Billentyűzési ritmus alapú azonosítás és hitelesítés érintőképernyős mobileszközökön

Keystroke dynamics based identification and authentication for touchscreen-based mobile devices

ANTAL Margit, SZABÓ László Zsolt, LÁSZLÓ Izabella

Sapientia EMTE, Műszaki és Humántudományok kar, Marosvásárhely

manyi@ms.sapientia.ro, lszabo@ms.sapientia.ro, la_bella_6@yahoo.com

Abstract

Currently touchscreen-based smartphones and tablets are widespread. People store more and more sensitive information on these devices, therefore it is essential to strengthen the existing authentication mechanisms. Typing pattern, formally known as keystroke dynamics is useful to enhance the security of password-based authentication. Keystroke dynamics has a long research history. Typing patterns are usually represented using time-based features. Touchscreen-based devices allow the addition of touchscreen-based features such as pressure and finger area. In this paper we examine the effect of these additional touchscreen features on the identification and verification performance through our dataset of 42 users. Performance measurements are reported as accuracies in the case of identification system and using Detection Error Tradeoff (DET) curves for authentication measurements. Results show that these additional features enhance the accuracy of both processes.

Keywords: security, mobile biometrics, touchscreen, authentication, Android.

Összefoglaló

Napjainkban egyre inkább elterjedté válik az érintőképernyős okostelefonok és tabletek használata. A felhasználók is egyre több - biztonsági szempontból érzékeny - adatot tárolnak ilyen típusú eszközökön, ezért nagyon fontos, hogy megfelelő biztonsági mechanizmusokkal lássuk el ezen eszközöket. Az egyik leggyakrabban használt biztonsági mechanizmus a felhasználónév/jelszó páros. Ez a biztonsági mechanizmus megerősíthető, ha nemcsak a jelszó helyességét vizsgáljuk, hanem azt is, hogy a jelszó begépelésének ritmusa megegyezik-e a tulajdonos gépelési ritmusával.

A számítógépes gépelési ritmus egy régóta kutatott terület, viszont nagyon kevés tanulmány létezik, amely ennek alkalmazhatóságát mobiltelefonos biztonsági mechanizmusokban elemzi. Dolgozatunkban azt vizsgáljuk meg, hogy az érintőképernyőről származó jellemzők milyen mértékben javítják a billentyűzési ritmus alapú biometrikus rendszerek teljesítményét. A teljesítmények méréséhez különböző osztályozó algoritmusok segítségével kialakított biometrikus rendszereket használunk, az eredményeket a biometriában szokásos Detection Error Tradeoff (DET) görbék segítségével szemléltetjük. A dolgozatunk fő következtetése, hogy az érintőképernyős jellemzők, mint például a nyomás és az érintett felület nagysága, nagymértékben javítja a billentyűzési ritmus alapú biometrikus rendszerek teljesítményét.

Kulcsszavak: biztonság, mobil biometria, érintőképernyő, autentikáció, Android.

1. Bevezetés

A számítógépes biztonság egyik alterületének tekinthető a mobil eszközök biztonsága, amelynek fontossága egyre inkább növekszik a mobil eszközök elterjedtségének köszönhetően. Mivel nagyon sokan használnak mobil eszközt érzékeny műveletek lebonyolítására, ezért nagyon fontos, hogy ezeket megfelelő biztonság mellett végezhessék. Ha például idegen kézbe kerül egy okostelefon, amely történetesen tárolt valamely rendszerbe való belépéshez szükséges jelszót, akkor a készülék

jelezze, hogy nem a tulajdonos kezében van az eszköz és akadályozza meg a rendszer használatát.

A billentyűzési ritmus egy régóta kutatott terület. Már a 19. században a távirdászok tudták egymást azonosítani a gépelési ritmusuk alapján. Az ember gépelési mintája egy viselkedési jellemző, amely egy bizonyos tanulási idő után stabillá válik. A kutatások nagy része klasszikus számítógép billentyűzeten vizsgálta a billentyűzési ritmus alkalmazhatóságát felhasználó azonosítására, illetve hitelesítésére. A billentyűzési mintákból különböző idő-alapú jellemzőket vontak ki, illetve különböző statisztikai és gépi tanulási modelleket használtak a felhasználók azonosítására. Erről egy részletes áttekintést a Teh és tsai. tanulmányában [10] olvashatunk.

A billentyűzési ritmus alkalmazhatóságát mobileszközökre még nagyon kevés tanulmány vizsgálta. A létező tanulmányok főként PIN kódok gépelésének ritmusát tanulmányozták. Kutatásunkban arra a kérdésre keressük a választ, hogy az érintőképernyő-alapú jellemzők, mint a nyomás vagy az érintett felület mérete, javítják-e a billentyűzési ritmus alapú biometrikus rendszerek pontosságát?

2. Biometria és billentyűzési ritmus

A biometrikus rendszerek két üzemmódban működhetnek: azonosítási, illetve hitelesítő (autentikáció, ellenőrzés) üzemmódban.

Hitelesítés esetében a rendszer azt ellenőrzi, hogy a személy tényleg az-e, akinek vallja magát. Ebben az esetben a személytől származó aktuális mintát csak a hitelesítést végző személy tárolt mintáival hasonlítja össze a rendszer (egy-az-egyhez hasonlítás), majd ennek alapján eldönti, hogy elfogadja, illetve elutasítja a személyt. Nagyon sok biometrikus rendszer ezt a feladatot egy két osztályos osztályozással valósítja meg. Ebben az esetben a pozitív osztályt a felhasználótól származó minták alkotják, a negatívat pedig egy külvilágot modellező osztály alkotja, amelyet leggyakrabban az összes többi felhasználó bizonyos kritérium szerint kiválasztott mintái alkotják.

Azonosítási üzemmódban a rendszer a személytől származó mintát az összes felhasználóval összeveti, majd eldönti, hogy mely ismert felhasználóhoz hasonlít leginkább a minta. Ebben az esetben egy-a-sokhoz hasonlítás történik, tulajdonképpen az aktuális személytől érkező mintát a rendszer besorolja valamely, a rendszer által ismert osztályba (több osztályos osztályozási feladat)

Biometrikus rendszerek hatékonyságának mérésére több mutatót is használnak. A két legfontosabb mutató a *hibás elfogadási arány* (FAR –False Acceptance Rate), amely azt jelzi, hogy az illetéktelen behatolókat a rendszer milyen valószínűséggel fogadja el, illetve a *hibás visszautasítási arány* (FRR – False Rejection Rate), amely a hiteles felhasználó elutasításának valószínűségét jelzi.

A billentyűzési ritmus alapú biometrikus rendszerek két nagy előnnyel is rendelkeznek a többi biometrikus rendszerrel szemben. Az első és legfontosabb előny az alacsony költség, hiszen nem szükséges drága hardver eszköz a minták vételezéséhez. A másik nagy előnyük, hogy nemcsak belépési ponthoz kötött rendszerekben lehet például a jelszó gépelési ritmusát figyelni, hanem folyamatos ellenőrzésre is használható bizonyos, gépelést igénylő alkalmazások esetén.

A legfontosabb két jellemző, amit gépelési mintákból nyernek ki, a billentyű leütése és felengedése között eltelt idő, illetve két egymást követő billentyű leütése között eltelt idő. Ezen kívül szokás még használni három vagy több billentyű leütése közötti időt, amit n-graph-nak neveznek, ahol n az egymást követő billentyűk száma. Nagyon sok formafelismerésben használt statisztikai, illetve gépi tanuláson alapuló módszert alkalmaztak billentyűzési ritmus felismerésére is.

Dolgozat	#személyek	#minták/személy	Jelszó	Jellemzők	Eredmény(ek)
[9]	10	100	4 számjegy	$H+P_{min}+P_{max}$	EER: 15.2%
[11]	152	10	17számjegy	H+P+FA+UD	FAR: 4.19%; FRR: 4.59%
[5]	13	-	-	H+ P+FA+ egyebek	FAR: 14.0%; FRR: 2.2%
[8]	10	30	10 számjegy	P	EER: 1%

1. táblázat Érintőképernyős eszközöket használó biometrikus rendszerek hibái (a jellemzők magyarázata a 2. táblázatban)

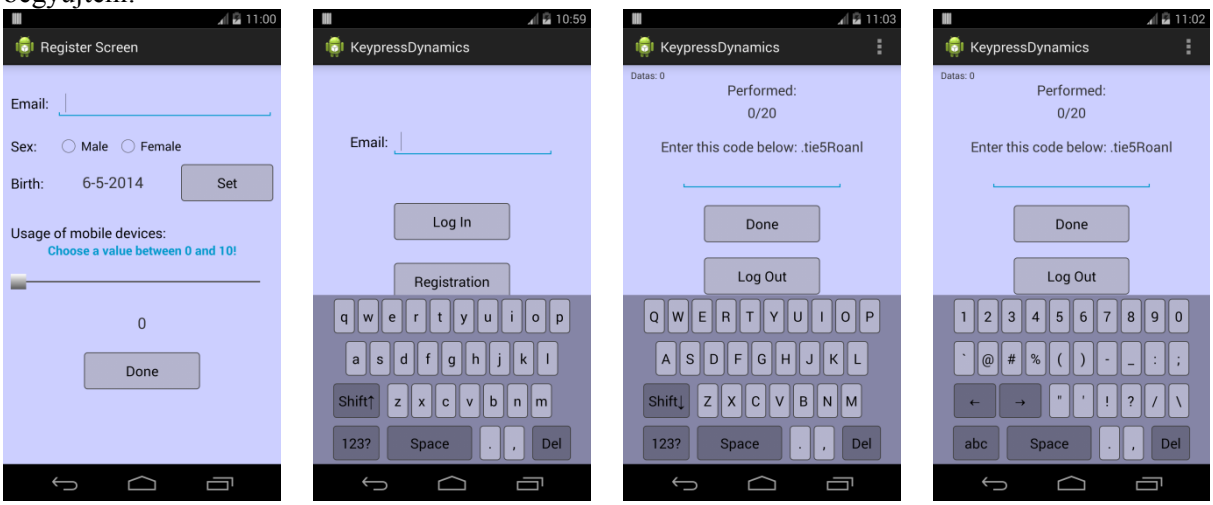
Az 1. táblázatban összefoglaltuk azon tanulmányokat, amelyek érintőképernyős eszközökön gyűjtött adatokon végezték a méréseket. Látható, hogy a tanulmányok csak számjegyeket tartalmazó

jelszavakat használtak, tehát valós jelszó esetében még nem mérték az érintőképernyős jellemzők (nyomás, érintési terület nagysága) hatását a billentyűzési ritmus alapú hitelesítési rendszerek teljesítményére vonatkozóan.

3. Módszertan

3.1 Adatgyűjtés

Az adatgyűjtést egy saját fejlesztésű Android alkalmazással végeztük, amely saját szoftverbillentyűzetet használt a gépelési adatok naplózásához. Az alkalmazás legfontosabb képernyői az 1. ábrán láthatók Az adatgyűjtésben 42 személy vett részt, mindenki beleegyezett abba, hogy a gépelési mintáit felhasználjuk kutatási célokra. A 42 személy közül 24 férfi és 16 nő volt, az átlagéletkor pedig 22.2 év. A gépelési ritmust több tényező is befolyásolja, ezért ajánlott a mintákat több szesszióban gyűjteni. Az adathalmaz gyűjtése két hetet vett igénybe és minden felhasználó legalább két különböző alkalommal szolgáltatott adatot. Felhasználónként 51 helyes mintát sikerült begyűjteni.



1. ábra (a) Regisztráció; (b) Bejelentkezés; (c) Adatgyűjtés – alfabetikus billentyűzet; (d) Adatgyűjtés – numerikus billentyűzet.

Két típusú eszközt használtunk az adatok begyűjtésére: egy Nexus 7 típusú tabletet és egy LG Optimus L7 II P710 típusú okostelefont. 37 felhasználó a tableten, illetve 5 felhasználó a telefonon szolgáltatott adatot. Minden felhasználó ugyanazt az erősnek minősített jelszót gépelte be: **.tie5Roanl**. Ezt a jelszót használta Killourhy és tsai. is egy 2009-ben megjelent tanulmányban [7]. Célszerű, ha minden felhasználó ugyanazt a jelszót gépeli be, mert így lehetővé válik hamis elfogadási arány (FAR) típusú hiba mérése is. A jelszó begépelése 14 billentyű leütését tette szükségessé, amelyek a következők voltak: 8 betű, egy számjegy, a . karakter, kétszer a Shift billentyű (nagy- és kisbetű közötti váltás), illetve kétszer a numerikus billentyűzetre váltó billentyű lenyomása (numerikus/alfabetikus billentyűzet közötti váltás). A billentyű lenyomásakor naplóztuk az időbélyeget, a nyomást és az érintett felület nagyságát, felengedéskor pedig csak az időbélyeget. A mérésekhez használt jellemzőket a 2. táblázat szemlélteti.

Jellemző	Magyarázat	Jellemzők száma
H: hold time	Ugyanazon billentyű leütése és felengedése között eltelt idő.	14 (összesen 14 billentyű leütését tette szükségesé a jelszó begépelése)
DD: down-down time	Két egymást követő billentyű leütése között eltelt idő.	13 (14 billentyű leütése között összesen 13 eltelt időt tudunk mérni)
UD: up-down time	Két egymást követő billentyű leütési és felengedési időpontjai között eltelt idő.	13 (14 billentyű leütése között összesen 13 eltelt időt tudunk mérni)
P: key hold pressure	A billentyű lenyomás pillanatában mért nyomás.	14
FA: finger area	Az érintési felület nagysága a billentyű lenyomás pillanatában.	14
AH: aveargе hold time	A 14 lenyomási idő átlaga.	1
AP: average pressure	A 14 nyomásérték átlaga.	1
AFA: average finger area	A 14 értintési felület átlaga.	1
Összesen		**71**

2. táblázat A felhasznált jellemzők és ezek jelentése.

3.2. Osztályozó algoritmusok

Az osztályozási algoritmusokhoz a Weka (3.7. verzió) [12] gépi tanulási programcsomag osztálykönyvtárait használtuk. Az osztályozási algoritmusok paramétereit a Weka keresési algoritmusaival optimalizáltuk.

A billentyűzési ritmusról készült publikációkban sokféle osztályozási algoritmussal kísérleteztek. Egy áttekintést ezekről az algoritmusokról és az ezekkel elért eredményekről a Teh és tsai. tanulmányában [10] olvashatunk. Dolgozatunkban azokkal a Weka programcsomagban implementált osztályozási algoritmusokkal dolgozunk, amelyek adathalmazunkra a legjobb osztályozási eredményeket nyújtották.

A Naive Bayes egy valószínűségi osztályozó, amely Bayes valószínűségi tételén alapszik. Ez az osztályozó feltételezi, hogy a minták jellemzői egymástól függetlenek, ami a valóságban általában nem igaz. Ennek ellenére, ezzel az osztályozóval számos valós osztályozási problémában jó eredményeket lehet elérni.

A Bayes háló szintén egy valószínűségi modell, amely a valószínűségi változók halmazát és ezek közötti feltételes függőségeket egy irányított, körmentes gráffal ábrázolja [2]. A gráf csomópontjai a valószínűségi változóknak felelnek meg, míg élek kötik össze azon csomópontokat, amelyek között függőségi viszony van.

A k-NN (k-legközelebbi szomszéd, IBk a Weka programcsomagban) egy minta alapú osztályozó, amelyben egy minta besorolása a k legközelebbi szomszédok osztályai alapján történik. A mi adathalmazunkra k=1 értékre kaptuk a legjobb eredményeket.

A döntési fák nagyon népszerű osztályozók, de billentyűzési ritmus felismerésére csak az utóbbi években kezdték alkalmazni. Az algoritmus előnyei közé tartozik, hogy mind a tanítási, mind pedig a tesztelési időigény alacsony. Dolgozatunkban a C4.5 (Weka J48) algoritmust használtuk, 0.2 konfidencia tényezővel, illetve legkevesebb 4 minta/levél beállításokkal.

A Random forests (véletlen erdők) [3] egy együttes tanulási algoritmus (ensemble learning method), amely több döntési fát épít. Egy minta besorolása többségi szavazáson alapszik. Az algoritmus paraméterezhető a használt fák számával, amelyre mi 100-at használtunk, illetve a döntési fák maximális mélységével is.

Az SVM (tartóvektor-gép) osztályozó egy hipersíkot határoz meg, amely segítségével az adatok lineárisan elválaszthatók. Amennyiben az adatok lineárisan nem szeparálhatók, egy transzformációt végez az adatokra úgy, hogy a transzformált térben az adatok lineárisan szeparálhatók legyenek. Dolgozatunkban a LibSVM implementációt [4] használtuk, RBF típusú kernellel. A kernel paramétereinek becslését rács-alapú kereséssel határoztuk meg. A paraméterek a 41 jellemzős adathalmazra $C=10.55$, $\gamma=1.86$, míg a 71 jellemzős adathalmazra $C=7.46$, $\gamma=0.25$. Mérés előtt a jellemzőket normalizáltuk.

Az MLP (multilayers perceptrons, többrétegű perceptron modell) olyan mesterséges neuronháló, amelyet backpropagation eljárással tanítunk. Az osztályozó egyik paramétere a szintek száma, amelynek a Weka programcsomagban beállított alapértelmezett értékét használtuk: (attribútumok száma + osztályok száma)/2.

4. Mérési eredmények

4.1. Azonosítás (osztályozás)

Annak érdekében, hogy megválaszolhassuk a kutatási kérdést, a méréseket két adathalmazon végeztük. Az első adathalmaz 41 jellemzőt (H+DD+UD+AH), illetve a második 71 jellemzőt (H+DD+UD+AH+P+FA+AP+AFA) tartalmaz. Látható, hogy a második részhalmazként tartalmazza az elsőt, kiegészítve azt érintőképernyős jellemzőkkel. A mérésekhez használt két adathalmaz nyilvánosan is elérhető a http://www.ms.sapientia.ro/~manyi/keystroke.html címen.

Az osztályozási mérési eredményeket a 3. táblázatban foglaltuk össze. Az eredmények átlagértékek, amelyeket 10 darab 10 részes rétegzett keresztvalidációs (10-fold stratified cross-validation) futtatásból számítottunk. Ez összesen 100 futtatást jelent. Az átlagos pontosság mellett zárójelben feltüntettük a szórást is.

Osztályozó	Pontosság átlag (szórás) H+DD+UD+AH (41 jellemző)	Pontosság átlag (szórás) H+DD+UD+AH+P+FA+AP+AFA (71 jellemző)
Naive Bayes	50.15% (2.86)	78.93% (2.63)
Bayes háló	75.95% (2.65)	91.94% (1.73)
C4.5 (J48)	54.79% (3.84)	69.02% (3.32)
k-NN (IBk)	41.07% (2.83)	72.98% (2.25)
SVM (LibSVM)	61.71% (3.22)	88.33% (1.87)
Random forests	82.53% (2.53)	93.04% (1.65)
MLP	53.01% (3.39)	86.26% (2.19)

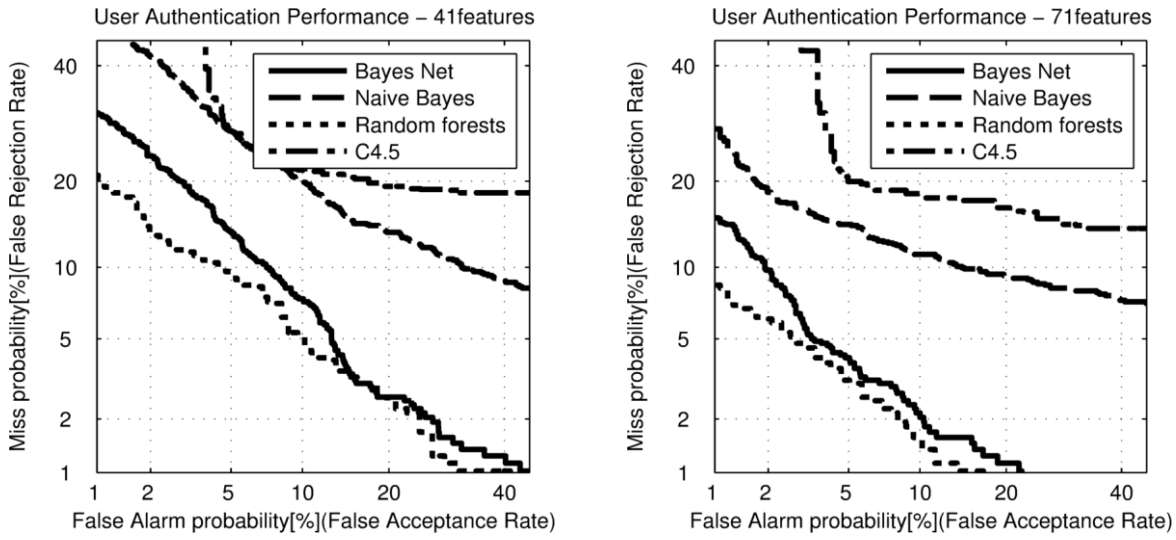
3. táblázat Osztályozási pontosságok a 41, illetve a 71 jellemzős adathalmazokra.

4.2. Hitelesítés (autentikáció)

Hitelesítés során a rendszer a bemeneti minta alapján a felhasználót vagy elfogadja, vagy pedig elutasítja. Amennyiben osztályozási feladatként tekintjük a kérdést, akkor ez egy két osztályos osztályozás, amelyben a pozitív osztály az adott felhasználó osztálya, a negatív osztály pedig a külvilág. Az elfogadás azt jelenti, hogy a mintát a pozitív osztályba, az elutasítás pedig azt, hogy a mintát a negatív osztályba sorolja be az osztályozó.

Az autentikációs mérésekhez két-osztályos osztályozásokat végeztünk. Minden egyes felhasználóra külön tanító- és teszt-adathalmazokat alkottunk. A tanító halmaz a felhasználótól származó minták 3/5-ét tartalmazta, illetve az összes többi felhasználó mindenikétől 3 mintát. A teszt halmaz tartalmazta a maradék 2/5 adatot a felhasználótól, illetve az adathalmazban szereplő összes többi felhasználó mindenikétől 2 mintát. A kiválasztott felhasználó a pozitív osztály, az összes többi felhasználó lesz a második osztály a kétosztályos osztályozásnál, ez képviseli a negatív osztályt, vagyis a külvilágot. Mivel összesen 42 felhasználónk van és 51 minta felhasználónként, a tanító halmaz 30+41*3= 153 mintát tartalmazott és a teszt halmaz pedig 21+41*2=103 mintát.

Bizonyos osztályozók nemcsak besorolják az osztályozandó mintát a legmegfelelőbb osztályba, hanem azt is megadják, hogy a minta milyen valószínűséggel tartozik a pozitív, illetve a negatív osztályokhoz. Felhasználva ezen adatokat, ROC típusú görbéket (Receiver Operating Curve) [6] állíthatunk elő, amelyek szépen szemléltetik a rendszerek kétféle hibáját, illetve az ezek közötti viszonyt is. A biometrikus rendszerek esetén ROC görbe helyett DET (Detection Error Tradeoff) [1] típusú görbéket szokás használni, amely a FAR és FRR típusú hibák közötti viszonyt ábrázolja, továbbá a görbéről leolvasható az EER is, amely a DET görbe és az első negyed szögfelezőjének metszéspontjánál van. A hibagörbék alapján sokkal jobban összehasonlíthatók a különböző biometrikus rendszerek, mint például ha csak az osztályozási hibát (vagy pontosságot) használnánk.



2. ábra DET hibagörbék (a) 41-jellemzős adathalmaz (b) 71-jellemzős adathalmaz

A 2. ábra a 41 és a 71 jellemzős adathalmazok hibagörbéit szemlélteti 4 különböző osztályozó használata mellett. Látható, hogy a hibagörbék alapján is ugyanaz a sorrend az osztályozó algoritmusok között, mint az előző alfejezetben kapott többosztályos osztályozási pontosság alapján. A legjobbnak a Random forests bizonyult, ezt követi a Bayes háló alapú osztályozó. Az is látható, hogy az érintőképernyő alapú jellemzők ezen két osztályozó esetében az egyenlő hibaarányt (EER) felére csökkentik (41 jellemző: 8%, 71 jellemző: 4%).

5. Következtetések

Dolgozatunkban azt vizsgáltuk, hogy milyen hatással van az érintőképernyős jellemzők bevezetése a billentyűzési ritmus alapú azonosításra és autentikációra. Az azonosítási és autentikációs méréseket a Weka programcsomagra épülő programokkal végeztük, amelyekben 7-féle osztályozót alkalmaztunk: SVM, C4.5, Random forests, MLP, k-NN, Bayes net, Naive Bayes. A kipróbált osztályozók közül a Random forests és a Bayes háló alapúak nyújtották a legjobb teljesítményeket. Mind az azonosítási, mind pedig az autentikációs mérések azt igazolják, hogy ezen jellemzők jelentősen megnövelik a billentyűzési ritmus alapú rendszerek pontosságát.

Köszönetnyilvánítás

A publikáció elkészítését a TÁMOP-4.2.2.C-11/1/KONV-2012-0004 számú projekt támogatta.

Hivatkozások

[1] Alvin F. Martin, George R. Doddington, Terri Kamm, Mark Ordowski, Mark A. Przybocki: The DET curve in assessment of detection task performance. EUROSPEECH 1997

[2] Bouckaert R R. Bayesian Network Classifiers in Weka. 2008. http://www.cs.waikato.ac.nz/~remco/weka.bn.pdf

[3] Breiman. L. Random forests, Machine Learning 45(1) , 2001, p. 5-32.

[4] Chang CC, Lin C.-J. LIBSVM : a library for support vector machines. ACM Transactions on Intelligent Systems and Technology, 2:27:1--27:27, 2011.

[5] Draffin B, Zhu J, Zhang J, KeySens. Passive User Authentication through Micro-behavior Modeling of Soft Keyboard Interaction, In: Lecture Notes of the Institute for Computer Sciences, Social Informatics and Telecommunications Engineering Volume 130, 2014, p. 184-201.

[6] Fawcett T, An introduction to ROC analysis. Pattern Recognition Letters 27(8), 2006, p. 861-874.

[7] Killourhy K S, Maxion R A. Comparing anomaly detection algorithms for keystroke dynamics, in Proceedings of the IEEE/IFIP International Conference on Dependable Systems and Networks (DSN '09), 2009, p. 125–134.

[8] Saevanee H, Bhattarakosol P, Authenticating user using keystroke dynamics and finger pressure, 6th IEEE Consumer Communications and Networking Conference, 2009, p. 1-2.

[9] Sen S, Muralidharan K. Putting Pressure on Mobile Authentication, Seventh International Conference on Mobile Computing and Ubiquitous Networking (ICMU), 2014, p. 56-61.

[10] Teh P S, Teoh A B J, Yue S. A Survey of Keystroke Dynamics Biometrics, The Scientific World Journal, vol. 2013, Article ID 408280, 2013, 24 pages.

[11] Trojahn M, Arndt F, Ortmeier F. Authentication with Keystroke Dynamics on Touchscreen Keypads - Effect of different N-Graph Combinations. In: MOBILITY 2013, The Third International Conference on Mobile Services, Resources, and Users, 2013, p. 114-119.

[12] Witten I H, Frank E, Hall M, Data Mining: Practical machine learning tools and techniques, Morgan Kaufmann Publishers, 2011.