В. Ф. Очков,
Национальный исследовательский университет МЭИ, Москва

MATHCAD И КРИПТОГРАФИЯ

Аннотация

В статье рассмотрена программа, позволяющая составлять частотные характеристики текстов и помогающая расшифровывать тексты.

Ключевые слова: расшифровка текста, Mathcad.

Контактная информация

Очков Валерий Федорович, доктор тех. наук, профессор, Национальный исследовательский университет МЭИ, Москва; *адрес:* 111250, г. Москва, Красноказарменная ул., д. 14; *телефон:* (495) 362-71-71; *e-mail:* ochkov@twt.mpei.ac.ru

V. F. Ochkov,
National Research University MPEI, Moscow

MATHCAD AND CRYPTOGRAPHY

Abstract

A program that allows you to make the frequency characteristics of texts and helps to decrypt texts is considered in the article.

Keywords: cryptography, Mathcad.

Информатика — это наука о работе с информацией. В частности, о ее защите от посторонних взглядов. Защищать письменные тексты помогает раздел информатики под названием ***криптография*** или ***тайнопись***.

Расшифровка одного такого текста красочно описана в рассказе Эдгара По «Золотой жук». Этого американского писателя по праву называют родоначальником детективного жанра в литературе.

Сюжет рассказа незамысловат. Один человек находит на берегу моря пергамент, приносит его домой и случайно оставляет у огня. На пергаменте от нагрева проступают таинственные символы:

53‡‡†305))6*;4826)4‡.)4‡);806*;48†8¶60))85;1‡(;:‡*8†83(88)5*†;46(;88*96*?;8)*‡(;485);5*†2:*‡(;4956*2(5*—4)8¶8*;4069285);)6†8)4‡‡;1(‡9;48081;8:8‡1;48†85;4)485†528806*81(‡9;48;(88;4(‡?34;48)4‡;161;:188;‡?;

Подобные цепочки «таинственных» символов мы иногда можем видеть на дисплее компьютера, когда выбрана неверная кодировка текста.

Человек, нашедший пергамент, расшифровывает текст и находит по нему запрятанный пиратами клад. Шифр был основан на замене букв исходного текста другими буквами или символами. Это один из самых древних и самых простых способов шифрования текстов.

К настоящему времени разработано огромное количество систем шифрования различной степени сложности. Новый импульс этому процессу дало развитие компьютерных информационных технологий. Пересылая информацию даже в пределах компьютера, мы ее «шифруем» — переводим привычные буквы и цифры в нули и единицы: в цепочки «таинственных символов». Компьютеры, пересылая информацию друг другу по локальным сетям или по Интернету, могут использовать специальные алгоритмы шифрования для того, чтобы передаваемую информацию не мог прочесть тот, кому это не положено.

Расшифровка текста, основанного на замене символов, состоит из двух этапов — рутинного и творческого. Рутинный этап — это подсчет частоты использования тех или иных символов в исходном тексте. Герой рассказа Эдгара По сделал это вручную. Мы же можем подсчитать это на компьютере, написав несложную программу на языке Mathcad (рис. 1).

В переменную Текст (рис. 1) вставлена (скопирована из Интернета) цепочка таинственных символов из рассказа Эдгара По. Показано только начало этой цепочки, в которой 203 символа. Подсчет числа символов в переменной, хранящей строку, ведет встроенная в Mathcad функция strlen (str — string, строка; len — length, длина). Нам понадобится еще одна строковая функция Mathcad — функция substr (подстрока), возвращающая часть строки, начиная с нужного места (второй аргумент функции substr) и нужной длины (третий аргумент).

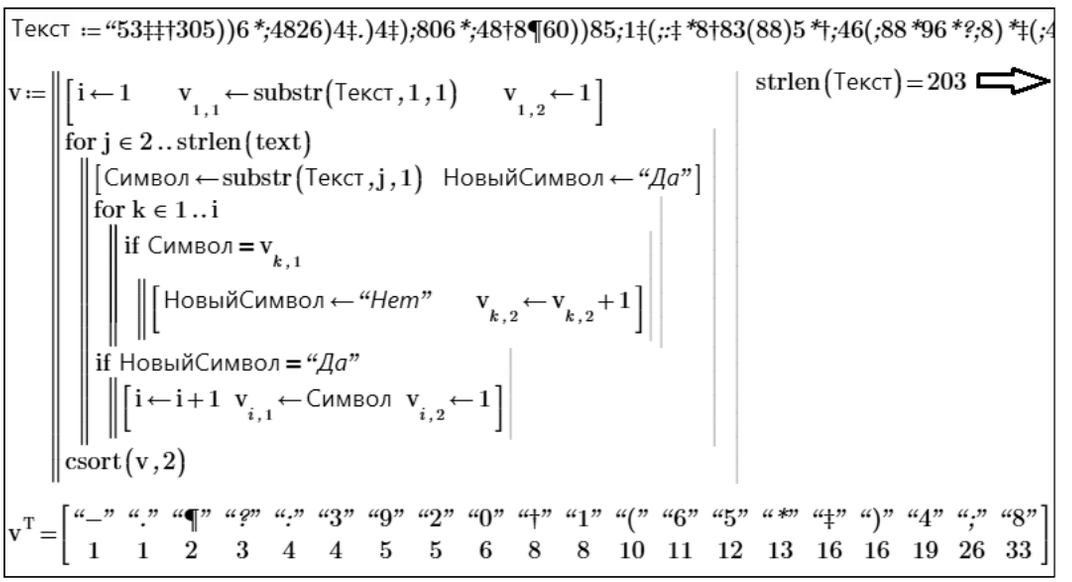

Рис. 1. Программа подсчета частоты встречи символов в тексте

В программе, показанной на рисунке 1, в матрице v будут формироваться два столбца — сам символ и частота его появления в тексте. Первой строкой программы записывается первая строка этой матрицы — в нашем случае цифра 5 и единица. Далее в программе циклом с параметром j ведется перебор переменной-строки Текст со второго его элемента (первый уже зафиксирован в матрице v) до последнего. Принимается, что очередной символ текста, зафиксированный в переменной Символ, ранее не встречался: в переменную НовыйСимвол заносится значение «Да». Во втором (вложенном) цикле с параметром k ведется проверка этого предположения перебором ранее сформированных строк матрицы v. Если окажется, что очередной символ уже встречался в тексте (Символ = $v_{k,1}$) и наше предположение станет ошибочным (переменная НовыйСимвол принимает значение «Нет»), то значение элемента матрицы $v_{k,2}$ увеличивается на единицу. Если же очередной символ из переменной Текст в матрице v еще не зафиксирован (переменная НовыйСимвол сохранила значение «Да» по выходу из цикла с параметром k), то в матрице v создается новая строка ($i \leftarrow i + 1$), куда заносится сам новый символ ($v_{i,1} \leftarrow$ Символ) и единица ($v_{i,2} \leftarrow 1$). Возвращает программа отсортированную и транспонированную матрицу v — см. нижнюю часть рисунка 1.

Наш шифрованный текст, как уже было подсчитано, содержит 203 знакоместа. Отдельных же символов в тексте всего 20, из которых наиболее часто (33 раза) встречается восьмерка. Такой короткий текст можно, конечно, проанализировать и без компьютера, как это сделал герой рассказа Эдгара По (рассказ опубликован впервые в 1848 г.). Но через нашу программу можно пропустить более объемные тексты. Это позволит узнать, какие буквы в том или ином языке используются чаще, а какие реже. Эту же работу можно провести, беря за основу не буквы, а отдельные слова. В английском языке, на котором был написан рассказ «Золотой жук», самая «частая» буква — это буква «е», а самое частое слово — это определенный артикль «the». Это и послужило ключом к расшифровке текста в рассказе Эдгара По (творческая часть дешифровки). Как?! Прочтите или перечитайте рассказ еще раз: http://lib.ru/INOFANT/POE/Goldbug.txt

Рассказ Эдгара По «Золотой жук» в свое время (середина девятнадцатого века) вызвал большой интерес общества к криптографии. Тут можно упомянуть и знаменитый рассказ Артура Конан Дойла «Пляшущие человечки» (http://lib.ru/AKONANDOJL/sh_dancm.txt), сюжет которого также основан на расшифровке цепочек символов.

Но если раньше тайнописью занимались в основном военные и дипломаты, то в наше время с этим важным разделом информатики прямо или косвенно приходится сталкиваться всем и повсеместно, разговаривая по сотовому телефону или снимая деньги в банкомате...

Мы надеемся, что наша простенькая программа, помогающая дешифровке простого кода, станет некоей стартовой точкой изучения более сложных алгоритмов криптографии. А пока на уроках информатики можно поиграть в такую игру. Двое разрабатывают для переписки код, основанный на замене одних символов другими, а третий, «перехватывая» и анализируя эти сообщения, раскрывает «шифровки». (В таких сообщениях, как правило, опускаются пробелы и знаки препинания. Это было сделано в шифровке рассказа Эдгара По. В рассказе же Артура Конан Дойла пляшущий человечек с флажком в руке отмечал начало слова.) Сколько понадобится для этого перехваченных сообщений и времени — это отдельный вопрос. Вернее, задание на создание программы для раскрытия ключа шифра. Часть этой программы показана на рисунке 1.

Интернет-источники

1. *Конан Дойл А.* Пляшущие человечки. http://lib.ru/AKONANDOJL/sh_dancm.txt

2. Криптография // Образовательный сайт А. Н. Варгина. http://www.ph4s.ru/book_kripto.html

3. *По Э. А.* Золотой жук. http://lib.ru/INOFANT/POE/Goldbug.txt