Задача 9: Шпион

Накорнеева Юлия, Шишко Тимофей, Астахов Александр
Лицей БНТУ

Минск 2021

Резюме:

Решены пункты: 1.1, 1.2, 1.3, 1.4, 2, частично решен пункт 3. Предложено обобщение. При исследовании данной задачи получены следующие основные результаты:

1. Для пункта 1.1 время дешифровки равно 32с.

2. Было доказано, что шифр может быть расшифрован для любых натуральных t и N. Найдена формула для нахождения времени дешифровки R для пунктов 1.2, 1.3:

$$R = \frac{N}{\gcd(t, N)} - 1.$$

3. Для пункта 1.4 было доказано, что шифр может быть расшифрован при соблюдении следующего условия:

$$\gcd(a, N) = 1.$$

   Была найдена формула для нахождения времени дешифровки R:

$$R = \operatorname{lcm}\left(P_N(a), P_{\frac{N(a-1)}{\gcd(b,N(a-1))}}(a)\right) - 1.$$

4. Для пункта 2 доказано, что шифр может быть расшифрован для любых натуральных N и для дешифровки необходимо затратить $R = 1$ секунду.

5. Для пункта 3 было доказано, что шифр может быть расшифрован для любых натуральных k и N. Оценено время дешифровки:

$$R = \operatorname{lcm}[a_1, a_2 \dots a_{n-1}, a_n] - 1.$$

   Где $\{a_i\}$, $i \in [1, n]$ – множество всех периодов символов одного слова.

6. Обобщен пункт 2.

Постановка задачи:

Шпион ищет секретную информацию. Увидев любой текст, он может мгновенно понять, является ли этот текст той информацией, которую он ищет. К сожалению, он практически ничего больше делать не умеет. Тем не менее, он получил доступ к компьютеру, на котором есть искомая информация. Но она один раз зашифрована шифром, который называется шифр Ш, и выведена на экран. Также компьютер за 1 секунду может шифровать любой текст шифром Ш и вывести результат на экран. Больше компьютер ничего существенного также делать не умеет. Во всех следующих пунктах надо решить задачу: "Сможет ли шпион узнать искомую информацию? В случае положительного ответа, найдите точное время, за которое он это сделает или оцените его."

1. В этом пункте шифрование текста побуквенное.

   1.1 Шифр Ш сдвигает русский алфавит на 2 по кругу.

   1.2 Шифр Ш сдвигает алфавит из N символов на 2 по кругу.

   1.3 Шифр Ш сдвигает алфавит из N символов на t по кругу.

   1.4 Шифр Ш берет номер буквы n(нумерация ведется с 0) в алфавите из N символов, вычисляет $an + b$ и выдает букву с таким же остатком от деления на N , что и вычисленное выражение. Здесь a, b — заданные числа (не меняются в процессе шифрования).

2. Шифр Ш шифрует текст длины N^2 следующим образом: текст построчно сверху вниз записывается в квадрат $N \times N$. Затем он выписывается по столбцам слева направо. Это и есть результат.

   Вначале решите задачу для $N \in \{3, 4, 5\}$. Ниже приведен пример для $N = 3$ и текста "123456789".

$$123456789 \implies \begin{matrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{matrix} \implies 147258369$$

3. Шифр Ш зависит от числа k. Текст длины N записывается в строку следующим образом:

   - первый символ записывается на первую строку, затем каждый последующий символ записывается ниже предыдущего на следующую строку до тех пор, пока не будет записан символ на последнюю строку;
   - далее направление записи символов меняется — каждый последующий символ записывается выше предыдущего на соответствующую строку, пока не будет записан символ на первой строке
   - и снова направление записи меняется...

$$123456789 \implies \begin{matrix} 1 & 5 & \downarrow & 9 \\ 2 & 4 & 6 & 8 \\ 3 & \uparrow & 7 & \uparrow \end{matrix} \implies 159246837$$

   3.1 Решите задачу для $k = 2$ и $N = 2t$.

   3.2 Затем и для произвольного N.

   3.3 Рассмотрите другие значения k.

4. Предложите свои направления исследования в этой задаче и изучите их.

Решение:

Замечание: В данной работе за R примем в секундах время дешифровки, которое необходимо затратить шпиону, что в свою очередь,равно количеству повторений шифрования сделанное компьютером. Введем $\gcd(a,b) = \text{НОД}(a,b)$ и $\text{lcm}(a,b) = \text{НОК}(a,b)$.

1.3 Шифр Ш сдвигает алфавит из N символов на t по кругу.

Пусть символ g под номером $a \in [0; N-1]$ такой, что после z повторений шифра он вернётся в исходный, тогда:

$$\forall\ a \in [0; N-1]\ \exists\ g_a : tz + a \equiv a\ (\text{mod}\ N) \implies tz \equiv 0\ (\text{mod}\ N) \implies tz = Nb.$$

Так как $(tz + a \equiv a\ (\text{mod}\ N) \iff tz \equiv 0\ (\text{mod}\ N)) \implies \forall\ z_a = \text{const}$. Тоесть каждый символ предет в себя за одинаковое количество повторений шифра.

Поделим обе части на $\gcd\ (N,t)$, тогда получим следующее равенство

$$\frac{tz}{\gcd(N,t)} = \frac{Nb}{\gcd(N,t)} \implies zt^{'} = bN^{'}.$$

Где $t^{'} = \frac{t}{\gcd(N,t)}$ и $N^{'} = \frac{N}{\gcd(N,t)}$; $N^{'}$ и $t^{'}$ - взаимно простые.

Выразим z:

$$z = \frac{N^{'}b}{t^{'}}.$$

Так как $\gcd(N^{'}, t^{'}) = 1$ и $z \in \mathbb{N}$, значит $b \equiv 0\ (\text{mod}\ z)$. Тогда для получения наименьших из возможных z соблюдается условие $\frac{b}{t^{'}} = 1$.

Получим:

$$z = N^{'} = \frac{N}{\gcd(N,t)}.$$

Ввиду того, что одно повторение шифра было выполнено компьютером первоначально, то необходимое количество повторений R Шифра Ш равно:

$$R = \frac{N}{\gcd(N,t)} - 1.$$

Имея количество повторений, которые необходимы для дешифровки шифра, и зная, что компьютер затрачивает одну секунду на шифровку, с помощью полученной формулы мы можем оценить время, которое шпион затратит на дешифровку. И оно будет равно количеству повторений R.

1.1 Шифр Ш сдвигает русский алфавит на 2 по кругу.

Для пункта 1.1 имеем $N = 33$ и $t = 2$, тогда количество повторений R шифра равно:

$$R = \frac{33}{\gcd(33,2)} - 1 = 32.$$

Тогда шпион затратит 32 секунды на дешифровку.

Ответ: 32 секунды.

1.2 **Шифр Ш сдвигает алфавит из N символов на 2 по кругу.**

Так как пункт 1.2 является частным случаем пункта 1.3, имеем:

$$R = \frac{N}{\gcd(N, 2)} - 1.$$

1.4 **Шифр Ш берёт номер буквы (нумерация ведется с 0) в алфавите из N символов, вычисляет $an + b$ и выдаёт букву с таким же остатком от деления на N, что и вычисленное выражение. Здесь a, b — заданные числа (не меняются в процессе шифрования).**

Для того чтобы шпион смог расшифровать текст, после шифрования каждого символа не должно быть два и более одинаковых символов, полученных из разных исходных символов. Исследуем это условие. Допустим, что есть такие два символа n и $n^{'}$ такие, что:

$$an + b \equiv an^{'} + b \pmod N \implies an \equiv an^{'} \pmod N.$$

Рассмотрм случай, когда $\gcd(a, N) \neq 1$:

$$N > \frac{N}{\gcd(a, N)} > 0$$

Пусть $n = 0$ и $n^{'} = \frac{N}{\gcd(a,N)}$, тогда:

$$0 \equiv \frac{aN}{\gcd(a, N)} \pmod N$$

Тогда получаем, что n и $n^{'}$ переходят в один символ. Значит шпион не сможет расшифровать текст при $\gcd(a, N) \neq 1$.

Рассмотрм случай, когда $\gcd(a, N) = 1$.

Предположим, что: $n_i \not\equiv n_j \pmod N$, где $n_i \in [1, N-1]$ и $n_j \in [1, N-1]$, и при этом $an_i \equiv an_j \pmod N$, но из того, что $an_i \equiv an_j \pmod N$ следует, что $n_i \equiv n_j \pmod N$(так как $\gcd(a, N) = 1$), тоесть получили противоречие, значит шпион сможет расшифровать текст при $\gcd(a, N) = 1$.

Таким образом мы определили ограничения на N и a. Далее, пусть $n_1 \in [0\ ;\ N-1]$ — исходный символ, тогда после первого повторения шифра получим символ n_2:

$$an_1 + b \equiv n_2 \pmod N.$$

После второго повторения получим:

$$an_2 + b = a(an_1 + b) + b = a^2 n_1 + b(1 + a) \equiv n_3 \pmod N.$$

После третьего повторения:

$$a^3 n_1 + b(1 + a + a^2) \equiv n_4 \pmod N.$$

Пусть после m повторений мы вернёмся в исходный символ, тогда:

$$a^m n_1 + b(1 + a + a^2 + \ldots + a^{m-1}) \equiv n_1 \pmod N.$$

Заметим, что в скобках геометрическая прогрессия (для $a \geqslant 2$). Тогда для $a = 1$ возвращаемся к пункту 1.3. Воспользовавшись формулой суммы n-членов геометрической прогрессии, получим:

$$a^m n_1 + \frac{b(a^m - 1)}{a - 1} \equiv n_1 \pmod{N} \implies (a^m - 1)n_1 + \frac{b(a^m - 1)}{a - 1} \equiv 0 \pmod{N}.$$

Заметим, что:

$$\frac{b(a^m - 1)}{a - 1} = \text{const}.$$

Пусть:

$$\frac{b(a^m - 1)}{a - 1} = -c \implies (a^m - 1)n_1 \equiv c \pmod{N}.$$

Так как $n_1 \in [0\ ;\ N - 1]$, то при $n_1 = 1$ и $n_1 = 2$ получим следующие сравнения:

$$(a^m - 1) \equiv c \pmod{N},$$

$$2(a^m - 1) \equiv c \pmod{N}.$$

Найдём разность:

$$(a^m - 1) \equiv 0 \pmod{N} \implies \frac{b(a^m - 1)}{a - 1} \equiv 0 \pmod{N}.$$

Рассмотрим $\frac{b(a^m-1)}{a-1} \equiv 0 \pmod{N}$:

$$b(a^m - 1) = N(a - 1) \cdot d.$$

Поделим обе части на $\gcd(b, N(a - 1))$:

$$(a^m - 1) \cdot \frac{b}{\gcd(b, N(a-1))} = \frac{N(a-1)}{\gcd(b, N(a-1))} \cdot d,$$

$$\gcd\left(\frac{b}{\gcd(b, N(a-1))}, \frac{N(a-1)}{\gcd(b, N(a-1))}\right) = 1.$$

Поскольку нам необходимо найти наименьшее из возможных m, имеем:

$$a^m - 1 \equiv 0 \ \left(\mod \frac{N(a-1)}{\gcd(b, N(a-1))}\right).$$

Так как $a^m - 1 \equiv 0 \pmod{N}$, получаем систему:

$$\begin{cases} a^m - 1 \equiv 0 \ \left(\mod \frac{N(a-1)}{\gcd(b,N(a-1))}\right), \\ a^m - 1 \equiv 0 \pmod{N}. \end{cases}$$

Имеем $\gcd(a, N) = 1$ — условие расшифровки, и $\gcd\left(a, \frac{N(a-1)}{\gcd(b,N(a-1))}\right) = 1$ — следует из условия расшифровки. Для нахождения наименьшего возможного m воспользуемся мультипликативным порядком целого числа по модулю. Обозначим через P.

$$\begin{cases} m \vdots P_{\frac{N(a-1)}{\gcd(b,N(a-1))}}(a) \\ m \vdots P_N(a) \end{cases} \implies m \vdots \operatorname{lcm}\left(P_N(a), P_{\frac{N(a-1)}{\gcd(b,N(a-1))}}(a)\right)$$

Так как рассматривается наименьшее m имеем:

$$m = \operatorname{lcm}\left(P_N(a), P_{\frac{N(a-1)}{\gcd(b,N(a-1))}}(a)\right).$$

Ввиду того, что одно повторение шифра было выполнено компьютером первоначально, то необходимое количество повторений R Шифра Ш равно:

$$R = \operatorname{lcm}\left(P_N(a), P_{\frac{N(a-1)}{\gcd(b,N(a-1))}}(a)\right) - 1.$$

Имея количество повторений, которые необходимы для дешифровки шифра, и зная, что компьютер затрачивает одну секунду на шифровку, с помощью полученной формулы мы можем оценить время, которое шпион затратит на дешифровку. И оно будет равно количеству повторений R.

2. **Шифр Ш шифрует текст длины N^2 следующим образом: текст построчно сверху вниз записывается в квадрат $N\times N$. Затем он выписывается по столбцам слева направо. Это и есть результат.**

   При шифровании текста длиной N^2 все символы записываются построчно в таблицу $N\times N$, примем её за матрицу $N\times N$. Тогда, так как при шифровке происходит замена столбцов на строки — транспонирование матрицы, после двойного транспонирования матрицы получаем исходную матрицу.

$$\begin{bmatrix} k_{11} & k_{12} & \cdots & k_{1N} \\ k_{21} & k_{22} & \cdots & k_{2N} \\ \vdots & \vdots & \ddots & \vdots \\ k_{N1} & k_{N2} & \cdots & k_{NN} \end{bmatrix} \Longrightarrow \begin{bmatrix} k_{11} & k_{21} & \cdots & k_{N1} \\ k_{12} & k_{22} & \cdots & k_{N2} \\ \vdots & \vdots & \ddots & \vdots \\ k_{1N} & k_{2N} & \cdots & k_{NN} \end{bmatrix} \Longrightarrow \begin{bmatrix} k_{11} & k_{12} & \cdots & k_{1N} \\ k_{21} & k_{22} & \cdots & k_{2N} \\ \vdots & \vdots & \ddots & \vdots \\ k_{N1} & k_{N2} & \cdots & k_{NN} \end{bmatrix}$$

   k — символ шифра Ш.
   То есть необходимо совершить 2 повторения шифра для его дешифровки. Так как одно повторение уже совершено компьютером, то нашему шпиону остаётся выполнить ещё одно повторение. Тогда шпион затратит $R = 1$ секунду на дешифровку.

   Для $N \in \{3, 4, 5\}$ время дешифровки равно 1 секунде, как было доказано выше.

3. **Шифр Ш зависит от числа k. Текст длины N записывается в строку следующим образом:**

   - первый символ записывается на первую строку, затем каждый последующий символ записывается ниже предыдущего на следующую строку до тех пор, пока не будет записан символ на последнюю строку;
   - далее направление записи символов меняется — каждый последующий символ записывается выше предыдущего на соответствующую строку, пока не будет записан символ на первой строке;
   - и снова направление записи меняется...

$$123456789 \Longrightarrow \begin{matrix} 1 & 5 & \downarrow & 9 \\ 2 & 4 & 6 & 8 \\ 3 & \uparrow & 7 & \uparrow \end{matrix} \Longrightarrow 159246837$$

3.3 N **и** k **– произвольные:**

Докажем, что шифр можно расшифровать. Допустим у нас есть два слова, такие, что после шифрования они перейдут в одно слово:

$$\begin{matrix} a_1, a_2 \dots a_n \\ b_1, b_2 \dots b_n \end{matrix} \xrightarrow{\text{Ш}} c_1, c_2 \dots c_n.$$

Тогда:

$$(\forall\ i,\ j\ a_i \xrightarrow{\text{Ш}} c_j\ \wedge\ b_i \xrightarrow{\text{Ш}} c_j) \implies (a_1, a_2 \dots a_n = b_1, b_2 \dots b_n).$$

Значит:

$$a_1, a_2 \dots a_n = b_1, b_2 \dots b_n.$$

То есть не существует двух различных слов, которые перейдут в одно. Из этого следует, что шифр можно расшифровать.

Рассмотрим некоторое слово:

Известно, что два разных слова после шифрования шифром Ш будут разными шифрованными словами.

Пусть шифруется слово длины n, тогда рассмотрим букву под номером k. После некоторого числа повторений a наш символ k станет под тем же номером, что и в начале. Тогда пусть a_k – наименьшее количество повторений шифра Ш, необходимых для того, чтобы символ под номером k вернулся в себя (будем называть это периодом символа под номером k). Предположим, что наименьшее чило повторений $a_k > n$. Тогда по принципу Дирихле найдутся две одинаковые буквы, что стояли на номере k, которые возвращаются в себя за меньшее число повторений, получаем противоречие. Тогда:

$$a_k \leqslant n.$$

Значит символ под номером k станет на место под номером k не больше, чем за $a_k \leqslant n$ итераций шифра Ш, применненых к слову.

$\{a_i\}$, где $i \in [1, n]$ – множество всех периодов символов одного слова. Значит наименьшее количество применений Шифра Ш к исходному слову:

$$R = \text{lcm}[a_1, a_2 \dots a_{n-1}, a_n] - 1.$$

А оно в свою очередь не больше, чем $\text{lcm}[1, 2 \dots n-1, n] - 1$.

3.1 $k = 2$ **и** $N = 2t$**.**

Пункт 3.1 является частным случаем пункта 3.3.

3.2 $k = 2$ **и** N**— произвольное.**

Пункт 3.2 является частным случаем пункта 3.3.

Обобщение:

Шифр Ш шифрует текст длины N^n следующим образом: в первый квадрат текст записывается построчно сверху вниз, затем аналогично в последующие квадраты. После из них составляются кубы и так до n-мерного гиперкуба. Теперь обозначим координатные оси гиперкуба как $x_1, x_2, x_3 \dots x_n$. Затем переведём оси следующим образом: $x_1, x_2, x_3 \dots x_n \to x_n, x_1, x_2 \dots x_{n-1}$ и выпешем из получившегося гиперкуба текст также, как записывали изначально. Это и есть результат.

Рассмотрим элемент $k_{j_1 \ldots j_n}$:

После первой итерации шифр переходит в элемент:

$$k_{j_n, j_1 \ldots j_{n-1}}$$

После второй итерации получим:

$$k_{j_{n-1}, j_n, j_1 \ldots j_{n-2}}$$

Значит после $n-1$ итераций шифра, наш текст перейдет в:

$$k_{j_2, j_3 \ldots j_1}$$

Применив ещу одну итерацию получим исходный символ. То есть наш шифр расшифровывается за n повторений. Тогда, так как одно повторение выполнено компьютером первоначально, то время дешифровки равно:

$$R = n - 1$$

Для пункта 2 $n = 2$, что является частным случаем обобщения.