информации.

Определены математические модели, позволяющие использовать процедуру символьного расщепления для защиты и восстановления информации, описана модель симметричной защиты символа, а также проведен вероятностный анализ стойкости защиты при использовании символьного расщепления.

Приведено сравнение некоторых потоковых методов защиты с методом, основанным на расщеплении, а также дано сравнение известных абсолютно стойких методов защиты с расщеплением. Проведено сравнение между традиционными методами замены, основанными на операциях модульной арифметики, и методом символьного расщепления. Показаны отличия между расщеплением и китайской теоремой об остатках. В заключение показаны иллюстративные примеры работы расщепления.

Alhussain Amanie (Russia)

PROBABILISTIC ANALYSIS OF THE ASYMPTOTIC SECRECY OF INFORMATION ENCRYPTION BASED ON THE INTEGER SPLITTING METHOD

In the thesis a new procedure, called integer splitting, is proposed as a new way of applying modular arithmetic in the field of information protection. The basic definitions and the concepts of this procedure are presented. The mathematical models of the corresponding transformation with this procedure are described, also, their properties are investigated, and the main theorems showing the applicability of the splitting procedure in the field of information protection are proven.

The mathematical models of symmetric integer splitting method are described and the probabilistic analysis of the asymptotic secrecy of integer splitting encryption method is provided.

Some other streaming encryption methods are compared with the proposed splitting encryption method. A comparison of traditional perfect secrecy methods with the splitting one is presented as well as a comparison between replacement encryption methods based on modular arithmetic operations and the encryption method of symbolic splitting is given. As well, the comparison between the splitting and the Chinese remainder theorem is provided. In addition, some illustrations of the splitting encryption method in the field of information security are shown.