А.Х. Алхуссайн, В.Л. Стефанюк

Вероятностные свойства процедуры расщепления[1]

Аннотация. В работе дается вероятностный анализ процедуры защиты информации путем расщепления данных, предложенной ранее авторами статьи. Определено, что вероятности величин, получаемых при расщеплении, сохраняют требуемые свойства независимости и несовместности возникающих случайных переменных. На этом основании сделан вывод, что предложенная ранее процедура расщепления данных с последующим гаммированием остается абсолютно стойкой. Доказана теорема о том, что с ростом глубины расщепления вероятность несанкционированного восстановления символа на приемном конце убывает по экспоненте, что позволяет говорить об асимптотической стойкости расщепления самого по себе. Показаны важные достоинства использования обобщенного расщепления, отсутствующие в традиционном гаммировании, и которые затрудняют вскрытие исходного текста по его содержанию при замене истинно вероятностного источника на генератор псевдослучайных чисел.

Ключевые слова: целочисленное расщепление, псевдослучайные числа, ГПСЧ, гаммирование, обобщенное расщепление, семантическое восстановление, абсолютная стойкость, теорема К. Шеннона, асимптотическая стойкость расщепления.

Введение

Метод целочисленного расщепления, о котором идет речь в данной статье, при потоковой передаче означает замену каждого символа в потоке на цепочку целых чисел. Он предложен нами в публикациях [1-3] в качестве одного из способов применения модульной арифметики в задачах обеспечения требуемого контроля уровня защиты информации. В работе [1] этот метод описан в деталях и приведены его основные определения и понятия, когда каждое целое число заменяется (по базе другого целого числа) на последовательность k целых чисел, что названо расщеплением k-го уровня.

Если в работах [1-3] приведены достаточно строгие утверждения относительно процедуры расщепления, как детерминированного алгоритма, то в настоящем исследовании производится анализ процедуры с точки зрения теории вероятностей. Предпринятый здесь подход строится в значительной степени по аналогии с соответствующими работами К. Шеннона [4], относящимися к принципам действия «секретных систем связи».

Прежде чем напомнить основной результат К. Шеннона, относящийся к защите информации методом гаммирования, отметим, что вероятностный анализ с тех пор прочно вошел в практику исследователей, работающих в различных областях. В качестве примера укажем недавнюю публикацию [5], в которой изучается время отклика системы, опирающейся на облачные вычисления. Кстати, облачное хранение информации предлагается в [1] одним из возможных объектов использования процедуры расщепления. В работе [6] приводятся основные теоремы и алгоритмы расчета *вероятностных характеристик* для целого ряда моделей, так называемых мультисервисных сетей.

Другим примером может служить публикация [7]. В ней особо подчеркивается потребность рассматривать стохастические явления не как внешние по отношению к системе, но как присущие самой системе. В работе изучаются

[1] Исследования проводились на кафедре информационных технологий Российского университета дружбы народов. Работа выполнена при финансовой поддержке по программе президиума РАН 1.5 П и фонда РФФИ (грант № 15-07-07486-а).

известные процессы «рождения-гибели». Нам показалось особенно важным то, что предложенный в статье [7] метод позволяет организовать «правильное сочетание детерминированных и стохастических явлений» в системе. Сразу подчеркнем, что подобное сочетание детерминированных и вероятностных явлений в единой системе является также одной из наиболее характерных черт изучаемой системы с расщеплением, предложенной нами ранее в публикациях [1-3].

Однако вероятностные явления, на которых сконцентрированы усилия в настоящей работе, весьма близки к тем, что были изучены К. Шенноном [4], который показал, что защита данных с применением метода гаммирования, использующего логическую процедуру *XOR,* может характеризоваться абсолютной стойкостью. Поскольку гаммирование также используется в нашей процедуре защиты, опирающейся на расщепление, то, прежде всего, естественно проверить выполнение тех требований к используемым случайным величинам, которые лежат в основе результатов К. Шеннона.

С этой целью в разделе 1 кратко представлена математическая модель метода целочисленного расщепления по публикации [1], что необходимо для анализа вероятностных свойств, характерных для этой процедуры. В разделе 2 рассматриваются строгие утверждения, связанные с методом целочисленного расщепления, с точки зрения защиты данных, и приводятся строгие доказательства соответствующих свойств. В первой части этого раздела предполагаются выполненными те требования к вероятностным процессам, которые накладывались К. Шенноном, а именно требования независимости и несовместности вероятностей событий, участвующих в работе системы. Во второй части этого раздела рассматриваются особенности метода расщепления, которые ведут к повышению уровня защиты данных по сравнению с известными методами и в случае, когда такое предположение, вообще говоря, не выполняется, поскольку в системе защиты случайные величины создаются с помощью генератора псевдослучайных величин. В разделе 3 рассматривается и доказывается теорема о том, что при сделанных допущениях вероятность восстановления передаваемого символа экспоненциально быстро убывает с ростом глубины расщепления, свидетельствуя об *асимптотической стойкости* самой процедуры расщепления, вне зависимости от процедуры гаммирования.

1. Алгоритм защиты данных с применением целочисленного расщепления и последующего гаммирования

От источника случайных чисел (или генератора псевдослучайных чисел) поступает величина *r(t)*, необходимая как при передаче путем расщепления, так и при восстановлении информации. В результате расщепления в момент *t* создаются целые числа $\delta^{(2)},\delta^{(3)},.....,\delta^{(k-1)},\delta^{(k)},q^{(k)}$. Затем поступают случайные величины *r(t+1), r(t+2),...., r(t+k+1),* используемые для дополнительной защиты при передаче этого символа путем гаммирования. Предполагается, что величина $r(t) > 0$ превосходит максимальное значение символов по выбранной кодовой таблице.

Математическая модель этапа защиты символа S(t) с кодом a(t):

Результат защиты при расщеплении

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

В публикации [1] приведено определение процедуры, названной *обобщенным расщеплением*, при котором на каждом шаге формирования величин, показанных в (1), используется новая величина $r(t) > 0$ с указанным выше свойством.

Результат защиты при гаммировании

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

Математическая модель этапа восстановления символа:

Результат восстановления после гаммирования

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

Результат восстановления расщеплённого символа

$$\begin{cases} r(t)\oplus Y & при \quad k=1 \\ \dfrac{\left(r(t)-\delta^{(i)}\right)}{q^{(i)}},\ где\ \ i=k,k\text{-}1,...,3,2\ при\ k>1 \end{cases} \quad (4)$$

2. Теоретические результаты

Необходимость использования псевдослучайных чисел вместо обычных вероятностей связана с организацией системы, позволяющей законному адресату восстановить исходное сообщение. Тем не менее, предварительный анализ возможности раскрытия текста посторонним лицом, не имеющим ключа сокрытия текста, является актуальным, поскольку позволяет оценить надежность зашиты в теоретическом аспекте.

Так, в классической работе К. Шеннона [4] было показано, что при определенных требованиях к «истинно» вероятностным процессам достижима абсолютная стойкость гаммирования, т.е. невозможность вскрытия текста посторонним лицом, не располагающим ключом защиты информации. Приводимая ниже теорема свидетельствует о достижении абсолютной стойкости предложенной нами ранее процедуры расщепления [1-3], в которой предусмотрено последующее применение процедуры гаммирования.

Предварительно следует подчеркнуть, что вероятностный анализ встречается и в других областях, далеких от проблем передачи информации. Например, в публикации [7] особо отмечается роль вероятностных процессов, которые происходят именно «внутри» исследуемой системы, а не вне ее. В случае использования процедур расщепления и гаммирования для защиты информации подобные процессы играют особо важную роль и на этапе теоретического анализа, и на этапе реального функционирования системы защиты информации методом расщепления данных.

Следующий результат сближает изучаемую в настоящей статье систему защиты информации с той, что изучалась К. Шенноном [4], который продемонстрировал принципиальную роль независимости и несовместности вероятностных событий, характеризующих работу системы гаммирования.

Теорема 1. *Элементы цепочки, получаемой в случае обобщенного расщепления* [1], *являются независимыми и несовместными, при условии, что вероятностные числа, используемые в расщеплении, являются независимыми и несовместными.*

Доказательство. Случайные числа, используемые при исследовании гаммировании, по предположению удовлетворяют требованиям [8-13]:

$$\Pr(A * B * C * ....) = \Pr(A) * \Pr(B) * \Pr(C) * ..., \quad (5)$$

$$\Pr(A + B + C + ...) = \Pr(A) + \Pr(B) + \Pr(C) + ... \quad (6)$$

В нашем случае в силу алгоритма создания расщепления (1) также имеем, что вероятности чисел $\Pr(a_i)$, составляющие цепочку из k чисел, полученных при расщеплении (Splitting или кратко SPL), удовлетворяют соотношению:

$$\Pr(a_1 * a_2 * .... * a_k) = \prod_{i=1}^{k} \Pr(a_i), \quad (7)$$

т.е. числа, получаемые в результате обобщенного расщепления, оказываются независимыми по вероятности, несмотря на то, что их вычисление производится последовательно вдоль цепочки.

Далее, поскольку события, т.е. числа цепочки, получаемые в результате обобщенного расщепления, являются несовместными, то суммарная вероятность всей цепочки оказывается равной:

$$\Pr(a_1 + a_2 + ... + a_k) = \sum_{i=1}^{k} \Pr(a_i). \quad (8)$$

Формулы (7) и (8) составляют суть этой теоремы. Они говорят о том, что при использовании обобщенного расщепления, опирающегося на вектор независимых и несовместных случайных величин $\vec{r} = (r_1, r_2, ..., r_l)$, элементы цепочки оказываются независимыми и несовместными.

***Следствие* 1.** Из Теоремы 1 следует, что система с расщеплением и последующим гаммированием обладает в условиях этой теоремы *абсолютной стойкостью* в смысле работы [4].

Кроме того, на каждом шаге расщепления каждое частное $q^{(i)}$ расщепляется на два целых числа $\left(\delta^{(i+1)}, q^{(i+1)}\right)$, в соответствие с соответствующими уравнениями $\delta^{(i+1)} = r_j \bmod q^{(i)}$ и $q^{(i+1)} = \left\lfloor \frac{r_j}{q^{(i)}} \right\rfloor$. В силу этого можно заключать, что они не связаны друг с другом, хотя зависят от значения $q^{(i)}$. Важно, что эти величины определяются единственном образом, согласно теореме 0 из работы [1].

Тогда, на каждом шаге расщепления будут создаваться два несвязанных числа, после этого одно из них будет расщепляться на два новых несвязанных числа и т.д. Конечный результат состоит из цепочки, которая содержит целые

числа от каждого уровня и два несвязанных числа, сгенерированные на последнем уровне.

Можно заключать что, в любом случае, если используется расщепление с одним случайным числом r_j или обобщенное расщепление с использованием вектора $\vec{r} = (r_1, r_2, ..., r_l)$, результат расщепления будет состоять из элементов цепочки, которые не связаны друг с другом.

Теорема 2. *Процедура расщепления существенно затрудняет семантическое восстановление исходного текста несанкционированным пользователем.*

Доказательство. В отличие от изолированного гаммирования, в результате его применения *после* процедуры расщепления, достигается повышенная стойкость даже при использовании ГСПЧ. Это происходит благодаря дополнительной защите данных в результате расщепления, поскольку вместо символов текста в канале связи или при защищенном хранении будут передаваться символы, не входящие в обычный алфавит для естественного языка в выбранной кодовой таблице. Кроме того, в результате расщепления в канале могут возникать символы, принадлежащие алфавиту, но не имеющие никакого смыслового оправдания. Поэтому распространенные методы несанкционированного раскрытия текста, основанные на семантическом обнаружении регулярностей при передаче, вообще говоря, перестают действовать.

В Табл. 1 приводится несколько примеров для передаваемого символа при одиночном гаммировании и при гаммировании после расщепления. В таблице исходные псевдослучайные числа были созданы с помощью линейного конгруэнтного генератора [1] со следующими параметрами: начальное значение-3, модуль-500, множитель-3, приращение-8, объем генерации-40. Кроме того, использовался детерминированный генетический алгоритм [1-3], в котором количество поколений-3, длина хромосомы-10. Видно, что многие возникающие еще до применения гаммирования символы, либо являются непечатными, либо никак не оправданными с точки зрения содержания, т.е. семантики текста.[2]

Табл. 1. Результат защиты текста в системе с расщеплением для исходного символа S

Уровень расщепления k	**Результат системы с расщеплением**	**Соответствующие символы**
$k = 1$	279	**S**
$k = 2$	**322 277**	
$k = 3$	322 279 685	
$k = 4$	322 279 713 710	
$k = 5$	322 279 713 708 291	

Таким образом, при расщеплении могут непредсказуемым образом появляться «осмысленные» символы, которые на самом деле никак не связаны с содержанием текста [3].

Кроме того, в отличие от логической процедуры *XOR*, предлагаемая процедура *SPL* не допускает посимвольного применения при восстановлении исходного текста. Поэтому приведенные в Табл. 1 данные получены с помощью достаточно громоздкой процедуры восстановления, возможность которого, однако, гарантируется теоремами 1-3, доказанными еще в работе [1]. Для сравнения с процедурой расщепления полезно привести сводку различных систем, в которых применяется модульная арифметика.

В Табл. 2 представлены некоторые математические модели методов защиты данных на основе операции модульной арифметики для сравнения с математической моделью целочисленного расщепления, которое описано в уравнениях (1) и (2).

Наблюдаются два отличия алгоритма целочисленного расщепления от перечисленных традиционных алгоритмов защиты, применяющих операции модульной арифметики:

Первое - операция с применением модуля используется в традиционных методах только один раз для каждого символа. В предлагаемом алгоритме целочисленного расщепления эта операция используется *k-1* раз.

Второе – модули, используемые во всех традиционных способах защиты (Цезарь, Виженер, Аффинный, Хилла…), основанных на операции модульной арифметики, совпадают с размером соответствующего алфавита, тогда как в алгоритме целочисленного расщепления,

[2] В канале связи или в файле, хранимом, например, в облаке [1-3], разумеется, используется бинарное кодирование. В Табл. 1 приведены результаты кодирования и «символы», которые могут быть восстановлены интеллектуальным агентом, опираясь на известное свойство процедуры *XOR*.

[3] Отметим, что отбрасывать «непечатные» символы без предварительных сведений о содержании передаваемого материала также нельзя.

Табл. 2. Некоторые математические модели методов замены, основанные на операциях модульной арифметики

Шифры	Математическая модель шифрования	Математическая модель дешифрования	Примечание
Цезаря	$y = x + k \pmod n$	$x = y - k \pmod n$	Модуль n- размер алфавита
	где x - символ открытого текста, y - символ шифрованного текста, n - мощность алфавита (кол-во символов), k - ключ. [8-10]		
Виженера	$c_j = m_j + k_j \pmod n$	$m_j = c_j - k_j \pmod n$	Модуль n - размер алфавита
	где m_j - буквы открытого текста, c_j - буквы шифрованного текста, n - количество букв в алфавите, k_j - буквы ключа.		
Аффинный	$E(x) = (ax + b) \mod n$	$D(x) = a^{-1}(x - b) \mod n$	Модуль n - размер алфавита
	где модуль n - размер алфавита, а пара a *И* b - ключ шифра. [8]		
Хилла	$C = E(K,P) = K * P \mod n$	$P = D(K,C) = K^{-1} * C \mod n$	Модуль n - размер алфавита
	где P - вектор-столбцы высоты m, представляющий открытый текст, C - вектор-столбцы высоты m, представляющий зашифрованный текст, K - матрица m x m, представляющая ключ шифрования, модуль n - размер алфавита.		

величина модуля изменяется на каждом шаге работы системы и не связана с размером алфавита. Перечисленные свойства затрудняют семантическое восстановление исходного текста несанкционированным пользователем.

3. Асимптотическая стойкость процедуры расщепления

Возвращаясь к теоретическому исследованию вероятностных особенностей процедуры расщепления, как таковой, следует подчеркнуть некоторое ее специфическое свойство, связанное с глубиной расщепления *k.*

Пусть для определенности речь идет о расщеплении символа S с кодом представления *a* в некоторой кодовой таблице. Теоремы 0 и 1 из работы [1] утверждают, что в принципе по вектору *c* компонент расщепления можно однозначно вычислить *a,* пользуясь теми детерминированными соотношениями, по которым осуществляется расщепление этого символа.

Однако возникает следующий естественный вопрос. Можно ли по компонентам расщепления в явном виде восстановить символ *a,* если в процессе расщеплении участвуют вероятностные, а не только детерминированные события и операции*?* Оказывается, такое восстановление *a* также возможно, но лишь с определенной вероятностью.

Исчерпывающий ответ на поставленный вопрос дает следующая теорема:

Теорема 3. *Вероятность восстановления символа a по результату расщепления с экспоненциально быстро убывает с ростом k, согласно выражению:*

$$\Pr(a) = \left(\frac{1}{2}\right)^{k-1} \times \prod_{i=1}^{k-1}\left(\frac{1}{2}\right)^{i}, k = 2,\ldots \qquad (8)$$

Доказательство. В условиях теоремы примем, что участвующие в расщеплении величины *r(t+1), r(t+2),...., r(t+k+1),* являются несовместными случайными величинами, выпадающими с некоторыми независимыми друг от друга вероятностями. (Отличие от действующей системы защиты символьной информации состоит в том, что это «настоящие» случайные числа, а не псевдослучайные величины, как в работах [1-3].) В силу этого предположения относительно случайных величин, вектор $c^{(k)} \equiv (\delta^{(2)}, \delta^{(3)}, ..., \delta^{(k-1)}, \delta^{(k)}, q^{(k)})$ содержит случайные компоненты, причем, логическая сумма вероятностей составляющих этого вектора и дает вероятность вычисления конкретного расщепленного символа, в данном случае, символа *a*, т.е. $\Pr(a)$.

Начнем рассмотрение со случая $k = 2$, опираясь на (1), поскольку при $k = 1$ расщепление фактически не производится. При вычислении

вероятности $\Pr(a)$ следует учесть, что в силу независимости и несовместности случайных величин *r(t+1), r(t+2),...., r(t+k+1)* и в силу того, что компоненты вектора c вычисляются по детерминированным алгебраическим правилам (выше в статье или [1]) в порядке слева направо, так что, например, $\delta^{(t+1)}$ вычисляется по $\delta^{(t)}$ и $r(t+1)$, то эти компоненты оказываются независимыми по вероятности.

Поясним этот факт на примере $k=2$, который является первым шагом расщепления, создающим две величины, а именно $\delta^{(2)}, q^{(2)}$, на основе «выпадения» случайной величины *r(2)*. Поэтому они являются независимыми случайными величинами, имеющими равные вероятности, в сумме составляющие единицу. Отсюда получаем: $\Pr(\delta^{(2)}) = \Pr(q^{(2)}) = \frac{1}{2}$. В результате вероятность вектора $c^{(2)} = (\delta^{(2)}, q^{(2)})$ равна:

$$\Pr(c^{(2)}) = \Pr(\delta^{(2)} \wedge q^{(2)}) = \Pr(\delta^{(2)}) \times \Pr(q^{(2)}) = \\ = \frac{1}{2} \times \frac{1}{2} = \frac{1}{4} \quad (9)$$

Рассмотрим теперь случай $k=3$. Здесь первый шаг расщепления порождает такие же случайные величины $\delta^{(2)}, q^{(2)}$ с теми же вероятностями, но величина $q^{(2)}$ теперь будет использована на втором шаге расщепления, чтобы создать новую пару величин $\delta^{(3)}, q^{(3)}$ на основе случайной величины *r(3)*. По аналогии с предыдущим эти новые случайные величины будут иметь равные друг другу вероятности «выпадения», составляющие в сумме единицу. При этом вычисляя полную вероятность событий $\delta^{(3)}$ и $q^{(3)}$, следующих из события $q^{(2)}$, получаем:

$$\Pr(\delta^{(3)}) = \Pr(q^{(3)}) = \frac{1}{2} \times \frac{1}{2} = \frac{1}{4}.$$

Вычисляя полную вероятность реализации всех трех компонент вектора в виде тройки событий $c^{(3)} = (\delta^{(2)}, \delta^{(3)}, q^{(3)})$, имеем:

$$\Pr(c^{(3)}) = \Pr(\delta^{(2)} \wedge \delta^{(3)} \wedge q^{(3)}) = \\ = \Pr(\delta^{(2)}) \times \Pr(\delta^{(3)}) \times \Pr(q^{(3)}) = \frac{1}{2} \times \frac{1}{4} \times \frac{1}{4} = \frac{1}{32}. \quad (10)$$

Табл. 3. Вероятности восстановления при использовании расщепления вплоть до *8*-го уровня глубины

Уровень расщепления k	**Вероятность результата защиты расщеплением** $\Pr(c)$
$k=2$	0.25
$k=3$	0.03125
$k=4$	0.00195
$k=5$	0.000061
$k=6$	0.0000009537
$k=7$	0.00000000745
$k=8$	0.000000000029

Продолжая этот анализ путем индукции, получаем выражение (8).

Теорема доказана. Она свидетельствует о том, что процедура расщепления порождает *асимптотическую стойкость* защиты информации, поскольку вероятность восстановления расщепленной величины стремится к нулю при $k \to \infty$. Её можно рассматривать как некоторый асимптотический аналог теоремы К. Шеннона об абсолютной стойкости защиты информации гаммированием [4].

В Табл. 3 приведены числовые значения для вероятности защиты расщеплением, вытекающие из Теоремы 3. т.е. вероятности восстановления символа при различных значениях *k*. На Рис. 1 показан график поведения выражения (8) для тех же уровней расщепления.

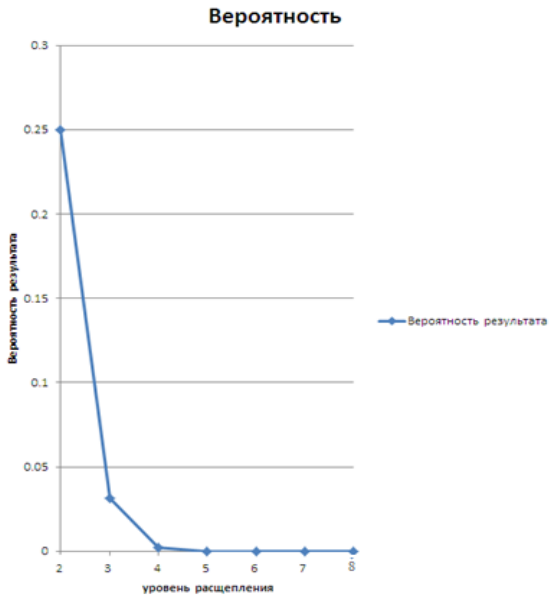

Рис 1. График вероятностей восстановления символа для $k = 2,\ldots,8$

Из выражения (8) видно, что при $k \to \infty$ вероятность вычисления защищенного символа a стремится к нулю.

Теорема 3 и иллюстративные материалы к ней показывают, что процедура расщепления, взятая сама по себе, уже является серьезной защитой при передаче одиночного символа. Доказанная теорема является совершенно новой, но по смыслу она аналогична теореме К.Шеннона об абсолютной стойкости гаммирования [11-13], поскольку расщепления в асимптотике обеспечивает абсолютную стойкость [14, 15]. При этом доказанная Теорема 2 показывает, что процедура расщепления в значительной степени ослабляет возможность раскрытия передаваемого текста за счет учета его содержания.

Заключение

Из доказанных Теорем 1-3, можно сделать вывод, что расщепление обеспечивает более глубокую защиту передаваемой информации от действий злоумышленников различного рода по сравнению с традиционными алгоритмами защиты информации, основанными на операциях модульной арифметики.

Кроме того, гаммирование, при его применении к результату расщепления, остается абсолютной стойким, поскольку результат защиты методом расщепления не содержит никакой информации о значениях символов, к которым применяется гаммирование, причем требования к независимости вероятностей символов и их несовместности, сохраняются.

Особый интерес представляет теорема об *асимптотической стойкости* процедуры расщепления, взятой самой по себе, с ростом глубины расщепления. Аналога такому свойству в утверждениях К. Шеннона нет, и этот результат является совершенно новым.

В отличие от применения изолированного гаммирования при его применении *после* процедуры расщепления достигается повышенная стойкость даже при использовании псевдослучайных чисел от ГПСЧ. Это происходит благодаря дополнительной защите данных в результате расщепления, поскольку кроме символов текста в канале связи (или при защищенном хранении) будут передаваться либо символы, не входящие в обычный алфавит, либо символы, не входящие в состав передаваемого текста, либо символы, являющиеся не имеющим смысла повторением символов такого текста. Поэтому традиционные методы несанкционированного раскрытия текста, основанные на семантике и статистическом обнаружении регулярностей при передаче, вообще говоря, перестают действовать.

Литература

1. Стефанюк В.Л., Алхуссайн А.Х. Контроль степени защиты информации методом целочисленного расщепления// Искусственный интеллект и принятие решений. № 4.2016. С.86-91.
2. Стефанюк В.Л., Алхуссайн А.Х., Симметричное шифрование на основе метода расщепления// Естественные и технические науки. № 3. 2016. С.130-133.
3. Stefanyuk V.L., Alhussain A.H. Symmetric Encryption on the Base of Splitting Method// Bulletin of PFUR, Series Mathematics. Information Sciences. Physics. 2016. № 2. pp.53-61.
4. Шеннон К. Теория связи в секретных системах. Работы по теории информации и кибернетике. М.: Издательство иностранной литературы. 1963.С. 333-402.
5. Горбунова А.В., Зарядов И.С., Матюшенко С.И., Самуйлов К.Е., Шоргин С. Я. Аппроксимация времени отклика системы облачных вычислений. Информация и ее применение. **9**:3.2015. С.32–38.
6. Башарин Г.П., Самуйлов К.Е. , Яркина Н.В. , Гудкова И.А. Новый этап развития математической теории телетрафика// Автоматика и телемеханика. № 12. 2009. С.16–28.
7. Sevastianov L.A. The probability scheme of constructing the mathematical model of shadowed spattering.// Comp. Phys. Comm., 2000, V.130, № 1-2, P.41-46.
8. Романец Ю.В., Тимофееф П.А., Шаньгин В.Ф. Защита информации в компьютерных системах и сетях. М.: Радио и связь. 1999. 328с.
9. Рябко Б.Я., Фионов А.Н. Основы современной криптографии для специалистов в информационных технологиях// Москва: научный мир. 2004. 173с.
10. Саломаа А. Криптография с открытым ключом// Москва: научный Мир. 1996. 318с.
11. José Luis Gómez Pardo, Introduction to Cryptography with Maple// Springer Science & Business Media. 2012. 706p.
12. Douglas R. Stinson, Cryptography: Theory and Practice, Third Edition// CRC Press. 2005. 616p.
13. Mikhail J. Atallah, Algorithms and Theory of Computation Handbook//CRC Press. 1998. 1312p.
14. Serge Vaudenay, A Classical Introduction to Cryptography: Applications for Communications Security//Springer Science & Business Media. 2006. 336p.
15. Hazewinkel M. Encyclopaedia of Mathematics: Coproduct – Hausdorff – Young Inequalities//Springer. 2013. 963 p.

Алхуссайн Амани Хасн. Аспирант Российского университета дружбы народов (РУДН). Окончила РУДН в 2013 году. Количество печатных работ: нескольких десятков. Область научных интересов: теория систем обеспечения безопасности передачи и хранения информации, практическое программирование соответствующих систем, применение детерминированного генетического алгоритма для улучшения вероятностных свойств псевдослучайных последовательностей, анализ авторского метода целочисленного расщепления, изучение вероятностных характеристик защиты информации в различных системах. E-mail: amanie.alhussain@mail.ru

Стефанюк Вадим Львович. Ведущий научный сотрудник Института проблем передачи информации РАН (ИППИ), профессор Кафедры информационных технологий Российского университета дружбы народов. Окончил МГУ им. В.М. Ломоносова в 1962 году. Доктор технических наук. Количество печатных работ: более 250. Область научных интересов: искусственный интеллект, коллективное поведение автоматов, мобильная радиосвязь, мета-экспертные системы, аксиоматика сбора свидетельств, транзакционный анализ в обучении, динамические экспертные системы и их использование, теоретико категорные методы, цепи Маркова-Стефанюка, творческое решение задач, семиотическая интроспекция, целочисленная (модульная) арифметика, безопасная передача символьной информации, метод расщепления, интеллектуальная система подготовки международных конференций. E-mail: stefanuk@iitp.ru

Probability Properties of Splitting Procedure

A.H. Alhussain, V.L. Stefanuk

Abstract. The paper contains a probability analysis of information safety providing with data splitting method that has been proposed before by the present authors. The research was made in analogy to the Claude Shannon's proof of perfect secrecy of gamming procedure under certain properties of gammas. The present paper shows that the probabilities of values obtained after splitting procedure have the required properties of independence and incompatibility. Due to this fact the paper concludes that our splitting procedure, followed by gamming one, remains absolutely safe. Besides, the paper provides a theorem on asymptotic safety of splitting alone as the splitting depth tends to infinity. Independently the present paper shows that the splitting procedure has certain advantages which differ it from traditional gamming, and which make it especially difficult to perform statistical recovery of the original text on the base of its content even in the case when real randomness is replaced with pseudorandom numbers.
Keywords: numerical splitting, generalized splitting, gamming, pseudorandom numbers, GPRN, text semantics, absolute safety, asymptotic absolute safety.

References

1. V.L.Stefanuk, A.H.Alhussain. 2016. Kontrol stepenju zashity informatsii methodom tselochislennogo rascheplenia [Control the Level of Protection the Information by the Usage of Integer Splitting]. Iskusstvenny Intellkt I Prinyatie Reshenii [Artificial Intelligence and Decision Making] 4:86-91.
2. Stefanuk V.L., Alhussain A.H.. 2016. Symmetrichnoe shifrovanie na osnove metoda rascheplenia [Symmetric ciphering based on splitting metod]. Estestvennye i technicheskie nauki [Natural and technical sciences] 93:3:130-133
3. V.L. Stefanyuk, A.H. Alhussain, Symmetric Encryption on the Base of Splitting Method// Bulletin of PFUR, Series Mathematics, Information Sciences, Physics. 2016: 2: 53-61.
4. Shannon C. Communication theory of secrecy systems. Bell System Techn. J., 28: 4: 656-715, 1949.
5. Gorbunova A.V., Zaryadov I.S., Matyushenko S.I., Samyilov K.E., Shorgin S.Ya. 2015. Approximatsia vremeny otklika sistemy oblachnykh vychislenii [Response time approximation for cloud computing]. Informatsia i yeye primenenie [Information and its application], **9**:3:32–38.
6. Basharin G.P., Samyilov K.E., Yarkina N.V., Gudkova I.A. 2009 Novyi etap razvitiya matematicheskoy teorii teletrafika [New stage of development of nathematical theory of teletraffic]. Avtomatica i telemechanika [Automation and Remote Control]. 12: 6–28.
7. Sevastianov L.A. The probability scheme of constructing the mathematical model of shadowed spattering. 2000// Comp. Phys. Comm., 130: 1-2: 41-46.
8. Romanets Yu.V., Tomofeev P.A., Shan'gin V.F. 1999. Zashita informatsii v komputernych systemach I setyach [Safety of information in computer systems and nets] // M: Radio i svyz' [Moscow: Radio and Communication]. 328p.
9. Ryabko B.Ya., Phionov A.N. Osnovy sovremennoi kriptographii for specialistov v informacionnych technologiyach [The basis for modern cryptography for information technology specialists] // Moskva: Nauchnyi Mir [Moscow: Scientific World], 2004. 173p.
10. Salomaa A., Public-Key Criptography// N.Y.: Springer-Verlag, 1990. 318p.
11. José Luis Gómez Pardo, Introduction to Cryptography with Maple// Springer Science & Business Media, 2012. 706p.
12. Douglas R. Stinson, Cryptography: Theory and Practice, Third Edition// CRC Press, 2005. 616 p.

13. Mikhail J. Atallah, Algorithms and Theory of Computation Handbook//CRC Press, 1998. 1312p.
14. Serge Vaudenay, A Classical Introduction to Cryptography: Applications for Communications Security//Springer Science & Business Media, 2006. 336p.
15. M. Hazewinkel, Encyclopaedia of Mathematics: Coproduct – Hausdorff – Young Inequalities//Springer–2013. 963 p.

Alhussain Amanie Hasn. Postgraduate of Russian University for Peoples' Friendship; 117198, Moscow, Miklucho-Maklaya str. 3, Russian Federation. E-mail: amanie.alhussain@mail.ru

Stefanuk Vadim Lvovitch. Leading Researcher of the Institute for Information Transmission Problems of Russian Academy of Sciences; 101447, Moscow, B. Karetny per. 19, Russian Federation. Professor of Russian University for Peoples' Friendship; 117198, Moscow, Miklucho-Maklaya str. 3, Russian Federation; Doctor of Sciences, Senior Researcher, Academician of Russian Academy for Natural Sciences, ECCAI Fellow, E-mail: stefanuk@iitp.ru