Tehnici de CRIPTARE

Claudiu Soroiu

În acest număr continuăm prezentarea tehnicilor de criptare cu descrierea metodelor Vigenere si XOR. De asemenea, veți putea citi câteva informații referitoare la criptanaliză și la securitatea datelor.

Criptarea Vigenere

Tehnica de criptare *Vigenere* este o tehnică polialfabetică, în sensul că, prin criptare, două sau mai multe caractere diferite ale textului inițial pot fi codificate folosind același caracter în textul criptat.

Pentru a cripta un text folosind această metodă se alege mai întâi un alfabet, astfel încât toate caracterele conținute în textul inițial să se afle printre caracterele alfabetului și apoi o parolă formată din caractere ale alfabetului. Nu este necesar ca alfabetul ales să conțină numai caracterele textului inițial.

În continuare este prezentat modul în care se realizează criptarea unui text folosind tehnica *Vigenere*.

Algoritmul de criptare

Considerăm alfabetul format din literele: **A B C D E F**. Primul pas al algoritmului constă în aranjarea literelor alfabetului într-o matrice pătratică cu L linii și L coloane, unde L este numărul de simboluri ale alfabetului.

Prima linie a matricei este formată din literele alfabetului în ordine lexicografică, iar următoarele L - 1 linii sunt obținute fiecare din linia precedentă printr-o permutare la stânga cu o poziție.

În cazul alfabetului ales matricea obținută este următoarea:

	1	2	3	4	5	6
1	A	B	C	D	E	F
2	B	C	D	E	F	A
3	C	D	E	F	A	B
4	D	E	F	A	B	C
5	E	F	A	B	C	D
6	F	A	B	C	D	E

Se observă că un caracter nu apare decât o singură dată pe o linie sau o coloană și fiecare linie sau coloană conține toate literele alfabetului considerat.

Următorul pas constă în alegerea unei parole formată din simboluri ale alfabetului. Să presupunem că dorim să criptăm textul **FAAD**, folosind parola **BAB**.

Pentru ca această tehnică să fie eficientă, ar fi ideal ca lungimea parolei să fie mai mare sau cel puțin egală cu lungimea textului care urmează a fi criptat. Datorită faptului că o astfel de parolă nu poate fi reținută de către utilizator, se alege o parolă mai scurtă asupra căreia se aplică diverse metode prin care ea este extinsă pentru ca lungimea ei să depășească lungimea textului. Cea mai simplă metoda utilizată în acest sens constă în concatenarea parolei cu ea însăși până în momentul în care lungimea sa este cel puțin egală cu lungimea textului.

Pentru exemplul luat, parola, prin concatenare cu ea însăși, devine **BABBAB**.

La ultimul pas, fiecare al k-lea caracter din textul inițial este criptat astfel:

- se determină numărul de ordine i, corespunzător poziției caracterului în alfabetul considerat;
- se determină numărul de ordine j, corespunzător poziției în alfabetul considerat a celui de-al k-lea caracter din parolă;
- se scrie caracterul care se află pe linia i și coloana j în matrice.

Textul considerat, prin criptare, folosind parola extinsă **BABBAB**, devine **AABE**. Pentru $k = 1$, caracterul de pe prima poziție a textului este **F** ($i = 6$), caracterul de pe prima poziție a parolei este **B**, deci ($j = 2$).

Așadar, caracterul care va apărea pe prima poziție a textului criptat se află pe coloana a doua a celei de-a șasea linii din matrice: **A**.

Acest procedeu continuă pentru toate valorile k mai mici sau egale cu lungimea textului.

Analizând textul inițial și cel criptat se observă că primele două caractere **F** și **A**, devin prin criptare, același caracter, adică **A**, iar cele două caractere **A** de pe pozițiile 2 și 3 sunt codificate diferit în textul criptat.

Algoritmul de decriptare

Algoritmul folosit pentru decriptarea unui mesaj codificat utilizând tehnica *Vigenere* nu este foarte complex.

Primul pas care trebuie parcurs pentru a decripta un text constă în construirea parolei extinse pe baza parolei folosite. Procedeul prin care se realizează extinderea parolei în procesul de decriptare trebuie să fie același cu cel folosit pentru a extinde parola în procesul de criptare.

Pentru a decripta un text se aplică următorul algoritm: pentru fiecare al *k*-lea caracter din parola extinsă se găsește numărul de ordine *j* al acestuia în cadrul alfabetului considerat, se caută pe coloana *j* din matrice al *k*-lea caracter din textul criptat și se scrie caracterul care are numărul de ordine *i*, unde *i* este numărul liniei pe care a fost găsit caracterul.

Considerăm alfabetul definit anterior; în procesul de decriptare se ține cont de matricea atașată acestui alfabet.

Pentru textul **AABE** care a fost criptat folosind parola **BAB** (extinsă prin concatenare cu ea însăși, această parolă devine **BABBAB**) prin decriptare se obține textul **FAAD** astfel: pentru primul caracter din parola extinsă *k* este egal cu 1; pe coloana *j* = 2 corespunzătoare numărului de ordine al caracterului **B** din parolă se caută caracterul **A** din text și se găsește pe linia *i* = 6 și în final se scrie caracterul **F** care are numărul de ordine 6 în cadrul alfabetului. Acest procedeu se aplică pentru toate caracterele textului criptat.

Concluzii

Se observă că în cazul extinderii unei parole cu ea însăși, tehnica de criptare *Vigenere* este similară cu criptarea lui *Cezar* pe grupuri de elemente.

Dacă parola inițială are lungimea *n* și considerăm un număr strict pozitiv $k < n$, atunci, dacă extragem caracterele de pe pozițiile k, $n + k$, $2 \cdot n + k$, ... etc. din textul inițial, se observă că pentru criptarea lor s-a folosit același deplasament.

Criptarea XOR

Tehnica de criptare ***XOR*** se bazează pe operatorul binar ***XOR*** (*e**X**clusive **OR** - sau exclusiv*). Reamintim faptul că rezultatul aplicării acestui operator asupra a două valori binare identice este 0, în timp ce rezultatul aplicprii sale asupra a două valori binare diferite este 1. Rezultatele aplicării acestui operator sunt sintetizate în figura următoare.

⊕	0	1
0	0	1
1	1	0

Algoritmul de criptare

Pentru a cripta un mesaj folosind această metodă este nevoie de o parolă și de un alfabet. În acest caz, numărul de simboluri ale alfabetului trebuie să aibă forma 2^n, unde *n* este un număr natural.

Și de această dată, ar fi ideal ca lungimea parolei să fie mai mare sau egală cu lungimea textului care urmează a fi criptat. Ca urmare a acestui fapt, primul pas al algoritmului de criptare constă în obținerea parolei extinse, la fel ca în cazul criptării *Vigenere*.

După obținerea parolei extinse, procedeul de criptare este: pentru fiecare al *k*-lea caracter din text și al *k*-lea caracter din parolă se caută în cadrul alfabetului numerele de ordine *i* și *j* ale celor două caractere, se aplică operatorul *XOR* asupra reprezentării binare a celor două numere (se obține reprezentarea binară a unui număr *n*) și, în final, se scrie caracterul care are numărul de ordine *n* în cadrul alfabetului.

Fie alfabetul format din simbolurile setului de caractere ***ASCII*** extins, care are 2^8 elemente. Dacă dorim să criptăm un text, și presupunem că pe poziția *k* avem caracterul **A** care are numărul de ordine 65, iar în parola extinsă avem pe poziția *k* caracterul **B** cu numărul de ordine 66, atunci, prin aplicarea operatorului *XOR* între reprezentările binare ale celor două numere, se obține caracterul care are numărul de ordine 3. În cadrul setului de caractere *ASCII* extins, caracterele cu numere de ordine mai mici decât 32 poartă numele de *caractere de control*, și sunt considerate caractere neimprimabile. De obicei, dacă se folosește standardul *ASCII*, după criptarea cu această metodă, foarte puține caractere sunt imprimabile.

Aplicarea operatorului *XOR* asupra reprezentării binare a numerelor 65 și 66 este ilustrată în cele ce urmează:

```
0 1 0 0 0 0 0 1 = 65
0 1 0 0 0 0 1 0 = 66
0 0 0 0 0 0 1 1 =  3
```

Algoritmul de decriptare

Algoritmul de decriptare este asemănător cu cel de criptare, singura deosebire constă în faptul că mesajul criptat ia locul textului inițial. Să presupunem că în textul criptat avem pe poziția *k* caracterul cu numărul de ordine 3 și în parola extinsă avem pe poziția *k* caracterul cu numărul de ordine 66 (**B**), prin aplicarea operatorului *XOR* asupra reprezentării binare ale celor două numere se obține caracterul cu numărul de ordine 65 (**A**), astfel:

```
0 0 0 0 0 0 1 1 =  3
0 1 0 0 0 0 1 0 = 66
0 1 0 0 0 0 0 1 = 65
```

În cazul în care nu se cunoaște parola cu care a fost criptat un mesaj, dar se cunoaște faptul că acesta conține foarte multe caractere imprimabile, se presupune că parola are cel puțin $n \geq 3$ caractere și se încearcă aplicarea operatorului *XOR* între textul criptat și un caractere imprimabil astfel încât, pentru toate numerele întregi strict pozitive $k < n$, în mulțimea formată cu caracterele de pe pozițiile k, $n + k$, $2 \cdot n + k$, ..., etc., să se obțină un maxim de caractere imprimabile care pot să facă parte din textul inițial. În cazul în care nu s-au obținut rezultatele dorite numărul *n* se incrementează cu 1 și se reia pasul anterior.

După cum se poate observa, și această tehnică de criptare este polialfabetică.

Dacă asupra unui text se aplică succesiv mai multe tehnici de criptare diferite, atunci procesul de decriptare este îngreunat foarte mult datorită faptului că dacă un criptanalist încearcă să decripteze textul pe baza frecvenței de apariție a literelor, el are foarte puține șanse de reușită, mai

ales dacă una dintre tehnicile de criptare utilizate operează cu două alfabeturi distincte și de lungimi diferite.

Criptanaliză

Criptanaliza este știința care se ocupă cu studiul decriptării mesajelor când nu se cunoaște cheia de criptare.

Există mai multe procedee de analiză a textelor criptate. Cele mai des utilizate sunt analiza frecvenței de apariție a caracterelor și analiza probabilistică. De exemplu, în limba engleză, caracterul cel mai frecvent folosit este **E** iar grupurile de două caractere care apar cel mai des sunt **TH** și **HE**.

Metodele matematice care stau la baza criptanalizei și criptografiei moderne sunt foarte complexe.

Pentru a decripta un text este nevoie de cel puțin două informații. Prima informație constă în cunoașterea metodei generale de criptare utilizată sau categoria de tehnici de criptare din care face parte, iar a doua informație constă din cunoașterea unor parametri implicați în procesul de criptare (informații asupra modului de utilizare al cheii de criptare).

În practică, utilizatorii criptografiei dețin în cele mai multe cazuri, echipamente pentru criptarea și decriptarea datelor pentru o anumită tehnică sau categorie de tehnici de criptare.

Securitatea datelor în zilele noastre

Dacă dorim să accesăm un cont de pe un *server* avem nevoie de o parolă. Confidențialitatea mesajelor și a datelor transmise de la *client* la *server* și invers depinde de gradul de complexitate al metodei de criptare folosite de cele două sisteme *client*, respectiv *server* și de complexitatea parolei alese.

Chiar dacă se folosește o parolă complexă, utilizarea unui cod unic de acces prezintă unele dezavantaje. În cazul în care parola este interceptată de anumite persoane, complexitatea sa nu mai are o importanță foarte mare. Cea mai sigură modalitate de acces la un *server* ar fi ca la fiecare accesare a contului să se utilizeze o altă parolă. Acest tip de metode poartă numele de ***one time pass*** (*parolă folosită o singură dată*).

În cele ce urmează vom prezenta un exemplu de realizare a unui sistem ce se încadrează în tipul *one time pass*.

Prezentare

În cazul parolelor care se folosesc o singură dată, trebuie utilizată o tehnică de criptare a datelor. De obicei, parolele sunt constituite din caractere imprimabile, deci alfabetul utilizat trebuie să fie o submulțime a setului de caractere imprimabile din întregul set de caractere utilizat de calculator.

În general, dacă dorim să criptăm un text T, ale cărui litere aparțin unui alfabet A, avem nevoie de:

- o funcție de criptare f definită pe mulțimea părților textului T;
- o parolă sau cheie de criptare.

De obicei, caracterele din textul inițial și cel criptat aparțin aceluiași alfabet.

Dintr-un alfabet pot lipsi semnele obișnuite de punctuație și spațiile, cuvintele fiind scrise unul după altul, lucru care îngreunează foarte mult munca celui care încearcă să decripteze mesajul criptat în eventualitatea în care nu deține parola folosită în procesul de codificare a mesajului.

Textul inițial T și textul criptat C sunt împărțite în mai multe grupuri compuse din una, două sau mai multe litere, în funcție de tehnica de criptare utilizată. În tehnicile prezentate anterior cele două texte erau împărțite în grupuri de câte o singură literă, deoarece transformările erau aplicate la un moment dat, asupra unui singur caracter din text.

Funcția de criptare f este definită pe mulțimea tuturor grupurilor de litere care se formează din textul inițial, și, de obicei, este o funcție bijectivă. Inversa ei, f^{-1}, se folosește pentru a decripta mesajul.

Fie f_P funcția de criptare asociată parolei P. Pentru o parolă indrodusă către utilizator procesul de criptare este dat de formula $f_P(T) = C$.

Pentru a decripta textul C, de obicei, se folosește funcția inversă f_P^{-1} asociată parolei P, și în procesul de decriptare se folosește formula $f_P^{-1}(C) = T$.

Un *sistem one time pass* este un sistem în care conexiunea dintre *client* și *server* se face prin folosirea unui model *one time pass*. În cazul unui sistem *one time pass* se alege o parolă inițială P_0 care este stocată pe *server* și un număr întreg strict pozitiv n suficient de mare. Numărul n poate să aibă o valoare implicită, fără a mai necesita alegerea lui de către utilizator. Funcția de criptare, pe care atât clientul cât și *server*-ul o aplică parolei P_0, trebuie să fie bijectivă, iar inversa ei trebuie să fie foarte greu de calculat folosind procedee matematice.

Clientul deține, la rândul său, parola inițială P_0 și numărul n. Când se realizează conexiunea dintre *client* și *server*, clientul aplică funcția de criptare f de n ori astfel:

$$P_i = f(P_{i-1}),\ i = 1, ..., n.$$

Parola P_n care se obține este transmisă *server*-ului. La rândul său *server*-ul creează, utilizând același procedeu, propria sa parolă P_n, și dacă coincide cu cea pe care a trimis-o clientul se realizează conexiunea și numărul n este decrementat automat cu 1 atât pe stația *server* cât și pe stația *client*. În acest mod, la următoarea accesare a contului de pe *server*, se va folosi o altă parolă.

Modelul *one time pass* presupune faptul că toate cele n parole care se obțin prin aplicarea succesivă a funcției f, sunt distincte. În momentul în care s-au epuizat toate cele n parole distincte, trebuie înlocuită cheia (parola) P_0.

Dacă funcția f ar fi ușor inversabilă, atunci, dacă parola curentă este interceptată de o persoană neautorizată, prin aplicarea funcției f^{-1} se obține parola care va fi utilizată la următoarea accesare a contului de pe *server*.

Funcția f trebuie aleasă astfel încât, chiar dacă expresia sa intră în posesia unor persoane neautorizate, expresia funcției f^{-1} să fie foarte greu de calculat.

Claudiu Soroiu este redactor al GInfo. Poate fi contactat prin e-mail la adresa csoroiu@yahoo.com.