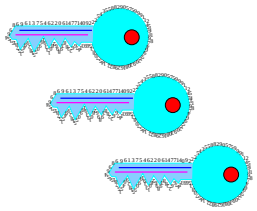
KRIPTOGRAFIJA IN RAČUNALNIŠKA VARNOST

Aleksandar Jurišić

IMFM

http://valjhun.fmf.uni-lj.si/~ajurisic



Uvod

1. poglavje

Klasična kriptografija

2. poglavje

Shannonova teorija:

3. poglavje

Simetrični kriptosistemi

4. poglavje

RSA kriptosistem in faktorizacija

5. poglavje

Drugi javni kriptosistemi

6. poglavje

Sheme za digitalne podpise

7. poglavje

Zgoščevalne funkcije

8. poglavje

Distribucija ključev

in uskladitev ključev

9. poglavje

Identifikacijske sheme

10. poglavje

Kode za overjanje

11. poglavje

Sheme za deljenje skrivnosti

12. poglavje

Generator psevdo-naključnih števil

13. poglavje

Dokazi brez razkritja znanja

Priloga A

Dokaz izreka o gostoti praštev

Uvod

Odkar so ljudje pričeli komunicirati,
pa naj si bo to preko govora, pisave, radija,
telefona, televizije ali računalnikov,
so želeli tudi *skrivati* vsebino svojih sporočil.

Ta nuja, oziroma že kar obsedenost
po *tajnosti*, je imela dramatičen vpliv
na vojne, monarhije in seveda tudi
na individualna življenja.

Vladarji in generali so odvisni od uspešne
in učinkovite komunikacije že tisočletja,
hkrati pa se zavedajo posledic, v primeru,
če njihova sporočila pridejo v napačne roke,
izdajo dragocene skrivnosti rivalom ali
odkrijejo vitalne informacije nasprotnikom.

Danes vse to velja tudi za moderna vodstva
uspešnih podjetij in tako postaja

"informacijska/računalniška varnost"

eno izmed najbolj pomembnih gesel
informacijske dobe.

Vlade, industrija ter posamezniki,
vsi hranijo informacije v *digitalni obliki*.

Ta medij nam omogoča številne prednosti
pred fizičnimi oblikami:

- je zelo kompakten,
- prenos je takorekoč trenuten,

hkrati pa je omogočen tudi

- organiziran dostop do raznovrstnih podatkovnih baz.

Z razvojem

- telekomunikacij,
- računalniških omrežij in
- obdelovanja informacij

pa je precej lažje prestreči in spremeniti
digitalno (elektronsko) informacijo kot pa
njenega *papirnega predhodnika*.

Zato so se povečale zahteve po **varnosti**.

Informacijska in računalniška varnost

opisuje vse preventivne postopke in sredstva
s katerimi preprečimo nepooblaščeno uporabo
digitalnih podatkov ali sistemov,
ne glede na to ali gre pri ustreznih podatkih kot sta

digitalni denar (nosilec vrednosti) in
digitalni podpis (za prepoznavanje)

za

- razkritje,
- spreminjanje,
- zamenjavo,
- uničenje,
- preverjanje verodostojnosti.

Predlagani so bili številni ukrepi, a niti eden med njimi
ne zagotavlja *popolne varnosti*.

Med preventivnimi ukrepi, ki so na voljo danes, nudi

kriptografija

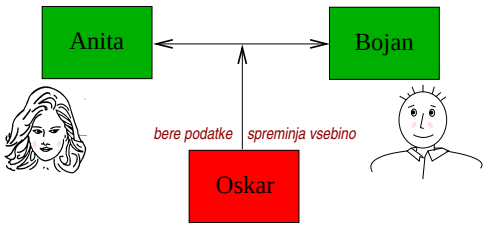
(če je seveda pravilno implementirana ter uporabljana)

največjo stopnjo varnosti

glede na svojo prilagodljivost digitalnim medijem.

Kaj je kriptografija?

Kriptografija je veda o komunikaciji v prisotnosti
aktivnega napadalca.



Primer:

pošiljanje papirnih dokumentov po p

Kakšna zagotovila varnosti so na voljo? In kak

- **Fizična varnost:** zapečatene kuverte.
- **Zakonska infrastruktura:**
  ročni podpis je zakonsko sprejeto sredstvo,
  zakoni proti odpiranju/oviranju pošte, itd.
- **Poštna infrastruktura:**
  varni in sprejeti mehanizmi za
  dostavljanje pošte širom po svetu.

Primer: digitalni podatki

- **ZA**: hranjenje je enostavno in poceni, hiter in enostaven transport.
- **PROTI**: enostavno kopiranje; transportni mediji niso varni (npr. pogovor po mobilnem telefonu, internetna seja, ftp seja, komunikacija s pomočjo elektronske pošte).
- **Vprašanje**: Kako lahko omogočimo/ponudimo enake možnosti za papirni kakor tudi digitalni svet?

Odšifriranje (razbijanje) klasičnih šifer

Kriptografske sisteme kontroliramo s pomočjo ključev, ki določijo transformacijo podatkov.
Seveda imajo tudi ključi digitalno obliko (binarno zaporedje: 01001101010101...).

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

pozna kriptosistem oziroma algoritme, ki jih uporabljamo, ne pa tudi ključe, ki nam zagotavljajo varnost.

Vohunova dilema

Bilo je temno kot v rogu, ko se je vohun vračal v grad po opravljeni diverziji v sovražnem taboru.

Ko se je približal vratom, je zaslišal šepetajoč glas:

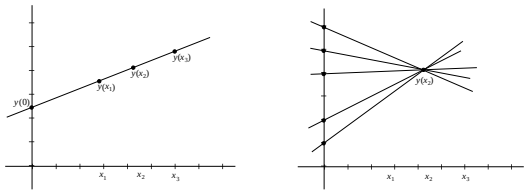
Kako vohun prepriča "stražarja", da pozna geslo, ne da bi ga izdal morebitnemu vsiljivcu/prisluškovalcu?

Deljenje skrivnosti

Problem: *V banki morajo trije direktorji odpreti trezor vsak dan, vendar pa ne želij zaupati kombinacijo nobenemu posamezni Zato bi radi imeli sistem, po katerem lahk odpreta trezor poljubna dva med njimi.*

Ta problem lahko rešimo z $(2,3)$-stopenjsko sh

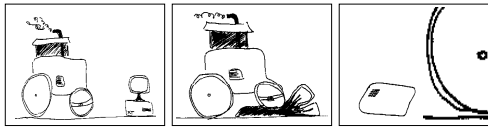
Stopenjske sheme za deljenje skrivnosti sta le neodvisno odkrila **Blakey in Shamir**.



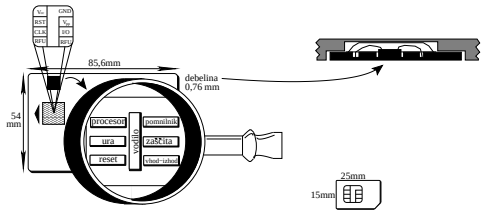
Vsak dobi le y-koordinato svoje točke.
Program v trezorju ima še ustrezne od 0 različne x- koordinate, zato lahko izračuna ključ $y(0)$.
Vsaki točki natanko določata premico in s tem ključ.

Če imamo eno samo točko, ne moremo ugotoviti, kateri ključ je pravi, saj so vsi videti enako dobri.

Pametne kartice



Po računski moči so pametne kartice primerljive z originalnim IBM-XT računalnikom,
kartice s **kripto koprocesorjem** pa v nekaterih opravilih prekašajo celo 50 Mhz 486 računalnik.



Velikost pametne kartice ustreza ISO 7810 standardu, sestavljajo pa jo mikroprocesor, pomnilnik (ROM, RAM, EEPROM), vhodno/izhodna enota (I/O).

Zakaj pametna kartica

Gotovo je najbolj pomembna razlika med p kartico in magnetno kartico

varnost.

Pametna kartica ima svoj **procesor**, ki kontro vse interakcije med od zunaj nedostopnim spo in različnimi zunanjimi enotami.

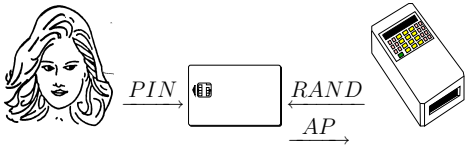
Dodaten, pomemben, del pametne kartice je **non-volatile spomin (ROM)**,
t.j. spomin, ki se ga ne da spremeniti in ostane prisoten tudi po prekinitvi napajanja.

Zagotovitev varnosti

Identifikacija se opravi v dveh delih:

(a) kartica mora biti zares prepričana, da jo uporablja njen lastnik (lokalno overjanje),

(b) kartica komunicira (varno) z računalnikom (dinamično overjanje).

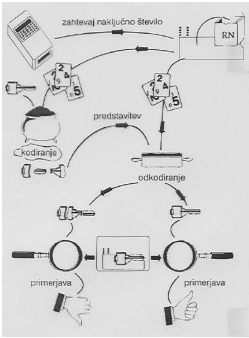


Biometrični testi

Pametna kartica zgenerira naključno število, ter ga pošlje čitalniku.

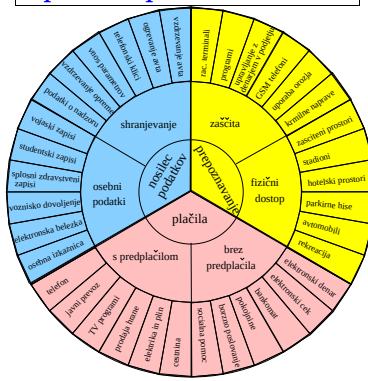
Ta ga zašifrira z zasebnim ključem in rezultat pošlje pametni kartici.

Če pametna kartica uspešno odšifrira naključno število z javnim ključem, potem je prepričana o pristnosti čitalnika.

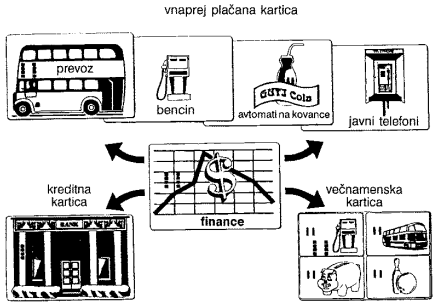


Enak proces poteka v nasprotni smeri.

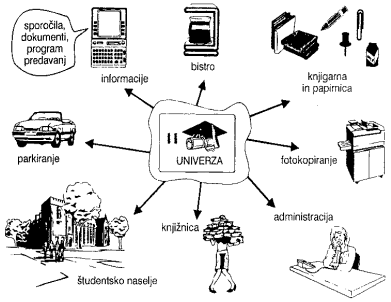
Uporaba pametnih kartic



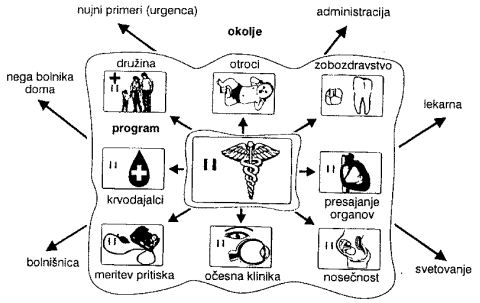
Plačilne, kreditne in večnamenske kartice, ki se uporabljajo na področju *financ*.



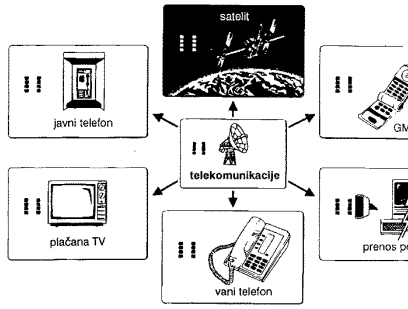
Uporaba pametnih kartic na *univerzi/fakulteti*, ki je ponekod mesto v malem.



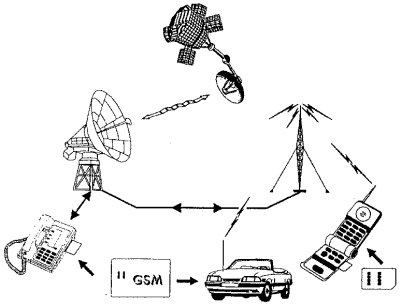
Področja v *zdravstvu*, kjer se uporabljajo pametne kartice.



Uporaba pametne kartice v *telekomunikac* uporabniški elektrotehniki.



GSM (globalni sistem za prenosno komuniciranje)



Javna kriptografija

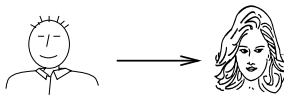
Glede na pomembnost podatkov, ki jih varujemo, se moramo odločiti za ustrezno obliko zaščite:

- Geslo (PIN) in zgoščevalne funkcije predstavljajo osnovno zaščito,
- AES (Advanced Encryption Standard) simetrični kriptosistemi nudijo srednji nivo,
- javna kriptografija (Public Key Scheme) pa visok nivo zaščite.

Odlična uvodna knjiga o moderni kriptografiji je: Albrecht Beutelspacher, **Cryptology**, MAA, 1994.

Koncept javne kriptografije

Bojan pošlje Aniti pismo, pri tem pa si želi, da bi pismo lahko prebrala le ona (in prav nihče drug) **[zaščita]**.



Anita pa si poleg tega želi biti prepričana, da je pismo, ki ga je poslal Bojan prišlo prav od njega **[podpis]**.

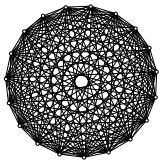
Predpostavimo, da se Anita in Bojan prej dog
za **skupen ključ**, ki ga ne pozna nihče drug
(simetričen kriptosistem).

Če Bojan z njim zašifrira pismo, je lahko prep
da ga lahko odklene le Anita.

Hkrati pa je tudi Anita zadovoljna, saj je prep
da ji je pismo lahko poslal le Bojan.

Tak pristop je problematičen vsaj iz dveh razlogov:

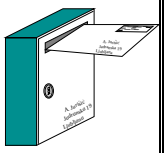
1. Anita in Bojan se morata *prej* dogovoriti za skupen ključ,
2. upravljanje s ključi v omrežju z n uporabniki je kradratne zahtevnosti $\binom{n}{2}$, vsak uporabnik pa mora hraniti $n-1$ ključev.



Leta 1976 sta Whit **Diffie** in Martin **Hellman** predstavila koncept kriptografije z javnimi ključi.

Tu ima za razliko od sim. sistema vsak uporabnik **dva** ključa, podatke *zaklepa*, drugi pa jih *odklepa*.

Pomembna lastnost tega sistema:
ključ, ki zaklepa, ne more odklepati
in obratno,
ključ, ki odklepa, ne more zaklepati.



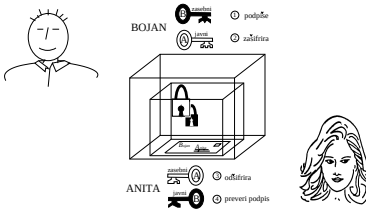
To omogoči lastniku, da en ključ objavi, drugega pa hrani v tajnosti (npr. na pametni kartici). Zato imenujemo ta ključa zaporedoma **javni** in **zasebni**.

Ta pristop omogoča veliko presenetljivih načinov uporabe, npr. omogoča ljudem varno komuniciranje, ne da bi se predhodno srečali zaradi izmenjave/dogovora o tajnem ključu.

Vsak uporabnik najprej objavi svoj javni ključ, zasebnega pa zadrži zase. Vsak lahko nato z javnim ključem zašifrira pismo, bral (odšifriral) pa ga bo lahko le lastnik ustreznega zasebnega ključa.

Bojan pošlje Aniti podpisano zasebno pismo:

(1) podpiše ga s svojim zasebnim ključem Z_B
(2) zašifrira z Anitinim javnim ključem J_A.



(3) Anita ga s svojim zasebnim ključem Z_A od
(4) z Bojanovim javnim ključem J_B preveri po

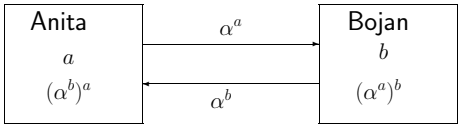
V razvoju javne kriptografije je bilo predlaganih in razbitih veliko kriptosistemov.

Le nekaj se jih je obdržalo in jih lahko danes smatramo za varne in učinkovite.

Glede na matematični problem na katerem temeljijo, so razdeljene v tri skupine:

- **Sistemi faktorizacije celih števil**
  npr. RSA (Rivest-Shamir-Adleman).
- **Sistemi diskretnega logaritma**
  npr. DSA.
- **Kripto sistemi z eliptičnimi krivuljami**
  (Elliptic Curve Cryptosystems)

Izmenjava ključev
(Diffie-Hellman)



Anita in Bojan si delita skupni element grupe: α^{ab}.

Končne grupe so zanimive zato, ker računanje potenc lahko opravimo učinkovito, ne poznamo pa vedno učinkovitih algoritmov za logaritem (za razliko od $\mathbb{R}$).

Kaj je kriptografija

- cilji kriptografije
- širši pogled na kriptografijo
- gradniki kriptografije

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

Cilje kriptografije bomo dosegali z matematičnimi sredstvi.

Cilji kriptografije

1. **Zasebnost/zaupnost/tajnost:**
   varovanje informacij pred tistimi, ki jim vp
   dovoljen, dosežemo s šifriranjem.
2. **Celovitost podatkov:**
   zagotovilo, da informacija ni bila sprem
   nedovoljenimi sredstvi (neavtoriziranimi sr

3. **Overjanje sporočila (ali izvora podatkov):**
   potrditev izvora informacij.
4. **Identifikacija:**
   potrditev identitete predmeta ali osebe.
5. **Preprečevanje tajenja:**
   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

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).

Širši pogled na kriptografijo – varnost informacij

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

(a) Varnost računalniškega sistema

tj. tehnična sredstva, ki omogočajo računalniškega sistema, ki lahko pomeni sa računalnik z več uporabniki, lokalno mrežo Internet, mrežni 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

(b) 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

Širši pogled na kriptografijo – varnost informacij

1. *varnost računalniškega sistema*
2. *varnost na mreži*

NAUK: Kriptografija je samo majhen del varnosti informacij.

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 kriminalce*

NAUK: Uporabna kriptografija je več kot samo zanimiva matematika.

1. poglavje

Klasična kriptografija

- zgodovina (hieroglifi, antika, II. svetovna vojna)
- zamenjalna šifra

Klasične šifre in razbijanje

- prikrita, zamenjalna (pomična, afina), bločna (Vigenerjeva, Hillova)
- Kerckhoffov princip in stopnje napadov
- napad na Vigenerja (Kasiski test, indeks naključja)
- napad na Hillovo šifro
- tokovne šifre

Zgodovina

Kriptografija ima dolgo in zanimivo zgodovino:

– Hieroglifi, Špartanci, Cezar, ...

D. Kahn, **The Codebreakers**
(The Story of Secret Writing),
hrvaški prevod: (K. and M. Miles),
Šifranti protiv špijuna,
Centar za informacije i Publicitet, Zagreb 1979.
(429+288+451+325=1493 strani).

Hieroglifi

Razvili so jih antični Egipčani.
Komunicirali so v jeziku
sestavljenemu iz sličič namesto besed.

Najbolj izobraženi ljudje so jih razumeli,
toda v religioznemu kontekstu
– npr. napisi na grobovih –
so njihovi duhovniki uporabljali tajne
kriptografske verzije znakov, da bi bila vsebina
več vredna (saj je šlo za božje besede) in bolj mistična.

Mnoge religije so uporabljale tajne znake,
ki so jih razumeli le določeni izbranci.

Razbijalci šifer

Obstajajo od kar poznamo šifriranje.

L. 1799 so v Egipčanski Rosetti
našli skoraj 2.000 let star kamen.
Na njemu so bili trije teksti:

- hieroglifi,
- pisava egiptčanov (demotic) in
- starogrščina.

Ko je bil končan prevod iz Grščine,
je bilo možno razvozlati tudi hieroglife,
iz katerih smo izvedeli o zgodovini antičnega E

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 bilo 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 **prikrite šifre**,
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 odšifriranja.

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 olajšan
Sir John Trevanion, njemu pa je vsekakor
pobegniti, sicer bi ga v gradu Colcester gotovo u
prav tako, kot so Sir Charlesa Lucasa ter Sir
Lislea.

Druga svetovna vojna

- Enigma (Nemčija),
- Tunny (Nemčija),
- Purple (Japonska),
- Hagelin (ZDA).

Zamenjalna šifra

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! = 403291461126656355840000000
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 pred 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: `03CWC%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- 03CWC%J(- 4-DC
!CW-?C NJDJ D-?+- 03CWC%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:

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

(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 (
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 in naloga je bila rešena (poskusite sami!).

Podobna naloga je v angleščini dosti lažja, s
tem jeziku veliko členov THE, A in AN, ver
zato običajno najprej izpustimo presledke iz te
ga želimo spraviti v tajnopis.

V angleščini imajo seveda črke drugačno gostot
slovenščini.

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.