Informacija in kodi

UN2-1-AV 2024/2025

Tajno kodiranje I

Simon Dobrišek

november 2024

Teme predavanja

Uvod

Vigenerjev kriptografski sistem

Popolna tajnost kriptografskega sistema

Vernamov kriptografski sistem

Odpornost kriptografskega sistema

Sodobni kriptografski sistemi

Sodobni kriptografski sistemi s tajnim ključem

Uvod

Tajno kodiranje - *tajnopisje, kriptografijo* - so poznali že stari Egipčani 2 000 let pred našim štetjem, pa tudi stari Grki in stari Rimljani.

Že takrat so ljudje želeli prikrivati svoja sporočila pred tistimi ljudmi, ki jim njihova sporočila niso bila namenjena. Namen tajnega kodiranja se do danes ni spremenil.

S tajnim kodiranjem dano sporočilo M, navadno je to niz znakov iz dane končne množice znakov Θ, prikrijemo z nekim *šifrirnim* postopkom E.

Uvod

Za ta šifrirni postopek mora obstajati tudi dešifrirni postopek D, s katerim lahko razkrijemo prvotno prikrito sporočilo - kriptogram - $E(M)$. To je

$$D\big(E(M)\big) = M.$$

V preteklosti sta bila šifrirni in dešifrirni postopek tajna, ker je bilo le tako možno preprečiti, da bi ljudje, pred katerimi je bilo potrebno skriti pomen sporočila, prikrito sporočilo razkrili.

Uporaba takih postopkov je bila omejena na manjše skupine uporabnikov, saj bi ob izstopu enega uporabnika iz skupine vsi preostali uporabniki morali preiti na uporabo drugega postopka.

Uvod

Moderna kriptografija rešuje ta problem s pomočjo *ključev*.

Ključ, označimo ga s K, je parameter funkcije ali algoritma šifriranja oziroma dešifriranja.

Sedaj sta algoritma šifriranja in dešifriranja lahko znana tudi izven skupine ljudi, ki se tajno sporazumevajo, pod pogojem, da izven skupine nihče ne pozna ključa.

Prikriti niz znakov, pravimo mu tudi *tajnopis* ali *kriptogram*, zapišemo sedaj kot

$$C = E(M, \mathrm{K}).$$

Uvod

Seveda mora tudi sedaj dešifrirni postopek $D(C, \mathrm{K})$, ki razkrije tajnopis, izpolnjevati pogoj

$$D\big(E(M, \mathrm{K}), \mathrm{K}\big) = M.$$

Zato lahko sistem, ki omogoča tajno komunikacijo - *kriptografski sistem*, zapišemo s trojico $(\mathrm{M}, \mathrm{K}, \mathrm{C})$, kjer je:

- **M** množica odprtih sporočil,
- **C** množica prikritih sporočil - kriptogramov in
- **K** končna množica parametrov algoritmov šifriranja in dešifriranja - ključev,

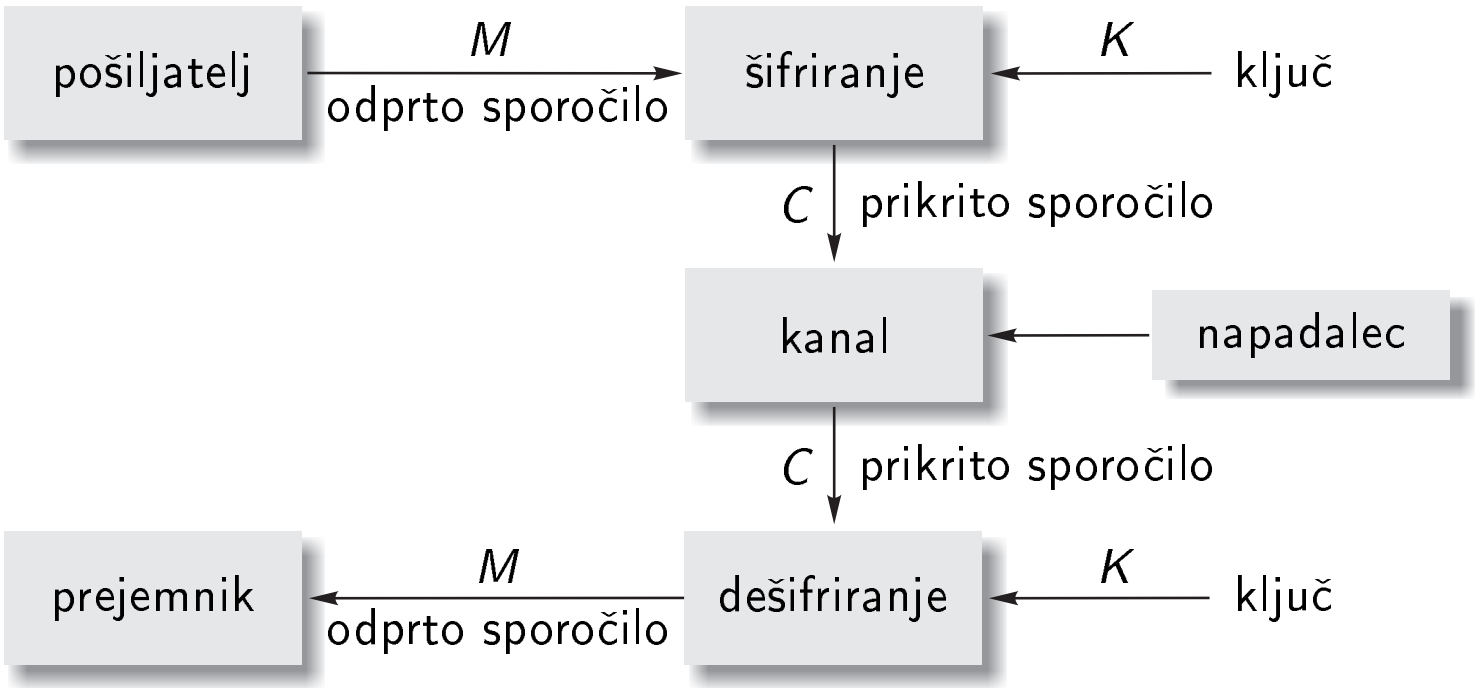
pod predpostavkami, da obstajata takšni funkciji oziroma algoritma

$$E : \mathrm{M} \times \mathrm{K} \to \mathrm{C} \qquad \text{in} \qquad D : \mathrm{C} \times \mathrm{K} \to \mathrm{M},$$

da za vsak $(M, \mathrm{K}) \in \mathrm{M} \times \mathrm{K}$ velja

$$D\big(E(M, \mathrm{K}), \mathrm{K}\big) = M.$$

Uvod



Slika: Zasnova kriptografskega sistema s ključem.

Uvod

Poskus razkrivanja tajnopisa s strani oseb, ki jim sporočilo ni namenjeno, imenujemo *napad* na kriptografski sistem ali *kriptoanaliza*.

Napadalec na kriptografski sistem navadno pozna:

- šifrirni in dešifrirni algoritem,
- pogosto tudi jezik sporočila in statistične lastnosti tega jezika,
- temo sporočila in podobno ter seveda tudi
- tajnopis C, ki ga želi razkriti, ali celo
- tajnopis in poljubno število parov prejšnjih tajnopisov in razkritih sporočil (M, C).

Da bi razkril tajnopis, potrebuje le ključ.

V nadaljevanju bomo najprej opisali nakaj (starejših) kriptografskih sistemov s ključem.

Zamenjava črk

Množica znakov Θ bo v vseh primerih abeceda črk slovenskega jezika, ki ji bomo ponekod dodali še presledek (prazen znak).

Ključ K postopka je določena permutacija π črk abecede Θ. Funkcija šifriranja zamenja vsako črko sporočila M z njeno sliko v permutaciji π.

Primer 6.1

Na primer, če sta:

$$\Theta = \{A,B,C,Č,D,E,F,G,H,I,J,K,L,M,N,O,P,R,S,Š,T,U,V,Z,Ž\},$$

$$\mathrm{K} = \mathrm{LMRŠBTDEAFGHINPJKSOŽCUVZČ},$$

šifriramo sporočilo *JAN JE LEP* tako, da najprej izpustimo presledke med besedami, nato pa niz črk slovenske abecede *M=JANJELEP* preslikamo v tajnopis *C=GLPGTITK*. □

Zamenjava črk

Tako zakrito sporočilo napadalci skušajo odkriti z ugotavljanjem pogostnosti znakov v tajnopisu.

Ob predpostavki, da je znan jezik, v katerem je napisano sporočilo, zamenjajo znake v tajnopisu z enako pogostimi znaki v jeziku sporočila. Vendar se kratka sporočila oziroma tajnopisi tako ne dajo odkriti.

Lahko bi poskusili ustvariti vse variacije 8 črk izmed 25 črk slovenske abecede (za primer 6.1), vendar bi za to, z računalnikom, ki zmore 320.000 MIPS-ov (Intel Core i7 6950X), potrebovali približno teden dni (za vse variacije 10 črk pa približno 10 let, vendar le 1000 sekund na super-računalniku Sunway MPP s 100 PFLOPS).

Vigenerjev kriptografski sistem

Vigenerjev sistem temelji na seštevanju znakov sporočila in znakov ključa, ki je sestavljen iz d znakov.

Te ponavljamo pod znaki sporočila in jih seštevamo po modulu θ, kjer je θ moč množice Θ.

Vigenerjev kriptografski sistem

Za slovensko abecedo, ki ji dodamo še presledek, *šifrirni postopek* zapišemo kot[1]

$$E : c_i \equiv (m_i + k_i) \mod 26,$$

kjer je:

- m_i kod i-tega znaka (odprtega) sporočila iz Θ,
- k_i kod i-tega znaka v naključnem nizu znakov iz Θ, to je ključa,
- c_i kod i-tega znaka tajnopisa,

Dešifrirni postopek je določen kot

$$D : m_i \equiv (c_i - k_i) \mod 26.$$

[1] Izraz $u \equiv v \mod m$, ki ga beremo: *števili u in v sta kongruentni po modulu m*, pove, da je razlika $u - v$ deljiva s številom m.

Vigenerjev kriptografski sistem - primer

Primer 6.2

Vzemimo, da črke slovenske abecede, ki jim dodamo še presledek, kodiramo s pomočjo spodnje razpredelnice:

A	B	C	Č	D	E	F	G	H	I	J	K	L
0	1	2	3	4	5	6	7	8	9	10	11	12
M	N	O	P	R	S	Š	T	U	V	Z	Ž	
13	14	15	16	17	18	19	20	21	22	23	24	25

Vzemimo, da želimo šifrirati sporočilo *JAN JE LEP*. Vzemimo tudi, da je $d = 2$ in ključ *LA*.

Razpredelnica ponazarja Vigenerjevo šifriranje danega sporočila:

M:	J	A	N		J	E		L	E	P
K:	L	A	L	A	L	A	L	A	L	A
C:	V	A	A		V	E	K	L	R	P

□

Cezarjev kriptografski sistem

Ko je $d = 1$, imenujemo Vigenerjev kriptografski sistem po Juliju Cezarju, ki je ta kriptografski sistem uporabljal za prikrivanje svojih sporočil.

Popolna tajnost kriptografskega sistema

Vzemimo, da je dan kriptografski sistem $(\mathrm{M}, \mathrm{K}, \mathrm{C})$ ter da sta množici sporočil $\mathbf{M}$ in ključev $\mathbf{K}$ končni.

Sedaj lahko predpostavimo, da je

1. $P(M_i)$ verjetnost, da odpošljemo sporočilo M_i $(i = 1, 2, ..., \mu)$ in $\sum_i P(M_i) = 1$, ter
2. $P(\mathrm{K}_j)$ verjetnost, da uporabimo za šifriranje ključ K_j $(j = 1, ..., \kappa)$ in $\sum_j P(\mathrm{K}_j) = 1$. Predpostavimo tudi, da izbor ključa za šifriranje ni odvisen od sporočila, ki ga pošiljamo.

Ti dve porazdelitvi verjetnosti določata porazdelitev verjetnosti nad množico tajnopisov $\mathbf{C}$.

Popolna tajnost kriptografskega sistema

Verjetnost tajnopisa C_k $(k = 1, 2, ..., \zeta)$ je enaka

$$P(C_k) = \sum P(M_i)P(\mathrm{K}_j), \tag{1}$$

kjer poteka seštevanje preko vseh dvojic (M_i, K_j), za katere velja $D(C_k, \mathrm{K}_j) = M_i$.

Popolna tajnost kriptografskega sistema

Porazdelitve $P(M_i)$, $P(\mathrm{K}_j)$ in $P(C_k)$ pa določajo:

- pogojni porazdelitvi nad množico $\mathrm{M} \times \mathrm{C}$:

$$P(C_k \mid M_i) = \sum P(\mathrm{K}_j), \tag{2}$$

kjer seštevamo po vseh ključih K_j, za katere velja $D(C_k, \mathrm{K}_j) = M_i$, ter[2]

$$P(M_i \mid C_k) = \frac{P(M_i)P(C_k \mid M_i)}{P(C_k)} \text{ in} \tag{3}$$

[2]Sledi iz Bayesove formule $P(A \mid B) = \frac{P(A)P(B|A)}{P(B)}$.

Popolna tajnost kriptografskega sistema

- pogojni porazdelitvi nad množico $(\mathrm{K} \times \mathrm{C})$:

$$P(C_k \mid \mathrm{K}_j) = \sum P(M_i), \tag{4}$$

kjer seštevamo po vseh sporočilih M_i, za katere velja $D(C_k, \mathrm{K}_j) = M_i$, ter

$$P(\mathrm{K}_j \mid C_k) = \frac{P(\mathrm{K}_j)P(C_k \mid \mathrm{K}_j)}{P(C_k)}. \tag{5}$$

Če z M, K in C označimo naključne spremenljivke z zalogami vrednosti M, K in C, lahko iz danih porazdelitev določimo entropije sporočil, ključev in tajnopisov, kakor tudi vezane in pogojne entropije sporočil, ključev in tajnopisov.

Spremenljivki M in K sta neodvisni, spremenljivka C pa je funkcija M in K.

Popolna tajnost kriptografskega sistema

DEFINICIJA 6.1 *Kriptografski sistem* $(\mathrm{M}, \mathrm{K}, \mathrm{C})$ *zagotavlja popolno tajnost, oziroma nezlomljivost, če je pogojna verjetnost, da je oddano sporočilo* M_i*, ko je sprejet tajnopis* C_k*, enaka apriorni verjetnosti tega sporočila. To je, če je*

$$P(M_i \mid C_k) = P(M_i) > 0 \tag{6}$$

za vsak $C_k \in \mathrm{C}$ *in vsak* $M_i \in \mathrm{M}$*.*

POSLEDICA 6.1 *Kriptografski sistem* $(\mathrm{M}, \mathrm{K}, \mathrm{C})$ *zagotavlja popolno tajnost tedaj in le tedaj, če je pogojna verjetnost, da je pri šifriranju bil uporabljen ključ* K_j*, ko je sprejet tajnopis* C_k*, enaka apriorni verjetnosti ključa* K_j*. To je, če je*

$$P(\mathrm{K}_j \mid C_k) = P(\mathrm{K}_j) > 0 \tag{7}$$

za vsak $\mathrm{K}_j \in \mathrm{K}$ *in* $C_k \in \mathrm{C}$*.*

Popolna tajnost kriptografskega sistema - primer

Primer 6.3

Vzemimo, da je dan kriptografski sistem $(\mathrm{M}, \mathrm{K}, \mathrm{C})$, kjer so: $\mathrm{M} = \{M_1, M_2\}$, $\mathrm{K} = \{\mathrm{K}_1, \mathrm{K}_2, \mathrm{K}_3\}$ in $\mathrm{C} = \{C_1, C_2, C_3, C_4\}$.

Vzemimo, da sta verjetnosti, da odpošljemo sporočili M_1 in M_2 enaki $P(M_1) = 1/4$ in $P(M_2) = 3/4$, verjetnosti, da za šifriranje uporabimo ključe K_1, K_2 in K_3 pa $P(\mathrm{K}_1) = 1/2$, $P(\mathrm{K}_2) = 1/4$ in $P(\mathrm{K}_3) = 1/4$.

Predpostavimo, da smo tajnopise ustvarili s pomočjo naslednjih parov (M_i, K_j): $C_1 = E(M_1, \mathrm{K}_1)$, $C_2 = E(M_2, \mathrm{K}_1)$, $C_2 = E(M_1, \mathrm{K}_2)$, $C_3 = E(M_2, \mathrm{K}_2)$, $C_3 = E(M_1, \mathrm{K}_3)$ in $C_4 = E(M_2, \mathrm{K}_3)$.

Popolna tajnost kriptografskega sistema - primer

Iz (1) – (5) lahko izračunamo:

- verjetnosti tajnopisov:

$$\begin{aligned} P(C_1) &= 1/8, \quad P(C_2) = P(M_2)\cdot P(\mathrm{K}_1) + P(M_1)\cdot P(\mathrm{K}_2) = 7/16, \\ P(C_3) &= P(M_2)\cdot P(\mathrm{K}_2) + P(M_1)\cdot P(\mathrm{K}_3) = 1/4, \quad P(C_4) = 3/16, \end{aligned}$$

- pogojni porazdelitvi nad množico $\{M_1, M_2\} \times \{C_1, C_2, C_3, C_4\}$:

$$\begin{aligned} P(C_1 \mid M_1) &= 1/2, & P(C_1 \mid M_2) &= 0, \\ P(C_2 \mid M_1) &= 1/4, & P(C_2 \mid M_2) &= 1/2, \\ P(C_3 \mid M_1) &= 1/4, & P(C_3 \mid M_2) &= 1/4, \\ P(C_4 \mid M_1) &= 0, & P(C_4 \mid M_2) &= 1/4, \end{aligned}$$

$$\begin{aligned} P(M_1 \mid C_1) &= 1, & P(M_2 \mid C_1) &= 0, \\ P(M_1 \mid C_2) &= 1/7, & P(M_2 \mid C_2) &= 6/7, \\ P(M_1 \mid C_3) &= 1/4, & P(M_2 \mid C_3) &= 3/4, \\ P(M_1 \mid C_4) &= 0, & P(M_2 \mid C_4) &= 1, \end{aligned}$$

Popolna tajnost kriptografskega sistema - primer

- pogojni porazdelitvi nad množico $\{\mathrm{K}_1, \mathrm{K}_2, \mathrm{K}_3\} \times \{C_1, C_2, C_3, C_4\}$:

$$
\begin{array}{lll}
P(C_1 \mid \mathrm{K}_1) = 1/4, & P(C_1 \mid \mathrm{K}_2) = 0, & P(C_1 \mid \mathrm{K}_3) = 0, \\
P(C_2 \mid \mathrm{K}_1) = 3/4, & P(C_2 \mid \mathrm{K}_2) = 1/4, & P(C_2 \mid \mathrm{K}_3) = 0, \\
P(C_3 \mid \mathrm{K}_1) = 0, & P(C_3 \mid \mathrm{K}_2) = 3/4, & P(C_3 \mid \mathrm{K}_3) = 1/4, \\
P(C_4 \mid \mathrm{K}_1) = 0, & P(C_4 \mid \mathrm{K}_2) = 0, & P(C_4 \mid \mathrm{K}_3) = 3/4,
\end{array}
$$

$$
\begin{array}{lll}
P(\mathrm{K}_1 \mid C_1) = 1, & P(\mathrm{K}_2 \mid C_1) = 0, & P(\mathrm{K}_3 \mid C_1) = 0, \\
P(\mathrm{K}_1 \mid C_2) = 6/7, & P(\mathrm{K}_2 \mid C_2) = 1/7, & P(\mathrm{K}_3 \mid C_2) = 0, \\
P(\mathrm{K}_1 \mid C_3) = 0, & P(\mathrm{K}_2 \mid C_3) = 3/4, & P(\mathrm{K}_3 \mid C_3) = 1/4, \\
P(\mathrm{K}_1 \mid C_4) = 0, & P(\mathrm{K}_2 \mid C_4) = 0, & P(\mathrm{K}_3 \mid C_4) = 1.
\end{array}
$$

Opazimo, da dan kriptografski sistem ne zagotavlja popolne tajnosti. □

Nezlomljiv kriptografski sistem

Naslednja izreka 6.1 in 6.2 definirata pogoje, ki jim mora zadostiti *nezlomljiv* kriptografski sistem.

IZREK 6.1 *Potreben pogoj, da zagotavlja kriptografski sistem* $(\mathrm{M}, \mathrm{K}, \mathrm{C})$ *popolno tajnost, je, da razpolaga z vsaj toliko ključi, kolikor je sporočil.*

Dokaz: Vzemimo, da imamo μ sporočil in en sam ključ K. S pomočjo ključa K (v danem vrstnem redu) bijektivno (povratno enolično) preslikamo sporočila $M_1, ..., M_\mu$ v tajnopise $C_1, ..., C_\mu$. Ker sistem zagotavlja popolno tajnost, za vsako sporočilo M_l $(l = 1, 2, ..., \mu)$ velja

$$P(M_l \mid C_k) = P(M_l) > 0.$$

Sledi, obstajati mora vsaj μ ključev, zato da dobimo pozitivne verjetnosti za vsak $l = 1, 2, ..., \mu$. ■

Nezlomljiv kriptografski sistem

IZREK 6.2 *Kriptografski sistem* $(\mathrm{M}, \mathrm{K}, \mathrm{C})$ *zagotavlja popolno tajnost, če razpolaga z ravno toliko ključi, kolikor je sporočil in imajo vsi ključi enako verjetnost, da so uporabljeni za šifriranje sporočil.*

Dokaz: Izberimo nek $C_k \in \mathrm{C}$ in označimo ključe $\mathrm{K}_1, \mathrm{K}_2, ..., \mathrm{K}_\kappa$, tako da lahko pišemo $E(M_i, \mathrm{K}_i) = C_k$ za $i = 1, 2, ..., \kappa$. Zato lahko

$$P(M_i \mid C_k) = \frac{P(M_i)P(C_k \mid M_i)}{P(C_k)}$$

pišemo kot

$$P(M_i \mid C_k) = \frac{P(M_i)P(\mathrm{K}_i)}{P(C_k)}.$$

Nezlomljiv kriptografski sistem

Za kriptografski sistem, ki zagotavlja popolno tajnost velja (6), zato

$$P(\mathrm{K}_i) = P(C_k), \qquad i = 1, ..., \kappa.$$

Vsi ključi K_i imajo enako verjetnost $(= P(C_k))$, da jih uporabimo za šifriranje sporočil, in ker je število vseh ključev κ, je

$$P(\mathrm{K}_i) = \frac{1}{\kappa}, \qquad i = 1, ..., \kappa.$$

■

Vernamov kriptografski sistem

Vernamov sistem, podobno kot Vigenerjev, temelji na seštevanju znakov sporočila in znakov naključnega ključa, ki pa mora biti vsaj tako dolg kot sporočilo.

Pri tem je vsak znak v nizu znakov, ki predstavljajo ključ, neodvisno in z enako verjetnostjo 1/26 (za slovensko abecedo, ki ji je dodan presledek) izbran iz abecede Θ.

Slabost Vernamovega sistema je v tem, da ni primeren za prenos sporočil velikih dolžin, saj zahteva enako dolžino ključa, ki ga je potrebno po zaupni poti dostaviti osebi, s katero se želimo tajno sporazumevati.

Vernamov kriptografski sistem - primer

Primer 6.4

Vzemimo, da želimo šifrirati sporočilo, dolgo 10 znakov, *JAN JE LEP*.

Tudi ključ mora biti dolg 10 znakov. Vzemimo, da je ključ, ki sicer ni naključen, *TRALALALAA*.

Črke slovenske abecede kodiramo v desetiškem številskem sistemu s pomočjo razpredelnice iz primera 6.2.

Razpredelnica ponazarja Vernamovo šifriranje danega sporočila:

M:	J	A	N		J	E		L	E	P
K:	T	R	A	L	A	L	A	L	A	A
C:	D	R	N	K	J	R		Ž	E	P

Vernamov kriptografski sistem

Vidimo, da ima v Vernamovem kriptografskem sistemu množica ključev moč $\kappa = \theta^n$, kjer je n dolžina ključa (in sporočila), θ pa število znakov abecede. Ker so vsi ključi enako verjetni, je

$$P(\mathrm{K}) = \frac{1}{\theta^n}. \tag{8}$$

Ker je ključ niz n neodvisnih znakov iz Θ, je tudi množica tajnopisov $\mathbf{C}$ sestavljena iz θ^n enakoverjetnih tajnopisov, zato velja

$$P(\mathrm{K} \mid C) = \frac{1}{\theta^n} \tag{9}$$

za vsak $\mathrm{K} \in \mathrm{K}$.

Iz (7) ter (8) in (9) sledi, da zagotavlja Vernamov kriptografski sistem popolno tajnost, to je

$$P(K \mid C) = P(K) = 1/\theta^n.$$

Odpornost kriptografskega sistema

Pri napadih na kriptografske sisteme, ki ne zagotavljajo popolne tajnosti, je pomembno, da kriptoanalitik razpolaga z dovolj dolgim tajnopisom.

Odpornost kriptografskega sistema je podana z dolžino tajnopisa C, ki zagotavlja, da je obravnavani tajnopis zgrajen z določenim parom sporočilo - ključ (M, K).

Odpornost kriptografskega sistema - primer

Primer 6.5

Vzemimo, da imamo kriptografski sistem, ki temelji na Cezarjevih šifrah.

Nadalje vzemimo, da kodiramo črke slovenske abecede in presledek, tako kot to ponazarja razpredelnica v primeru 6.2.

Vzemimo, da razpolaga kriptoanalitik s tajnopisom

$$LCPBLGBRGP.$$

Ključ K_0 lahko poiščemo tako, da najprej naključno izberemo eno črko iz Θ in tajnopis dešifriramo z odštevanjem koda ključa od kodov tajnopisa po modulu 26.

Če nismo dobili smiselnega besedila, nadaljujemo z drugo črko iz Θ in tako naprej, dokler ne dobimo besedila, ki nas pomensko zadovolji.

Odpornost kriptografskega sistema

Vzemimo, da izberemo najprej $\mathrm{K}_0 = D$. Prvih nekaj znakov kriptograma nam da *NDSČ...*, zato nadaljujemo z naključno izbiro ključev.

Vzemimo, da sedaj izberemo $\mathrm{K}_0 = N$. Dobimo *APDO...*.

Ker je tudi to nesmiselno, izberemo $\mathrm{K}_0 = C$. Dobimo smiselno sporočilo *JAN JE LEP*.

Vendar v tem trenutku še ne vemo, da je to edini ključ, ki omogoča dešifriranje danega tajnopisa v smiselno sporočilo.

Z nadaljevanjem postopka naključne izbire ključev iz množice Θ bi se tudi v to hitro prepričali. □

Odpornost kriptografskega sistema

Iz primera je razvidno, kako lahko uspešno napademo dan kriptografski sistem. Vendar, če bi imeli samo dva znaka v tajnopisu, ključa kriptografskega sistema ne bi mogli odkriti.

Predno poiščemo odgovor na vprašanje, kako dolg mora biti tajnopis, da lahko odkrijemo ključ kriptografskega sistema, dokažimo izrek, ki govori o povprečni informaciji, ki jo tajnopis C vsebuje o ključu K.

IZREK 6.4 *V kriptografskem sistemu* $(\mathrm{M}, \mathrm{K}, \mathrm{C})$ *je pogojna entropija ključev glede na tajnopise (dvoumnost ključev)* $H(\mathrm{K} \mid C)$ *enaka*

$$H(\mathrm{K} \mid C) = H(M) + H(\mathrm{K}) - H(C), \tag{10}$$

kjer so $H(M)$ *entropija sporočil,* $H(\mathrm{K})$ *entropija ključev, ter* $H(C)$ *entropija tajnopisov.*

Odpornost kriptografskega sistema

Dokaz: Dokaz temelji na enačbah::

$$H(X_1, X_2) = H(X_1) + H(X_2 \mid X_1) \quad \text{in}$$

$$H(X_1, X_2, X_3) = H(X_1) + H(X_2 \mid X_1) + H(X_3 \mid X_2, X_1).$$

Velja namreč

$$H(M, \mathrm{K}, C) = H(C \mid \mathrm{K}, M) + H(\mathrm{K}, M).$$

Ko sta sporočilo M in ključ K znana, je $H(C \mid \mathrm{K}, M) = 0$, zato $H(M, \mathrm{K}, C) = H(\mathrm{K}, M)$.

Ker sta sporočilo M in ključ K tudi neodvisna, je

$$H(M, \mathrm{K}, C) = H(M) + H(\mathrm{K}).$$

Odpornost kriptografskega sistema

Ko sta tajnopis C in ključ K znana, je $H(M \mid \mathrm{K}, C) = 0$, zato $H(M, \mathrm{K}, C) = H(\mathrm{K}, C)$.

Sedaj lahko pišemo

$$\begin{aligned} H(\mathrm{K} \mid C) &= H(\mathrm{K}, C) - H(C) \\ &= H(M, \mathrm{K}, C) - H(C) \\ &= H(M) + H(\mathrm{K}) - H(C). \end{aligned}$$

S tem je izrek dokazan. ■

Odpornost kriptografskega sistema

Primer 6.6

Za kriptografski sistem iz primera 6.3 lahko izračunamo:

- entropijo sporočil

$$H(M) = -\sum_{i=1}^{2} P(M_i) \log_2 P(M_i) = -(1/4 \log_2 1/4 + 3/4 \log_2 3/4) = 0{,}81,$$

- entropijo ključev

$$H(\mathrm{K}) = -\sum_{j=1}^{3} P(\mathrm{K}_j) \log_2 P(\mathrm{K}_j) = -\big(1/2 \log_2 1/2 + 2(1/4 \log_2 1/4)\big) = 1{,}50,$$

- entropijo tajnopisov

$$H(C) = -\sum_{k=1}^{4} P(C_k) \log_2 P(C_k) = -(1/8 \log_2 1/8 + 7/16 \log_2 7/16 + \\ + 1/4 \log_2 1/4 + 3/16 \log_2 3/16) = 1{,}85 \quad \text{in}$$

Odpornost kriptografskega sistema

- pogojno entropijo ključev glede na tajnopise

$$
\begin{aligned}
H(\mathrm{K} \mid C) &= -\sum_{j=1}^{3}\sum_{k=1}^{4} P(C_k)P(\mathrm{K}_j \mid C_k)\log_2 P(\mathrm{K}_j \mid C_k) \\
&= -\big(1/8(\log_2 1+0+0)+ \\
&\quad +7/16(6/7\log_2 6/7+1/7\log_2 1/7+0)+ \\
&\quad +1/4(0+3/4\log_2 3/4+1/4\log_2 1/4)+ \\
&\quad +3/16(0+0+\log_2 1)\big) \\
&= 0{,}46\,.
\end{aligned}
$$

Ker je $H(M)+H(\mathrm{K})-H(C) = 0,81+1,50-1,85 = 0{,}46 = H(\mathrm{K} \mid C)$, vidimo, da (10) (to je, trditev izreka 6.2) velja. □

Odpornost kriptografskega sistema

Vzemimo sedaj, da je dan kriptografski sistem $(\mathrm{M}, \mathrm{K}, \mathrm{C})$, ki ne zagotavlja popolne tajnosti, ter s C_n označimo tajnopis, z M_n pa sporočilo dolžine n znakov.

Ker iz (10) sledi

$$H(\mathrm{K} \mid C_n) = H(M_n) + H(\mathrm{K}) - H(C_n),$$

definiramo odpornost n_0 kot tisto vrednost n (če obstaja), pri kateri postane $H(\mathrm{K} \mid C_n) = 0$. Iščemo torej $n > 0$, tako da

$$H(M_n) + H(\mathrm{K}) - H(C_n) = 0. \tag{11}$$

Odpornost kriptografskega sistema

Predpostavimo:

1. Za naravni jezik, ki ga obravnavamo, lahko vzamemo

$$H(M_n) \simeq n \cdot H,$$

kjer je H entropija vira (povprečna informacija, ki jo vsebuje znak, oddan iz vira).

2. Kriptografski sistem je takšen, da so v tajnopisu vsi nizi znakov dolžine n enako verjetni:

$$H(C_n) \simeq n \cdot K \log_d \theta,$$

kjer je θ moč množice Θ. (To je razumna predpostavka, ker vsak dober kriptografski sistem želi 'izravnati' statistične lastnosti jezika.)

Odpornost kriptografskega sistema

Če kriptografski sistem zadošča zgornjima predpostavkama, iz (11) sledi

$$n_0 \cdot H + H(\mathrm{K}) - n_0 \cdot K \log_d \theta = 0,$$

to je

$$n_0 = \frac{H(\mathrm{K})}{K \log_d \theta - H} \ . \tag{12}$$

Navadno predpostavimo, da imajo vsi ključi enako verjetnost, da so izbrani za šifriranje sporočila.

Odpornost kriptografskega sistema

V tem primeru je število znakov tajnopisa, ki je potrebno za enolično določitev ključa, enako

$$n_0 = \frac{K \log_d \kappa}{K \log_d \theta - H} \ , \tag{13}$$

kjer so κ moč množice ključev, θ moč abecede znakov, ki smo jo uporabili pri pisanju sporočila, in H entropija jezika, v katerem smo sporočilo napisali.

Odpornost kriptografskega sistema - primer

Primer 6.7

Vzemimo, da imamo kriptografski sistem, ki temelji na zamenjavi črk, ki jih je 25. V tem primeru je moč množice ključev[3] enako 25!.

Če dodatno predpostavimo, da je entropija slovenskega jezika 2 bita na črko abecede, iz enačbe (13) sledi

$$n_0 = \frac{\log_2(25!)}{\log_2 25 - 2} = \frac{83,68}{2,64} \approx 32\,.$$

Za dan kriptografski sistem lahko torej par sporočilo - ključ enolično določimo iz tajnopisa, dolgega vsaj 32 znakov. □

[3] Število permutacij brez ponavljanja $P_n = n!$.

Sodobni kriptografski sistemi

Zaradi izredno hitrega razvoja informacijskih sistemov v pričetku 70-ih let prejšnjega stoletja je zelo naraslo zanimanje za tajnost podatkov, ki so se (med računalniki) prenašali na daljavo ali pa shranjevali v elektronske pomnilnike.

Sodobni kriptografski sistemi uporabljajo postopke, ki *nalogo odpiranja tajnopisa brez poznavanja ključa prevedejo na kakšno zahtevno računsko nalogo, ki za svoje reševanje zahteva zelo veliko računskega časa*.

Sodobne kriptografske sisteme delimo na sisteme s:

- TAJNIM KLJUČEM in na sisteme z
- JAVNIM KLJUČEM.

Sodobni kriptografski sistemi s tajnim ključem

Spoznali bomo algoritem DES (Data Encryption Standard), na katerem temelji klasični primer sistema s tajnim ključem.

Čeprav je ameriški institut za standardizacijo in tehnologijo (NIST-National Institute of Standard and Technology) že izbral njegovega naslednika, je algoritem DES iz leta 1976 še vedno prisoten v številnih sodobnih kriptografskih sistemih.

Naslednika je NIST izbral na mednarodnem razpisu leta 2001 in sicer algoritem AES (Advanced Encryption Standard).

Poleg navednih algoritmov so različni razvijalci predlagali še vrsto drugih, kot so Blowfish, IDEA, NewDES, SAFER, CAST5 in FEAL.

Varnost algoritma DES je možno še izboljšati s trikratnim kodiranjem TripleDES (3-DES, TDES) z različnimi vmesnimi ključi.

Kriptografski sistem DES

Sistem DES (Data Encryption Standard) s pomočjo ključa dolžine 56 dvojiških znakov preslika blok sporočila dolžine 64 dvojiških znakov v blok tajnopisa dolžine 64 dvojiških znakov.

Blok sporočila je najprej izpostavljen začetni permutaciji, nato 16-im iteracijam izračunavanja nelinearne transformacije v odvisnosti od ključa in inverzni permutaciji, ki da na izhodu blok tajnopisa.

Dešifrirni postopek je inverzen šifrirnemu, pri čemer uporabimo isti ključ.

Kriptografski sistem DES - postopek šifriranja

Sporočilo vnašamo v šifrirni algoritem po blokih dolžine 64 dvojiških znakov.

Za nadaljnjo obravnavo označimo posamezne znake bloka sporočila z indeksi $1, 2, 3, ..., 64$, ki označujejo njihova mesta v bloku.

- Blok sporočila najprej izpostavimo začetni permutaciji **IP**, ki jo podaja razpredelnica:

58	50	42	34	26	18	10	2
60	52	44	36	28	20	12	4
62	54	46	38	30	22	14	6
64	56	48	40	32	24	16	8
57	49	41	33	25	17	9	1
59	51	43	35	27	19	11	3
61	53	45	37	29	21	13	5
63	55	47	39	31	23	15	7

Iz razpredelnice sledi, da 58-i znak bloka sporočila postane po permutaciji prvi znak, 50-i znak bloka sporočila postane drugi znak in tako naprej, 7-i znak bloka sporočila postane zadnji znak.

Kriptografski sistem DES - postopek šifriranja

- Permutirani blok sporočila označimo z L_0R_0 in ga razdelimo na dva bloka dolžine 32 znakov L_0 in R_0:

L_0					R_0				
1	2	...	31	32	33	34	...	63	64
58	50	...	16	8	57	49	...	15	7

Oba bloka izpostavimo naslednjemu iterativnemu računanju:

$$\begin{aligned} L_n &= R_{n-1} \\ R_n &= L_{n-1} + f(R_{n-1}, \mathrm{K}_n), \end{aligned}$$

kjer sta K ključ in $n = 1, 2, ..., 16$ iteracijski indeks.

Kriptografski sistem DES - postopek šifriranja

- Rezultat zadnje iteracije sta bloka R_{16} in L_{16}, ki ju združimo in izpostavimo inverzu začetne permutacije IP^{-1}, ki ga podaja razpredelnica:

40	8	48	16	56	24	64	32
39	7	47	15	55	23	63	31
38	6	46	14	54	22	62	30
37	5	45	13	53	21	61	29
36	4	44	12	52	20	60	28
35	3	43	11	51	19	59	27
34	2	42	10	50	18	58	26
33	1	41	9	49	17	57	25

- Rezultat zadnje permutacije je blok tajnopisa dolžine 64 dvojiških znakov.

Kriptografski sistem DES - postopek dešifriranja

- Blok tajnopisa (64 dvojiških znakov) najprej izpostavimo začetni permutaciji **IP**.
- Na izhodu začetne permutacije dobimo blok dolžine 64 znakov $R_{16}L_{16}$, ki ga razdelimo na bloka dolžine 32 znakov R_{16} in L_{16}.
- Bloka R_{16} in L_{16} sta sedaj izpostavljena naslednjemu postopku računanja:

$$\begin{aligned} R_{n-1} &= L_n \\ L_{n-1} &= R_n + f(L_n, \mathrm{K}_n), \end{aligned}$$

  kjer sta K_n ključ in $n = 1, 2, ..., 16$ iteracijski indeks.
- Rezultat zadnje iteracije sta bloka L_0 in R_0, ki ju združimo in izpostavimo inverzu permutacije IP^{-1}, ki dokončno odpre tajnopis.
- Rezultat je blok sporočila dolžine 64 znakov.

Kriptografski sistem DES - izračun funkcije $f(R, \mathrm{K})$

- Blok R dolžine 32 znakov preslikamo s pomočjo funkcije E v blok dolžine 48 znakov. Funkcijo E podaja razpredelnica:

32	1	2	3	4	5
4	5	6	7	8	9
8	9	10	11	12	13
12	13	14	15	16	17
16	17	18	19	20	21
20	21	22	23	24	25
24	25	26	27	28	29
28	29	30	31	32	1

- Temu bloku prištejemo po modulu 2 ključ K dolžine 48 dvojiških znakov.
- Blok dolžine 48 znakov, ki ga dobimo na ta način, razdelimo na 8 blokov B_i z dolžino 6 znakov:

$$E(R) + \mathrm{K} = (B_1 B_2 B_3 B_4 B_5 B_6 B_7 B_8).$$

Kriptografski sistem DES - izračun funkcije $f(R, \mathrm{K})$

- Bloke B_i transformiramo v bloke dolžin 4-ih znakov $\mathrm{S}_i(B_i)$ s pomočjo funkcij S_i; $i = 1, ..., 8$, ki so podane z naslednjimi razpredelnicami:

S_1:

14	4	13	1	2	15	11	8	3	10	6	12	5	9	0	7
0	15	7	4	14	2	13	1	10	6	12	11	9	5	3	8
4	1	14	8	13	6	2	11	15	12	9	7	3	10	5	0
15	12	8	2	4	9	1	7	5	11	3	14	10	0	6	13

S_2:

15	1	8	14	6	11	3	4	9	7	2	13	12	0	5	10
3	13	4	7	15	2	8	14	12	0	1	10	6	9	11	5
0	14	7	11	10	4	13	1	5	8	12	6	9	3	2	15
13	8	10	1	3	15	4	2	11	6	7	12	0	5	14	9

Kriptografski sistem DES - izračun funkcije $f(R, \mathrm{K})$

S_3:

10	0	9	14	6	3	15	5	1	13	12	7	11	4	2	8
13	7	0	9	3	4	6	10	2	8	5	14	12	11	15	1
13	6	4	9	8	15	3	0	11	1	2	12	5	10	14	7
1	10	13	0	6	9	8	7	4	15	14	3	11	5	2	12

S_4:

7	13	14	3	0	6	9	10	1	2	8	5	11	12	4	15
13	8	11	5	6	15	0	3	4	7	2	12	1	10	14	9
10	6	9	0	12	11	7	13	15	1	3	14	5	2	8	4
3	15	0	6	10	1	13	8	9	4	5	11	12	7	2	14

S_5:

2	12	4	1	7	10	11	6	8	5	3	15	13	0	14	9
14	11	2	12	4	7	13	1	5	0	15	10	3	9	8	6
4	2	1	11	10	13	7	8	15	9	12	5	6	3	0	14
11	8	12	7	1	14	2	13	6	15	0	9	10	4	5	3

Kriptografski sistem DES - izračun funkcije $f(R, \mathrm{K})$

S_6:

12	1	10	15	9	2	6	8	0	13	3	4	14	7	5	11
10	15	4	2	7	12	9	5	6	1	13	14	0	11	3	8
9	14	15	5	2	8	12	3	7	0	4	10	1	13	11	6
4	3	2	12	9	5	15	10	11	14	1	7	6	0	8	13

S_7:

4	11	2	14	15	0	8	13	3	12	9	7	5	10	6	1
13	0	11	7	4	9	1	10	14	3	5	12	2	15	8	6
1	4	11	13	12	3	7	14	10	15	6	8	0	5	9	2
6	11	13	8	1	4	10	7	9	5	0	15	14	2	3	12

S_8:

13	2	8	4	6	15	11	1	10	9	3	14	5	0	12	7
1	15	13	8	10	3	7	4	12	5	6	11	0	14	9	2
7	11	4	1	9	12	14	2	0	6	10	13	15	3	5	8
2	1	14	7	4	10	8	13	15	12	9	0	3	5	6	11

Kriptografski sistem DES - izračun funkcije $f(R, \mathrm{K})$

- Prvi in zadnji dvojiški znak bloka B_i uporabimo kot indeks vrstice v razpredelnici funkcije S_i, drugi, tretji, četrti in peti dvojiški znak pa uporabimo kot indeks stolpca.
- Vrstica in stolpec določata element razpredelnice S_i, ki v dvojiškem zapisu predstavlja izhod $\mathrm{S}_i(B_i)$.

Kriptografski sistem DES - primer

Primer 6.8

Vzemimo, da je $B_1 = (110110)$. Prvi in šesti znak določata indeks vrstice: $10_2 = 2_{10}$, drugi, tretji, četrti in peti znak pa indeks stolpca: $1011_2 = 11_{10}$.

V razpredelnici S_1 je v tretji vrstici in dvanajstem stolpcu (najmanjša vrednost indeksa vrstice in stolpca je nič!) element $7_{10} = 0111_2$. Torej, $S_1(B_1) = (0111)$.

Kriptografski sistem DES - primer

Izhodi iz funkcij S_i skupaj sestavijo blok dolžine 32 znakov

$$S_1(B_1)S_2(B_2)S_3(B_3)S_4(B_4)S_5(B_5)S_6(B_6)S_7(B_7)S_8(B_8),$$

ki ga vodimo na permutator **P**:

16	7	20	21
29	12	28	17
1	15	23	26
5	18	31	10
2	8	24	14
32	27	3	9
19	13	30	6
22	11	4	25

Izhod iz permutatorja **P** je blok dolžine 32 znakov

$$\mathrm{P}\Big(S_1(B_1)S_2(B_2)...S_8(B_8)\Big).\square$$

DES - postopek izbire ključev $K_1, ..., K_{16}$

Ključ sestavlja 64 dvojiških znakov, vendar jih dejansko uporabimo le 56, ostalih 8 pa lahko uporabimo za odkrivanje napak pri gradnji, porazdeljevanju in shranjevanju ključev. Z znaki 8, 16, ... 64 zagotovimo liho število enic v zlogih.

Ključ dolžine 64 znakov zato najprej vodimo na funkcijo permutirane izbire PC_1:

$$\begin{array}{ccccccc} 57 & 49 & 41 & 33 & 25 & 17 & 9 \\ 1 & 58 & 50 & 42 & 34 & 26 & 18 \\ 10 & 2 & 59 & 51 & 43 & 35 & 27 \\ 19 & 11 & 3 & 60 & 52 & 44 & 36 \end{array} \quad C_0$$

$$\begin{array}{ccccccc} 63 & 55 & 47 & 39 & 31 & 23 & 15 \\ 7 & 62 & 54 & 46 & 38 & 30 & 22 \\ 14 & 6 & 61 & 53 & 45 & 37 & 29 \\ 21 & 13 & 5 & 28 & 20 & 12 & 4 \end{array} \quad D_0,$$

ki preoblikuje ključ dolžine 64 znakov v ključ dolžine 56 znakov.

DES - postopek izbire ključev $K_1, ..., K_{16}$

Iz razpredelnice je razvidno, da znaki ključa na mestih 8, 16, 24, 32, 40, 48, 56 in 64 niso uporabljeni.

Prvi del razpredelnice določa funkcijo C_0 (28 znakov), drugi del pa funkcijo D_0 (28 znakov). Krožni premik C_0 in D_0 za 1 znak v levo določa funkciji C_1 in D_1.

Nato iz niza znakov C_1D_1 določimo ključ K_1 z uporabo funkcije permutirane izbire PC_2:

14	17	11	24	1	5
3	28	15	6	21	10
23	19	12	4	26	8
16	7	27	20	13	2
41	52	31	37	47	55
30	40	51	45	33	48
44	49	39	56	34	53
46	42	50	36	29	32

DES - postopek izbire ključev $\mathrm{K}_1, \ldots, \mathrm{K}_{16}$

Ključ K_2 določimo iz niza C_2D_2, prav tako z uporabo funkcije permutirane izbire PC_2. Niz C_2D_2 smo dobili tako, da smo niza C_1 in D_1 krožno premaknili za 1 znak v levo.

Opisani postopek ustvarjanja ključev nadaljujemo do K_{16}, vendar število premikov v levo za določitev nizov C_i in D_i ni v vseh iteracijah enako 1[4]. Število krožnih premikov v levo za določitev C_1 do C_{16} ter D_1 do D_{16} določa naslednja razpredelnica:

A	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
B	1	1	2	2	2	2	2	2	1	2	2	2	2	2	2	1

A ... korak iteracije

B ... število krožnih premikov v levo

[4] Na primer, C_{11} in D_{11} dobimo iz C_{10} in D_{10} z dvema krožnima premikoma v levo. Z uporabo funkcije permutirane izbire PC_2 nad $C_{11}D_{11}$ pa dobimo ključ K_{11}

DES - napad na kriptografski sistem

Če kriptoanalitik razpolaga s parom tajnopis – razkrito sporočilo (C, M), lahko z *napadom surove sile* poišče ključ K, za katerega velja $C = E_{\mathrm{DES}}(M, \mathrm{K})$, v povprečju z 2^{56-1} poskusi šifriranj in primerjanj tajnopisov.

Na tekmovanju iz kriptoanalize *Challenge III* leta 1999 so s tovrstnim napadom našli ključ v 22-ih urah in 15-ih minutah na omrežju 100 000 računalnikov, ki je lahko preverjalo 245 milijard ključev v sekundi.

V naslednjih letih sta bila zato predlagana dva, na kriptografske napade bolj odporna kriptografska sistema: 3-DES in AES.

Kriptografski sistem 3-DES

Kriptografski sistem 3-DES s pomočjo treh ključev, vsak dolžine 56 dvojiških znakov, preslika blok sporočila M dolžine 64 dvojiških znakov v blok tajnopisa C dolžine 64 dvojiških znakov s trikratno uporabo kriptografskega sistema DES.

Kriptografski sistem 3-DES

Preslikavo bloka sporočila v blok tajnopisa lahko opravimo na dva načina:

- *EEE način* – sporočilo M šifriramo s pomočjo ključa K_1, dobljen tajnopis nato ponovno šifriramo s pomočjo ključa K_2, novi tajnopis pa še enkrat s pomočjo ključa K_3, to je

$$C = E_{\mathrm{DES}}\Big(E_{\mathrm{DES}}\big(E_{\mathrm{DES}}(M, \mathrm{K}_1), \mathrm{K}_2\big), \mathrm{K}_3\Big),$$

- *EDE način* – sporočilo M šifriramo s pomočjo ključa K_1, dobljen tajnopis nato dešifriramo s pomočjo ključa K_2, rezultat pa še enkrat šifriramo s pomočjo ključa K_3, to je

$$C = E_{\mathrm{DES}}\Big(D_{\mathrm{DES}}\big(E_{\mathrm{DES}}(M, \mathrm{K}_1), \mathrm{K}_2\big), \mathrm{K}_3\Big).$$

Trenutno velja ocena, da tajnopise, dobljene s šifrirnim postopkom 3-DES, ne bo mogoče razkriti najmanj do leta 2030.

Vprašanja

- Kaj je kriptografski sistem in kako ga zapišemo?
- Opišite Vigenerjev kriptografski sistem!
- Kdaj kriptografski sistem zagotavlja popolno tajnost?
- Kaj določa odpornost kriptografski sistem?
- Kako delimo sodobne kriptografske sisteme?
- Opišite osnovno zamisel kriptografskega sistema DES!
- Naštejte nekaj kriptografskih standardov, sorodnih DES!