Информатика и информационные технологии

УДК 519.72

СИСТЕМА КРИПТОГРАФИЧЕСКОГО ПРЕОБРАЗОВАНИЯ ЧИСЕЛ ЛИНЕЙНЫМИ РЕКУРРЕРТНЫМИ ФОРМАМИ.

А.В. Анисимов

Киевский национальный университет имени Тараса Шевченка

Рассматривается двухступенчатая система кодирования чисел, основанная на представлении чисел в виде $aP_n + bQ_n$, где P_n и Q_n линейные рекуррентные последовательности. Последовательности P_n и Q_n определяются разложением в цепные дроби квадратичных иррациональностей вида $\frac{a+\sqrt{b}}{c}$. В системах симметричной криптографии числа a, b и c является ключами.

Ключевые слова: линейные формы, цепные дроби, недетерминированная криптография

Розглядається двоступенева система кодування чисел, заснована на представленні чисел у вигляді $aP_n + bQ_n$, де P_n та Q_n лінійні рекурентні послідовності. Послідовності P_n і Q_n визначаються розкладанням в ланцюгові дроби квадратичних ірраціональностей виду $\frac{a+\sqrt{b}}{c}$. У системах симетричної криптографії числа a, b і c є таємними ключами.

Ключові слова: лінійні форми, ланцюгові дроби, недетермінірована криптографія

Введение

Рассматривается двухступенчатая система симметрического криптографического числового преобразования, основанная на представлении чисел в виде линейных комбинаций соседних пар линейных рекуррентных последовательностей. На первом этапе исходное число x недетерминировано раскладывается в сумму, $x = aP_n + bQ_n$ где P_n и Q_n соответственно числитель и знаменатель подходящей цепной дроби квадратичной иррациональности вида $\frac{a+\sqrt{b}}{c}$. Выбор n не детерминирован. На втором этапе тройка чисел (a, b, n) взаимно — однозначно кодируется при помощи линейной рекуррентной последовательности P : $P_{m+1} = a_{m+1}P_m + P_{m-1}$, которая также определяется другой квадратичной иррациональностью. Числа a, b и c являются симметричными ключами системы шифрования. Доказывается абсолютная стойкость системы

в случае пассивных атак (наблюдение только потока шифротекстов без доступа к машине шифрования). Дополнительно возможно первоначальное недетерминированное изменение числа x при помощи внесения случайных параметров типа OAEP [1].

1. Линейные формы числовых последовательностей

Все числа, рассматриваемые в данном разделе, принадлежат натуральному ряду. В дальнейшем это предполагается по умолчанию.

Пусть $U = u_1, u_2, \ldots$ и $V = v_1, v_2, \ldots$ — две заданные бесконечные последовательности положительных натуральных чисел. Предполагаем, что последовательность U неограниченна сверху. Это означает, что для любого числа c существует такой номер n, что для всех $m, m > n$ выполняется неравенство $u_m > c$.

Выражение $au_n + bv_n$, $u_n \in U$, $v_n \in V$ называем линейной (U,V) формой.

Представление натурального числа x в виде:

$$x = au_n + bv_n, \tag{1}$$

где $a > 0$, называем линейным (U,V) представлением. Число n называем рангом представления (1).

Представление (1) определяется тройкой чисел (a, b, n).

Если для всех n члены последовательности u_n и v_n взаимно просты, то последовательности U и V называем ортогональными. В этом случае используем обозначение $U \perp V$.

В случае $U \perp V$ для любого числа x и любого ранга n существует линейное представление (1) такое, что $0 < a \le v_n$ (Приложение, Лемма 5.1).

При фиксированном ранге n представление (1) называем лево-каноническим, если $a > 0$ и коэффициент a — минимальный.

Очевидно, что если $U \perp V$ то, лево-каноническое представление (1) фиксированного ранга единственное.

В качестве примера рассмотрим линейные формы Фибоначчи.

В этом случае последовательность U задается последовательностью чисел Фибоначчи:

$$F_0 = 0, F_1 = F_2 = 1, \ F_n = F_{n-1} + F_{n-2}, \ n \ge 3.$$

Последовательность V получается сдвигом U на один шаг влево $u_n = F_n$, $v_n = F_{n+1}$. Рассматриваем представления чисел вида $x = aF_n + bF_{n+1}$.

Примеры разложения чисел в максимальные линейные формы Фибоначчи:

$$31 = 3F_5 + 2F_6, 123 = 2F_9 + F_{10}, 197 = 23F_6 + F_7 = 10F_6 + 9F_7.$$

Заметим, что для числа 197 существует два представления максимального ранга. Для числа 197 представление с наименьшим первым коэффициентом $197 = 10F_6 + 9F_7$ лево-каноническое.

Если $a > 0,\ b \geq 0,$ то представление (1) называем положительно-определенным.

Положительно-определенное представление (1) максимального ранга называем максимальным.

Если $a > 0,\ b < 0,$ то представление (1) называем отрицательно-определенным.

Отметим следующие свойства отрицательно-определенного линейного представления (1). Предполагаем, что $U \perp V$.

1. Для любого положительного числа x и любого ранга n существует отрицательно-определенное представление (1).

2. Для произвольного числа x и произвольного фиксированного ранга n количество отрицательно-определенных представлений (1) бесконечно.

Выбирая многообразие последовательностей U и V, получаем многообразие форм линейного представления чисел. Для разных прикладных задач возможен выбор специальных соответствующих последовательностей U и V [2–6].

2. Линейные рекуррентные последовательности и линейные формы.

Пусть $A = a_3, a_4, \ldots$ — последовательность положительных натуральных чисел. Последовательность A определяет линейную рекуррентную последовательность P следующего вида:

$$\mathrm{P}: \quad P_1 = P_2 = 1,\ P_n = a_n P_{n-1} + P_{n-2},\ n > 2. \tag{2}$$

Рассматриваем представление натуральных чисел в виде положительно-определенных линейных форм следующего вида:

$$x = aP_n + bP_{n+1}, \tag{3}$$

где $a > 0,\ b \geq 0.$

Последовательность P называем базисом представления. Для представления (3) выполняются следующие свойства:

1. Для всех индексов n числа P_n и P_{n+1} взаимно просты.

2. Для любого положительного натурального числа x существует максимальное лево-каноническое представление (3). В таком представлении выполняются неравенства $0 < a < P_{n+1}$.

3. Если для числа x существует положительно-определенное представление (3) ранга n, $n > 2$, то тогда для x существует положительно - определенные представления меньших рангов $n - k,\ 0 < k \leq n - 2$.

4. Если n ранг максимального представления (3), то выполняется неравенство $x < P_{n+1}P_{n+2}$.

Менее очевидны следующие два свойства максимальных линейных форм, которые сформулируем в виде теорем.

Теорема 1. Положительно-определенное представление (3) максимальное лево-каноническое тогда и только тогда, когда выполняются условия:

$$\frac{b}{a_{n+2}} < a < P_{n+1}. \tag{4}$$

Следствие 1. Предположим, для натуральных чисел a и b, $a > 0$ выполняется неравенство $a + b < P_{n+1}$. Тогда для любого номера m, $m \geq n$, линейная форма:

$$y = (a+b)P_m + bP_{m+1} \tag{5}$$

максимальная лево-каноническая.

Доказательство. Выполняются неравенства:

$$\frac{b}{a_{m+2}} < a + b < P_{n+1} < P_{m+1}.$$

Следовательно, согласно теореме 1, линейная форма (5) — максимальная лево-каноническая.

Отметим, что декомпозиция числа x в максимальное лево-каноническое представление (3) требует выполнения $0(\log x)$ арифметических операций.

3. Кодирование пар натуральных чисел одним числом

Хорошо известна Канторовская нумерация пар натуральных чисел. Пара (a, b) взаимно-однозначно кодируются числом:

$$(a,b) \leftrightarrow \frac{(a+b)(a+b+1)}{2} + b.$$

Для целей настоящей работы мы предлагаем альтернативную кодировку. Предлагаемая кодировка не является взаимно-однозначным соответствием. Для каждой пары (a, b) множество кодовых номеров бесконечно. Пара (a, b) однозначно восстанавливается по любому ее кодовому номеру m за $O(\log m)$ арифметических операций.

Пусть задана рекуррентная последовательность (2). Паре $(a,b),\ a > b,\ b \geq 0$ сопоставляем число c:

$$(a,b) \to c = (a+b)P_n + bP_{n+1}, \tag{6}$$

где n — любой номер такой, что выполняется неравенство $(a+b) < P_{n+1}$.

Согласно следствию 1, представление (6) задает максимальную лево-каноническую форму. Поэтому пара (a, b) восстанавливается очевидным способом путем декомпозиции числа c в максимальную лево-каноническую форму и восстанавливаем числа a при помощи вычитания второго коэффициента из первого коэффициента соответствующей линейной формы.

Получаем, пара (a, b) однозначно кодируется бесконечным количеством чисел вида $(a+b)P_m + bP_{m+1}, \quad m \geq n, \quad a+b < P_{n+1}$.

Если выбирать для кода пары (a, b) только наименьшее n такое, что $P_n \leq a+b < P_{n+1}$, то такое кодирование будет инъекцией в множество натуральных чисел.

Через $l(x)$ обозначим битовую длину числа x.

Отметим, что (6) определяет неравенства:

$$P_n^{\,2} < c < 2P_{n+1}^2.$$

Это означает, что битовая длина $l(c)$ всегда минимум в два раза превышает битовую длину числа $l(\max(a,b))$.

Если возникает необходимость кодировать одним числом пары (a, b), где допустимо равенство $a = 0$, то (6) необходимо заменить на соответствие $(a,b) \to c = (a+b+1)P_n + bP_{n+1}, \quad a+b+1 < P_{n+1}$.

Обозначим через $C_A(a,b)$ результат кодирования пары (a, b) при помощи базисной рекуррентной последовательности P (2), определяемой последовательностью A. $C_A(a,b)$ — одно из кодовых чисел вида $(a+b)P_m + bP_{m+1}, a+b < P_{m+1}$.

Выбор параметра m недетерминирован.

Отметим, что для пар натуральных чисел кодирование линейными формами задает бесконечное множество кодирующих функций, определяемых выбором последовательности P.

Кодирование кортежей натуральных чисел можно свести к последовательному кодированию пар чисел. При таком кодировании происходит экспоненциальный рост длины кода, т.к. кодирование по формуле (6) при выборе минимального номера m удваивает длину кода по сравнению с длиной числа max (a, b). Для целей сокращения длины кода кортежей типа $(m_1, m_2, ..., m_k), m > 0, i = 1...k$ можно воспользоваться следующим приемом.

Предположим, все числа записываются в выбранной стандартной системе счисления по основанию M. Для удобства можно считать, что выбрана двоичная система записи чисел.

Пусть S — строка в алфавите из M цифр, a — число. Запись $a \parallel s$ обозначает число, получаемое из a путем дописывания к a справа строки S.

Предположим задано число c:

$c = (c_m c_{m-1} ... c_0)_M, 0 \leq c_i < M, i = 0, ..., m-1. \quad 0 < c_m < M$.

Положим $\tilde{c} = 0c_{m-1} ... c_0$.

Тройке чисел $(a,b,c), a > 0, b \geq 0, c > 0, \quad y = AP_n + BP_{n+1}$ сопоставляем число, где $A = (a \parallel \tilde{c}) + (b \parallel c,) \quad B = b \parallel c, \quad A < P_{n+1}$.

Разложив y в максимальную линейную форму в базисе P, мы определим числа $a_1 = a \parallel \tilde{c}$ и $b_1 = b \parallel c$. Просматривая цифровую запись a_1 и b_1 справа налево, определим первое несовпадение соответствующих цифр. В b_1 это

будет позиция, занимаемая цифрой c_m, а в a_1 соответствующую позицию занимает 0.

Таким образом однозначно восстанавливается число c. Затем, очевидным образом находим a и b.

В общем случае для кода кортежа $(m_1, m_2, ..., m_k), m_1 > 0, m_2 \geq 0, m_i > 0, i = 3...k$, к числу a_1 дописываем справа $\tilde{m}_3 \| \tilde{m}_4 \| ... \| \tilde{m}_k$, а к числу m_2 дописываем справа $a_3 \| a_4 \| ... \| a_k$.

Получаем числа $A = a_1 \| \tilde{a}_3 \| ... \| \tilde{a}_k$ и $B = a_2 \| \tilde{a}_3 \| ... \| \tilde{a}_k$.

Кортеж $(a_1, a_2, ..., a_k)$ кодируется числом:

$$y = (A + B)P_n + BP_{n+1}, \tag{7}$$

где $A + B < P_{n+1}$.

Очевидно, аналогично выше рассмотренному случаю кодирования троек чисел по цифровой записи A и B, однозначно восстанавливаются числа $a_1, a_2, ..., a_k$.

Получаем, что при кодировании кортежей максимальными линейными формами по формуле (7) в случае ограниченности чисел a_i из (2) длина кода для кортежа $(m_1, m_2, ..., m_k)$ будет порядка $2(l(m_1) + l(m_2) ... + l(m_k))$.

4. Недетерминированная криптография

Предлагаем следующую схему кодирования положительных натуральных чисел.

Пусть $A = a_0, a_1, ...$ бесконечная последовательность целых положительных чисел.

Последовательность A задает иррациональное число α, представленное бесконечной цепной дробью $\alpha = [a_0; a_1, ...]$.

Рассматриваем множество подходящих дробей $\alpha_n = [a_0; a_1, ..., a_n] = \frac{P_n}{Q_n}$,

где P_n числитель, а Q_n — знаменатель подходящей дроби n-го порядка.

Положим $U = Q_0, Q_1, ...$ и $V = P_0, P_1, ...$. Очевидно $U \perp V$. Рассматриваем линейные представления чисел (1) в таком базисе (U, V).

Кодирование числа x происходит в два этапа. Сначала число x кодируется соответствующей тройкой (a, b, n), где a и b коэффициенты отрицательно-определенной линейной формы $x = aQ_n - bP_n$ ранга n, $a > 0, b < 0$. На втором этапе тройка (a, b, n) кодируется по формуле (7) задающей код троек чисел.

Кодирующий алгоритм $E_{A,B}(x)$.

Вход: натуральное число x, последовательности натуральных положительных чисел:

$A = a_0, a_1, ...;$ $B = b_3, b_4, ...;$

Результат: $y = код\,(x)$.

Начало

Этап 1

1. Недетерминировано случайным (псевдослучайным) образом выбрать ранг n.

2. Число x недетерминированным образом разложить в отрицательно-определенную линейную форму ранга n, $x = aQ_n - bP_n$.

3. $z \leftarrow (a,b,n)$.

Этап 2

4. Сформировать линейную рекуррентную последовательность $\mathcal{B}$.

$$\mathcal{B}: \qquad B_1 = B_2 = 1,\ B_{i+1} = b_{i+1}B_i + B_{i-1},\ i > 2.$$

5. По $z = (a,b,n)$ вычислить:

$$y = C_B(a \,||\, \tilde{n} + b \,||\, n,\ b \,||\, n) = ((a \,||\, \tilde{n}) + (b \,||\, n))B_m + (b \,||\, n)B_{m+1},$$

где $(a \,||\, \tilde{n}) + (b \,||\, n) < B_{m+1}$.

Результат: $\leftarrow y$.

Конец.

Недетерминированность шага 2 заключается в случайном выборе числа k, $a = a_0 + kP_{n+1}, b = b_0 + kP_n$, где $x = a_0Q_n - b_0P_n$ — лево-каноническое отрицательно-определенное представление числа x. Недетерминированность шага 5 заключается в недетерминированном выборе ранга m.

Обозначим через $C_{A,B}(x)$ заключительный результат вышеописанного кодирования числа x, $y = C_{A,B}(x)$.

Декодирующий алгоритм $D_{A,B}(y)$

Начало

Вход: $y = C_{A,B}(x)$; последовательности A, B.

Результат $= x$: $E_{A,B}(x) = y$.

Декодирование происходит детерминировано в обратном порядке.

1. В базисе $\mathcal{B}$ разложить число y в максимальную лево-каноническую линейную форму $y = AB_m + BB_{m+1}$;

2. Вычислить $b||n = B,\ a||\tilde{n} = A - B$;

3. По $b||n$ и $a||\tilde{n}$ вычислить n, b и a;

4. Вычислить исходное значение $x = aQ_n - bP_n$.

5. *Результат* $\leftarrow x$.

Конец. $D_{A,B}(y)$

Последовательности A и B считаем секретными симметричными ключами.

Вышеописанную схему кодирования обозначим $LRC(A, B)$ (*LRC* — Linear Recurrence Coding).

Отметим, что на каждом шаге алгоритма кодирования недетерминированный выбор соответствующих параметров может быть

реализован случайным образом путем использования параметров любых доступных хаотичных процессов.

Предлагаемая схема кодирования обладает определенной доказуемой криптостойкостью. А именно, наблюдение любого количества выходных значений кодера без знания ключевых параметров не дает никакой информации о входных значениях. Более точно, справедлива следующая теорема.

Теорема 3. Пусть в системе $LRC(A, B)$ задана последовательность кодовых чисел $y_1 = C_{A,B}(x_1),..., y_k = C_{A,B}(x_k)$.

Существует бесконечная последовательность A' такая, что $y_1 = C_{A',B}(x_1'),..., y_k = C_{A',B}(x_n')$, где $x_1',...,x_n'$ некоторые числа, отличные от $x_1,...,x_n$. Множество таких последовательностей A' бесконечно (континиум).

Доказательство. Пусть $\alpha' = [a_0'; a_1', ...]$ любое иррациональное число, которое монотонно меньше α, $\frac{P_n'}{Q_n'} = \beta_n = [a_0'; a_1', ..., a_n']$ подходящая к α рациональная дробь *n*-го порядка P_n' — числитель дроби, Q_n' — знаменатель.

Числа $x_i' = a_i Q_{n_i}' - b_i P_{n_i}', i = 1,..., k$ положительны и их кодирование тройками (a_i, b_i, n_i) в базисе α' совпадает с соответствующими тройками чисел x_i в базисе α'.

Пусть $A' : a_0', a_1', ...$ последовательность, определяемая дробью α'.

Таким образом, в системе кодирования $LRC(A',B)$ положительные числа $x_i' = a_i Q_{n_i}' - b_i P_{n_i}'$ получат такие же значения кодов, что и исходные числа x_i в системе $LRC(A, B)$. Это означает, что, не имея никакой информации о ключевой последовательности *A*, невозможно получить никакой информации об исходном наборе $x_1,...x_k$.

Теорема доказана.

Для ускорения процесса вычислений для выбора ключевой последовательности *A* можно использовать только значения $a_i \in \{1, 2\}$. Даже в этом упрощенном варианте выполняются условия теоремы 3 при условии, что $\alpha = [a_0; a_1, ...]$ отлично от числа $\frac{1+\sqrt{3}}{2} = [1, 2, 1, 2, ...]$.

В системе симметричного кодирования $LRC(A, B)$ последовательности *A* и *B* являются секретными ключами. Их удобно задавать разложением в цепные дроби квадратичних иррациональностей вида $\frac{a+\sqrt{b}}{c}$. Таким образом фактичными ключами являются числа (a, b, c).

Для усиления криптографической стойкости при помощи линейных форм возможно внесение в исходное сообщение дополнительных случайных параметров. Возможны такие варианты.

Пусть n исходное числовое сообщение. Выбираем случайное число $r, l(r) \le l(m)$. Кодируем m числом $y = (m + r)P_n + rP_{n+1}$, где $m + r < P_{n+1}$. Другие возможные комбинации максимальных линейных форм также могут быть использованы.

Например: $(m \oplus r)P_n + rP_{n+1}$, $(m \| r)P_n + rP_{n+1}$, $(m \oplus H(r))P_n + rP_{n+1}$

где $H(r)$ хеш функция.

1. Bellare M., Rogaway P.: Optimal Asymetric Encryption In: De Santis, A. (ed) Advances of Cryptology: Proceedings of EURO-CRYPT '94 , LNCS, 1995. vol. 950. pp. 92–111.
2. Анисимов А. В. Кодирование данных линейными формами числовых последовательностей. *Кибернетика и системный анализ*. 2003. № 1. С. 3–15.
3. Анисимов А. В. Представление чисел в смешанном базисе (2,3). *Кибернетика и системный анализ.* 2009. № 4. С. 3–18.
4. Анисимов А. В. Представление чисел в двухбазисных системах. *Кибернетика и системный анализ.* 2013. №4. С. 1–14.
5. Anisimov A. V. Prefix Encoding by Means of the (2,3) — Representation of Numbers. *IEEE Transactions on Information Theory.* 2013. vol. 59. №4. pp. 2359–2374
6. Анисимов А. В., Завадский И. А. Помехоустойчивое префексное кодирование с помощью нижнего (2,3) - представления чисел. *Кибернетика и системный анализ.* 2014. №2. С.1–15.

UDC 519.72

SYSTEM OF CRYPTOGRAPHIC TRANSFORMATIONS OF NUMBERS BY MEANS OF LINEAR RECURRENT FORMS

A.V. Anisimov

Taras Shevchenko National University of Kiev, Ukraine

Introduction. Two-level system of encoding integers by linear forms $aP_n + bQ_n$, where P_n and Q_n are linear recurrent sequences. These sequences are defined by factoring quadratic irrationalities into continued fractions. Firstly, a number x is represented as a form $x = aA_n + bB_n$, where A_n / B_n is a convergent to some fixed quadratic irrationality. At the second stage the triple (a, b, n) is encoded by a maximal linear form of another linear recurrent sequence $(a, b, n) \to cP_n + dP_{n+1}$. The sequences A_n, B_n, P_n are considered as hidden symmetric keys given by coefficients of corresponding quadratic irrationalities. Properties of such encodings are established.

The purpose of the article is to develop and study a nondeterministic system of cryptographic integer encoding by means of linear recurrent sequences.

Methods. We used methods of continued fractions, properties of linear forms, and bijective encoding of natural numbers.

Results. We proved as a theorem that such a system of encoding is absolutely resistant to passive crypto-attacks. With some further additions it is also resistant to stronger types of attacks.

Conclusion. The proposed system of integer encoding is easy to construct, and it has some proven properties that allows using it as a primitive basic procedure for light weighted cryptography.

Keywords: Linear forms, continued fractions, nondeterministic cryptography.

1. Bellare M., Rogaway P.: Optimal Asymetric Encryption In: De Santis, A. (ed) Advances of Cryptology: Proceedings of EURO-CRYPT '94 , LNCS, 1995. vol. 950, pp. 92–111.
2. Anisimov A. V. Data Coding by Linear Forms of Numerical Sequences. Cybernetics and Systems Analysis. 2003. N 1. P. 3–15.
3. Anisimov A. V. Integer representation in the mixed base (2,3). Cybernetics and Systems Analysis. 2009. № 4. P. 3–18.
4. Anisimov A. V. Two-base numeration systems. Cybernetics and Systems Analysis. 2013. №4. P. 1–14 .
5. Anisimov A. V. Prefix Encoding by Means of the (2,3)-Representation of Numbers. *IEEE Transactions on Information Theory.* 2013. vol. 59. № 4. P. 2359–2374
6 Anisimov A. V. Zavadskyi I. A. Robust Prefix Encoding Using Lower (2,3) Number Representation. Cybernetics and Systems Analysis. 2014. № 2. P.1–15.

Получено 3.10.16