(Digitalni) ključi

Aleksandar Jurišić in Klemen Klanjšček

→ **Šifre oziroma bolj splošno kriptosisteme pogosto kontroliramo s ključi, ki določajo pretvorbo podatkov. Seveda imajo tudi *ključi* lahko digitalno obliko. Ko bomo primerjali njihovo dolžino, bo to kar binarno zaporedje: 01001101010101 ... Pri tem se držimo *Kerckhoffsovega principa*, ki pravi:**

> »Napadalec« pozna kriptosistem oziroma postopke (algoritme), ki jih uporabljamo, ne pa tudi ključev.

Preden se poglobimo v kriptosisteme, in njihovo razbijanje, pomislimo na vsakdanje življenje. Najbolj pogosto uporabljamo posebne vrste ključev, ki jih imenujemo **gesla** in **PIN-i** (kratica pomeni Personal Identification Number, tj. osebno identifikacijsko število). Pri prvih radi uporabljamo čim širši nabor znakov, pri slednjih pa samo števke od 0 do 9 (saj običajno v tem primeru na tipkovnici ni drugih znakov).

V starih časih (tudi pr. n. št.), glej sliko 1, so imele mestne utrdbe stražarje, ki so varovali mestna vrata. Opremljeni so bili z orožjem in v nemirnih časih niso spustili noter nikogar, ki ni poznal pravilnega gesla.

Napadi na gesla

Gesla nam omogočajo kontrolo dostopa v večino računalnikov (danes sem štejemo tudi telefone). Skorajda ni osebe v modernem svetu, ki ne uporablja gesel vsak dan in praktično povsod, npr. za prijavo v e-pošto, račun za družbena omrežja, do našega delovnega prostora, za dostop do shranjenih podatkov, denarja itd. Problem vohunove dileme pa ostaja v bistvu nespremenjen, glej sliko 2.

SLIKA 1.
Kaj se je zgodilo s tistimi, ki so povedali napačno geslo? Pravilo je bilo: »Poskrbite zanje!«

> Kako vemo, ali lahko zaupamo tistemu, ki želi od nas geslo? Nas morda kdo posluša ali celo snema?

Zaupanje lahko dosežemo z digitalmi podpisi (in potrdili), glej Presekov članek [2]. Odgovor na drugo vprašanje predstavlja patentirana tipkovnica 4×3 (prikazana na zaslonu na dotik), ki vsakič naključno premeša številke na tipkah za vnos PIN-a, hkrati pa omogoča, da lahko te številke preberemo le pod skoraj pravim kotom, glej [4]. Enostavnejši odgovor na obe vprašanji bomo predstavili v razdelku Predstavljanje in overjanje.

Če si geslo težko zapomnimo, je bolj verjetno, da ga bomo kmalu pozabili in s tem izgubili dostop. Nekateri sistemi nam sicer omogočajo ponastavitev gesla, a je ta pot lahko precej zamudna, skorajda v vseh primerih pa smo pri tem izpostavljeni določenim ranljivostim. Včasih si pomagamo tako, da si težko zapomnljiva gesla zapišemo, s čimer se seveda izpostavimo novim nepotrebnim tveganjem (zapis lahko nekdo najde in prebere).

Ko nastavljamo geslo na računalniku, lahko naletimo na veliko omejitev, kot npr.:

- imelo naj bi vsaj 8 znakov,
- vsebovalo naj bi vsaj eno veliko in eno malo črko ter
- vsebovalo naj bi vsaj eno številko itd.

Razlogi zanje so očitni iz tabele 1.

> Dobro geslo si je
>
> - enostavno zapomniti in
> - težko uganiti.
>
> Ta dva kriterija si seveda pogosto nasprotujeta.

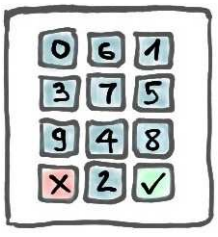
Napadalec utegne imeti na voljo tudi večjo računsko moč, kot je naš telefon ali računalnik, zato kaže še malo premisliti o možnostih (glej tabeli 2 in 3).

Ena od možnih obramb je, da po treh napačnih vnosih gesla sistem zamrzne za kratek čas.

primer	število	zahtevnost	dolžina gesla	čas za razbijanje
urpro	5	25 (majhne črke)	24 bitov	manj kot 2 sekundi
ke1a7Az	7	62 (črke in številke)	42 bitov	5 dni
GK,X1 b@V+	10	95 (znaki na tipkovnici)	66 bitov	233977 let

TABELA 1.
Uganjevanje gesel, pri čemer smo upoštevali, da naš računalnik (npr. pametni telefon) preveri 10 milijonov gesel na sekundo.

SLIKA 2.
Vohunova dilema. Bilo je temno kot v rogu, ko se je vohun vračal v grad po opravljeni diverziji v sovražnem taboru. Blizu obzidja je iz teme zaslišal glas. Kako vohun prepriča »stražarja«, da pozna geslo, ne da bi ga izdal morebitnemu skritemu vsiljivcu?



SLIKA 3.
Premešana številčnica.

Statistika gesel

Zgoraj omenjene zahteve sistemov pri izbiri gesel so bile precej pogoste, vendar bomo spoznali, da niso bistveno prispevale k varnosti (posebej, če si je uporabnik geslo res napisal na listič in ga pozabil ob računalniku).

Vseeno si oglejmo, kaj se dogaja, kadar ni nobenih omejitev (takšnih naj bi bilo skoraj 80% oblačnih storitev), glej npr. [5]. Pogosto se namreč zgodi, da je v javnost prišla kakšna večja baza gesel, ki nam omogoči vpogled v naše navade.

Pri `LinkIn` je imelo npr. preko 80 % uporabnikov dolžino gesel 6–10 znakov, od tega več kot 30 % dolžino 8. Statistika gesel pri `Rockyou.com`, kjer smo lahko proučili 32M nezaščitenih gesel, je odkrila, da je bilo eno najbolj pogostih gesel `123456` (0,11 %). Sledila sta `12345` (0,04 %), `123456789` (0,04 %) in nato še `1234567` (0,02 %), `1234568` (0,02 %). Med prvimi desetimi so bili še `password` (0,03 %), `iloveyou` (0,03 %), `princess` (0,02 %), `rockyou` (0,02 %), `abc123` (0,02 %). 40 % gesel je vsebovalo samo male črke, 15 % samo številke, 35 % samo male črke in številke, medtem ko jih je le 4 % vsebovalo tudi posebne znake. Znana je tudi uporaba variacij na temo enostavnih gesel, npr. razne kombinacije besede »geslo«, števila 123 in trojice »abc«. Pogosta so tudi imena mesecev in dnevov v tednu. Na koncu gesla znamo najti ločila – npr. pike, številke ali celo letnice.

> Geslo za izstrelitev nuklearne rakete je bilo 20 let »00000000«. *Facebook* je za glavno geslo dostopa do vseh profilov uporabljal »Chuck Norris«.

Pri PIN-ih zaznamo podobne, a vseeno še hujše težave. Skoraj polovica od vseh v neki kompromitirani bazi je bilo štirimestnih (glej `www.datagenetics.com/blog/september32012/`). V tem primeru so bili najbolj popularni naslednji: `1234` (11 %), `1111` (6 %), `0000` (2 %), `1212` (1,2 %), `7777` (0,7 %), `1004`, `2000` (vsak po 0,6 %), `4444`, `2222`, `6969`, `9999` (vsak po 0,5 %), `3333`, `5555`, `6666`, `1122` (vsak po 0,4 %), `1313`, `8888`, `4321`, `2001`, `1010` (vsak po 0,3 %). Torej pravkar omenjenih 20 PIN-ov pokrije več kot četrtino vseh. Sledile so dolžine PIN-ov 6 (18 %), 8 (11 %), 5 (9 %), 7 (7 %), 9 (3 %), 10 (1,5 %) itd, vse s podobnimi težavami.

Da je stanje še slabše, skoraj 30 % uporabnikov uporablja ista gesla oz. PIN-e na številnih napravah ali računih. Slednje ni nič čudnega, saj 70 % uporabnikov vsaj 10 računov ščiti le z gesli. To število bo samo še rastlo, zato prihaja do uporabe aplikacij

posameznik, 1 računalnik	2^{27} – 2^{33} ključev/sek.
akademsko omrežje, 256 računalnikov	2^{35} – 2^{41} ključev/sek.
velika spletna skupina, 10M računalnikov	2^{50} – 2^{55} ključev/sek.
zelo velika spletna skupina (npr. Bitcoin omrežje)	2^{58} – 2^{68} ključev/sek.
vojaška obveščevalna organizacija	2^{50} – 2^{70} ključev/sek.

TABELA 2.
Napadi z grobo silo – koliko ključev lahko pregledamo v eni sekundi (za šifro AES-128).

dolžina ključa (v bitih)	**posameznik**	**akademsko omrežje**	**velika spletna skupina**	**zelo velika spletna skupina**	**vojaška obveščevalna organizacija**
40	minute	sekunde*	sekunde*	sekunde*	sekunde*
56	leta	dnevi	sekunde*	sekunde*	sekunde*
64	stoletja	leta	ure*	minute*	sekunde*
80	∞	stotisočletja	leta	ure*	ure*
128	∞	∞	∞	∞	∞

TABELA 3.
Napadi z grobo silo – koliko časa potrebujemo za različne dolžine ključev.
* zaradi organizacije in komunikacije med računiškimi sistemi oz. ljudmi bodo napadi običajno trajali dlje.

za upravljanje gesel. Vendar, če v ta namen uporabljate kar vaš priljubljeni brskalnik, morate imeti obvezno nastavljeno tudi **glavno geslo** (angl. master password). Če pa govorimo o podobnih geslih na različnih napravah, se delež poveča na 50 %. 66 % ljudi pa uporablja 1 ali 2 gesli za vse svoje račune. Oseba običajno zamenja geslo na vsake 2,5–3 leta.

Kako si zapomnimo varnejša gesla?

Geslom, ki vsebujejo več kot eno pomenljivo besedo, pravimo **skrivna fraza** (angl. passphrase). Ni treba, da ta fraza nekaj pomeni (slika 4).

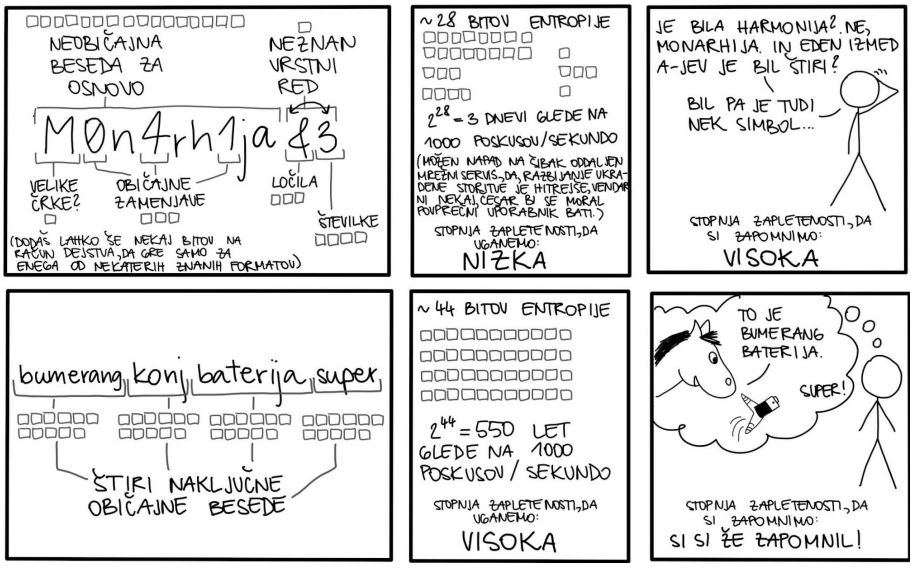
Predstavljanje in overjanje

Če poenostavimo, je **identifikacija** dokaz uporabnikove identitete, **overjanje** pa je postopek (za nas bo pogosto kriptografski protokol), pri katerem uporabnik želi potrditi, da je res tisti, za kogar se predstavlja.

Za potrjevanje uporabnikove identitete se uporablja podmnožico naslednjih treh lastnosti:

- nekaj, kar uporabnik **ve** – npr. gesla, številke (PIN),
- nekaj, kar uporabnik **je** – to so biometrični podatki, ki temeljijo na fizičnih lastnostih uporabnika, kot so prstni odtis, šarenica, obraz, ali način hoje, glej sliko 7,
- nekaj, kar uporabnik **ima** – t. i. varnostni žetoni, kot so fizični ključi, osebna izkaznica, vozniško dovoljenje, bančna kartica, mobilni telefon itd.

V primeru plačila z bančno kartico ali telefonom se npr. običajno uporabljata dve lastnosti: bančna kartica ali telefon (nekaj, kar uporabnik ima) ter PIN številka (nekaj kar, uporabnik ve). V zadnjem času pa se pri plačevanju s telefonom vse bolj pogosto uporablja biometrika.



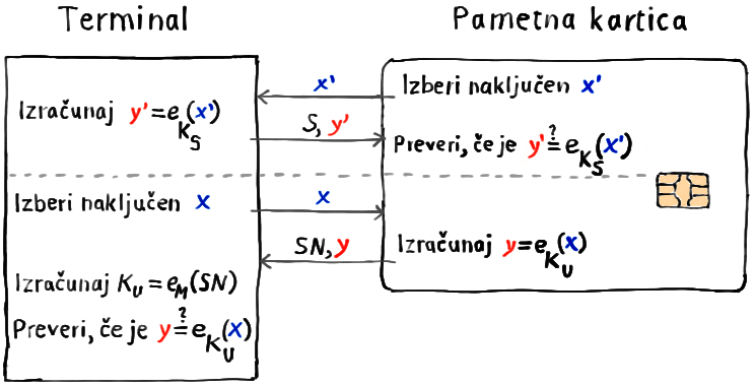
SLIKA 4.
Več kot 20 let so se sistemski administratorji trudili, da bi natrenirali uporabnike k uporabi gesel, ki si jih težko zapomnijo, računalniki pa so jih kljub temu hitro uganili (idejo iz XKCD [3] sta v slovenščino prenesli Neja Cirar in Ajda Markič).

Anita in Bojan delita tajni (skrivni) ključ K, ki ga uporabljata za šifriranje.

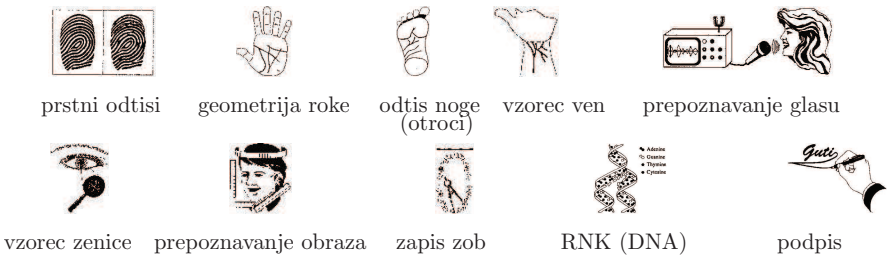
1. Bojan izbere 64-bitni izziv X in ga pošlje Aniti.
2. Anita izračuna $Y = K \oplus X$ in ga pošlje Bojanu.
3. Bojan izračuna $Y' = K \oplus X$ in preveri, če je $Y = Y'$.

Primer. (4 biti)
$K = 10$, $X = 6$,
$Y = K \oplus X = 10 \oplus 6$
$= 1010_2 \oplus 0110_2$
$= 1100_2 = 12$.

SLIKA 5.
Identifikacijski protokol z izzivom in odgovorom. Operacijo $\oplus$ (ekskluzivni ali nad istoležnimi biti) bi lahko zamenjali s kakšno drugo funkcijo, tj. $y = e_K(x)$ ter $y' = e_K(x)$ in še vedno preverjali $y = y'$. Skoraj vse sheme uporabljajo protokole z izzivom in odgovorom, vendar pa najbolj koristne ne uporabljajo skupnih ključev. Ker presegajo ta sestavek in jih pustimo za kdaj drugič.



SLIKA 6.
Primer konkretne uporabe protokola izziv in odgovor. Kako zdravstvena kartica nekega uporabnika in terminal zdravstvenega sistena preverita pristnost drug drugega? Najprej morata imeti tako terminal kot kartica možnost generiranja naključnih števil. Kako pa je s ključi? Terminali imajo glavni ključ M, in skupinski ključ K_S, nimajo pa prostora, da bi hranili ključe vseh uporabnikov. Zato je vsakemu uporabniku U s serijsko številko SN dodeljen ključ $K_U := e_M(\mathrm{SN})$ (K_U določimo s pomočjo glavnega ključa M in uporabnikove serijske številke SN), ter vsi skupinski ključi, ki so potrebni za prepoznavanje različnih skupin (poleg skupine terminalov so tu še zdravniške kartice, pa kartice različnih lekarn itd).



SLIKA 7.
Biometrični testi. Poudariti je treba, da ni idealnega testa. Prstne odtise npr. puščamo na raznih predmetih, ki se jih dotaknemo. Če smo dolgo v vodi ali na mrazu, se prstni odtis malce spremeni. Še večji problem nastane, če si prst poškodujemo. Si predstavljate, da bi morali na bankomatu oddati ugriz? Se kdo spomni dejanskega preverjanja podpisa, le-to bi moral itak opraviti izurjen grafolog.

Upravljanje ključev

Postavimo si nekaj osnovnih vprašanj:

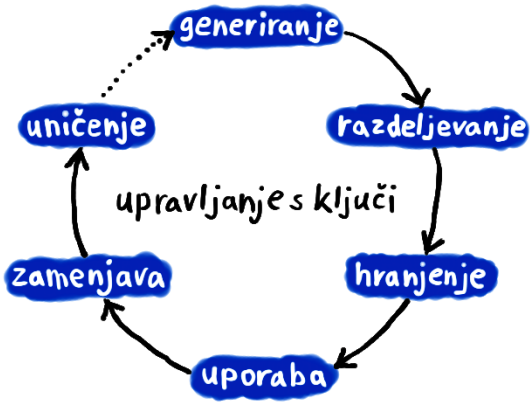
- Od kod dobimo ključe?
- Zakaj zaupamo ključem?
- Kako vemo, čigav ključ imamo?
- Kako omejiti uporabo ključev?
- Kaj se zgodi, če je izgubljen/ukraden zasebni ali tajni ključ? Kdo je odgovoren?
- Kako preklicati ključ?
- Kako omogočimo servis preprečitve zanikanja (ko nekdo ne prizna določenega dejanja)?

Za bolj izkušene naj omenimo, da ta vprašanja veljajo tako za simetrične (tajne) ključe kot tudi za javne in zasebne ključe, glej Presekov članek [1].

Osnovni cilj je vzdrževanje **zasebnosti** in **celovitosti** ključev. Prvo pomeni, da ostanejo ključi tajni, drugo pa, da jih nihče ne more zamenjati ali pokvariti. Ti dve lastnosti morata biti zagotovljeni skozi ves življenjski cikel ključa, se pravi od začetka generiranja ključa, pa vse do takrat, ko ključ ni več v uporabi in je bil uničen, glej sliko 8.

Tako rekoč vedno je ključ zamenjan z novim ključem, kar pomeni, da gre za cikličen proces, v katerem uničenju sledi namestitev/vključitev novega ključa. Vendar pa je novi ključ zgeneriran, dodeljen in shranjen, še preden je stari ključ uničen. V nekaterih sistemih imamo še dodatno zahtevo, da ključe arhiviramo.

Vzvodi za nadzor so potrebni skozi celoten življenjski cikel ključa. To vsebuje od neke vrste revizijskih postopkov (angl. audit trail) do dnevnikov



SLIKA 8.
Glavne stopnje življenjskega kroga prikazuje naslednji diagram.

uporabe ključev, ki seveda niso smiselni brez nadzora ter pooblastil za ustrezen odgovor/ukrep v primeru, da je bil ključ ukraden/izgubljen. Zato je pogosto (posebno v večjih sistemih) zaželeno, da imajo ključi ustrezne lastnike, ki so odgovorni za njihovo hranjenje/zaščito.

Pri algoritmih za kriptosisteme lahko uporabimo poljubno zaporedje bitov (občasno tudi drugih znakov), kar pomeni, da lahko večina uporabnikov sama **generira** svoje ključe. Biti morajo nepredvidljivi: pogosto uporabimo

- ročne metode (npr. metanje kovanca/kocke, ali bolj realistično, premik miške),
- izpeljavo iz osebnih podatkov (npr. PIN) ali
- (psevdo) generatorje naključnih števil (RNG).

Generiranje velikih praštevil zahteva »napredno«

1. Ne izbiraj pogostih (generičnih) gesel ali ključev posebne oblike.
2. Ne hrani gesel ali ključev na manj varnih (preveč očitnih/izpostavljenih) mestih.
3. Ne uporabljaj istih ali podobnih gesel in ključev v različnih programih oz. računih (ne recikliraj gesel in ključev).
4. Ne kopiraj gesel in ključev (kopiraj/prilepi prek odložišča – angl. clipboard).
5. Ne pošiljaj ali deli z nikomer gesel in ključev.

TABELA 4.

Popravi 5 slabih navad, ki vplivajo na varnost gesel oz. bolj splošno ključev.

1. Gesla in ključi naj bodo dovolj dolgi.
2. Pogosto posodabljaj/obnavljaj/zamenjaj svoja gesla in ključe (in stori to takoj, če/ko izveš, da je bil tvoj račun ukraden/izgubljen).
3. Skrivne fraze so dobra strategija za pomnenje gesel (vsaj tistih glavnih).
4. Uporabljaj večstopenjsko (faktorsko/dejavniško) overjanje identitete (avtentikacijo) za pomembne/kritične račune (če je le možno).
5. Uporabljaj dober program za upravljanje ključev in gesel.

TABELA 5.

Pet dobrih navad za varnost gesel oz. bolj splošno ključev.

matematično obdelavo in lahko zahteva precejšno moč/prostor/čas. Uporabniki so včasih prisiljeni zaupati zunaj generiranim ključem oziroma programski opremi. Za ilustracijo: v primeru kriptosistema RSA je njegova varnost povezana z napadalčevo zmožnostjo razcepa velikega števila. Če bi proces generiranja ključev dal le omejeno število praštevil, bi napadalec lahko generiral ista praštevila in nato za vsakega izmed njih preveril, ali delijo dano število.

Razdeljevanje ključev in varno **hranjenje** sta osrednjega pomena. Problemi, na katere naletimo in njihove rešitve, so večinoma podobni, zato jih obravnavamo skupaj. Razlogi za uporabo močnih kriptografskih algoritmov je preprečiti napadalcu, da izračuna ključ. Seveda vse to nima smisla, če jih napadalec lahko najde nezaščitene nekje na sistemu.

Kako veliko je VELIKO?

Živimo le 2–3 milijarde sekund. Za kilo K (10^3), mega M (10^6), giga G (10^9) in tera T (10^{12}), ki smo se jih že kar navadili, sledijo enote peta P (10^{15}), exa E (10^{18}), zeta Z (10^{21}), jota Y (10^{24}), ronna R (10^{27}) in quetta Q (10^{30}). Slednja je v rangu števila $\mathbf{2^{100}}$ (zakaj?), kar je velikost prostora, iz katerega radi v kriptografiji naključno izberemo varen ključ ali kak drug skriven parameter. Varnostni strokovnjaki so dolžni slediti priporočilom za izbiro velikosti ključev za posamezno skupino kriptosistemov, ki se s časom seveda nenehno spreminjajo, glej npr. Wikipedijo [6].

Naloge

1. Koliko je različnih znakov na tvoji tipkovnici (ne pozabi na male in velike črke)? Bi znal vse našteti?
2. Koliko zrn riža je povprečno v enem kilogramu riža (gotovo ga imate kaj doma)? (*Namiga:* Zakaj je v vprašanju beseda povprečno? Ali to pomeni, da boš moral šteti večkrat?)
3. Koliko bi bilo možnih gesel, če bi izbirali za geslo besedo v slovenskem jeziku? Kaj pa, če bi si izbral dve ali tri besede? In kaj, če se ne bi omejili s slovenščino? Koliko je različnih jezikov?
4. Kakšna je verjetnost, da uganemo štirimestni PIN neke bančne kartice, če bi imeli možnost poskusiti 20-krat (upoštevaj, da je bil PIN izbran naključno)? Kaj pa, če bi imeli opravka s 6- ali 8- mestnim PIN-om. Ali bi bilo smiselno dati možnost večjega vnosa napačnih PIN-ov pri daljših PIN-ih?
5. Pri štirimestnih PIN-ih so izredno popularne letnice (rojstev), s katerima številkama bi torej začel ugibanje? (Tudi sicer se največ PIN-ov začne z 1, sledijo pa tisti z 0 in 2, glej sliko na naslovnici.)
6. Ali imaš načrt B za primer, ko pozabiš geslo/ključ? Lahko predlagaš kakšnega?
7. Koliko različnih računov in gesel uporabljaš (oz. ocenjuješ, da jih boš v naslednjih 10 letih)? Zelo verjetno je, da napadalec pozna nekaj tvojih gesel. Če obstaja skupno jedro teh gesel, oceni število potrebnih ugibanj za odkritje drugih gesel.
8. Naštej nekaj dobrih in slabih lastnosti obnavljanja gesel preko e-pošte.
9. Predlagaj še kakšno rešitev za vohunovo dilemo (glej sliko 2).
10. Kako točno poteka preverjanje gesla? Ali se lahko izognemo, da bi gesla nekje hranili (kajti kot smo omenili, zna takšna baza uiti izpod nadzora, pa če se še tako trudimo)? Nekatera podjetja imajo sezname gesel za svoje zaposlene. Kje naj jih hranijo?
11. Predlagaj še kakšno funkcijo $e_K(x)$ (glej sliko 5) in premisli, kaj so njene dobre in slabe lastnosti.
12. Poišči kakšen odgovor na osnovna vprašanja glede upravljanja ključev, ki smo jih zastavili na začetku razdelka Upravljanje ključev.
13. Naštej nekaj dobrih in slabih lastnosti posameznih biometričnih testov (glej sliko 7).
14. Naštej nekaj prednosti in pomanjkljivosti zgornjih slabih in dobrih navad za izbiro gesel/ključev (tabeli 4 in 5). Bi znal dopolnit oba seznama še s tremi priporočili?
15. Na koga bi se obrnil, če bi moral implementirati varno preverjanje gesel v svoji aplikaciji? Kako bi se lahko izognil implementaciji gesel?
16. V razdelku Kako si zapomnimo varnejša gesla? smo omenili, da si za skrivno frazo izberemo npr. 4 naključne besede. Predlagaj kakšen enostaven postopek, kako priti do teh besed z uporabo igralne kocke.
17. Oceni dolžino povprečno dolžino skrivnih fraz, ki si jih uporabljal ti ali tvoji prijatelji doslej. Kako velika je množica, iz katere izbiraš besede za tvorjenje fraz?

Doslej se nam je uspelo izogniti zgoščevalnim funkcijam, kar pomeni, da se bomo morali pomeniti o geslih še enkrat.

Literatura

[1] A. Jurišić, *Poštni nabiralnik in kriptografski protokoli*, 3. del, *Presek* **36** (2008–2009), 22–26.

[2] A. Jurišić in J. Veselinović, *Zaupati ali ne zaupati: digitalni podpis v kriptografiji*, 4. del. *Presek* **38** (2010–2011), 25–29.

[3] R. Munroe, *Spletni strip XKCD*, `https://xkcd.com/936/`

[4] Y. S. Ryu, D. H. Koh, B. L. Aday, X. A. Gutierrez, J. D. Platt, *J. of usibility studies* **5**/2, February 2010, 65–75, dostopno na `https://uxpajournal.org/usability-evaluation-of-randomized-keypad/`, ogled 1. 4. 2023.

[5] J. H. Simmons, *Passwords – 25 Fun Facts and Statistics!*, dostopno na `https://www.vpncrew.com/passwords-facts-and-statistics/`, ogled 1. 4. 2023.

[6] *Security level*, dostopno na `https://en.wikipedia.org/wiki/Security_level`, ogled 1. 4. 2023.

×××