# ARCSECURE – CENTRALIZED HUB FOR SECURING A NETWORK OF IOT DEVICES

Project ID: 2020-086

Project Proposal Report

IT17126816 – C.N Samarasekara

B.Sc. (Hons) Degree in Information Technology

Specializing in Cyber Security

Department of Information Systems Engineering

Sri Lanka Institute of Information Technology

Sri Lanka

February 2020

# ARCSECURE – CENTRALIZED HUB FOR SECURING A NETWORK OF IOT DEVICES

Project ID: 2020-086

Project Proposal Report

IT17126816 – C.N Samarasekara

Mr. Kavinga Yapa Abeywardena

B.Sc. (Hons) Degree in Information Technology

Specializing in Cyber Security

Department of Information Systems Engineering

Sri Lanka Institute of Information Technology

Sri Lanka

February 2020

## Declaration

We declare that this is our own work and this proposal does not incorporate without acknowledgment any material previously submitted for a degree or diploma in any other university or institute of higher learning and to the best of our knowledge and belief it does not contain any material previously published or written by another person except where the acknowledgment is made in the text.

| Name | Student ID | Signature |
|---|---|---|
| C.N Samarasekara | IT17126816 | |
| | | |

The above candidates are carrying out research for the undergraduate Dissertation under my supervision.

Signature of the supervisor: Date:

## Abstract

IoT has provided an answer for people who were constantly on the lookout for a solution that enabled them to interact with the day to day devices they use. As a result, the IoT market has seen a surge over recent years with the rise in devices expected to increase by the year for the foreseeable future. Anything connected to the Internet is susceptible to threats, IoT devices are no different. With the market for IoT blossoming and it still being in its juvenile days, vendors are more concerned with putting out a profitable product to rival their competitors than worrying about security. The target audience for our proposed device is smart homes and SOHO networks. The target audience was chosen taking into account that they are the most vulnerable as most lack basic security features. Most of them have their routers as the first line of defence. Unlike our home Wi-Fi/ethernet connections which at best can result in our home PC or laptop getting compromised, whereas an IoT network getting compromised is equivalent to giving up your entire household. While the primary focus of our research is to address this issue. This paper would focus specifically on one key area out of the big picture which is user authentication to the device. As a resolution method, a multi-factor authentication mechanism is proposed besides utilizing keystroke dynamics to identify possible malicious user behavior.

**Keywords** – IoT, keystroke dynamics, multi-factor, authentication, biometrics

## Table of Contents

## List of Figures



## List of Tables

## List of Abbreviations

| | |
|---|---|
| IoT | Internet of Things |
| SOHO | Small Office Home Office |
| SSO | Single Sign-On |
| MFA | Multi-Factor Authentication |
| PIN | Personal Identification Number |
| SVM | Support Vector Machine |
| GA | Genetic Algorithm |
| App | Application |
| Et al. | ‘and others’ |

## 1. Introduction

IoT is rapidly progressing with each passing year with novel devices being introduced and the existing systems going through a wave of innovation. There is hardly any doubt that IoT has completely changed the landscape on how we would go about interacting with technology in our day to day lives. With progression in innovation comes a spike in threats. Already the cyberspace consists of e-services such as e-commerce and e-banking that have enabled hackers to tap into billions of dollars each year [1]. To top that off with IoT coming to the forefront in recent years, most of the day to day devices that we use are now linked to the internet. As convenient as it may sound it has now further broadened the attack surface for hackers and has now put them in pole position to not just compromise those devices but to eventually take control of entire IoT networks. Avast in their smart home security report for 2019 has highlighted some key statistics that show the severity of the issue in hand:

- At least 5 devices are connected to the internet in an average household.
- 40.8% of homes have at least 1 device that is vulnerable to cyber-attacks.
- 69.2% of those vulnerable devices have weak credentials [2].

The SSL Store offered the following stats on IoT vulnerabilities and how these devices are used in cyber-attacks.

- 75% of infected devices in IoT attacks are routers.
- IoT devices typically attacked within 5 minutes.
- 76% of risk professionals think IoT leaves them at risk of cyber-attacks.
- IoT Malware Attacks Skyrocketed in 2018, Trend Continues to 1H 2019 [3].

This is where ARCSECURE comes in. It is designed keeping the above-mentioned issues in mind. ARCSECURE is a simple plug and play device that will be connected to the router and it would act as the first line of defence before traffic can reach the router, playing a similar role to a perimeter firewall. This paper solely focusses on one important aspect of ARCSECURE which is implementing a secure user authentication mechanism.

User authentication is not a novel concept. Posing a challenge and granting access upon the correct answer has been around for long as anyone can remember and is still relied upon to this very day [3]. Passwords have been the go-to option for authentication and have been the most used mechanism [4]. Over the years requirements and complexity of passwords have increased to offer better security [4]. The challenge with simple challenge-response authentication has always been in convincing the user to utilize a strong password. Passwords can be a burden if its complex it becomes hard to remember and if it's simple it can easily be found out and not to forget having to remember multiple passwords just adds to the misery.

## 2. Background & Literature survey

This section would explore the proposed system and its functionalities and review the pre-existing research and analysis previously carried out in this area of study. The focus of this paper as mentioned previously would be on a secure user authentication mechanism for users using IoT in a smart home or a SOHO environment.

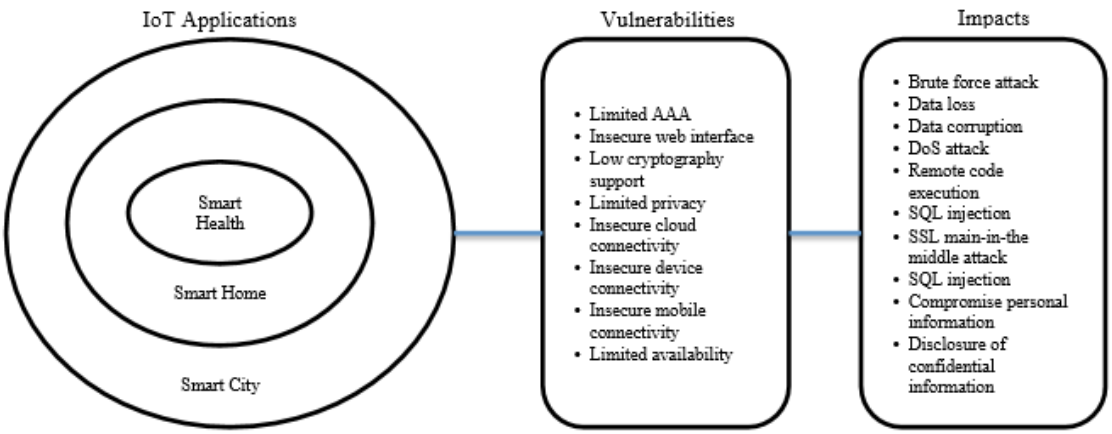


Figure 4.1: IoT Application, vulnerabilities, and the impacts

Source: [10]

Keystroke dynamics have been researched for a long time. In 2002 Bergadano et al [5] researched keystroke dynamics for user authentication using the volunteer's self-collected dataset, data collected using the same text for all individuals, resulting in only 0.01 percent passing the authentication impostor. In 2003, Yu and Cho [6] performed preliminary experimental research on the selection of a function subset selection of keystroke dynamics identity verification and found that GA-SVM yielded good accuracy and speed of learning. Revett et al. [7] researched user authentication in 2007 and began researching authentication for a dynamic keystroke. The author has suggested that biometrics are robust, specifically fingerprints, but they can be easily spoofed.

In 2009, to classify consumers, Zahid et al. [8] performed a study on complex keystrokes on the mobile phone. For the back-end, they used a fuzzy particle swarm optimization classifier in the front-end and genetic algorithm to distinguish 3 different features to be

used in user identification: key hold time, the time difference between pressing and releasing a key. Diagraph time, the time difference between removing one key and pressing the next one. Error rate, the backspace key is released several times. Besides, 5 classifiers are used to train these features: Baive Bayes, Back Propagation Neural Network (BPNN), Radial Basis Function Network (RBFN), Kstar, J48.

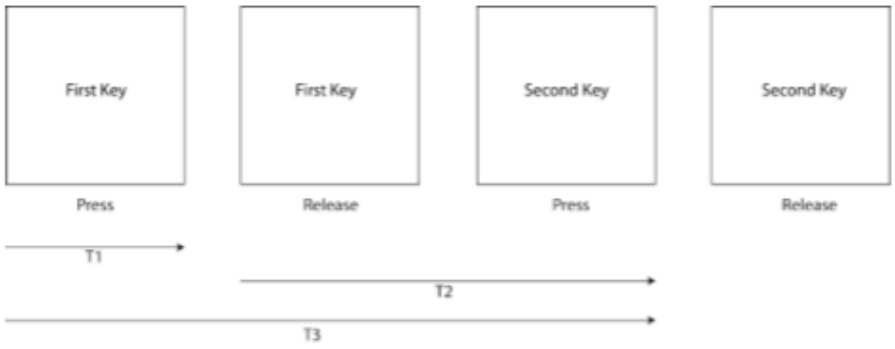


Figure 4.2: Keystroke Timing

Source: [11]

## 3. Research Gap

A smart home or SOHO IoT network would include several IoT devices connected to the Internet via a single router. Most of these IoT devices can be from different vendors as purchasing an entire network of devices from the same vendor can prove to be expensive. Also purchasing all devices from the same vendor prevents the flexibility in choosing devices of the same type that has better performance or functionality over the other. For those reasons, most users tend to purchase devices that are from various vendors. Therefore, in such a scenario where IoT devices from various vendors are in use an issue arises, which is that the user must access them separately. This becomes an unnecessary burden for the user.

Keystroke dynamics is a field of science that studies a rhythm or sequence that occurs in the style of typing of individuals. Some of the information that can be extracted from its features is the time when keys on the keyboard are pressed, keys on the keyboard are lifted and keystrokes are moved from one keypad to another [9]. Compared with other solutions to behavioral authentications the biometric keystroke dynamic system is relatively unexplored. Compared with other biometric schemes, this is coupled with the small number of studies performed. Although typically lower than other biometric behavioral authentication schemes, dynamic keystroke has several benefits and offers continuous monitoring capability [10].

## 4. Research Problem

For IoT devices in a smart home or a SOHO network, the router acts as the first line of defence from external threats. Having a router as the first line of defence is more the first mistake. Avast points out the following stats when it comes to routers in smart homes:

- 59.7% are with either weak credentials or with some form of vulnerability.
- 59.1% of users have never logged onto their routers neither have they been bothered to update their routers' firmware [2].

We shall further explore authentication and keystroke dynamics user identification for IoT through research conducted previously. Authentication is among the highest vital aspects for consideration towards the look of secure IoT communication. Authentication is often rendered because of it being the first phase towards access control, and it is often device authentication or user authentication [11], even more. However, the availability of a lightweight and distributed authentication scheme for total security solutions towards IoT applications remains one of the toughest challenges [12]. Device authentication is critical and a challenging task for the emerging IoT [13].

Person identification supported keystroke dynamics may be considered as a classification task, that various classifiers could also be applied, like neural networks [14], data evolution methods [15] and classifiers supported Kolmogorov-Smirnov test [16]. While ensemble techniques [17,18] are found to be powerful for various classification tasks, their recent applications include person identification supported keystroke dynamics data [19]. during this paper we consider the task of person identification supported keystroke dynamics as a time-series classification problem, that nearest neighbor classifiers with dynamic time warping (DTW) distance [20] are shown to be competitive [21,22] with various, more complex models like neural networks [23], Hidden Markov Models [24] or “super-kernel fusion scheme” with these empirical results are supported by studies that specialize in the theoretical aspects of classification [25,26].

# 5. Objectives

## 5.1. Main Objectives

The main objective of this research component elaborated within this paper is to prepare a secure user authentication mechanism for the plug and play device ARCSECURE. The expected end-product is to allow a user to securely access their IoT devices through a portable device. This will be achieved by taking into consideration the problems that are currently faced by IoT users as discussed in the Research Gap and Research problem sections. The idea is to maximize security as much as possible by adding various layers of protection that a potential attacker may need to bypass to eventually access the device. All of this is expected to be carried out while trying the level best not to compromise on performance and user experience.

The following are the main objectives to be covered to meet the expected outcome:

- Create an SSO feature that can provide access to all devices on the network from just a single authentication process.
- Use of MFA mechanisms to verify the user to the IoT network.
- Creating a more flexible and user-friendly access control feature.
- Use of keystroke dynamics as a means of authenticating the user.
- Use of biometric keystroke dynamics to identify possible malicious user behavior.

## 5.2. Specific Objectives

The following is a breakdown of sub-objectives:

- Create a captive portal.
- Create a token-based multi-factor authentication mechanism.
- Create a mobile application.
- Create a mechanism for device authentication.

- Create a fingerprint scanner in the mobile app.
- Have the mobile application pop a simple yes/no to validate the user upon signing in.
- Implement an access control feature that governs the authority of a user.
- Developing a machine learning algorithm for keystroke dynamics.
- Train the algorithm to identify keystrokes.
- Create a feature to temporarily block a user's IP upon being proven false.
- Create a mechanism to alert the user on possible false intrusion attempts.

## 6. Methodology

This section will explore the expected approach that will be taken to arrive at our final goal which is ultimately ensuring that a user authenticating via the ARCSECURE plug and play portable device gains secure access to their IoT network devices.

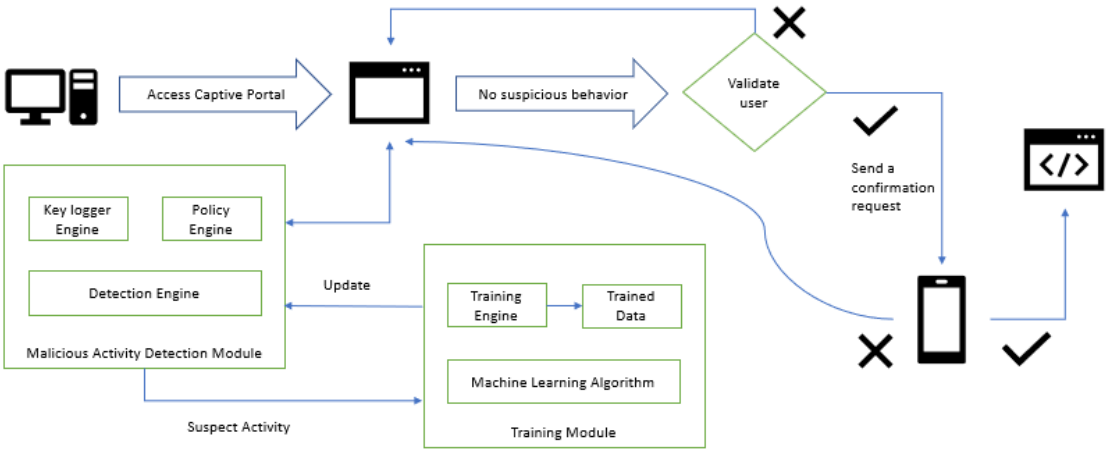


Figure 10.1: System Diagram

As the above diagram interprets a user trying to gain access to the IoT network would first be greeted with a captive portal where he/she would be required to enter their user credentials. Since this is a SSO feature a user who is successfully granted access to the network would have complete access to all devices at once. This is a critical issue as it can lead to a single point of failure. To overcome this issue a simple user credential validation will not suffice. To raise the level of security a notch MFA will be used. The first section of this MFA process utilizes keystroke dynamics and the other utilizing a combination of biometrics and token-based authentication.

### 6.1. Keystroke Dynamics for Authentication

The user will be required to type in some text for the machine learning algorithm to figure out if the typing pattern matches with the initial pattern it learned from the user. The distinctive, behavioral characteristics measured by Keystroke Recognition include the subsequent variables:

- The cumulative typing speeds.
- The time which elapses between consecutive keystrokes.
- Dwell Time – the duration a key is held down for.
- The frequency during which the num pad or function keys are used.
- The timing for capitalization sequence (which key is released first shift or the letter key).
- Flight Time – The time taken to press a key release it and then press on a key again.
- Any Error Rates, like using the Backspace or Delete keys.

If a near similar match is found, then the user can progress further into the authentication process. Else the user will be temporarily blocked, and the main user will be alerted who in turn would be able to revoke changes if he/she desires.

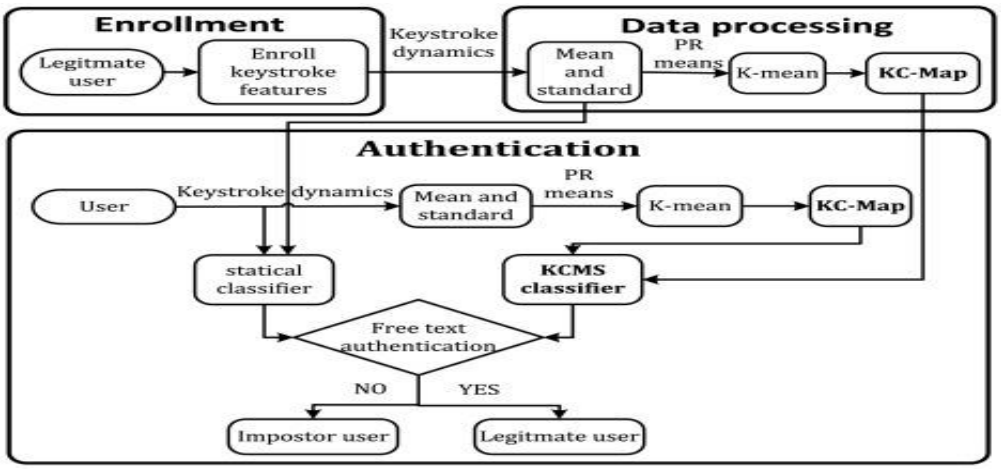


Figure 6.1.1: Proposed model for keystroke dynamics implementation

Source: [20]

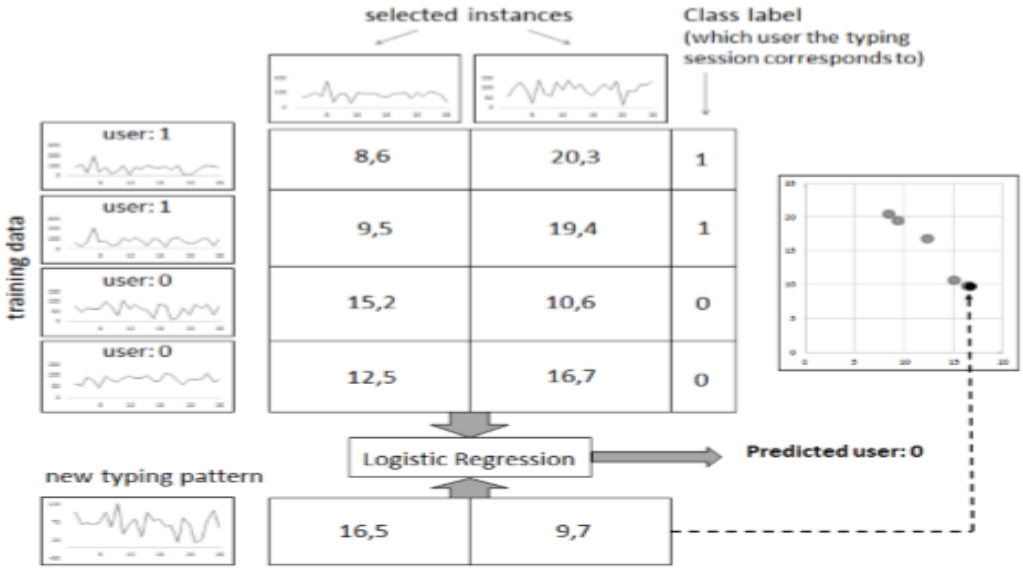


Figure 6.1.2: Projection-based classification of keystroke dynamics data.

Source: [27]

### 6.2. MFA Process

As part of the MFA process, the user upon successful validation of user credentials will have to accept a simple Yes/No response which will be sent to a specially made app in the user's mobile phone. The app upon opening will request the user to authenticate to the app using a fingerprint scanner. Once validated he/she can confirm the Yes/No request sent to the specific mobile app. The user will now be able to access the IoT network based on their authorization right that the user is granted.

# 7. Project Requirements

## 7.1. Functional Requirements

- Access all IoT devices from a single authentication.
- Record typing patterns.
- Distinguish between typing patterns.
- Identify accessing device (device authentication).
- Identify accessing device IP.
- Identify malicious behavior.
- Block IP’s identified as malicious.

## 7.2. Non-Functional Requirements

- Performance
- Security
- Availability
- Reliability
- Maintainability
- Usability

## 8. Budget and Justification

Estimated budget for the physical devices and equipment.

| Device | Price | Quantity | Total |
|---|---|---|---|
| Raspberry Pi | Rs. 16,000.00 | 1 | Rs. 16,000.00 |
| IoT devices | Rs. 4000.00 | 2 | Rs. 8000.00 |
| Ethernet Cables | Rs. 500.00 | 1 | Rs. 500.00 |
| | | | Rs. 24,500.00 |

Table 12.1: Physical equipment budget

Estimated Budget for other expenses.

| Items | Price | Total |
|---|---|---|
| Stationary | Rs. 2500.00 | Rs. 2500.00 |
| Printing | Rs. 2000.00 | Rs. 2000.00 |
| Other | Rs. 2000.00 | Rs. 2000.00 |
| | | Rs. 6500.00 |

Table 12.2: Other expenses budget

Total Expenses

| Items | Price |
|---|---|
| The estimated budget for the networking Devices | Rs. 24,500.00 |
| Estimated Budget for other expenses | Rs. 6500.00 |
| Total estimated cost for the year | Rs. 31,000.00 |

Table 12.3: Total budget

## 9. Reference List

[1] (https://hostingtribunal.com/blog/hacking-statistics/, n.d.)

[2] (https://cdn2.hubspot.net)

[3] (https://www.thesslstore.com)

[4] (https://hackernoon.com, n.d.)

[5] Bergadano, F., Gunetti, D., Picardi, C.. User authentication through keystroke dynamics. ACM Transactions on Information and System Security (TISSEC) 2002;5(4):367–397.

[6] Yu, E., Cho, S. Ga-SVM wrapper approach for feature subset selection in keystroke dynamics identity verification. In: Neural Networks, 2003. Proceedings of the International Joint Conference on; vol. 3. IEEE; 2003, p. 2253–2257.

[7] Revett, K., Gorunescu, F., Gorunescu, M., Ene, M., Magalhaes, S., Santos, H.. A machine learning approach to keystroke dynamics-based user authentication. International Journal of Electronic Security and Digital Forensics 2007;1(1):55–70.

[8] Zahid, S., Shahzad, M., Khayam, S.A., Farooq, M.. Keystroke-based user identification on smartphones. In: International Workshop on Recent Advances in Intrusion Detection. Springer; 2009, p. 224–243.

[9] Epp, C., Lippold, M., Mandryk, R.L. Identifying emotional states using keystroke dynamics. In: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems. ACM; 2011, p. 715–724.

[10] Ali, M.L., Monaco, J.V., Tappert, C.C., Qiu, M.. Keystroke biometric systems for user authentication. JournalofSignalProcessingSystems 2017;86(2-3):175–190.

[11] S. Shaju, “BISC Authentication Algorithm : An Efficient New Authentication Algorithm Using Three Factor Authentication for Mobile Banking,” 2016.

[12] P. Mahalle, “Identity Authentication and Capability Based Access Control (IACAC) for the Internet of Things,” J. Cyber …, vol. 1, pp. 309–348, 2013.

[13] D. Chen et al., “S2M: A Lightweight Acoustic FingerprintsBased Wireless Device Authentication Protocol,” IEEE Internet Things J., vol. 4, no. 1, pp. 88–100, 2017.

[14] Wong, F.W.M.H., Supian, A.S.M., Ismail, A.F., Kin, L.W., Soon, O.C.: Enhanced user authentication through typing biometrics with artificial neural networks and k-nearest neighbor algorithm. 35th IEEE Asilomar Conference on Signals, Systems and Computers, vol. 2, 911–915 (2001)

[15] Kozierkiewicz-Hetmanska, A., Marciniak, A., Pietranik, A.: Data Evolution Method in the Procedure of User Authentication Using Keystroke Dynamics. Computational Collective Intelligence, Lecture Notes in Computer Science 9875, 379– 387 (2016)

[16] Ceffer A., Levendovszky J.: Kolmogorov-Smirnov test for keystroke dynamics-based user authentication. 17th IEEE International Symposium on Computational Intelligence and Informatics, DOI: 10.1109/CINTI.2016.7846387 (2016)

[17] Wozniak, M., Jackowski, K.: Fusers Based on Classifier Response and Discriminant Function – Comparative Study. Lecture Notes in Computer Science 5271, 361–368 (2008)

[18] Kurzynski, M., Wozniak, M.: Combining classifiers under probabilistic models: experimental comparative analysis of methods. Expert Systems 29(4), 374–393 (2012)

[19] Doroz, R., Porwik, P., Safaverdi, H.: The New Multilayer Ensemble Classifier for Verifying Users Based on Keystroke Dynamics. Computational Collective Intelligence, Lecture Notes in Computer Science 9330, 598–605 (2015)

[20] Sakoe, H., Chiba, S.: Dynamic programming algorithm optimization for spoken word recognition. IEEE Transactions on Acoustics, Speech and Signal Processing 26(1), 43–49 (1978)

[21] Xi, X., Keogh, E., Shelton, C., Wei, L., Ratanamahatana, C.A.: Fast time series classification using numerosity reduction. Proceedings of the 23rd ACM International Conference on Machine Learning, 1033–1040 (2006)

[22] Ding, H., Trajcevski, G., Scheuermann, P., Wang, X., Keogh, E.: Querying and mining of time series data: experimental comparison of representations and distance measures. Proceedings of the VLDB Endowment 1(2), 1542–1552 (2008)

[23] Nanopoulos, A., Alcock, R., Manolopoulos, Y.: Feature-based classification of timeseries data. International Journal of Computer Research, 10(3), 49–61 (2001)

[24] Kim, S., Smyth, P., Luther S.: Modeling waveform shapes with random effects segmental Hidden Markov Models. In Proceedings of the 20th Conference on Uncertainty in Artificial Intelligence, 309–316 (2004)

[25] Chen, G.H., Nikolov, S., Shah, D.: A latent source model for nonparametric time series classification. Advances in Neural Information Processing Systems

[26] 1088– 1096 (2013) 16. Devroye, L., Gyo¨rfi, L., Lugosi, G.: A probabilistic theory of pattern recognition, Springer Science & Business Media (1996)

[27] Neubrandt, Dora; Buza, Krisztian, Projection-based person identification, Advances in Intelligent Systems and Computing (2018)