ЧЕБЫШЕВСКИЙ СБОРНИК
Том 15 Выпуск 1 (2014)

УДК 519.72

МАТЕМАТИЧЕСКАЯ МОДЕЛЬ СИСТЕМЫ ЗАЩИТЫ ИНФОРМАЦИИ НА ОСНОВЕ ДИОФАНТОВА МНОЖЕСТВА

В. О. Осипян, А. В. Мирзаян (Краснодар),
Ю. А. Карпенко (Майкоп),
А. С. Жук, А. Х. Арутюнян (Краснодар)
rrwo@mail.ru

Аннотация

Развитие ассиметричной криптографии началось с появления разработки первой рюкзачной системы защиты информации, когда в 1978 году Ральф Меркель и Мартин Хеллман предложили использовать разные ключи для прямого и обратного преобразования данных при шифровании. Это была одна из первых криптосистем с открытым ключом, но она оказалась криптографически нестойкой. Позже Ади Шамир показал, что система Меркля-Хеллмана является ненадежной, и на данный момент эта модель, как и многие, основанные на ней были скомпрометированы. Как следствие, авторитет рюкзачных систем защиты информации был занижен. Тем не менее, некоторые из них, до сих пор считаются стойкими, например, модель, предложенная в 1988 году Беном Шором и Рональдом Ривестом. Несмотря на это попытки ее усовершенствования до сих пор не прекращаются, о чем свидетельствуют цикл работ Осипяна В.О. и других авторов. Более полный обзор работ в области анализа системы Меркля-Хеллмана и ее развития дан Б. Шнайером. Отметим особо, что все нестандартные задачи о рюкзаках K_G (обобщенная задача), K_U (универсальная или суперобобщенная задача), K_F (функциональная задача), впервые сформулированные и введенные Осипяном В.О., принадлежат классу NP-полных задач.

В данной работе обоснованы диофантовы трудности, возникающие при поиске уязвимостей в указанных системах защиты информации. На основе анализа ранее предложенных рюкзачных моделей выявлены качественные особенности нестандартных рюкзачных систем, повышающие их стойкость к известным атакам. Предлагается математическая модель полиалфавитной криптосистемы, в которой алгоритм обратного преобразования закрытого текста сводится к алгоритмически неразрешимой проблеме для аналитика. В статье красной нитью проходит идея К. Шеннона, который

считал, что наибольшей неопределённостью при подборе ключей обладают криптосистемы, содержащие диофантовы трудности.

Ключевые слова: рюкзачные системы защиты информации; стойкость алгоритма; диофантовы трудности; рюкзачный алгоритм; рюкзачный вектор; открытый текст; ключ; шифртекст.

MATHEMATICAL MODEL OF INFORMATION SECURITY SYSTEMS BASED ON DIOPHANTINE SETS

V. O. Osipyan, A. V. Mirzayan (Krasnodar),
Y. A. Karpenko (Maikop),
A. C. Zhuk, A. H. Arutyunyan (Krasnodar)
rrwo@mail.ru

Abstract

Development of the asymmetric cryptography started with the appearance of the first knapsack information protection system, when, in 1978, Ralph Merkel and Martin Hellman proposed to use different keys for forward and reverse mapping data for encryption. Now this model, like many based on are considered to be insecure. As a result the authority of knapsack systems was low. However, some of these systems are still considered persistent, for example, the model proposed in 1988 by Ben Shore and Ronald Rivest.

In the article stated and solved the problem of argumentation of cryptographic strength of the non-standard knapsack information security systems. Justified diophantine difficulties that arise in the study of vulnerabilities of the investigated information security systems. Revealed the qualitative features of non-standard knapsack systems that increase their resistance to known attacks. In this paper, we propose a mathematical model of polyalphabetic cryptosystem, in which the algorithm of inverse transformation of closed text is algorithmically unsolvable problem for the analyst. It's permeated with the idea K.Shennon, who believed that cryptosystems, containing Diophantine problems, have the greatest variation in the selection of key.

Keywords: knapsack information security system; resistance of algorithm; diophantine difficulties; knapsack algorithm; knapsack vector; plain text; key; ciphertext.

1. Введение

Исходя из теоретических положений [1] — [9] построения стойких и эффективных моделей систем защиты информации (СЗИ), отметим особо, что все математические задачи являются моделями сокрытия и защиты информации, а

решения этих задач соответствуют правильным ключам. Следовательно, выбор подходящей труднорешаемой задачи, в частности NP-полной задачи, позволяет смоделировать систему защиты информации на должном уровне. Особенно, если этот выбор, как отмечал К. Шеннон [1], связан с задачей, которая содержит диофантовы трудности. Заметим, что все нестандартные задачи о рюкзаках K_G (обобщенная задача), K_U (суперобобщенная задача), K_F (функциональная задача), впервые введенные автором [10], принадлежат классу NP-полных задач.

В работе, на основе многостепенных систем диофантовых уравнений [10, 11, 12],

$$(x_1, x_2, \ldots, x_m) \stackrel{n}{=} (y_1, y_2, \ldots, y_m) \tag{1}$$

предлагается математическая модель полиалфавитной криптосистемы, алгоритм обратного преобразования (дешифрования) закрытого текста которого сводится к алгоритмически неразрешимой проблеме для аналитика.

Так как при $n \geqslant m$ система (1) имеет лишь тривиальные решения [4, 5, 6]:

$$(a_1, a_2, \ldots, a_m) \stackrel{n}{=} (b_1, b_2, \ldots, b_m) \tag{2}$$

– совокупность $a_1, a_2, \ldots, a_m$ значений переменных $x_1, x_2, \ldots, x_m$ отличается от совокупности $b_1, b_2, \ldots, b_m$ значений переменных $y_1, y_2, \ldots, y_m$ лишь порядком следования значений, т. е. $\{a_1, a_2, \ldots, a_m\} = \{b_1, b_2, \ldots, b_m\}$, то исследуются только решения многостепенной системы диофантовых уравнений n-ой степени (1), когда $n < m$.

2. Алфавитная модель рюкзачной криптосистемы

Сопоставим каждому решению (2) системы (1) два рюкзака

$$A = (a_1, a_2, \ldots, a_m) \text{ и } B = (b_1, b_2, \ldots, b_m).$$

Числовые рюкзаки A и B размерности m будем называть равносильными между собой, имеющими степень n, если компоненты A и B являются решениями диофантово уравнения (1). Этот факт будем записывать следующим образом:

$$A \stackrel{n}{=} B \text{ или } (a_1, a_2, \ldots, a_m) \stackrel{n}{=} (b_1, b_2, \ldots, b_m). \tag{3}$$

Очевидно, введённое бинарное отношение (3) обладает следующими свойствами:

1. $A \stackrel{n}{=} A$ (рефлексивность);
2. Если $A \stackrel{n}{=} B$, то $B \stackrel{n}{=} A$ (симметричность);
3. Если $A \stackrel{n}{=} B$, $B \stackrel{n}{=} C$, то $A \stackrel{n}{=} C$ (транзитивность).

Так, например, следующие нормальные [10] двупараметрические рюкзаки размерности $m = 5$ равносильны между собой и имеют степень $n = 4$:

$$(19a + b, 15a + 5b, 11a + 9b, 3a + 17b, 2a + 18b) \stackrel{4}{=}$$

$$\stackrel{4}{=} (a + 19b, 5a + 15b, 9a + 11b, 17a + 3b, 18a + 2b)$$

В частности, при $a = 1$, $b = 2$ имеем следующие нормальные равносильные числовые рюкзаки степени $n = 4$:

$$(21, 25, 29, 37, 38) \stackrel{4}{=} (39, 35, 31, 23, 22).$$

Теперь рассмотрим диофантово представление семейства многостепенной системы

$$(x_1, x_2, \ldots, x_m) \stackrel{n}{=} (p_1, p_2, \ldots, p_m) \tag{4}$$

и определим множество W (диофантово множество) [13]

$$W = \{p_1, p_2, \ldots, p_m | x_1, x_2, \ldots, x_m \stackrel{n}{=} p_1, p_2, \ldots, p_m\}$$

– целых неотрицательных значений упорядоченных наборов $p_1, p_2, \ldots, p_m$, при которых уравнение (4) разрешимо относительно неизвестных
$x_1, x_2, \ldots, x_m$.

Так, например, можно описать множество тех значений a и b, для которых разрешима нормальная многостепенная система диофантовых уравнений степени $n = 5$:

$$(x_1, x_2, \ldots, x_6) \stackrel{5}{=} (2a + b, 3a + 6b, 11a + 7b, 13a + 17b, 21a + 18b, 22a + 23b).$$

В работе приведены диофантово представление многостепенных систем диофантовых уравнений для различных параметров m и n, причем по форме чаще всего общие параметрические решения задаются в виде систем линейных уравнений. Поэтому, для описания множества параметрических решений и определения диофантово представления семейства многостепенных систем, можно применить следующую основную теорему [9].

Теорема 1. *Пусть имеются две пары равносильных числовых рюкзаков, первая из которых представляет собой произвольное параметрическое решение многостепенной системы диофантовых уравнений n-ой степени*

$$(x_1, x_2, \ldots, x_m) \stackrel{n}{=} (y_1, y_2, \ldots, y_m) \tag{5}$$

а вторая – любое расширение первой:

$$(a_1, a_2, \ldots, a_m) \stackrel{n}{=} (b_1, b_2, \ldots, b_m),\ (\textit{или } A \stackrel{n}{=} B),\ 1 \leqslant n < m;$$

$$(c_1, c_2, \ldots, c_k) \stackrel{n+t}{=} (d_1, d_2, \ldots, d_k),\ (\textit{или } C \stackrel{n+t}{=} D),\ t \geqslant 1,\ 1 \leqslant n + t < k.$$

Тогда задача о равносильных числовых рюкзаках (A, v) (или (B, v)) разрешима, и её решение совпадает с решением для входа (C, v) (или (D, v)).

Доказательство. Пусть $A \overset{n}{=} B$ – два равносильных рюкзака, построенных на основе двупараметрических решений системы (3) в виде $x_i = x_i(a,b)$, $y_i = y_i(a,b)$, $i = 1 \dots m$. Очевидно, если вход (A, v) (или (B, v)) имеет решение для заданных значений параметров a и b, то имеет решение и вход (C, v) (или (D, v)), так как он является расширенным равносильным рюкзаком, полученным из рюкзака A.

Для построения эффективных систем защиты информации на основе диофантова представления многостепенных систем диофантовых уравнений, следует учитывать замечания.

1. Целесообразно рассмотреть модели открытых СЗИ, для которых числовые равносильные рюкзаки $C \overset{n+t}{=} D$ выступают в качестве открытых ключей, а соответственно рюкзаки $A \overset{n}{=} B$ – закрытых. Причём в отличие от стандартного сильного модульного умножения для рюкзачных систем, здесь относительно равносильных рюкзаков C и D можно предусмотреть и сложение по модулю $r > \max\{a, b\}$ с соответствующими ограничениями на операции по заданному модулю r, где $a > \max\{a_i\}$, $b > \max\{b_i\}$.

2. В зависимости от выбранных значений параметров a и b указанные рюкзаки могут быть как инъективными, так и не являться таковыми. В случае инъективных рюкзаков следует рассмотреть модели моноалфавитных СЗИ, а в противном случае – либо модели полиалфавитных СЗИ, либо одинаковые шифры исключить из рассмотрения.

3. Следует определять значения параметров a и b таким образом, чтобы рассмотренные рюкзаки были сверхрастущими или нормальными. В этом случае соответствующая задача о рюкзаке будет разрешима либо за линейное время, либо она будет принадлежать классу NP-полных задач.

Для удобства, представим математическую модель алфавитной криптосистемы в виде следующего кортежа [14]:

$$\sum\nolimits_0 = \langle M^*, E(m), D(s), S^* | V(E(m), D(s)) \rangle, \tag{6}$$

где M^* множество всех сообщений $m = m_1, m_2, \dots, m_k$ (открытых текстов) над буквенным или числовым алфавитом M. Здесь m_i, $i = 1 \dots k$ – элементарные сообщения (в частности, буквы или конкатенация букв из алфавита M); S^* – множество всех шифртекстов (криптограмм) $s = s_1, s_2, \dots, s_k$ криптосистемы (4); $E(m)$ – алгоритм прямого преобразования (шифрования) сообщения m в s; $D(s)$ – алгоритм обратного преобразования (дешифрования) шифртекста (криптограммы) s в $m \in M^*$. Подчеркнем, что алгоритмы $E(m)$ и $D(s)$ алфавитной криптосистемы связаны между собой таким образом – $V(E(m), D(s))$, что всегда произвольное сообщение $m = m_1, m_2, \dots, m_k \in M^*$ однозначно преобразовывается в соответствующую криптограмму $s = s_1, s_2, \dots, s_k \in S^*$ и наоборот: по криптограмме s всегда однозначно восстанавливается переданное сообщение m (для нас алфавит – непустое упорядоченное множество букв).

Альтернативным обозначением алгоритмов $E(m)$ и $D(s)$ для алфавитной криптосистемы (4) является K_E (или F_E) и K_D (или F_D) соответственно – как принято считать в классической криптографии [7, 8]. Мы их назовём иначе ключами (или функциями) шифрования и дешифрования соответственно.

Так, например, алфавитные модели рюкзачных криптосистем на основе конструктивного и закрытого рюкзаков можно представить как

$$\sum\nolimits_{D1} = \langle M^*, K_E(A,n), K_D(B,n), S^* | A \stackrel{5}{=} B \rangle$$

и

$$\sum\nolimits_{D2} = \langle M^*, K_E(A,X), K_D(B,Y), S^* | x_1^5 + x_2^5 + \ldots + x_5^5 = y^5 \rangle$$

соответственно.

Авторы не претендует на полноту освещения аналогичных математических моделей алфавитных криптосистем.

По техническим причинам удобнее рассмотреть построение математической модели полиалфавитной криптосистемы с открытым ключом на основе разработок автора [10, 14, 16], например, на основе обобщённого нормального рюкзака $\widetilde{A_P}$. Для простоты изложения в качестве элементарных сообщений рассмотрим буквы заданного алфавита, например, – заглавные буквы английского языка и пробел в качестве разделителя (см. табл. 1), причем для реализации алгоритмов E и D мы рассмотрим изоморфный к алфавиту M числовой алфавит $M_N = \{1, 2, \ldots, N\}$.

Таблица 1: Полиалфавитное соответствие между буквами открытого текста m и их числовыми эквивалентами

A	B	C	...	Z	_
$\{A\}$	$\{B\}$	$\{C\}$	...	$\{Z\}$	$\{\_\}$

Здесь $\{*\}$ означает множество шифров для символа $*$, при этом все множества шифров не пересекаются. В частности, в таблице 1 в качестве $\{*\}$ можно выбрать различные между собой числа заданной длины l из множества M_N. Очевидно, для подавления частоты встречаемости букв, длины могут быть разные для различных букв открытого текста.

В работе ко всем алфавитно-функциональным рюкзачным криптосистемам применяется раундовая функция: обобщенный аналог схемы Меркела-Хеллмана [2] и, красной нитью, проходит идея К. Шеннона, который считал, что наибольшей неопределённостью при подборе ключей обладают криптосистемы, содержащие диофантовы трудности [1].

Для реализации основной идеи автора [10] рассмотрим следующий иллюстрационный пример. Здесь отдельные детали функционирования практической криптосистемы опущены.

Предварительно буквы открытого текста $m = TRUST$ зашифруем таким образом ($s = s_1, s_2, \ldots, s_5$), как это показано в таблице 2.

Таблица 2: Полиалфавитное соответствие между буквами открытого текста m и их числовыми эквивалентами

m_i	T	R	U	S	T
s_i	2	4	5	6	3

Для указанного открытого текста m длины элементарных сообщений равны

$$l(T) = 2,\ l(R) = l(U) = l(S) = 1$$

соответственно, поэтому в качестве порогового значения p можно установить $p = 7$ и определить сверхрастущий обобщенный рюкзак $\widetilde{A_P}$ (или нормальный обобщенный рюкзак) как

$$\widetilde{A_P} = (3, 19, 133, 939, 6565),$$

и соответствующий шифртекст

$$S = \widetilde{A_P} \cdot S^T = (3, 19, 133, 939, 6565) \cdot (2, 4, 5, 6, 3)^T = 26076.$$

Очевидным образом по шифру $S = 26076$ однозначно восстанавливаются все буквы открытого текста m и их порядки следования, т.к.

$$26076 = 2 \cdot 3 + 4 \cdot 19 + 5 \cdot 133 + 6 \cdot 939 + 3 \cdot 6565.$$

Предлагается алфавитная модель рюкзачной криптосистемы, идея построения которой заключается в том, что легальный пользователь системы связи лишь по одному шифру самостоятельно определяет диофантово представление семейства многостепенной системы диофантовых уравнений. Сложность для аналитика заключается в том, чтобы найти само решение известной ему многостепенной системы диофантовых уравнений (1). Эту задачу легальный пользователь решает легко, так как он знает, как найти сами параметрические решения диофантово уравнения для высоких степеней.

Очевидно, рассматриваемый пример является лишь демонстрацией идеи приложения диофантовых и систем многостепенных диофантовых уравнений в области криптографии. Также очевидно, что указанная модель криптосистемы далека от практического приложения, так как многие аспекты прикладной криптографии здесь опущены ради реализации идеи К. Шеннона.

3. Заключение

В заключение отметим, что все рассмотренные СЗИ можно применить как криптосистемы с обнаружением и исправлением канальных и других ошибок – в силу того, что каждому элементарному сообщению можно сопоставить два шифра – по одному для каждого из двух равносильных рюкзаков заданной степени.

Отметим также, что приведённые автором [10] криптосистемы легко модифицируются для создания соответствующих систем электронной цифровой подписи.

СПИСОК ЦИТИРОВАННОЙ ЛИТЕРАТУРЫ

1. Shannon C. Communication theory of secrecy systems Bell System Techn. J. 28, № 4 — 1949. P. 656-715.

2. Diffie W., Hellman M. New directions in cryptography // IEEE Transactions on Information Theory. 1976. Vol. 22. pp. 644-654.

3. Rivest R.L., Chor B. A knapsack-type public key cryptosystem based on arithmetic in finite fields // IEEE Transactions on Information Theory. 1988. Vol. 34. No. 5. pp. 901-909.

4. Shamir A. A polynomial-time algorithm for breaking the basic Merkle - Hellman cryptosystem // Information Theory, IEEE Transactions. 1984. Vol. 30. No. 5. pp. 699–704.

5. Lenstra, Jr. H.W. Integer Programming with a Fixed Number of Variables // Mathematics of Operations Research. 1983. Vol. 8. No. 4. pp. 538-548.

6. Vaudenay S. Cryptanalysis of the Chor-Rivest cryptosystem // CRYPTO. 1998. pp. 243-256.

7. Саломаа А. Криптография с открытым ключом. М.: ИЛ, 1995. 380с.

8. А. П. Алфёров, А. Ю. Зубов, А. С. Кузьмин, А. В. Черемушкин Основы криптографии: учебное пособие для студентов ВУЗ. М.: Гелиос АРВ, 2002. 480 с.

9. В. О. Осипян, А. С. Арутюнян, С. Г. Спирина Моделирование ранцевых криптосистем, содержащих диофантовую трудность // Чебышевский сборник. 2010. Т. XI, вып. 1. С. 209–217.

10. Осипян В. О. Моделирование систем защиты информации содержащих диофантовы трудности. Разработка методов решений многостепенных систем диофантовых уравнений. Разработка нестандартных рюкзачных криптосистем: монография. LAP, 2012. 344 с.

11. Gloden A. Mehrgradide Gleichungen. Groningen, 1944.

12. Dickson L. E. History of the Theory of Numbers. Vol.2. Diophantine Analysis. N.-Y. 1971.

13. Матиясевич Ю. В. Диофантовы множества // Успехи мат. наук. 1972. Т. 27, вып. 5. С. 185–222.

14. Osipyan V. O. Buiding of alphabetic data protection cryptosystems on the base of equal power knapsacks with Diophantine problems // ACM, 2012, pp.124–129.

15. В. О. Осипян, К. В. Осипян Криптография в упражнениях и задачах. М.: Гелиос АРВ, 2004. 144 с.

16. Osipyan V. O. Different models of information protection system, based on the functional knapsack // ACM, 2011. pp 215–218.

17. В. О. Осипян, Ю. А. Карпенко, А. С. Жук, А. Х. Арутюнян Диофантовы трудности атак на нестандартные рюкзачные системы защиты информации // Известия ЮФУ. Технические науки. 2013. №12 С. 209–215.

Кубанский государственный университет
Поступило 20.02.2014