SLAPTAŽODŽIŲ, KAIP AUTORIZAVIMO PRIEMONĖS, SAUGUMO ANALIZĖ

Donatas Dervinis

Šiaulių valstybinė kolegija, Šiaulių universitetas

Anotacija

Slaptažodis, kaip prisijungimo priemonė, naudojama apie 89% atvejų, tačiau yra likusi pasenusi nuostata, kad jo ilgis pakankamas iš 6–8 simbolių su vienu–dviem skaičiais. Šiuo metu grafinių procesorių (GPU) masyvai gali perrinkti iki 350 milijardų žodžių per sekundę, todėl naudojami 8 simbolių slaptažodžiai sudaryti iš simbolių, skaičių ir kitų simbolių gali būti perrinkti per kelias valandas. Straipsnyje pateikiamos įvairios slaptažodžių kombinacijų analizės rodo, kad saugiu slaptažodžiu galima laikyti sudarytą iš 10–11 simbolių mažųjų ir didžiųjų raidžių, skaičių arba 9 simbolių ilgio – sudarytų iš 96 įvairių simbolių. Naudojant dėsningas kombinacijas: žodyno žodžius, šabloninius rinkinius – saugaus slaptažodžio ilgis ilgėja iki 11–13 simbolių ilgio.

Esminiai žodžiai: slaptažodžių ilgis, autorizacija, įsilaužimas, atakos**.**

Įvadas

Tapatybės, duomenų vagystės, šnipinėjimas – tai grėsmės su kuriomis susiduria ne tik IT specialistai, bet ir kiekvienas žmogus, bent minimaliai naudojantis IT priemonėmis: kompiuteriu, telefonu, išmaniaisiais daiktais (angl. *internet of things;* trump. *IoT*). Paskaičiuota, kad pasaulyje per 2016 metus įvykę elektroniniai nusikaltimai kainavo 450 milijardus JAV dolerių arba 60 JAV dolerių vienam žmogui. 53% verslo kompanijų tiesiogiai susidūrė su elektroniniais nusikaltimais ir nuo jų nukentėjo (Graham,2017).

Didėjant IT įrenginių kiekiui, jiems pingant, ir ypač didėjant išmaniųjų daiktų jungiamų prie interneto tinklo populiarumui, kiekvienas gamintojas stengiasi pagaminti nebrangų ir patrauklų vartotojui įrenginį. Neretai toks įrenginys jungiamas ir valdomas per interneto tinklą, tačiau jo saugumui ir autorizavimui skiriama labai mažai dėmesio arba jo visai neskiriama, t.y. neretai įrenginys neturi minimalių saugumo ir autorizavimo parametrų keitimo galimybės: nėra pradinio slaptažodžio, vartotojo vardo ar slaptažodis yra vienodas visiems įrenginiams, jo pakeitimas negalimas arba galimas tik slaptažodžio. Vartotojai, perkantys tokius įrenginius, taip pat dažnai neturi pakankamai žinių suprasti visų saugumo aspektų, todėl palieka nesaugų ar per trumpą slaptažodį. Tokie įrenginiai žalą daro ne tik pačiam vartotojui, bet ir kitiems interneto vartotojams pvz.: didelė dalis DDoS (angl. *Distributed Denial of Service*) atakų buvo atlikta naudojant išmaniuosius daiktais (Symantec, 2016).

Tyrimo objektas: autorizacijoje naudojami įvairaus ilgio simbolinis slaptažodis.

Tyrimo tikslas: išanalizuoti slaptažodžių, kaip autorizavimo priemonės patikimumą šiuolaikinėje IT aplinkoje.

Tyrimo uždaviniai:

1. Apžvelgti šiandienos slaptažodžių naudojimo tendencijas ir grėsmes.
2. Įvertinti ir išanalizuoti slaptažodžių patikimumą, atsižvelgiant į ilgį ir sandarą.

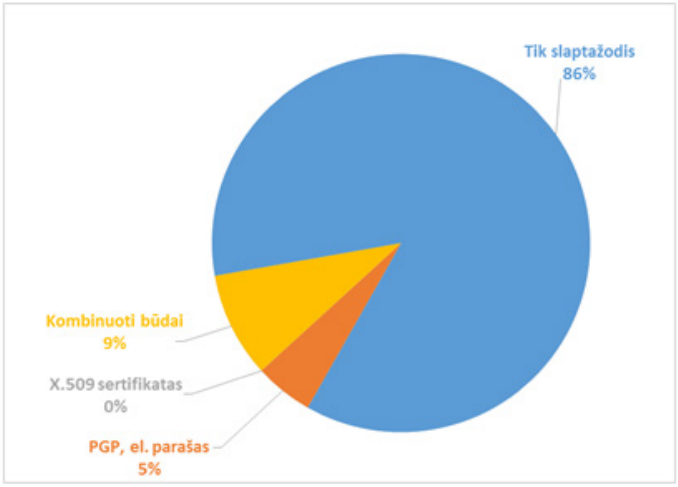
Metodika: Literatūros ir tarptautinės patirties analizė, duomenų analitinė analizė.

Situacijos apžvalga

Nors slaptažodis (angl. *password*) pirmą kartą kaip terminas buvo panaudotas dar Romos imperijos kariuomenėje, šiandien yra pati populiariausia autorizavimo priemonė palyginus su įvairiomis kitomis: elektroninis parašas, sertifikatai, mišrus autorizavimo būdas (Wei ir kt., 2016) ar biometrija (akies rainelės, piršto antspaudo ir pan. nuskaitymas). Įvairiuose šaltiniuose ir tyrimuose galima rasti slaptažodžio naudojimą, kaip pagrindinę autorizacijos priemonę nuo 80% iki 89% (Walker, 2015).

Prieigų ir įrenginių kiekis, kurie reikalauja autorizacijos, šiuolaikiniam žmogui yra labai didelis – kiekvienas turi kelis išmaniuosius įrenginius, keliasdešimt paskyrų internete. Jungiantis prie įvairių paslaugų pirmą kartą, prašomą sugalvoti ir įvesti prisijungimo vardą ir slaptažodį. Daugelis puslapių (tinklalapių) vietoje vartotojo vardo prašo įvesti savo elektroninio pašto adresą, tokiu atveju, vartotojui lieka sugalvoti tik slaptažodį. Tokios vartotojų prieigos tampa identifikuojamos ir susiejamos, t.y. skirtingose platformose galima lengvai identifikuoti tą patį vartotoją, todėl lemiamą saugumo vaidmenį atlieką tik slaptažodis. Deja, slaptažodžių sudarymo ir ilgio reikalavimai likę iš tų laikų, kai centrinis kompiuterio procesoriaus (angl. trump. *CPU)* veikimo greitis siekė 1GHz ir buvo naudojamas tik vienas branduolys. Dabar situacija situaciją pakeitė ne tik *CPU* spartos ar branduolių skaičiaus didėjimas, bet ir grafinio procesoriaus panaudojimo galimybės (angl trump.

GPU). Nauji GPU turi iki trijų tūkstančius branduolių – tai leidžia atlikti didelį kiekį nesudėtingų operacijų vienu metu. Naudojant GPU masyvą galima patikrinti iki $350x10^9$ žodžių per sekundę (Goodin, 2012).



1 pav. Slaptažodžio ir kitų priemonių naudojimo populiarumo statistika (Walker, 2015)

Jungtinės karalystės Ofcom komunikacijų bendrovės tyrimas rodo, kad net keturi iš dešimties (t.y. apie 40%) vartotojų naudoja tik vieną slaptažodį visose savo paskyrose (Cluley, 2016: 167), kiek senesnis IT saugos kompanijos BitDefender tyrimas parodė, kad šis skaičius yra dar didesnis – siekia 75% (BitDefender, 2010). Vartotojai, turintys kelis ar keliolika slaptažodžių, dažnai juos sudaro labai primityvius ir lengvai atspėjamus (Munson, 2014). Komunikacijų kompanija TeleSign atliko tarptautinį tyrimą ir nustatę, kad 21% vartotojų savo slaptažodžių nekeitė 10 ir daugiau metų, o 47% – 5 metus, tik 13% vartotojų slaptažodį keičia kas 6 mėnesius (TeleSign, 2015:15).

Slaptažodžių statistinis vertinimas

Paprastai, įvairūs interneto paslaugų tiekėjai slaptažodžio valdymą aprašo saugumo savo politikose: kokio ilgio slaptažodis turėtų būti, kiek ir kokių registrų (didžiosios raidės, mažosios raidės, skaičiai, standartiniais simboliai, specialus simboliai) naudoti (1 lentelė), kas kiek laiko reikia slaptažodį keisti.

1 lentelė. Įvairių kompanijų slaptažodžių reikalavimai

Kompanija	Slaptažodžio reikalavimai
Facebook	Simboliai6
Twitter	Simboliai6
LinkedIn	Simboliai6
Blogger	Simboliai8
WordPress	Simboliai6
Gmail	Simboliai8
Yahoo	Simboliai6
Hotmail	Simboliai8
Outlook	Simboliai8
Amazon	Simboliai6
Ebay	Simboliai6DM
Paypal	Simboliai8DMS

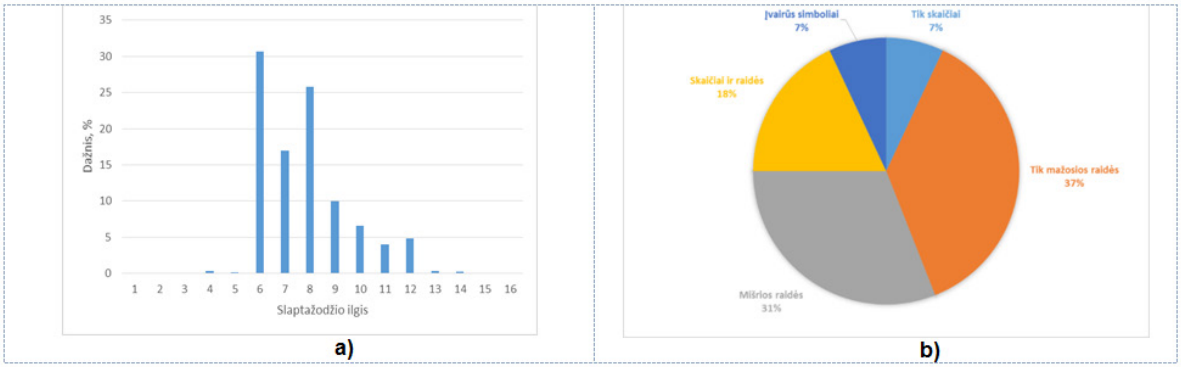
Žymėjimas lentelėje: *Simboliai* – bet koks simbolis, *skaičius* (čia 6 ar 8) – minimalus simbolių kiekis slaptažodyje, *DM* – būtina naudoti ir didžiąsias ir mažąsias raides, *DMS* – *DM* ir papildomai reikalingi skaičiai.

Kaip matoma 1 lentelėje, daugelio žinomų kompanijų reikalavimai autorizacijos slaptažodžiams nėra dideli – jie nėra ilgi ar sudėtingi – dažniausiai 6 ar 8 simboliai, retais atvejais reikalaujama kelių registrų (mažųjų ir didžiųjų raidžių ir /arba skaičių) kombinacijų. Tačiau šiose sistemose yra diegiami papildomi saugumo reikalavimai: įvedus neteisingą slaptažodį kelis kartus yra prašoma atlikti papildomus veiksmus tolesnei autorizacijai, tokius kaip patvirtinti, jog esi ne robotas, priverstinai prašoma patvirtinti autorizacija antriniu būdu (elektroniniu paštu, trumpąją žinute). Dažnas atvejis – tarp neteisingų bandymų įvedama autorizacijos pauzė,

kurios metu negalima atlikti autorizacijos, taip pat ji pailgėja po kiekvieno sekančio neteisingo bandymo. Retais atvejais autorizacijos prieiga yra blokuojama (užrakinama) visam laikui, o atblokavimas galimas tik alternatyviais būdais, tai atlieką žmogus pvz.: įmonės administratorius. Tačiau šis būdas taikomas organizacijos darbuotojams, jungiantis prie vidinių IT resursų arba finansinių įstaigų klientams.

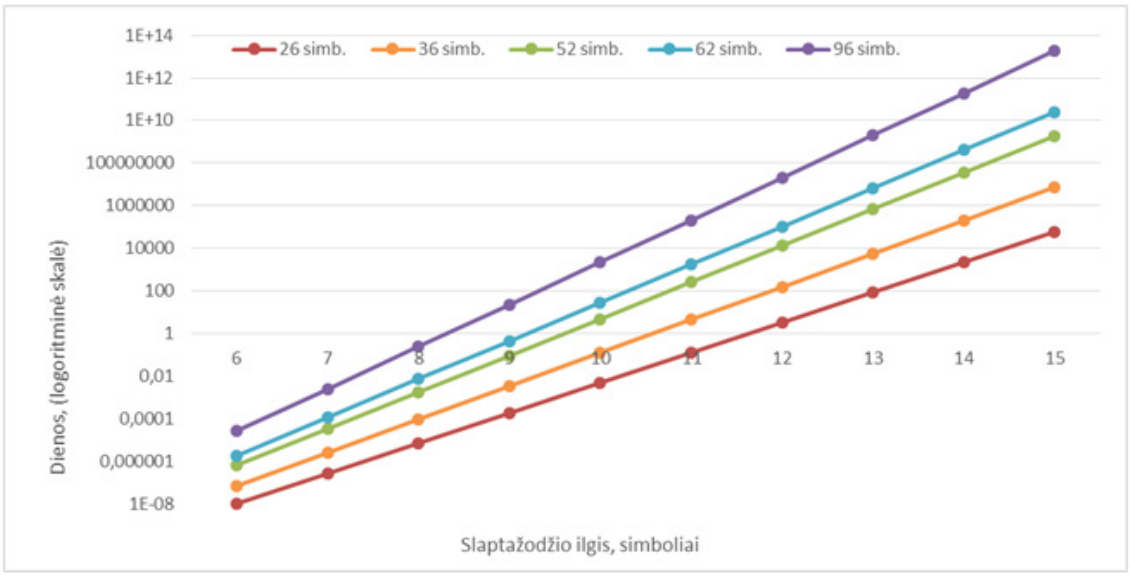
Didžiausia rizika yra prieigos prie koduotų bylų, dokumentų, įmonių ir asmeninių tarnybinių stočių, bei svetainių administravimo zonų, kurių autorizavimas neribojamas jokiais metodais: ribojamas neteisingų slaptažodžių kiekis, ilginamos pauzės laikas tarp neteisingų įvedimų ir pan. Tokiu atveju lengva naudoti įvairius slaptažodžio parinkimo metodus: visų įmanomų simbolių parinkimą (angl. *brute force attack*) arba žodyno ataką (angl. *dictionary attack*).

Teoriniam skaičiavimui buvo panaudota turima slaptažodžių duomenų bazė su 62 tūkstančiais nuasmenintų slaptažodžių rinkinių. Statistiškai apdorojus, buvo gautas slaptažodžių ilgio ir dažnio priklausomybė (2 pav.). Iš grafiko matoma, kad dažniausiai naudojamas slaptažodis yra 6–8 simbolių ilgio, o 74 % visų slaptažodžių yra iki 8 simbolių. Matoma tendencija, kad 37% naudojamų slaptažodžių turi tik vieno lygio registrus ir 31% – mažąsias ir didžiąsias raides. Raides ir skaičius naudoja – 18% procentų vartotojų. Ir tik 7 % prideda specialiuosius simbolius.



2 pav. Naudojamų slaptažodžių ilgio dažnis a) ir sandara b) (N=62 tūkst.)

Toliau pateikiami skaičiavimai ir sąlygos taikant visų įmanomų simbolių parinkimą (angl. *brute force attack*): skaičiavimui naudojamas, šiuo metu nesunkiai pasiekiamas, $350x10^9$ žodžių per sekundę perrinkimo greitis (Goodin, 2012). Taip pat laikoma, kad vienas raidžių registras yra po 26 lotyniškas raides arba abu (mažosios ir didžiosios) – 52 simboliai, skaičiai – 10 simbolių, visi kiti simboliai – 34 (pvz.: !#^+-~ [{?> ir t.t.), todėl galimos tokios kombinacijos: mažosios raidės (26 simb.), mažosios ir skaičiai (36 simb.), mažosios ir didžiosios (52 simb.), mažosios, didžiosios ir skaičiai (62 simb.), raidės, skaičiai ir kiti simboliai (96 simb.).



3 pav. Slaptažodžio iškodavimo laikas dienomis, taikant paprastą visų įmanomų kombinacijų parinkimo būdą

Priklausomai nuo duomenų svarbumo ir senėjimo, pakankamas slaptažodžio nustatymo laikas yra nuo kelių minučių (realaus laiko duomenims) iki kelių metų (ilgalaikės svarbios informacijos apsaugojimui).

Jei laikysime, kad orientacinis pakankamas laikas, per kurį nustatomas slaptažodis, yra 24 valandos, 26 simbolių slaptažodis tenkina stiprumo sąlygą tik sudarytas iš 12 mažųjų raidžių, naudojant raides ir skaičius (36 simbolius) – reikia 11 simbolių, arba 9 simbolių, jei naudojamos didžiosios, mažosios raidės ir skaičiai. Šiandien populiariausias 824 simbolių slaptažodis pakankamas tik tuo atveju, jei būtų naudojami visas įmanomas 96 simbolių rinkinys. Šie rezultatai yra patikimi tik tuo atveju, jei slaptažodis yra sudarytas iš atsitiktinių raidžių ir skaičių derinių pvz.: *d2Fr5ul4.*

Skaičiuojant slaptažodžio sudėtingumą arba išmaišymą yra skaičiuojama slaptažodžio entropija *H*:

$$H = \sum_{i=0}^{n} N \cdot \log_2(p_i) = \sum_{i=0}^{n} N \cdot \log_2(p_i), \quad (1)$$

čia *n* – simbolių kiekis, *N* – grupės simbolių skaičius; *p* – grupės galimų skirtingų simbolių kiekis (26,52...).

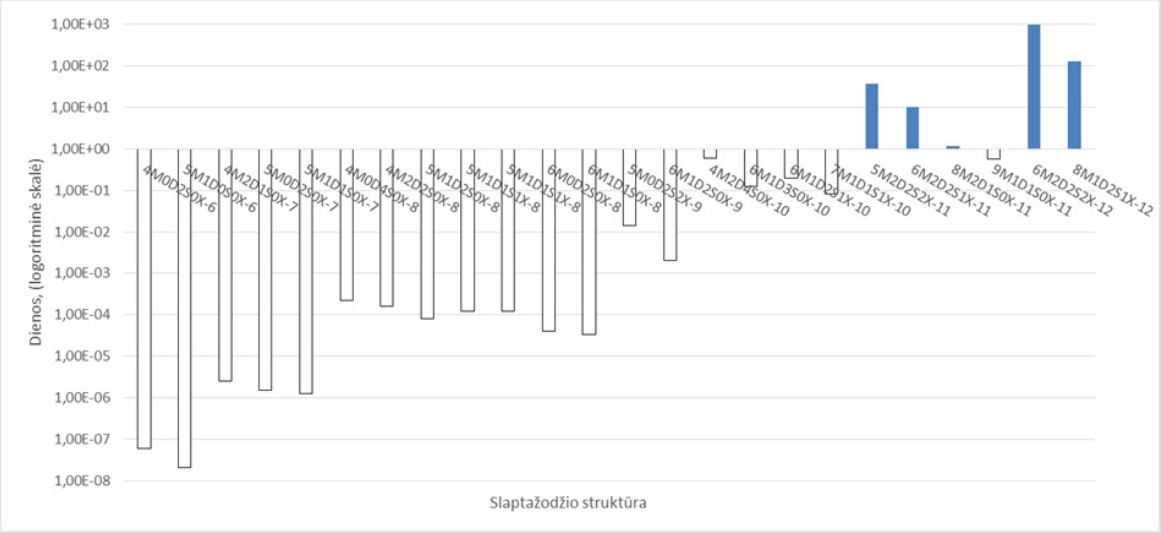
Nors entropija iki galo neparodo slaptažodžio sudėtingumo, bet galima santykinai palyginti dviejų slaptažodžių kombinacijų skaičių, o kartu ir reikiamą perrinkimo laiką. Darbe (2 pav. b) ir kitų autorių tyrimuose (Das ir kt., 2014) pastebėta, kad dažniausiai slaptažodžiai sudaromi naudojant mažąsias raides, įterpiant vieną ar dvi didžiąsias, bei 1-4 skaičius, ir tik retais atvejais, papildant kitais simboliais. 2 lentelėje apskaičiuota keliolika tokių šabloninių slaptažodžių entropijos.

2 lentelė. Įvairių slaptažodžių tipų entropijos

Slaptažodžio struktūra	Entropija
4M0D2S0X-6	30,71
5M1D0S0X-6	29,20
4M2D1S0X-7	36,16
5M0D2S0X-7	35,41
5M1D1S0X-7	35,16
4M0D4S0X-8	42,62
4M2D2S0X-8	42,11
5M1D2S0X-8	41,11
5M1D1S1X-8	41,74
5M1D1S1X-8	41,74
6M0D2S0X-8	40,11
6M1D1S0X-8	39,86
5M0D2S2X-9	48,58
6M1D2S0X-9	45,81
4M2D4S0X-10	54,02
6M1D3S0X-10	51,77
6M1D2S1X-10	52,40
7M1D1S1X-10	51,14
5M2D2S2X-11	59,98
6M2D2S1X-11	58,10
8M2D1S0X-11	54,96
9M1D1S0X-11	53,96
6M2D2S2X-12	64,68
8M1D2S1X-12	61,80

Slaptažodžio struktūros žymėjimas lentelėje ir toliau tekste: aMbDcSdX-Z – M–mažosios raidės; D – didžiosios raidės; S –skaičius; X – specialūs simboliai; a, b, c, d – sveikieji skaičiai parodo kiek nurodytų simbolių yra slaptažodyje, Z – rodo suminį slaptažodžio ilgį.

Pagal 2 lentelėje gautas entropijas suskaičiuotas perrinkimo laikas, priimant aukščiau aprašytas sąlygas. Iš 4 paveikslo galima matyti, kad tipiniai slaptažodžiai iš 10 ir mažiau simbolių yra nepakankamo ilgio – jų perrinkimo laikas mažesnis nei 24 valandos. Saugiais slaptažodžiais galima laikyti tik nuo 11 simbolių ir tik tuos, kurie turi bent 2 didžiąsias raides ir vieną skaičių (8M2D1S0X-11).



4 pav. Slaptažodžio dekodavimo laikas vertinant slaptažodžio struktūrą (struktūros kodavimas aMbDcSdX-Z aprašytas tekste aukščiau)

Žinoma, priklausomai nuo duomenų svarbos 24 valandos perrinkimo laikas gali būti per trumpas, todėl slaptažodžio ilgį reiktų didinti dar 1–2 simboliais iki 13–14.

Išvados

Naudojamų tekstinių slaptažodžių reikalavimai likę tokie pat kaip ir prieš 10 metų, nors yra populiariausia autorizacijos priemonė ir naudojama apie 86% pasaulio IT sistemų. Daugelis vartotojų naudoja ne ilgesnius kaip 6–8 simbolių, paprastus sudarytus tik iš mažųjų ir didžiųjų raidžių rinkinių slaptažodžius, o daugelis vartotojų paskutinį kartą slaptažodį keitė prieš 10 ar daugiau metų. Dažniausiai (74%) iki 8 simbolių ilgis naudojamas ilgis tinkamas tik kontroliuojamose sistemose, t.y. kur galima riboti neteisingų autorizacijų ir bandymų kiekį. Dokumentų, archyvų, asmeninių serverių autorizacijose, kur galima naudoti neribotą užklausų kiekį, slaptažodžiai iki 10 simbolių yra per trumpi. Laikant, kad pakankamas iškodavimo laikas yra 24 val. slaptažodžio ilgis turėtų būti nuo 11–12 simbolių, panaudojant mažiausiai 2 didžiąsias raides ir 1 skaičių.

Literatūra

1. BitDefender (2010) BitDefender Finds Exposed Social Media Credentials Often Provide Access to Email Accounts. Žiurėta 2017-04-02 per internetą: https://www.bitdefender.com/news/bitdefender-finds-exposed-social-media-credentials-often-provide-access-to-email-accounts-1682.html
2. Cluley G. (2016). “Adults’ Media Use and Attitudes Report 2016” ataskaita. Žiurėta 2017-04-02 per internetą: https://www.ofcom.org.uk/__data/assets/pdf_file/0026/80828/2016-adults-media-use-and-attitudes.pdf
3. Das, A., Bonneau, J., Caesar, M., Borisov, N., Wang, X. (2014). The tangled web of password reuse. Symposium on Advanced Information Systems Engineering– 28th International Conference, Springer.
4. Goodin D. (2012). 25-GPU cluster cracks every standard Windows password. Žiurėta 2017-04-02 per internetą: https://arstechnica.com/security/2012/12/25-gpu-cluster-cracks-every-standard-windows-password-in-6-hours/
5. Graham L. (2017) Cybercrime costs the global economy $450 billion: CEO. Žiurėta 2017-04-02 per internetą: http://www.cnbc.com/2017/02/07/cybercrime-costs-the-global-economy-450-billion-ceo.html
6. Munson L . (2014). Average person has 19 passwords – but 1 in 3 don’t make them strong enough. Žiurėta 2017-04-02 per internetą: https://nakedsecurity.sophos.com/2014/10/17/average-person-has-19-passwords-but-1-in-3-dont-make-them-strong-enough/
7. Symantec. (2016). Internet Security Threat Report. Žiurėta 2017-04-02 per internetą: https://www.symantec.com/content/dam/symantec/docs/reports/istr-21-2016-en.pdf
8. TeleSign (2015) Consumer Account Security Report. 2015. Žiurėta 2017-04-02 per internetą: https://www.telesign.com/wp-content/uploads/2015/06/TeleSign-Consumer-Account-Security-Report-2015-FINAL.pdf
9. Walker D. (2015) Authentication Methods Used in the RIPE Database.. Žiurėta 2017-04-02 per internetą: https://labs.ripe.net/Members/kranjbar/authentication-methods-used-in-the-ripe-database.
10. Wei J, Liu W, Hu X. (2016). Secure and Efficient Smart Card Based Remote User Password Authentication Scheme. International Journal of Network Security. Vol.18. No.4. 782-792.

Summary

THE SECURITY ANALYSIS OF TEXT-BASED PASSWORDS

Today’s many authentication for users (approx. 86%) are on text-based passwords. From past many users create passwords with length from 6 until 8 chars based major symbols: lowercase and number. Now already not enough – graphics processing unit (GPU) can to test 350 billion words per second. The password with 8 characters can be detected thru several hours. Many users (21%) changed own password 10 years ago. 40% of users use one password for all accounts. The object of research – text-based passwords. The aim of this research – analyze the security of password for users authentication. The data analysis (N=62 thous.) shown that 74% of password length until 8 symbols and 68% has only lowercase and uppercase. The experiment was made with 5 password groups: lowercase (total 26 symb.); lowercase and number (total 36 symb.); lowercase and uppercase (total 52 symb.); lowercase, uppercase and numbers (total 62 symb.); lowercase, uppercase numbers and other symbol (total 96 symb.). We set 24 hours threshold for “safe password”. The result shown that in brute force attack need minimum 12 symbols with lowercase password; 11 symbols with lowercase and number; 8 symbols length password is enough if are using all 96 possible symbols for password design. Next was calculated entropy of standard passwords which was based by mask. For example: 5 lowercase + 1 uppercase + 2 numbers + 2 other symbol – total password length 10 symbol. The result shown that “safe password” starting only from 11 sign length password with mask: 8 lowercase + 2 uppercase + 1 numbers. If need more secure password - detection time must increase 10 time, the password need increase by 1-2 extra symbols.

Keywords: length of passwords, authorization, hacking, attack.