УДК 681.3.04

ОРГАНИЗАЦИЯ ОБМЕНА ДАННЫМИ С ИСПОЛЬЗОВАНИЕМ СИСТЕМЫ ОСТАТОЧНЫХ КЛАССОВ

Магомедов Ш. Г., Мустафаев А. Г.

ФГБОУ ВПО «Дагестанский государственный технический университет», Махачкала, Россия (367015, г. Махачкала, проспект Имама Шамиля, 70), e-mail: arslan_mustafaev@mail.ru

Традиционная технология обеспечения защиты информации ограниченного доступа обычно предполагает использование методов шифрования. Однако требования, предъявляемые к подобным системам, достаточно строгие и, как следствие, громоздкие в реализации и затратные в эксплуатации, что часто делает их малопригодными и затруднительными для использования. В данной статье предложена процедура организации процесса приема-передачи данных ограниченного доступа с использованием системы остаточных классов, как для закрытия данных, так и для выполнения преобразований. Система остаточных классов обладает рядом существенных преимуществ, так как является непозиционной системой счисления, которая позволяет сократить разрядность обрабатываемых данных, представляя их в модулярном коде в системе взаимно простых модулей, и обладает способностью к обнаружению ошибок и самокоррекции. Разработаны алгоритмы, реализующие указанную процедуру.

Ключевые слова: система остаточных классов, алгоритм, прием и передача данных, защита информации, ключи шифрования.

ORGANIZATION OF DATA EXCHANGE USING A SYSTEM OF RESIDUAL CLASSES

Magomedov S. G., Mustafaev A. G.

FGBOU VPO "Dagestan State Technical University", Makhachkala, Russia (367 015, Makhachkala, Avenue Imam Shamil, 70), e-mail: arslan_mustafaev@mail.ru

Traditional technology of protect restricted information generally involves the use of encryption methods. However, the requirements for such systems, strong enough and, therefore, cumbersome to implement and costly to operate, which often makes them less suitable and difficult to use. This paper proposes a procedure for organizing the process of reception and transmission of data restricted to using the system of residual classes, as to close the data and to perform transformations. The system of residual classes has a number of significant advantages, since other number system, which reduces the bit data to be processed, representing them in the modular code system relatively simple modules, and has the ability to detect and correct errors. The algorithms that implement this procedure are developed.

Keywords: system of residual classes, algorithm, data reception and transmission, data protection, encryption keys.

Технология обеспечения защиты информации ограниченного доступа предполагает использование различных методов шифрования. Однако требования, предъявляемые к системам шифрования, достаточно строгие и, как следствие, не тривиальны в реализации и затратные в эксплуатации, что часто делает их малопригодными и затруднительными для использования. В данной работе предлагается процедура закрытия информации, опирающаяся на использование системы остаточных классов (СОК) [1, 2]. Хотя уровень стойкости к злоумышленным действиям предлагаемой процедуры ниже, чем у существующих методов шифрования, процедура имеет важное преимущество: она

достаточно проста в реализации и вполне приемлема для большинства служебных и приватных сообщений, передачи конфиденциальной информации. Безусловно, для передачи особо важных и секретных данных должны быть задействованы стойкие методы и процедуры, в частности, методы шифрования.

Другие достоинства предлагаемой процедуры по сравнению с существующими системами шифрования:

1) отсутствие необходимости обмена ключами (ключом) шифрования;

2) доступность для использования всеми физическими и юридическими субъектами, которые желают обменяться сообщениями именно с данным физическим или юридическим субъектом и знают или могут получить необходимые общедоступные параметры указанного субъекта; например, для юридического субъекта – наименование, юридический адрес, сфера деятельности и другие; для физического субъекта – фамилия, имя, отчество, возраст и другие.

Кратко в сжатой форме напомним основные положения СОК [3], одновременно описывая место каждой характеристики СОК в предлагаемой системе ограничения доступа к передаваемым данным. СОК – это непозиционная система счисления, числа в которой представляются остатками от деления на выбранную систему оснований $P_1, P_2, \ldots, P_n$, где числа в системе взаимно просты. Как будет ясно из дальнейшего пояснения, именно этот набор чисел и является основой системы обеспечения конфиденциальности данных. Диапазон представимых чисел от 0 до $P = P_1 \cdot P_2 \cdot \ldots \cdot P_n$. Передаваемая информация будет представлена в виде последовательности целых положительных чисел A_i. Каждое целое положительное число A представляется в виде набора наименьших положительных остатков (вычетов) от деления числа A на выбранные основания $P_1, P_2, \ldots, P_n$, результат можно записать в виде $A=(a_1, a_2, \ldots, a_n)$ – это и есть запись числа A в СОК. Основные операции, связанные с преобразованием чисел в процессе их передачи и приема, могут быть сведены к операциям сложения и умножения целых положительных чисел. При этом операции сложения, вычитания и умножения над числами в СОК производятся независимо по каждому основанию без переносов между разрядами (основаниями). Если исходные числа A, B, их сумма $A+B$ и их произведение $A \cdot B$ находятся в диапазоне $[0,P)$, то результаты операций сложения $A+B$ и умножения $A \cdot B$ могут быть однозначно представлены соответственно остатками g_i и r_i по тем же основаниям P_i, то есть $A=(a_1, a_2, \ldots, a_n)$, $B=(b_1, b_2, \ldots, b_n)$, $A+B=(g_1, g_2, \ldots, g_n)$, $A \cdot B = (r_1, r_2, \ldots, r_n)$.

После получения данных необходимо их обратно преобразовать из СОК в ППС. Указанная задача рассмотрена в [4]. Большинство известных алгоритмов преобразования из

СОК в ППС основаны на китайской теореме об остатках, в которой приводится следующее соотношение для обратного преобразования [1]: если число $d = (d_1, d_2, ..., d_n)$ в СОК по модулю $P_1, P_2, \ldots, P_n$, то

$$d = \sum_{i=1}^{n} m_i \delta_i d_i \pmod{P},$$

где $P = P_1 \cdot P_2 \cdot \ldots \cdot P_n$, а $\{m_i, i = \overline{1,n}\}$ есть решения уравнений $m_i \delta_i \equiv 1 \pmod{P_i}$.

Данная операция является одной из наиболее трудоемких и обычно выполняется после завершения всех вычислений и преобразований, связанных с приемом-передачей данных. Для повышения быстродействия во многих алгоритмах используются преимущества табличных методов. Характерная особенность известных алгоритмов – хранение констант СОК в памяти, таких как: модули, веса позиционных представлений, базисы и др.

Отметим, что при вычислении базисов СОК $P_1, P_2, \ldots, P_n$ наибольшие временные затраты связаны с операциями нахождения обратных весов в уравнении $m_i \delta_i \equiv 1 \pmod{P_i}$, где m_i – целое положительное число, называемое весом P_i, δ_i – остаток от деления полученной величины на модуль P_i.

Таким образом, основные характеристики определяются свойствами набора простых чисел $P_1, P_2, \ldots, P_n$. Предлагается следующая процедура их выбора.

1. Предварительно формируется база всех простых чисел, не превосходящих заданного числа *N*. Выбор числа *N* определяется особенностями предметной области, в которой решается проблема приема-передачи данных.

Рассмотрим вначале действия субъекта, передающего информацию.

2. Исходные данные о субъектах, передающем и принимающем информацию, записываются последовательно в виде строки; к этим данным могут быть добавлены (по договоренности сторон) конфиденциальные сведения, известные только передающей и принимающей сторонам, а также дописываются дата и интервал времени передачи с точностью до одной минуты. Полученная строка оцифровывается любым способом; например, путем сопоставления каждому знаку его ASCII кода. В результате получаем число *D,* однозначно соответствующее данной паре субъектов – передающего и принимающего информацию.

3. Полученное число *D* разбивается на блоки, каждый из которых как число не превосходит числа *N*. Все блоки для данного числа складываются по модулю *N*; в результате получается число *M*. Если $M \le \left[\frac{N}{2}\right]$, то *M* заменяется на $N - M$, так что всегда выполнено

неравенство $\left[\frac{N}{2}\right] \le M < N$. Очевидно, значение M зависит от атрибутов как передающего субъекта, так и принимающего.

4. Из базы простых чисел выбирается наибольшее простое число P, меньшее числа M. Полагаем $P_0 = M$, $P_1 = P$, вычисляем $R_1 = 1 - 4\frac{P_0 - P_1 - 1}{n_0 \cdot (P_0 - P_1)^2}$ и находим $Q_2 = P_0 \bmod(P_1 \cdot [R_1])$, где [ ] – знак целой части числа. (Коэффициент R_1 введен для того, чтобы, с одной стороны, число n простых чисел в окончательном наборе $P_1, P_2, \ldots, P_n$ не оказалось малым, а с другой стороны, соседние простые числа в наборе достаточно сильно различались. Дополнительные пояснения по выбору R_1 и n_0 приводятся ниже.) Выбираем из базы простых чисел наибольшее простое число P_2, меньшее Q_2. Процедура формирования простых чисел продолжается аналогичным образом по формулам: $Q_{j+1} = P_{j-1} \bmod(P_j \cdot [R_j])$,где $R_j = 1 - 4\frac{P_{j-1} - P_j - 1}{n_0 \cdot (P_{j-1} - P_j)^2}$ и P_{j+1} – наибольшее целое число, меньшее Q_{j+1}. Процедура продолжается до тех пор, пока либо достигнем $n = n_0$, либо при некотором $j = n+1$ получим $P_j = 1$.Так как $R_j < 1$, то очевидно $P_{j+1} < P_j$ для всех j.

5. Набор чисел $P_1, P_2, \ldots, P_n$ и есть искомый. Отметим, в полученном наборе простые числа расположены не в порядке возрастания, как описано выше, а в порядке убывания.

6. Передаваемое сообщение, аналогично пункту 3, записывается в виде текстовой строки, оцифровывается (получаем число C) и разбивается на блоки $C_1, C_2, ..., C_r$, каждый из которых по величине не превосходит значения $P = \prod_{i=1}^{n} P_i$. Затем каждое из чисел на основе одного из разработанных в {добавить ссылку} алгоритмов записывается в СОК. Получаем множество наборов $\Lambda = \{(c_{i1}, c_{i2}, ..., c_{in},), i = \overline{1,r}\}$, где (напомним) $c_{ij} = C_i \bmod P_j$. Сформированный набор Λ с приписанным именем передающей стороны и посылается оппоненту по процессу приема-передачи.

Рассмотрим теперь действия принимающей стороны. Принимающая сторона имеет все необходимые данные о себе и о партнере по связи, ей известен алгоритм формирования набора чисел $P_1, P_2, \ldots, P_n$, а также необходимые конфиденциальные данные, если таковые используются в сообщении. Напомним, что приемная сторона получает файл, содержащий $\Lambda = \{(c_{i1}, c_{i2}, ..., c_{in},), i = \overline{1,r}\}$ и наименование отправителя. Поэтому принимающая сторона выполняет следующие действия.

1. По имени отправителя формируются все данные об отправителе, необходимые для построения набора $P_1, P_2, \ldots, P_n$, который и формируется на основе описанного выше алгоритма – все данные, необходимые для его построения, у принимающей стороны имеются.

2. На основе обратного преобразования из СОК в ПСС восстанавливаются блоки C_i, $i = \overline{1,r}$ и формируется число $C = C_r C_{r-1} \ldots C_1$.

3. Полученное число C переводится в символьную форму на основе преобразования, обратного использованному в пункте 2 передающей стороны в процессе оцифровывания текста. Например, если использовались ASCII-коды для оцифровывания, то для восстановления текста число C записывается в двоичной (восьмеричной или шестнадцатиричной) форме, и каждый блок из 8 битов заменяется на соответствующий ASCII-символ. При использовании ЮНИКОДа версии UTF-16 двоичные блоки состоят из 16 битов. Полученный набор символов и есть исходный текст.

Отметим, что дата и время передачи добавлены в пункте 3 алгоритма для защиты от повторной передачи того же сообщении в другое время и в другой день. Интервал времени в одну минуту выделен в связи с тем, что в процессе подготовки передаваемого сообщения и его передачи время могло немного измениться.

Опишем возможный способ разбиения числа Cна блоки. $V_{i+1} = \left[\frac{V_i}{P_i}\right]$, $V_0 = C$, $C_i = V_{i-1} - V_i \cdot P_i$ – остаток от деления V_{i-1} на P_{i-1}. Аналогично разбивается на блоки число D.

Рассмотрим возможный способ выбора числа N. Основное требование: N должно быть достаточно большим, чтобы количество наборов $P_1, P_2, \ldots, P_n$ должно быть больше числа возможных субъектов, которые могут участвовать в процессе приема-передачи сообщений. Желательно также, чтобы числа $P_1, P_2, \ldots, P_n$ были в среднем достаточно большими (с точки зрения требований безопасности). По соображениям удобства компьютерной реализации целесообразно также, чтобы число N представлялось степенью двойки, то есть $N = 2^m$. Число возможных субъектов S не превзойдет в ближайшем будущем 50 миллиардов – пятикратное количество числа людей. Пусть $\pi(N)$ – число простых чисел, не превосходящих числа N. Тогда число наборов $P_1, P_2, \ldots, P_n$ пропорционально количеству подмножеств во множестве первых $\pi(N)$ простых чисел, то есть пропорциональна числу $2^{\pi(N)}$. По теореме Эйлера при больших N справедливо соотношение $\pi(N) \cong \frac{N}{\ln N}$. Таким образом, получаем следующую эвристическую оценку: $C \cdot 2^{\pi(N)} > S$, или $\frac{N}{\ln N} > \ln(S/C)$, где C – коэффициент

пропорциональности, отражающий желаемый уровень запаса всех возможных наборов простых чисел. Полагаем C=0.1. Наименьшее целое решение этого неравенства, представимое в виде $N = 2^m$, равно 256.Следовательно, требование наличия достаточного запаса наборов простых чисел выполняется для любого $N \geq 256 = 2^8$. По соображениям безопасности и удобства компьютерной обработки (размер двух байтов) предлагается взять $N = 65536 = 2^{16}$.

Поясним выбор выражений для коэффициентов R_j. За n итераций интервал выбора очередного простого числа уменьшится до величины $p < P_0 \cdot \prod_{j=1}^{n} R_j \leq P_0 \cdot (\overline{R})^n$, где $\overline{R} = \max\{R_j, j = \overline{1,n}\}$ и $\overline{R}$ можно представить в виде $\overline{R} = 1 - \frac{\Delta P - 1}{n_0 (\Delta P)^2}$, ΔP – разность между двумя соседними числами. Так как $p \geq 1$, то получаем следующее неравенство для возможных значений *n*:

$$P_0 \cdot \overline{R}^n > 1, \text{ или } n < \frac{-\ln(P_0)}{\ln \overline{R}} \approx -\ln(P_0)\left(\ln\left(1 - \frac{\Delta P - 1}{n_0 (\Delta P)^2}\right)\right)^{-1} \approx -\ln(P_0)\left(-\frac{\Delta P - 1}{n_0 (\Delta P)^2}\right)^{-1}.$$

Получаем: $n < \ln(P_0)\frac{n_0 (\Delta P)^2}{(\Delta P - 1)}$. Так как $\Delta P \geq 2$ и $\ln P_0 \geq \ln \frac{N}{2} \approx 10{,}4$, то ввиду возрастания по ΔP функции $\frac{(\Delta P)^2}{(\Delta P - 1)}$ получаем следующую оценку, справедливую для всех возможных значений ΔP: $n < 10{,}4 \cdot 2^2 \cdot n_0 = 41{,}6 \cdot n_0$.

Достаточно большая база данных по простым числам приведена в [5]. Эта база может использоваться при формировании оснований модулей в СОК.

Наконец, по поводу выбора n_0. Поскольку n_0 определяет число возможных параллельных преобразований, оптимальное значение n_0 должно соответствовать число ядер современного стандартного процессора. Исходя из этого, можно взять $n_0 = 8$.

Оценим стойкость предлагаемого метода закрытия передаваемых данных ограниченного доступа. Для этого произведем оценку числа вариантов, которые необходимо перебрать злоумышленнику, для нахождения истинных значений модулей $P_1, P_2, \ldots, P_n$ путем перебора.

Будем анализировать возможности злоумышленника для наиболее плохой ситуации, когда злоумышленник имеет возможности доступа к внутренним ресурсам компьютера. Именно, пусть злоумышленник может выполнить на компьютере вычисление и, имея доступ к внутренним ресурсам компьютера, может проконтролировать промежуточные значения в памяти компьютера, то есть ему известны остатки $r_1, r_2, ..., r_k$ от деления числа *R*на модули

$P_1, P_2, ..., P_k$ соответственно. Оценим среднее число наборов, которое должен перебрать злоумышленник для того, чтобы найти истинные значения $P_1, P_2, ..., P_k$, зная R и $r_1, r_2, ..., r_k$, при условии, что все сочетания равновероятны. Будем предполагать также, что длины всех модулей приблизительно одинаковы и лежат в пределах от $l-d$ до l.

Общее количество возможных значений каждого модуля равно $N = 2^l - 2^{l-d}$. Отсюда следует, вероятность обнаружения истинного набора значений $P_1, P_2, ..., P_k$ при одной попытке равна $p = \left(\frac{1}{N}\right)^k = 2^{-k(l-d)} \cdot (2^d - 1)^{-k}$, а среднее число испытаний при полном переборе равно: $\mu = \sum_{i=1}^{M} i(1-p)^{i-1} p$, где $M = N^k$ – общее число всех наборов. Здесь $(1-p)^{i-1} p$ есть вероятность того, что (i-1) попытки были неудачными, а на i-ой попытке злоумышленнику удалось найти истинный набор $P_1, P_2, ..., P_k$. Из формулы видно, что μ является средним значением для геометрического распределения, и по известной формуле получаем $\mu = \frac{1}{p} = (2^i - 2^{i-d})^k$. Отметим, что числа k и$l$связаны соотношением $k \cdot l \approx n$, где n– разрядность чисел, которые преобразуются на основе СОК; обычно это значение соответствует разрядности процессора.

Например, для 64-разрядного процессора можно взять k=8, l=8 и d=2. Тогда имеем: $\mu = (2^8 - 2^6)^8 = (256-64)^8 = 1846757322198614016 \approx 1846757322 \cdot 10^9$, то есть процессору, имеющему тактовую частоту 5 ГГц, потребуется время равное $t \approx 369351464$ сек только на перебор вариантов, что составляет более 11,7 лет непрерывной работы процессора. Последнее означает, что злоумышленник практически не имеет никаких реальных возможностей обнаружить истинный набор оснований $P_1, P_2, ..., P_k$.

Подводя итог, можно сказать, что предлагаемая процедура позволяет индивидуализировать передаваемые сообщения применительно к каждому субъекту, что обеспечивает достаточный уровень безопасности. Процедура не требует наличия специальных ключей шифрования, а опирается лишь на индивидуальные регистрационные атрибуты субъекта, что устраняет проблему доставки ключей и тем самым существенно упрощает в целом процесс организации обмена информацией.

Список литературы

1. Акушский Н. Я, Юдицкий Д. И. Машинная арифметика в остаточных классах. М.: Сов. радио, 1968. 439 с.

2. Shanks, D. Solved and Unsolved Problems in Number Theory, 4th ed. NewYork: Chelsea, 1993.

3. Гашков С. Б. Системы счисления и их применение. М.: МЦНМО, 2004. 52 с.

4. Исмаилов Ш-М.А., Магомедов Ш. Г. Алгоритмы и структуры преобразования числовых данных из позиционной системы счисления в систему остаточных классов // Научно-технические ведомости СПбГТУ. Информатика. Телекоммуникации. Управление. 2008. № 5(65). С. 159-169.

5. Виноградов И. М. Основы теории чисел. М.: Наука, 1972. 456 с.

Рецензенты:

Рехвиашвили С. Ш., д.ф-м.н., с.н.с., НИИ прикладной математики и автоматизации КБНЦ РАН, КБР, г. Нальчик.

Росс Г. В., д.т.н., профессор, зам. директора ФГУП ВНИИ ПВТИ, г. Москва.