SEMINAR IZ KRIPTOGRAFIJE

Uvod v kriptografijo

Aleksandar Jurišić

IMFM

16. februar 1998

Načrt

- predstavitev seminarja
- kaj je kriptografija
- zgodovina
- 20 let stara naloga iz Preseka
- klasična kriptografija
- razbijanje klasičnih tajnopisov

Predstavitev seminarja

1. predavanja
2. predstavitev vodje
3. namen seminarja
4. predznanje
5. učbenik
6. načrt dela
7. elektronsko obveščanje
8. domača stran
9. literatura
10. revije

Kaj je kriptografija?

Osnovna motivacija za naš študij je uporaba kriptografije v realnem svetu.

Cilje kriptografije bomo dosegali z matematičnimi sredstvi.

I. Cilji kriptografije:

1. **Zasebnost/zaupnost/tajnost:**
varovanje informacij pred tistimi,
ki jim vpogled ni dovoljen, dosežemo
s šifriranjem.

2. **Celovitost podatkov:**
zagotovilo, da informacija ni bila
spremenjena z nedovoljenimi sredstvi
(neavtoriziranimi sredstvi).

3. **Overjanje sporočila (ali izvora podatkov):**
   potrditev izvora informacij.

4. **Identifikacija:**
   potrditev identitete predmeta ali osebe.

5. **Preprečevanje nepriznavanja:**
   preprečevanje, da bi nekdo zanikal dano obljubo ali storjeno dejanje.

6. Drugi kriptografski protokoli:

1. grb/cifra po telefonu
2. mentalni poker
3. shema elektronskih volitev
   (anonimno glasovanje brez goljufanja)
4. (anonimni) elektronski denar

I. Cilji kriptografije:

1. zasebnost/zaupnost/tajnost
2. celovitost podatkov
3. overjanje sporočila (ali izvora podatkov)
4. identifikacija
5. preprečevanje nepriznavanja
6. drugi kriptografski protokoli

NAUK: Kriptografija je več kot samo širfiranje (enkripcija).

II. Širši pogled na kriptografijo – varnost informacij

Kriptografija je sredstvo, s katerim dosežemo varnost informacij, ki med drugim zajema:

Varnost računalniškega sistema

Tehnična sredstva, ki omogočajo varnost računalniškega sistema. Računalniški sistem lahko pomeni samo en računalnik z več uporabniki, lokalno mrežo (LAN), internet, web strežnik, bankomat, itd.

Med drugim obsega:

- varnostne modele in pravila, ki določajo zahteve po varnosti, katerim mora sistem ustrezati
- varen operacijski sistem
- zaščito pred virusi
- zaščito pred kopiranjem
- kontrolne mehanizme (beleženje vseh aktivnosti, ki se dogajajo v sistemu lahko omogoči *odkrivanje* tistih kršitev varnostnih pravil, ki jih ni mogoče preprečiti)
- analiza tveganja in upravljanje v primeru nevarnosti

Varnost na mreži

Zaščita prenašanja podatkov preko komercialnih mrež, tudi računalniških in telekomunikacijskih.

Med drugim obsega:

- protokole na internetu in njihovo varnost
- požarne zidove
- trgovanje na internetu
- varno elektronsko pošto

II. Širši pogled na kriptografijo – varnost informacij

1. varnost računalniškega sistema
2. varnost na mreži

NAUK: Kriptografija je samo majhen del varnosti informacij.

III. Gradniki kriptografije:

1. **matematika** (predvsem teorija števil)
2. **računalništvo** (analiza algoritmov)
3. **elektrotehnika** (hardware)
4. **poznavanje aplikacij** (finance,...)
5. **politika** (restrikcije, key escrow, NSA,...)
6. **pravo** (patenti, podpisi, jamstvo,...)
7. **družba** (npr. enkripcija omogoča zasebnost, a otežuje pregon kriminalcev)

NAUK: Uporabna kriptografija je več kot samo zanimiva matematika.

Povzemimo vse nauke:

Kriptografija je več kot samo širfiranje.

Kriptografija je samo majhen del varnosti informacij.

Uporabna kriptografija je več kot samo zanimiva matematika.

Zgodovina

Kriptografija ima dolgo in zanimivo zgodovino:

- Špartanci, mitološke reference

Še ena antična: o obriti glavi

Medtem, ko je bil genialni Histius na perzijskem sodišču, je hotel obvestiti Aristagorasa iz Grčije, da dvigne upor.

Seveda je bili pomembno, da nihče ne prestreže sporočilo.

Da bi zagotovil tajnost, je Histius obril sužnja, ki mu je nabolj zaupal, mu vtetoviral na glavo sporočilo [sužnju so rekli, da mu začenjajo zdraviti slepoto] in počakal, da mu zrastejo lasje.

Sužnju je bilo ukazano, da reče
Aristagorasu:
"*Obrijte mojo glavo in poglejte nanjo.*"

Aristagoras je nato zares dvignil upor.

To je primer **prikritega tajnopisa**, sporočilo je prisotno, a na nek način prikrito.

Poznamo mnogo takšnih primerov.

Varnost takega sporočila je odvisna od trika prikrivanja.

Tak trik je lahko odkriti, poleg tega pa ne omogoča hitrega šifriranja in dešifriranja.

To ne pride vpoštev za **resno uporabo**.

Anglija: Sir John dobi sporočilo: Worthie Sir John:- Hope, that is ye beste comfort of ye afflicted, cannot much, I fear me, help you now. That I would saye to you, is this only: if ever I may be able to requite that I do owe you, stand not upon asking me. 'Tis not much that I can do: but what I can do, bee ye verie sure I wille. I knowe that, if dethe comes, if ordinary men fear it, it frights not you, accounting it for a high honor, to have such a rewarde of your loyalty. Pray yet that you may be spared this soe bitter, cup. I fear not that you will grudge any sufferings; only if bie submission you can turn them away, 'tis the part of a wise man. Tell me, an if you can, to do for you anythinge that you wolde have done. The general goes back on Wednesday. Restinge your servant to command. - R.T.

Če vam uspe "med vrsticami" prebrati:

PANEL AT EAST END
OF CHAPEL SLIDES

verjetno ne boste občutili enakega olajšanja kot Sir John Trevanion, njemu pa je vsekakor uspelo pobegniti, sicer bi ga v gradu Colcester gotovo usmrtili prav tako, kot so Sir Charlesa Lucasa ter Sir Georga Lislea.

Druga svetovna vojna:

– Enigma (Nemčija),

– Purple (Japonska),

– Hagelin (ZDA).

Transpozicijski tajnopis

V transpozicijskem tajnopisu ostanejo črke originalnega sporočila nespremenjene, njihova mesta pa so pomešana na kakšen sistematičen način.

Primer: permutacija stolpcev

Te tajnopise zlahka prepoznamo, če izračunamo gostoto samoglasnikov (v angleščini je ta 40%, in skoraj nikoli ne pade zunaj intervala 35%–45%).

Težko jih rešimo, vendar pa se potrpljenje na koncu običajno izplača.

Zamenjalni (substitucijski) tajnopis

Primer: Cezar, Morzejeva abeceda,...

Domača zgodovina

Tomaž Pisanski, Skrivnostno sporočilo
Presek V/1, 1977/78, str. 40-42.

```
YHW?HD+CVODHVTHVO-!JVG:CDCYJ(JV/-V?HV(
-T?HVW-4YC4(?-DJV/-(?S-VO3CWC%J(-V4-DC
V!CW-?CVNJDJVD-?+-VO3CWC%J(-VQW-DQ-VJ+
V?HVDWHN-V3C:CODCV!H+?-DJVD-?=CV3JO-YC
```

(črko Č smo zamenjali s C, črko Ć pa z D)

Imamo $26! = 403291461126605635584000000$ možnosti z direktnim preizkušanjem, zato v članku dobimo naslednje nasvete:

(0) Relativna frekvenca črk in presledkov v slovenščini: presledek 173,

E	A	I	O	N	R	S	L	J	T	V	D
89	84	74	73	57	44	43	39	37	37	33	30

K	M	P	U	Z	B	G	Č	H	Š	C	Ž	F
29	27	26	18	17	15	12	12	9	9	6	6	1

(1) Na začetku besed so najpogostejše črke N, S, K, T, J, L.

(2) Najpogostejše končnice pa so E, A, I, O, U, R, N.

(3) Ugotovi, kateri znaki zagotovo predstavljajo samoglasnike in kateri soglasnike.

(4) V vsaki besedi je vsaj en samoglasnik ali samoglasniški R.

(5) V vsaki besedi z dvema črkama je ena črka samoglasnik, druga pa soglasnik.

(6) detektivska sreča

(0)	V	-	C	D	J	?	H	W	O	(	+	3
	23	19	16	12	11	10	9	7	6	6	5	4
	Y	4	!	/	Q	:	%	T	N	S	G	
	4	3	3	2	2	2	2	2	2	1	1	

Zaključek `V --> ' '` (drugi znaki z visoko frekvenco ne morejo biti).

Dve besedi se ponovita: `O3CWC%J(-`, opazimo pa tudi eno sklanjatev: `D-?+-` ter `D-?+C`.

Torej nadaljujemo z naslednjim tekstom:

```
YHW?HD+C ODH TH O-!J G:CDCYJ(J /- ?H
(-T?H W-4YD4(?-DJ /-(?S- O3CWC%J(- 4-DC
!CW-?C NJDJ D-?+- O3CWC%J(- QW-DQ- J+
?H DWHN- 3C:CODC !H+?-DJ D-?+C 3JO-YC
```

(3) Kanditati za samoglasnike e,a,i,o so znaki z visokimi frekvancami. Vzamemo:

$$\{\text{e,a,i,o}\} = \{\text{-,C,J,H}\}$$

(saj D izključi -,H,J,C in ? izključi -,H,C, znaki -,C,J,H pa se ne izključujejo)

Razporeditev teh znakov kot samoglasnikov izgleda prav verjetna. To potrdi tudi gostota končnic, gostota parov je namreč:

AV	CV	HV	JV	VO	?H	-D	DC	JM	W-	DJ	UC	CW	-?	VD
7	5	5	5	4	4	4	3	3	3	3	3	3	3	3

(5) Preučimo besede z dvema črkama:

```
   Samoglasnik na koncu
1) da ga na pa ta za (ha ja la)
2) če je le me ne se še te ve že (he)
3) bi ji ki mi ni si ti vi
4) bo do (ho) jo ko no po so to
5) ju mu tu (bu)
6) rž rt
```

```
    Samoglasnik na začetku
 1) ar as             (ah aj au)
 2) en ep              (ej eh)
 3) in iz ig
 4) on ob od os on     (oh oj)
 5) uk up uš ud um ur (uh ut)
```

in opazujemo besedi: /- ?H
ter besedi: J+ ?H.

J+ ima najmanj možnosti, + pa verjetno ni črka n, zato nam ostane samo še:

```
J+ ?H        DWHN-
/- ?H
iz te       (ne gre zaradi: D-?+C)
ob ta(e,o)  (ne gre zaradi: D-?+C)
od te       (ne gre zaradi: D-?+C)
```

tako da bo potrebno nekaj spremeniti
in preizkusiti še naslednje:
on bo; on jo; in so; in se; in je; in ta; en je; od tu ...

(6) Če nam po dolgem premisleku ne uspe najti rdeče niti, bo morda potrebno iskati napako s prijatelji (tudi računalniški program z metodo lokalne optimizacije ni zmogel problema zaradi premajhne dolžine tajnopisa, vsekakor pa bi bilo problem mogoče rešiti s pomočjo elektronskega slovarja).

Tudi psihološki pristop pomaga, je svetoval Martin Juvan

Podobna naloga je v angleščine dosti lažja, saj je v tem jeziku veliko členov THE, A in AN, vendar pa zato običajno najprej izpustimo presledke iz teksta, ki ga želimo spraviti v tajnopis. Po drugi strani pa imajo črke drugačno gostoto.

Razdelimo jih v naslednjih pet skupin:

1. E, z verjetnostjo okoli 0.120,
2. T, A, O, I, N, S, H, R, vse z verjetnostjo med 0.06 in 0.09,
3. D, L, obe z verjetnostjo okoli 0.04,
4. C, U, M, W, F, G, Y, P, B, vse z verjetnostjo med 0.015 in 0.028,
5. V, K, J, X, Q, Z, vse z verjetnostjo manjšo od 0.01.

Najbolj pogosti pari so (v padajočem zaporedju): TH, HE, IN, ER, AN, RE, ED, ON, ES, ST, EN, AT, TO, NT, HA, ND, OU, EA, NG, AS, OR, TI, IS, ET, IT, AR, TE, SE, HI in OF,

Najbolj pogoste trojice pa so (v padajočem zaporedju): THE, ING, AND, HER, ERE, ENT, THA, NTH, WAS, ETH, FOR in DTH.

Klasična kriptografija:

- zamenjalni (substitucijski) tajnopis
- afini tajnopis
- Vigenerejev tajnopis
- Hillov tajnopis

Kriptosistem je peterica $(\mathcal{P}, \mathcal{C}, \mathcal{K}, \mathcal{E}, \mathcal{D})$ za katero velja:

1. $\mathcal{P}$ je končna množica možnih čistopisov
2. $\mathcal{C}$ je končna množica možnih tajnopisov
3. $\mathcal{K}$ je končna množica možnih ključev.
4. Za vsak ključ $K \in \mathcal{K}$, imamo šifrirni postopek $e_K \in \mathcal{E}$ in ustrezen dešifrirni postopek $d_K \in \mathcal{D}$. $e_K : \mathcal{P} \longrightarrow \mathcal{C}$ in $d_K : \mathcal{C} \longrightarrow \mathcal{P}$ sta taki funkciji, da je $d_K(e_K(x)) = x$ za vsak $x \in \mathcal{P}$.

Zamični tajnopis: (shift cipher)
(poseben primer zamenjalnega tajnopisa)

`wewillmeetatmidnight`

```
22  4 22  8 11 11 12  4  4 19  0 19 12  8  3 13  8  6  7 19
 7 15  7 19 22 22 23 15 15  4 11  4 23 19 14 24 19 17 18  4
```

`HPHTWWXPPELEXTOYTRSE`

Kongruence: naj bosta a in b celi števili in m naravno število.

$$a \equiv b \pmod{m} \quad \Longleftrightarrow \quad m | b - a.$$

Afini tajnopis:

$$e(x) = ax + b \pmod{26} \qquad \text{za } a, b \in \mathbb{Z}_{26}$$

Za $a = 1$ dobimo zamični tajnopis. Funkcija je injektivna, če in samo če je $d(a, 26) = 1$.

Imamo $|\mathcal{K}| = 12 \times 26 = 312$ možnih ključev.

Za zamični tajnopis in afini tajnopis pravimo, da sta **monoabecedna**, ker preslikamo vsako črko v natanko določeno črko.

Vigenerejev tajnopis (1586):

Naj bo m neko naravno število in naj bo

$$\mathcal{P} = \mathcal{C} = \mathcal{K} = (\mathbb{Z}_{26})^m.$$

Za ključ $K = (k_1, k_2, \ldots, k_m)$ definiramo

$$\begin{aligned} e(x_1, \ldots, x_m) &= (x_1 + k_1, \ldots, x_m + k_m) \quad \text{in} \\ d(y_1, \ldots, y_m) &= (y_1 - k_1, \ldots, y_m - k_m), \end{aligned}$$

kjer sta operaciji "$+$" in "$-$" opravljeni po modulu 26.

To ni monoabecedni tajnopis.

Pravimo mu **poliabecedni tajnopis**.

Vigenerejev tajnopis in 26^m možnih ključev.

Za $m = 5$ je število 1.1×10^7 že preveliko, da bi "peš" iskali pravi ključ.

Hillov tajnopis (1929):

Naj bo m neko naravno število in naj bo

$$\mathcal{P} = \mathcal{C} = (\mathbb{Z}_{26})^m.$$

Za K vzemimo obrnljivo $m \times m$ matriko in definirajmo

$$e_K(x) = xK \quad \text{in} \quad d_K(y) = yK^{-1},$$

pri čemer so vse operacije opravljene v $\mathbb{Z}_{26}$.

Dešifriranje (razbijanje) klasičnih tajnopisov

Držali se bomo **Kerckhoffovega principa**, ki pravi, da "nasprotnik"

pozna kriptosistem, ki ga uporabljamo.

Ločimo naslednje nivoje napadov na kriptosisteme:

1. **samo tajnopis:** nasprotnik ima del tajnopisa
2. **poznani čistopis:** nasprotnik ima del čistopisa ter ustrezen tajnopis
3. **izbrani čistopis:** nasprotnik ima začasno na voljo šifrirno mašinerijo ter za izbrani $x \in \mathcal{P}$ konstruira $e(x)$
4. **izbrani tajnopis:** nasprotnik ima začasno na voljo dešifrirno mašinerijo ter za izbrani $y \in \mathcal{C}$ konstruira $d(y)$

Dešifriranje Vigenerejevega tajnopisa:

Test Kasiskega:
poiščemo dele tajnopisa, ki so identični in zabeležimo razdalje $d_1, d_2, \ldots$ med njihovimi začetki. Predpostavimo, da m deli največji skupni delitelj.

Indeks naključja (Wolfe Friedman, 1920):

Naj bo $x = x_1 x_2 \ldots x_n$ zaporedje n črk. Indeks naključja zaporedja x, označen z $I_c(x)$, je verjetnost, da sta dva naključna elementa zaporedja x enaka. Naj bodo $f_0, f_1, \ldots, f_{25}$ frekvence črk $A, B, C, D, \ldots, Z$ v zaporedju x. Potem je

$$I_c(x) = \sum_{i=0}^{25} \frac{f_i(f_i - 1)}{n(n-1)}.$$

Če so p_i pričakovane verjetnosti angleških črk, potem je

$$I_c(x) \approx \sum_{i=0}^{25} p_i^2 = 0.065.$$

Za povsem naključno zaporedje velja

$$I_c(x) \approx 26(1/26)^2 = 1/26 = 0.038.$$

Ker sta števili .065 in .038 dovolj narazen, lahko s to metodo najdemo dolžino ključa (ali pa potrdimo dolžino, ki smo jo uganili s testom Kasiskega).

Naj bosta $x = x_1 x_2 \dots x_n$ in $y = y_1 y_2 \dots y_n$ zaporedji n in n' črk. Vzajemen indeks naključja zaporedij x in y, označen z $MI_c(x,y)$, je verjetnost, da je naključni element v x enak naključnemu elementu v y. Potem je

$$MI_c(x,y) = \sum_{i=0}^{25} \frac{f_i f_i'}{nn'}.$$

Po drugi strani pa je

$$MI_c(x,y) \approx \sum_{i=0}^{25} p_{h-k_i} p_{h-k_j} = \sum_{i=0}^{25} p_h p_{h-s},$$

kjer je s relativen zamik $(k_i - k_j)$.

Izkaže se, da je $MI_c(x,y) \approx 0.065$ za $s = 0$ in $MI_c(x,y) \in [0.031, 0.045]$ za $s \neq 0$.

S pomočjo računalnika izračunamo 260 vrednosti $MI_c(y_i, y_j^s)$, kjer je $1 \leq i < j \leq 5$ in $0 \leq s \leq 25$, ter dobimo sistem enačb za $k_1, \ldots, k_m$.

Do naslednjega tedna

- preberite 1. poglavje (str. 1-39)

- rešite čim več nalog (str. 39-43)

Stinsonovega učbenika
Cryptography, Theory and Practice.