Dynamika psaní na klávesnici v kombinaci s klasickými hesly

Miloslav Hub
Ústav systémového inženýrství a informatiky, FES, Univerzita Pardubice

Abstract

Authentification as a data security instrument in our informational society is very important for keeping your data as safe as possible. There are 3 types of authentification: authentification by knowledge, by subjekt and biometric authentification. Each one has both advantages and disadvantages. They can be combined to increase the security of your information as well. There are some ideas that suggests reasonable combination of password and keystroke dynamic.

Specifika biometrické autentizace

Princip autentizace prostřednictvím biometriky spočívá v tom, že každá osoba má určité jedinečné vlastnosti svého těla. Těmito vlastnostmi jsou jednak parametry a jednak chování těla. Z tohoto důvodu lze biometriku použít pouze pro autentizaci osob. Narozdíl od ostatních způsobů autentizace osoba nemůže své biometrické charakteristiky zapomenout, ztratit, nemohou být odcizeny (pokud nebereme v úvahu např. úraz, jehož důsledkem je ztráta končetiny). Osoby své biometrické charakteristiky tedy stále „nosí“ s sebou. Biometrika je obecně považována za nejbezpečnější způsob prokázání identity, tedy s nejmenší pravděpodobností výskytu chyb. Biometrická zařízení však často jsou drahá a kladou na uživatele vysoké nároky.

Jinak je tomu v případě autentizace prostřednictvím dynamiky psaní na klávesnici. Jediné zařízení, které je potřeba je zpravidla běžná klávesnice. Uživatelé jsou zvyklí psát na klávesnici a často si ani nemusí uvědomovat, že je jejich dynamika tohoto psaní sledována.

Dosavadní výzkumy v oblasti dynamiky psaní na klávesnici

Užití charakteristik dynamiky psaní na klávesnici navrhl již v roce 1975 Spillan [10]. Prvním pokusem v této oblasti byl však až experiment Gainese a jeho spolupracovníků v roce 1980 [3]. Experiment byl proveden se 7-mi sekretářkami, které psaly texty o délce 300-400 slov, přičemž byla měřena prodleva mezi sousedními znaky (bigramy). Bylo zjištěno, že jednotlivé sekretářky dané bigramy píší podobnou rychlostí a to i v různých textech, což bylo i statisticky dokázáno. Tento pokus měl několik nedostatků, zejména malý počet testovaných osob.

Na Gainesův experiment navázalo v 90. letech mnoho dalších výzkumů [9], [8], [11]. Tyto výzkumy však stále pracovaly s dlouhým textem a měření bigramů. Zaměřeny byly zejména na zvýšení počtu osob, které se pokusu účastnily, na vyřazení nevhodných bigramů, na vytvoření šablony a ošetření extrémních hodnot.

Převratem byl až Gariciův patent [4] založený na jiném přístupu. Předpokladem bylo, že nejvhodnější data pro tuto autentizaci je samotné uživatelské jméno. Toto heslo bylo několikrát napsáno, zprůměrňováno, čímž byla vytvořena šablona (formou vektoru). Měřena byla prodleva mezi stiskem dvou kláves. Autentizovaný uživatel poté napsal své uživatelské jméno a tento vektor byl porovnán s vektorem šablony. Jako míra podobnosti byla zvolena Mahalanobisova vzdálenost. Pokud je vzdálenost vyšší než 100, předpokládá se, že se jedná o narušitele, pokud menší, než 50, tak o uživatele. Je-li vzdálenost 50 až 100, pak je třeba heslo napsat ještě jednou. Pravděpodobnost chyby přijetí je uváděna 0,01 % a chyba odmítnutí 50 %. Garcia se také zmiňuje, že lze v případě potřeby pravděpodobnost výskytu

chyb ovlivnit změnou těchto hladin. Garicia také navrhl i jiný, avšak podobný způsob. Každý uživatel by měl napsat 10-krát nejméně 1000 nejběžnějších slov. Poté by byl uživatel vyzván, aby napsal náhodně vybrané slovo. V praxi je však tento přístup pochopitelně nerealizovatelný.

Dalším přelomem byla práce Younga a Hammona, jejíž výsledkem byl další patent [12]. Na rozdíl od předešlých přístupů nebyly jako identifikační znaky brány pouze doby mezi stiskem jednotlivých kláves, ale také i doby stisku kláves. Porovnání vektoru šablony a přístupu je provedeno prostřednictvím Euklidovské vzdálenosti. Navíc doporučují šifrování těchto dat při přenosu mezi terminálem a serverem. Tito autoři se však zaměřili na průběžnou autentizaci.

Dále následovalo mnoho dalších výzkumů. Mezi nejcitovanější patří výzkumy R. Joice [7] a S. Blehy [1]. R. Joice navrhuje používat kombinaci uživatelské jméno, jméno, příjmení a heslo. Jako kritérium podobnosti volí součet absolutních vzdáleností jednotlivých složek vektorů (vektor šablony a vektor přístupu). S. Bleha navrhuje použití normalizované vzdálenosti a normalizovaného Bayesovského klasifikátoru. Kromě toho ve svém článku [1] rozebírá možnost použití dynamiky psaní na klávesnici k identifikaci. Podle jeho návrhu by uživatel napsal předem daný text a podle rytmiky psaní by systém určil, o jakého uživatele se konkrétně jedná. Lze diskutovat, zda je dynamika psaní na klávesnici pro identifikaci vhodná. Uživatel, který nechce být správně identifikován, může svoji rytmiku záměrně pozměnit. Pokud však chce být uživatel být identifikován správně, je evidentně vhodnější použít uživatelské jméno (login). Oba tito autoři používají jako identifikační znaky dobu mezi stiskem příslušných kláves.

V současné době je jedinou komerčně používanou aplikací pro statickou autentizaci prostřednictvím biometriky psaní na klávesnici program BioPassword [15], patentovaný [13] americkou firmou BioNet Systemes [14]. Tato aplikace používá jako identifikační znaky nejenom dobu mezi stiskem jednotlivých kláves, ale zároveň i dobu stisku příslušných kláves. Navíc umožňuje nastavení hranice kritéria podobnosti vektoru šablony a vektoru přístupu (10 stupňů) a tím regulovat poměr pravděpodobnosti chyby přijetí a chyby odmítnutí. R. Bragg [2] však upozorňuje na některé nedostatky tohoto softwaru.

Model rozhodování v biometrické autentizaci

Úkolem autentizace je rozhodnout na základě předložených identifikačních znaků, zda se jedná o oprávněného uživatele, či o narušitele, který se snaží vydávat za někoho jiného.
Nechť je každý autentizovaný subjekt S_i popsán třemi parametry:

$$S_i = (k_i, k_i^*, \vec{x}_i)$$

k…...skrytý nominální parametr nazývaný „skutečná identita subjektu“
k^*…..měřitelný nominální parametr nazývaný „tvrzení o identitě subjektu“
$\vec{x}$…..měřitelný parametr nazývaný „hodnoty identifikačních znaků“

Úkolem autentizace je rozhodnutí:

$$D : K^* \times \mathbf{X} \rightarrow \{pravda, nepravda\}$$

kde D je množina všech rozhodnutí, K^* je množina všech možných tvrzení o identitě k^* a $\mathbf{X}$ je množina všech možných hodnot identifikačních znaků $\vec{x}$. Rozhodnutí nabývá hodnoty „pravda“, pokud je předpokládáno, že $k_i^* = k_i$. V opačném případě nabývá hodnoty „nepravda“.

Autentizace je tedy rozhodování. Úlohu rozhodování však můžeme převést na úlohu rozpoznávání (konkrétně klasifikaci do dvou tříd) redukcí parametrů $S_i = (k_i, k_i^*, \vec{x}_i)$ na

$S_i = (C_i, \vec{x}_i)$, kde C_i je třída příslušnosti i-tého autentizovaného subjektu, a to $C_i = C^1$, pokud $k_i^* = k_i$ a $C_i = C^2$, pokud $k_i^* \neq k_i$.

Rozhodnutí, zda je subjekt tím, za který se vydává či nikoliv je prováděno na základě určení nepodobnosti vektoru předložených identifikačních znaků $\vec{x}$ a vektoru identifikačních znaků typických pro subjekt, za který se autentizovaný subjekt vydává – šablony $\vec{x}^{Š}$. Tuto nepodobnost označme $d(\vec{x}, \vec{x}^{Š})$. Šablona je vytvořena při evidenci uživatele v systému zpravidla zprůměrňováním několikanásobného předložení identifikačních znaků. Je-li tato nepodobnost vyšší, než d_{max}, pak je rozhodnuto, že se daný subjekt vydává za subjekt jiný, v opačném případě je identita subjektu potvrzena.

Identifikační znaky dynamiky psaní na klávesnici

Biometrika prostřednictvím dynamiky psaní na klávesnici je založena na předpokladu, že každý člověk píše na klávesnici odlišným způsobem, podobně jako je tomu například u vlastnoručního podpisu. Už v 19. století operátoři telegrafu podle této dynamiky dokázali rozeznat, kdo zprávu vysílá [6], [11].

Sledovaných identifikačních znaků během psaní na klávesnici může být mnoho. Snahou je, aby zvolené identifikační znaky co nejlépe odlišily daného uživatele od uživatelů ostatních.

Může to být:

- rychlost psaní na klávesnici
- doba stisku jednotlivých kláves
- doba mezi stiskem jednotlivých kláves
- doba napsání určitého bigramu, trigramu, apod.
- způsob používání speciálních kláves (Shift, Caps Lock,…)
- intenzita stisku jednotlivých kláves
- umístění prstu na jednotlivých klávesách

K zjištění posledních dvou druhů identifikačních znaků je třeba speciální klávesnice.

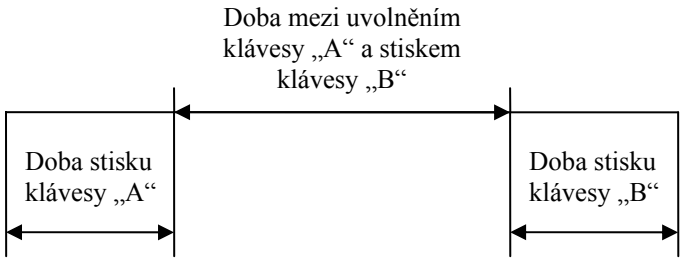

Schéma 1: Doba stisku kláves a doba mezi stiskem kláves hesla "AB"

Současné použití dynamiky psaní na klávesnici a klasických hesel

Použití vícefaktorové autentizace nabízí vyšší úroveň bezpečnosti, než v případě použití pouze jednoho faktoru. A to jak z kvantitativního hlediska, tak i z hlediska kvalitativního. Výzkum bude zaměřen na kombinaci znalostní autentizace prostřednictvím hesel a biometrické autentizaci prostřednictvím biometriky psaní na klávesnici. Jedná se o pozitivní, statickou, kooperativní, zjevnou, obvyklou autentizaci bez obsluhy ve standardním prostředí.

Samotné použití hesel, jak bylo prokázáno [5], poskytuje velmi nízkou úroveň bezpečnosti. Současné použití dynamiky psaní na klávesnici může tuto úroveň zvýšit, což je vhodné zejména pro systémy, v kterých je hodnota chráněných dat vysoká. Tato kombinace se jeví jako velice vhodná, zejména proto, že uživatelé jsou na používání klávesnice zvyklí a nebudou na ně tedy kladeny příliš zvýšené nároky. Jako identifikační znaky budou uvažovány jak doby stisku jednotlivých kláves, tak i doby mezi stiskem daných kláves (na rozdíl od předešlých výzkumů), jelikož se předpokládá, že tak bude získáno více identifikačních dat, než kdyby bylo uvažováno pouze jedno z toho. Další identifikační charakteristiky (např. intenzita stisku klávesy, umístění prstu na klávese,…) snímány nebudou, neboť by takový systém nebylo možno použít bez speciální klávesnice (např. v internetových kavárnách).

Snímání rytmiky psaní hesla však může mít svá úskalí. Uživatelé volí často krátká hesla a proto i množství dat pro rozhodnutí, zda se skutečně jedná o daného uživatele nebo za někoho, kdo se za něho vydává, může být velice malé. Lze tedy předpokládat, že toto rozhodnutí bude zatíženo větší chybou, než by tomu bylo v případě snímání rytmiky psaní dlouhého textu, jak je tomu v průběžné autentizaci. Větší výskyt těchto chyb však eliminuje nutnost znalosti příslušného hesla, tedy druhý faktor této vícefaktorové autentizace – znalost.

Závěr

Současné použití klasických hesel a snímání dynamiky psaní na klávesnici představuje možnost zvýšení bezpečnosti přístupových systémů bez enormního navýšení nákladů. Při snímání rytmiky psaní hesla lze sledovat mnoho charakteristik. Důležité je především měřit ty charakteristiky, které jsou pro danou osobu typické a které jsou odlišné od osob ostatních. Fakt, že hesla bývají zpravidla krátká vyvolává potřebu použít jiné metody, než je tomu v případě průběžné autentizace.

Literatura:

[1] Bleha S., Slivinsky CH., Hussien B.: *Computer-Access Security Systems Using Keystroke Dynamics*. IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 12, No. 12, December 1990.

[2] Bragg R.: *Biometric security products* [on-line]. [cit. 24-10-2004]. Dostupné z <http://www.mcpmag.com/Features/article.asp?EditorialsID=270>.

[3] Gaines R., Lisowski W., Press S., Shapiro, N.: *Authentication by keystroke timing: Some preliminary results*. Rand Report R-256-NFS. Rand Corporation, Santa Monica, CA. 1980

[4] Garica J.: *Personal identification apparatus*. Patent Number 4.621.334. U.S. Patent and Trademark Office. Washington, D.C., 1986

[5] Hub M.: *Bezpečnost znalostní autentizace.* Sborník příspěvků 3. mezinárodní konference doktorandů Partipácia doktorandov na vedecko-výskumnej činnosti, s. 53-57. Bratislava 2003. ISBN 80-225-1700-3.

[6] Ilonen J.: *Keystroke dynamics* [on-line]. Lappeenranta University of Technology, Finland [cit. 10-8-2004]. Dostupné z <http://www.it.lut.fi/kurssit/03-04/010970000/seminars/Ilonen.pdf >

[7] Joice R., Gupta G.: *Identity Authentication Based on Keystroke Latencies*. Technical report #5, Department of Computer Science, James Cook University, Australia. 1989.

[8] Leggett J. Williams, G., Umphress D.: *Verification of user identity via keyboard characteristics*. J.M. Carey, Ed. Ablex Publishing, Norwood, NJ., 1986

[9] Leggett J., Williams G.: *Verifing identity via keyboard characteristics*. Int. J. Man-Machine Studies 23. 1 (Jan. 1988). 67-76 .

[10] Leggett J, Wolliams G., Usnik M.: *Dynamic identity verification via keystroke characteristics*. International Journal of Man-Machine Studies, v36, s. 859-870, Sept. 1990

[11] Umphress D, Williams G.: *Identity verification through keyboard characteristics*. Int. J. Man-Machine Studies 23, 3 (Sept. 1985). 263-273

[12] Young J., Hammon R.W.: *Method and apparatus for verifyng an individual's identity*. Patent Number 4.805.222. U.S. Patent and Trademark Office. Washington, D.C., 1989.

[13] Zilberman A. G.: *Security method and apparatus empoying authentication by keystroke dynamics* (1998) United States Patent 6.442.692

[14] *About BioPassword, LLC*, BioNet Systems, LLC. [on-line]. [cit. 24-10-2004]. Dostupné z <http://www.bionetsystems.com>.

[15] *BioPassword*. [on-line]. [cit. 24-10-2004]. Dostupné z <http://www.biopassword.com>.

Kontaktní adresa:

Ing. Miloslav Hub
Ústav systémového inženýrství a informatiky
Fakulta ekonomicko-správní
Univerzita Pardubice
Studentská 84
532 10 Pardubice
E-mail: miloslav.hub@upce.cz

Recenzoval:

Ing. Jitka Komárková, Ph.D.
Ústav systémového inženýrství a informatiky
Fakulta ekonomicko-správní
Univerzita Pardubice
Studentská 84
532 10 Pardubice
E-mail: jitka.komarkova@upce.cz