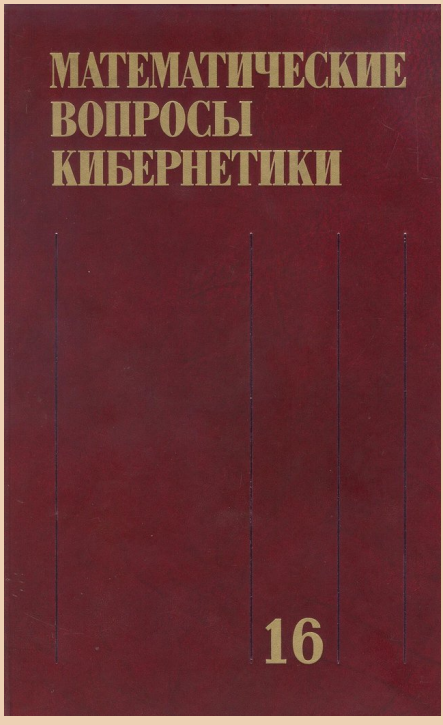
ИПМ им. М.В. Келдыша РАН • Электронная библиотека
Математические вопросы кибернетики • Выпуск 16



М. П. Минеев,
В. Н. Чубариков

Задача об искажении частоты появления знаков в шифре простой замены

Рекомендуемая форма библиографической ссылки:
Минеев М. П., Чубариков В. Н. Задача об искажении частоты появления знаков в шифре простой замены // Математические вопросы кибернетики. Вып. 16. — М.: Физматлит, 2007. — С. 242–245. URL: http://library.keldysh.ru/mvk.asp?id=2007-242

ЗАДАЧА ОБ ИСКАЖЕНИИ ЧАСТОТЫ ПОЯВЛЕНИЯ ЗНАКОВ В ШИФРЕ ПРОСТОЙ ЗАМЕНЫ

М. П. МИНЕЕВ, В. Н. ЧУБАРИКОВ

(МОСКВА)

§ 1. Введение

Классическим примером шифрограмм являются шифрованные сообщения, построенные с помощью способа простой замены, в частности, шифр Гая Юлия Цезаря (см., например, [1—9]).

Этот тип шифрования исходного текста получается использованием подстановки на множестве букв алфавита, причем различным символам шифрованного текста соответствуют различные буквы. В исходном тексте различные буквы встречаются с разной частотой. Существуют таблицы относительных частот встречаемости букв рассматриваемого алфавита (см., например, [8]). Поэтому, если текст сообщения достаточно длинный, то, подсчитывая частоту появления символов в криптограмме, можно с большой вероятностью восстановить буквы зашифрованного текста. Этот способ восстановления исходного текста называется частотным анализом. Отметим, что те части текста, где возникает неоднозначность, как правило, восстанавливаются по их смыслу.

Таким образом шифр простой замены определяется подстановкой σ всех букв или символов конечного алфавита A, т. е. при шифровании буквы исходного текста a, $a \in A$, заменяются на буквы шифрованного текста $\sigma(a)$. Эта подстановка σ задает ключ криптосистемы. Следовательно, количество всех ключей для шифров простой замены с алфавитом из n букв равно $n!$. Отметим, что исходный текст шифруется посимвольно и каждая буква шифрованного текста зависит только от соответствующей компоненты подстановки и от соответствующей буквы алфавита A.

Теоретически для посимвольного шифрования исходного текста с помощью замены можно использовать «бесконечный ключ» [2], т. е. последовательную замену каждой буквы исходного текста с помощью взаимно однозначной замены на букву шифрованного текста. Тогда исходный текст невозможно восстановить без знания ключа. Эффективную систему шифрования, которую можно рассматривать с позиции усовершенствования «бесконечного ключа», изобрел Вижинер. В основе системы шифрования Вижинера лежит «бесконечный ключ», отвечающий бесконечной периодической последовательности. Важным свойством этой системы шифрования является то, что частотные характеристики встречаемости букв в шифрованном тексте отличаются от среднестатистических частот встречаемости букв данного алфавита A, т. е. понятие подобной частоты как бы сглаживается.

Настоящая статья посвящена подобному эффекту искажения частот появления символов для шифра простой замены на основе «сжатия алфавита» за счет использования квадратичных вычетов и квадратичных невычетов в конечных полях.

§ 2. Искажение частот на примере русского алфавита

В качестве модельной ситуации рассмотрим русский алфавит, состоящий из 31 буквы (отождествляются буквы е, ё и ь, ъ). Известна таблица относительных частот встречаемости букв этого алфавита, упорядоченная в порядке убывания частот, в тексте на русском языке (см., например, [8]). Эта таблица помогает расшифровать криптограммы, полученные простой заменой. Расположим буквы в порядке убывания частот:

1) о — 0,090, **2)** е, ё — 0,072, **3)** а — 0,062, **4)** и — 0,062, **5)** н — 0,053, **6)** т — 0,053, **7)** с — 0,045, ..., **31)** ф — 0,002.

Поскольку число 31 — простое, все вычеты по модулю 31 можно разбить на три класса: квадратичные вычеты, квадратичные невычеты и вычет, отвечающий нулю. Как известно, количество квадратичных невычетов и количество квадратичных вычетов в полной системе вычетов по простому модулю одинаково, и в данном случае оно равно 15. Все квадратичные вычеты по модулю 31 исчерпываются следующими классами вычетов: $1, 2^2, 3^2, \ldots, 15^2$. Занумеруем все квадратичные вычеты по модулю 31 по убыванию их величины, а затем занумеруем квадратичные невычеты в порядке убывания

Если a — квадратичный вычет по модулю 31, то его решения $a_1 = b$ и $a_2 = 31 - b$ представляют собой два различных вычета по модулю 31.

Рассмотрим теперь некоторый открытый текст s и зашифруем его с помощью метода простой замены. Расположим буквы шифрованного текста в порядке убывания частот, нумеруя их от 1 до 31.

Каждой из первых пятнадцати занумерованных букв взаимно однозначно сопоставим квадратичные вычеты по модулю 31 в соответствии с их порядком нумерации, затем следующие пятнадцать букв взаимно однозначно отобразим в квадратичные невычеты по модулю 31 также в соответствии с их порядком нумерации и, наконец, оставшейся букве сопоставим нулевой вычет по модулю 31.

Далее продолжим шифровку следующим образом. Если буква α закодирована числом a и a — квадратичный вычет по модулю 31, причем вычеты a_1, a_2 решения сравнения $x^2 \equiv a \pmod{31}$, то последовательность чисел a в криптограмме заменяем на последовательность чисел $a_1, a_2, a_1, a_2, \ldots$. Например, если в криптограмме имеется 5 мест, на которых стоит число a, то заменяем в этих местах число a на следующую последовательность чисел a_1, a_2, a_1, a_2, a_1. Если же буква α закодирована числом a и a — квадратичный невычет по модулю 31 или 0, то в криптограмме это число оставляем без изменения.

Таким образом статистическая картина криптограммы изменена. Для восстановления первоначальной криптограммы надо все числа, отвечающие квадратичным вычетам по модулю 31, возвести в квадрат по модулю 31.

Остается передать получателю текста номера тех мест, на которых стоят квадратичные невычеты по модулю 31. Осуществим это следующим образом.

Последовательно обозначим места, на которых стоят квадратичные вычеты по модулю 31, — единицами, а места, на которых стоят квадратичные невычеты, — нулями. Полученную последовательность чисел $\varepsilon_1, \ldots, \varepsilon_n$, составленную из нулей и единиц, можно рассматривать как запись некоторого

числа m в двоичной системе счисления. Таким образом отправителю достаточно передать построенное число m, для того чтобы получатель знал места, на которых находятся квадратичные вычеты и на которых — квадратичные невычеты по модулю 31.

Итак, абонент $\mathscr{A}$ должен послать «секретное» число m абоненту $\mathscr{B}$ по каналу связи.

§ 3. Применение криптосистемы для передачи секретной информации

Для этого можно воспользоваться, например, известным алгоритмом А. Шамира для передачи секретной информации по каналу связи (см., например, [8, 9]). Приведем здесь этот алгоритм.

Абоненты $\mathscr{A}$ и $\mathscr{B}$ выбирают достаточно большое простое число p, $p>m$. Затем абонент $\mathscr{A}$ выбирает секретный ключ a, $1<a<p-1$, $(a, p-1)=1$, а абонент $\mathscr{B}$ — секретный ключ b, $1<b<p-1$, $(b, p-1)=1$. Для проверки условия взаимной простоты a и $p-1$ абонент $\mathscr{A}$ может использовать алгоритм Евклида. В случае, если числа a и $p-1$ не взаимно просты, то следует проверить на взаимную простоту числа $a+1$ и $p-1$ и т. д. Указанный процесс выбора ключа a оборвется через конечное число шагов, так как, например, $(p-2, p-1)=1$. Аналогичным образом может поступить и абонент $\mathscr{B}$. Далее абонент $\mathscr{A}$ находит натуральное число α такое, что

$$a\alpha \equiv 1 \pmod{p-1},\ 1<\alpha<p-1.$$

Аналогично поступает абонент $\mathscr{B}$. Он находит число β с условиями

$$b\beta \equiv 1 \pmod{p-1},\ 1<\beta<p-1.$$

Итак, абонент $\mathscr{A}$ имеет секретный ключ (a, α), а абонент $\mathscr{B}$ — секретный ключ (b, β).

Теперь абонент $\mathscr{A}$ пересылает число m абоненту $\mathscr{B}$ по открытому каналу за следующие четыре шага.

1-й шаг. Абонент $\mathscr{A}$ посылает абоненту $\mathscr{B}$ число

$$m_1 \equiv m^a \pmod{p-1},\ 0<m_1<p-1.$$

2-й шаг. Абонент $\mathscr{B}$ посылает абоненту $\mathscr{A}$ число

$$m_2 \equiv m_1^b \pmod{p-1},\ 0<m_2<p-1.$$

3-й шаг. Абонент $\mathscr{A}$ посылает абоненту $\mathscr{B}$ число

$$m_3 \equiv m_2^\alpha \pmod{p-1},\ 0<m_3<p-1.$$

4-й шаг. Абонент $\mathscr{B}$ находит число m с помощью секретного ключа β следующим образом:

$$m = m_4 \equiv m_3^\beta \pmod{p-1},\ 0<m_4<p-1.$$

Действительно, имеем

$$m_4 \equiv m^{ab\alpha\beta} = m^{a\alpha\cdot b\beta} \equiv m \pmod{p-1},$$

т. е. получаем $m_4 \equiv m \pmod{p-1}$, $0<m, m_4<p-1$.

Следовательно, $m=m_4$.

§ 4. Итерационная процедура для сжатия алфавита

Пусть n — натуральное число и q — простое число вида $q=2^n+1$. Тогда q является простым числом Ферма $F_m=2^{2^m}+1$, $m\geqslant 0$. На сегодняшний день известно пять простых чисел Ферма $F_0=3$, $F_1=5$, $F_2=17$, $F_3=2^8+1=257$ и $F_4=2^{16}+1=65537$. Мультипликативная группа поля F_q является циклической и состоит из $q-1=2^n$ элементов. Каждый из них имеет порядок 2^k, $0\leqslant k\leqslant 2^n$.

Пусть теперь алфавит A состоит из $q=2^n+1$ букв и символов. Тогда, используя процедуры, описанные в §2 и §3, в точности n раз, приходим к шифрованному тексту, алфавит которого отвечает только квадратичным невычетам и нулевому вычету по модулю q. Таким образом алфавит шифрованного текста будет состоять из $(q+1)/2$ буквы и символа.

СПИСОК ЛИТЕРАТУРЫ

1. Аршинов М. Н., Садовский Л. Е. Коды и математика. — М.: Наука, 1983.
2. Баричев С. Криптография без секретов. — http://www.artelecom.ru/library/books/swos/index/html.
3. Брассар Ж. Современная криптология. — М.: ПОЛИМЕД, 1999.
4. Гашков С. Б., Чубариков В. Н. Арифметика. Алгоритмы. Сложность вычислений. — М.: Наука, 1996..
5. Жельников В. Криптография от папируса до компьютора. — М.: ABF, 1996.
6. Коблиц Н. Курс теории чисел и криптографии. — М.: Научное изд-во ТВП, 2001.
7. Молдовян А. А., Молдовян Н. А., Советов Б. Я. Криптография. — СПб.: Изд-во «Лань», 2001.
8. Нечаев В. И. Элементы криптографии (Основы теории защиты информации): Учеб. пособие для ун-тов и пед. вузов. — М.: Высш.шк., 1999.
9. Чубариков В. Н. Элементы арифметики. — М.: Изд-во Механико-математического ф-та МГУ, 2007.

Поступило в редакцию 12 XI 2007