A Literature Survey on Authentication Using Behavioural Biometric Techniques

Rajvardhan Oak

Abstract With technological advancements and the increasing use of computers and internet in our day to day lives, the issue of security has become paramount. The rate of cybercrime has increased tremendously in the internet era. Out of the numerous crimes, identity theft is perhaps the one that poses the most dangers to an individual. More and more voices strongly declare that the password is no longer a reliable IT security measure and must be replaced by more efficient systems for protecting the computer contents. Behavioural biometrics is an emerging technology that resolves some of the major flaws of the previous scheme. This paper is the first stage of a project which aims to develop a novel authentication system using behavioural biometrics. It presents a comprehensive survey of various techniques and recent works in the respective fields.

Keywords Information security · Biometrics · Authentication mechanism

1 Introduction

Biometrics can be characterized basically as the estimation of human qualities. Biometric identifiers are then unmistakable, quantifiable qualities used to name and depict people [1]. Famous cases of biometric confirmation are retina examinations, fingerprint tests and DNA tests.

Biometrics verification is a strategy used by coordinating an individual's hereditary attributes or behavioural qualities with information that have already been learned, enlisted into a layout and organized in a framework database or on a token [2]. It can likewise be characterized as the idea of recognizing oneself by something that you know, something that you have, or something that you are [3]. That is, it identifies the innate qualities.

R. Oak (✉)
Department of Computer Engineering, Pune Institute of Computer Technology, Pune, India
e-mail: rvoak@acm.org

S. Bhalla et al. (eds.), *Intelligent Computing and Information and Communication*,
Advances in Intelligent Systems and Computing 673,
https://doi.org/10.1007/978-981-10-7245-1_18

For a parameter to be called a biometric identifier, it must satisfy some properties [1, 4]:

(1) Universality: Every person possesses that particular characteristic. For example, DNA can be called a biometric parameter, whereas birthmarks cannot.
(2) Uniqueness: The characteristic is different for every person. For example, while a fingerprint is a biometric, eye colour and blood group are not.
(3) Permanence: The characteristic does not disappear or change with time. For example, DNA is biometric but hormone levels are not.
(4) Collectability: It is possible to obtain readings of the characteristic using sensors in a feasible, fast and highly accurate manner. For example, voice can be called a biometric characteristic, but factors such as confidence and self-esteem cannot.
(5) Circumvention: The parameter is forgery-proof and it is nearly impossible to replicate it.

Biometric parameters may be classified into two types: Physiological and behavioural biometrics [1]. Physiological characteristics are those that are anatomic and biological properties of an individual. They include fingerprints, facial recognition, iris scan, voice recognition, palm veins, DNA, etc. [5]. These are the traditional means by which an individual's identity is verified.

Behavioural biometrics, on the other hand, refers to factors such as gait, GUI interaction, Haptics [6], programming style, registry access, system call logs, mouse dynamics, etc. [7, 8]. It depends on an individual's inward qualities and attributes [9].

The advantages of this technique over other customary biometric approaches are as follows [3, 5, 7]:

(1) It provides persistent security. The authentication process is not complete after the login, but there is continuous monitoring.
(2) Behaviours can be collected surreptitiously without alerting the user.
(3) No special hardware is necessary to identify the behaviours.
(4) It is difficult to replicate the behaviour, hence making identity theft less likely.

2 Classification of Behavioural Biometrics

Depending on the nature of the parameters collected and evaluated, behavioural biometric mechanisms can be subdivided into five types (Fig. 1).

(1) Authorship Based: It is based on the analysis of a work produced by the user. The system identifies styles and characteristics particular to a user as he writes/draws and verification is done based on the matches of these characteristics.
(2) HCI Based: It is based on the traits and mannerisms exhibited by the user while interacting with the system such as mouse movements or touchscreen strokes.

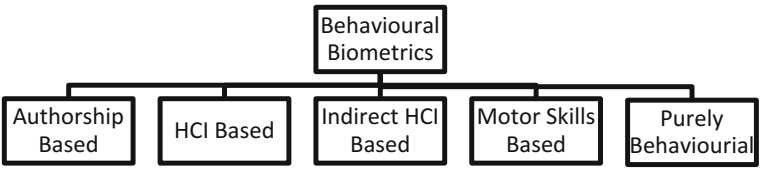

Fig. 1 Classification of behavioural biometrics

Every person applies different strategies, shortcuts and conventions while using a computer. A collection of such interaction particularities serves as a base for constructing a user identity and for authentication.

(3) Indirect HCI: It is very similar to HCI-based interaction. In indirect HCI-based systems, the system monitors the effects of the normal HCI actions. All user actions leave certain low-level digital evidence in the system in the form of system call traces, audit logs, execution traces, call stack analysis, etc.

(4) Motor-Skills based: It refers to the way in which the user utilizes the muscles. It is presumably the best inquired about of all behavioural biometric techniques. Human movements are a combination of muscle action, bone action and impulses travelling in the nervous system.

(5) Purely Behavioural: It is not based on the body part metrics or intrinsic behaviour. Rather, it is based on the fact that human beings utilize different strategies, innovative ideas, critical thinking and creative thinking. It attempts to quantify such traits and use them for authentication.

3 Evaluation Metrics

The following parameters are used as a measure of the effectiveness of the biometric systems:

(1) *FAR*: It is the false acceptance rate [5]. It is the ratio of the amount of attack instances incorrectly labelled as authentic to the total number of attack instances [5, 10]. Value of FAR must be as small as possible. It gives a measure of the percentage of attacks that could not be identified by the biometric system.

(2) *FRR*: It is the false rejection rate. It is the ratio of the amount of authentic interactions that were incorrectly classified as attacks and the total number of authentic, valid instances [10]. It gives a measure of the percentage of authentic interactions that were incorrectly classified as intrusive.

(3) *ROC*: It is the Receiver/Relative operating characteristic. It is a plot which represents a compromise between FAR and FRR. The matching algorithm of a biometrics mechanism has an established threshold which evaluates how close to the learned format an instance must be so that it qualifies as authentic [5]. Increasing the threshold reduces the FAR but increases the FRR. Decreasing

the threshold will reduce FRR but lead to higher FAR. The ROC helps to identify the optimum threshold so that both FAR and FRR are minimized.

(4) *EER/CER*: At this rate, both acceptance and rejection rates are numerically the same [5, 10]. ROC gives the value of EER. Ideally, the value of EER should be small [5].

4 Survey of Techniques

4.1 Keystroke Dynamics

It is a type of HCI-based behavioural biometric parameter. The keyboard is the primary input device used by humans to interact with a system. Different individuals have varied characteristics with respect to the speed of typing, error rate, use of certain key combinations, use of the touch typing method, etc. [10]. As a result of these differences, it is possible to verify the identity of the user. For example, the touch typing method is used by veteran typists, whereas novices use a hunt-and-peck technique which utilizes just two fingers. [7, 10]. Keystroke dynamics are based on the two important parameters: Flight Time and Dwell Time [8, 10]. Flight time is the time gap between releasing a key and pressing the next one, and the latter is the time for which a key is pressed. A large amount of research regarding keystroke dynamics for verification has been carried out. Bartolacci in 2005 and Curtin in 2006 have studied keystroke dynamics for long text analysis [4, 7, 10]. A study for email authorship identification was carried out by Gupta et al. in 2005. In [11], the author constructs digraphs (consisting of two adjacent characters) which he classifies into seven different categories and calculates mean latency for each category. This gives a measure of the programming experience of the user [11]. The authors in [12] have deduced that keystroke on integrating it with accelerometer biometrics, has a false acceptance rate of just 7%.

4.2 Mouse Dynamics

The mouse is probably the most important device after the keyboard which aids humans in interacting with the Graphical User Interface (GUI). In general, mouse dynamics refer to the characteristics of different individuals to use different pointing devices like mouse and light pen in different manners [13]. A number of different mouse gestures can be analyzed, such as single clicks, double clicks, scrolling, drag and drop and stillness [7]. To acquire the features, a software program intercepts the low-level events occurring because of mouse dynamics, along with associated attributes such as timestamps and cursor coordinates [10, 13]. At the high level, the gathered information would incorporate abnormal state itemized data about the

GUI-related activities of the client, for example, left the tap on the begin menu, double tap on explorer.exe, close notepad.exe window, and so forth [7]. Other factors such as velocities in a horizontal, vertical and tangential direction along with the angular direction, tangential acceleration and jerk also form a part of the captured data [10]. From this event log, various statistical and kinematic features are extracted which are used to build a user profile. Pusara and Brodley in 2004 have proposed an approach in which split the mouse event data is classified into movements of the mouse wheel and clicks on different entities on the screen [7]. Gamboa and Fred in 2003 have described an approach in which identification and authentication of humans are carried out by analyzing the human–mouse interaction in online gaming. In [14], mouse movements were captured as functions of timestamps and graphical coordinate values and analysis was done using support vector machines with an error rate of 1.3%.

4.3 Haptics

It identifies with the feeling of touch. Nowadays, intelligent cell phones are equipped with various sensing elements such as an accelerometer, gyroscope, computerized compass and high-resolution camera [6, 9]. As a result, it is possible to measure several physical quantities during use. The parameters measured are direction, pressure, force, angle and position of the user's interactions [7]. In [15], a biometric authentication system based on the haptics was built by integrating it with fuzzy logic. A combination of three factors such as hold-time, inter-key behaviour and finger pressure was proposed in [16]. As much as 30 behavioural features were proposed by Frank et al. (2012) [17]. Xu et al. (2014) have proposed a continuous and passive mechanism which achieved an error of less than 1% [15]. Sitova et al. (2015) introduced Hand Movement Orientation and Grasp (HMOG)-based system. Furthermore, Buriro et al. (2016) have developed a system which analyses micro-movements of a phone, and the exact points on the screen which are pressed [15]. In [18], a characteristic and consistent metric is known as Index of Individuality has been proposed which uses Gaussian Process Models to capture data.

4.4 Gait

It refers to a person's way of walking about. It is a muscle controlled biometric parameter. In gait-based biometrics, parameters such as kinematic patterns, knee ankle movements, moments, angles, hunch, etc. [19, 20]. It is a complex spatio-temporal activity which permits biometric identification of individuals at a distance generally via recorded video [7, 19]. The capture of data may be carried out by floor sensors, machine vision systems, or wearable sensors [20]. Gait is a

factor which is subject to several variations from person to person depending on age, gender, bone density, waddling, muscle strength, fat percentage and energy level [7]. As a result, this is one of the biometric parameters which is nearly impossible to replicate. Typical features may include arm swing, walking speed, stooping of the back, step size, head-foot distance and head-pelvis distance [7, 19]. By using three different approaches of signal correlation, frequency domain and data distribution statistics [15, 21], it was found by Mantyjarvi et al. (2005) that the lowest error of 7% was achieved with signal correlation method [21]. Gafurov et al. (2006) developed a method to identify the person using an accelerometer attached to the leg at an error rate of 10%. Derawi et al. (2010) attached the accelerometer to the hip and an error of 20% was seen. In [22], a gait-based WiFi signature system was proposed using a simple Naïve–Bayes classifier with a correct identification rate of 87%. Cola et al. [23] a device is worn on the wrist and identification was done with an error rate of 2.9%.

4.5 Log Files

Operating systems generally maintain exhaustive log files which contain records of every small activity initiated by the user. Such log file entries contain fields such as the identity of the user who fired a command, the timestamp, CPU usage and other associated parameters [7]. In a system based on the audit files, there is a high chance of false positives due to routine, legal activities such as adding new users, changing network settings or change in permissions [7]. Hence, in these systems, there is an overhead of informing the authentication program of such possibilities. Network level logs which maintain traffic and various attributes such as protocol, sequence number, length, correction checksum, etc. serve to identify intruders in the system. In the training phase, a profile is built which identifies certain behaviours as normal. A field known as ‘alert flag’ is set if any abnormal activity is detected. In [24], the authors describe a five-step process: (i) Formatting data, (ii) Compare degree of similarity, (iii) Clustering, (iv) Retranslation and (v) Detection, in order to identify insider threats using log entries. [25] suggests the creation of a distributed Control Flow Graph (CFG) by extracting template sequences from log files This is the normal, expected behaviour. All activities are analyzed by comparison with this CFG.

5 Comparative Study

(Table 1)

Table 1 A comparison of major behavioural biometric approaches

Behavioural approach	Typical parameters analysed	Results obtained
Keystroke dynamics	• Flight Time • Dwell Time • Error rate • Typing technique	FAR of 7% [11]
Mouse dynamics	• Mouse wheel velocity • Types of clicks	FAR of 1.3% [14]
Haptics	• Direction • Force • Pressure of touch	FER of 1% [7]
Gait	• Arm swing • Head-foot distance • Knee angle • Head-pelvis distance	FER of 2.9% [23]
Log Files	• Network traffic • CPU usage • Dump files	No quantified results yet

6 Conclusion

Behavioural biometrics analyses various parameters of a person's behaviour such as gait, stride, typing patterns, mouse patterns, etc. As it is very difficult to replicate, these systems have a high-security. systems with very low errors have been developed. Clearly, behavioural biometrics promises to usher in a new era in the domains of computer and information security.

References

1. L. Wang, X. Geng, L. Wang, and X. Geng. Behavioral Biometrics For Human Identification: Intelligent Applications. Information Science Reference—Imprint of: IGI Publishing, Hershey, PA, 2009
2. Michelle Boatwright, Xin Luo, "What Do We Know About Biometrics Authentication?", InfoSecCD '07: Proceedings of the 4th annual conference on Information security curriculum development 2007
3. James L. Wayman, Anil K. Jain, Davide Maltoni, and Dario Maio," Biometric Systems: Technology, Design and Performance Evaluation", Springer, ISBN: 1852335963
4. Tarik Mustafi´c, Arik Messerman, Seyit Ahmet Camtepe, Aubrey-Derrick Schmidt, Sahin Albayrak, "Behavioral Biometrics for Persistent Single Sign-On", Proceedings of the 7th ACM workshop on Digital identity management, 2011
5. K P Tripathi, "A Comparative Study of Biometric Technologies with Reference to Human Interface", International Journal of Computer Applications (0975 – 8887) Volume 14– No.5, January 2011
6. Wolff, Matt. Behavioral Biometric Identification on Mobile Devices. Foundations of Augmented Cognition. Springer Berlin Heidelberg, 783–791, 2013

7. Yampolskiy, R.V. and Govindaraju, V. (2008) 'Behavioural biometrics: a survey and classification', Int. J. Biometrics, Vol. 1, No. 1, pp. 81–113
8. Ioannis C. Stylios, Olga Thanou, Iosif Androulidakis, Elena Zaitseva, "A Review of Continuous Authentication Using Behavioral Biometrics", Proceedings of the SouthEast European Design Automation, Computer Engineering, Computer Networks and Social Media Conference 2016
9. Esther Vasiete, Vishal Patel, Yan Chen, Larry Davis, Ian Char, Rama Chellappa, Tom Yeh, "Toward a Non-Intrusive, PhysioBehavioral Biometric for Smartphones", MobileHCI 2014, Sept. 23–26, 2014, Toronto, ON, CA
10. Monika Bhatnagar, Raina K Jain, Nilam S Khairnar, "A Survey on Behavioral Biometric Techniques: Mouse vs Keyboard Dynamics", Proceedings of the International Conference on Recent Trends in engineering & Technology - 2013(ICRTET'2013)
11. Juho Leinonen, Krista Longi, Arto Klami, Arto Vihavainen, "Automatic Inference of Programming Performance and Experience from Typing Patterns", Proceedings of the 47th ACM Technical Symposium on Computing Science Education, 2016
12. Kyle R. Corpus, Ralph Joseph DL. Gonzales, Alvin Scott Morada, Larry A. Vea, "Mobile User Identification through Authentication using Keystroke Dynamics and Accelerometer Biometrics", Proceedings of the IEEE International Conference on Mobile Software Engineering and Systems, 2016
13. Zach Jorgensen, Ting Yu, "On mouse dynamics as a behavioral biometric for authentication", Proceedings of the 6th ACM Symposium on Information, Computer and Communications Security, 2011
14. Nan Zheng, Aaron Paloski, Haining Wang, "An Efficient User Verification System Using Angle-Based Mouse Movement Biometrics", ACM Transactions on Information and System Security, Vol 18, Issue 3
15. Andrea Kanneh, Ziad Sakr, "Biometric user verification using haptics and fuzzy logic", Proceedings of the 16th ACM international conference on Multimedia, 2008
16. Saevanee H., Bhatarakosol, P., (2008). User Authentication Using Combination of Behavioral Biometrics over the Touchpad Acting Like Touch Screen of Mobile Device. International Conference on Computer and Electrical Engineering, 2008. Page(s): 82–86
17. Frank, M., Biedert, R., Ma, E., Martinovic, I., Song, D., (2012). Touchalytics: On the Applicability of Touchscreen Input as a Behavioral Biometric for Continuous Authentication. IEEE Transactions on Information Forensics and Security. 2012. (Volume:8, Issue: 1). Page (s): 136–148
18. Daniel Buschek, Alexander De Luca, Florian Alt, "Evaluating the influence of Targets and Hand Postures on Touch-based behavioural biometrics", Proceedings of the ACM Conference on Human Factors in Computing Systems, 2016
19. Benabdelkader, C., Cutler, R., and Davis, L. 'Person identification using automatic height and stride estimation', IEEE International Conference on Pattern Recognition 2002
20. S Laxmi, Tata A S K Ishwarya, S Sreeja, "Exploring Behavioural Type Biometrics: Typing Rhythm, Gait, Voice", International Journal of Innovative Research in Computer and Communication Engineering, Vol.4, Issue 10, October 2016
21. Mantyjarvi, J., Lindholm, M., Vildjiounaite E., Makela, S.M., Ailisto, H. A. (2005). Identifying users of portable devices from gait pattern with accelerometers. IEEE International Conference on Acoustics, Speech, and Signal Processing, 2005. (Volume:2). Page(s): ii/973–ii/976 Vol. 2
22. Yan Li, Ting Zhu, "Gait-Based WiFi signatures for Privacy-Preserving", Proceedings of the 11th ACM on Asia Conference on Computer and Communications Security, 2016
23. Gugliemlo Cola, Marco Awenuti, Fabio Musso, Alessio Vecchio, "Gait-based Authentication using a wrist-worn device", Proceedings of the 13th International Conference on Mobile and Ubiquitous Systems: Computing, Networking and Services, 2016

24. Markus Wurzenberger, Florian Skopik, Roman Fiedler, Wolfgang Kastner, "Discovering Insider Threats from Log Data with High-Performance Bioinformatics Tools", Proceedings of the 8th ACM CCS International Workshop on Managing Insider Security Threats, 2016
25. Animesh Nandi, Atri Mandal, Shubham Atreja, Gargi B. Dasgupta, Subhrajit Bhattacharya, "Anomaly Detection Using Program Control Flow Graph Mining from Execution Logs", Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, 2016