Towards Keystroke Analysis using Neural Network for Multi-Factor Authentication of Learner Recognition in On-Line Examination

Mideth B. Abisado, Bobby D. Gerardo, and Arnel C. Fajardo

***Abstract*—Technology often tries to mimic nature. Therefore, the idea to recognize a user based on several of their traits, as it is done in real life, logically comes to mind. In this case, the motive of this study focuses on behavioral biometrics. The opportunity to use biometrics and pattern classification to develop a new solution using keystroke analysis and recognition to address online examination fraud and cheating issues. This framework could be a new non-intrusive recognition approach, taking a valuable part in the information system's security chain. User's keystrokes are recorded as they take the exam. The Multi-Layer Perceptron Neural Network is utilized to classify learner keystroke as they take an on-line examination.**

***Keywords*—keystroke analysis, multi-factor authentication.**

I. Introduction

A secure information system depends on successful authentication of legitimate users to prevent attacks from fraudulent persons. Traditional information security systems use single factor authentication. This means they can be easily accessed by unauthorized persons without access being noticed. This is one key issue on the use of Learning Management Systems (LMS) when implementing online examinations. These systems ask for student's username and passwords when accessing online content and taking examinations or any assessment task. Recognition of students are not verified. That's the challenge for another of the most significant online course providers, Coursera and Edx. The need to establish a reliable system to stop online cheating is fast becoming a mainstream concern. To address said issues, this paper designs the use of keystroke recognition for taking online course examination.

II. Background

There is a difference between authentication and nonrepudiation. The problem that the LMS recognize a sufficient level of trust and recognition on student profile. While this issue remains unanswered, Microsoft Corporation have devised a way to authenticate and recognize on-line exam takers by online proctored exam delivery [1]. This method still employs single factor authentication with interference of an online live proctor, a camera and security measures. In the field of blended learning, this cannot be facilitated by instructors. Costly proctor supervision provides only minimal assurance of academic integrity. Another solution provides eight Online Exam Control Procedures to prevent and detect cheating when professors use online exams without proctor supervision [2]. These online control procedures can be applied in either traditional residency courses using online exams or in courses conducted entirely online. The recommended control procedures help thwart student fraud by increasing the difficulties of online exam cheating. This paper suggests that the benefits of proctor supervision for online exams are less than the total direct and indirect explicit costs of proctoring. Thus, sufficient academic integrity can normally be achieved for online college courses without using proctor supervision. In his paper, Sarrayrih proposed a system that provides security to improve on-line examination by utilizing technologies such as biometric authentication using student's fingerprint [3]. In their research paper, they discussed the performance of student's online course exam with respect to security and main challenges faced by online course exams within the university. However, this will entail additional device on the end of the exam taker, as shown in Figure 1. The systems are connected using the star topology. The camera and finger print scanner inside the lab are connected to the security server; once the security server authenticates the biometrics of user, then the users can write the exam at the specific terminal provided to them. When an unauthorized user attempts to access the system from different location he is not allowed.

Fig. 1 System Architecture with Fingerprint scanner

Mideth B. Abisado, Bobby D. Gerardo, and Arnel C. Fajardo are with Technological Institute of the Philippines, Q.C., University of Sto. Tomas, Manila, Philippines

Biometrics measure individuals' unique physical or behavioral characteristics to recognize or authenticate their identities. Biometrics offer to inextricably link the authenticator to its owner, something passwords and tokens cannot do, since they can be lent or stolen. In terms of combining with existing systems, much research has gone into investigating the viability using the typing behavior present upon the entry of password credentials as an additional layer of authentication. This technique could potentially overcome the shortcomings of passwords, as not only must the password be known, but it must be entered in the manner of the legitimate user [4]. Use of behavioral biometrics shows that keystroke biometrics (KB) authentication systems are a less popular form of access control, although they are gaining popularity. In recent years, keystroke biometric authentication has been an active area of research due to its low cost and ease of integration with existing security systems [5]. Various researchers have used different methods and algorithms for data collection, feature representation, classification, and performance evaluation to measure the accuracy of the system, and therefore achieved different accuracy rates. Although recently, the support vector machine is most widely used by researchers, it seems that ensemble methods and artificial neural networks yield higher accuracy. Figure 2 lists the ontology of different authentication models.

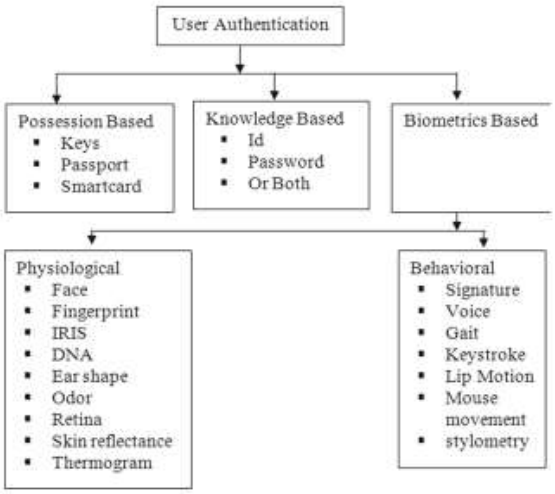

Fig. 2 Ontology of authentication models

A. Keystroke Analysis Using Neural Network

An Artificial Neural Network (ANN) is an information processing paradigm that is inspired by the way biological nervous systems, such as the brain, process information. The key element of this paradigm is the novel structure of the information processing system. It is composed of a large number of highly interconnected processing elements (neurons) working in unison to solve specific problems. ANNs, like people, learn by example. An ANN is configured for a specific application, such as pattern recognition or data classification, through a learning process. Learning in biological systems involves adjustments to the synaptic connections that exist between the neurons. This is true of ANNs as well. In this section, the explanation about the two neural networks: Multilayer perceptron (MLP) and classifier is mentioned.

B. Multilayer Perceptron (MLP) Network

A multilayer perceptron (MLP) is a feed forward artificial neural network model that maps sets of input data onto a set of appropriate outputs. Figure 3 shows the structure of the MLP network used in this paper. A MLP consists of multiple layers of nodes in a directed graph, with each layer fully connected to the next one. MLP utilizes a supervised learning technique called back propagation for training the network. MLP is a modification of the standard linear perceptron and can distinguish data that are not linearly separable. It consists of three main parts: an input layer, one or more hidden layers, and an output layer. The input layer distributes the input data to the processing elements in the next layer. The second stage is the hidden layer which incorporates the nonlinearity behavior and the last stage shows the output layer. Input and output are directly accessible, while the hidden layers are not. Each layer consists of several neurons. The goal of this type of network is to create a model that correctly maps the input to the output using historical data so that the model can then be used to produce the output when the desired output is unknown.

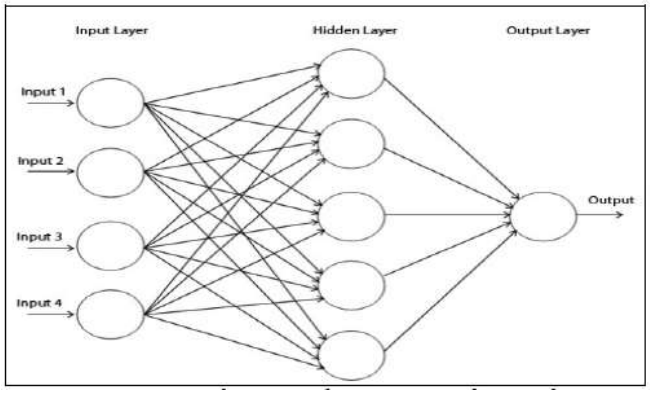

Fig. 3 Structure of the MLP network

C. Keystroke Analysis

Keystroke verification techniques can be classified as either static or continuous. Static verification approaches analyze keystroke verification characteristics only at specific times, for example, during the login sequence. Static approaches provide more robust user verification than simple passwords, but do not provide continuous security — they cannot detect a substitution of the user after the initial verification. Continuous verification, on the contrary, monitors the user's typing behavior throughout the course of the interaction. As early as 1980, researchers have been studying the use of habitual patterns in a user's typing behavior for identification. Keystroke dynamics is the process of analyzing the way a user types at a terminal by monitoring the keyboard inputs thousands of times per second in an attempt to identify users based on habitual typing rhythm patterns. It has already been shown that keystroke rhythm is a good sign of identity [6].

III. NEEDLEMAN-WUNSCH ALGORITHM (NM-W)

The Needleman-Wunsch algorithm handled the problem of how to authenticate users efficiently based on their keystroke behavior. The method creates a unique signature for each user using a membership function as a sequence of letters. Hence, we utilize the sequence alignment Needleman-Wunsch

algorithm to get more accurate value of authentication process. Furthermore, Blosum matrix is reconstructed to increase the similarity degree based on the convergence degree of letters in the keyboard. The experiments proved that Needleman is very promising in extracting user patterns with accuracy rate 80% and precision rate 90.3%. A comparison with other classifiers proved that the proposed approach achieves significantly better results [7]. Table I shows the different classifier results using Weka vis-à-vis the enhanced approach.

TABLE I: Classifier results using Weka tool vs enhanced approach

Method	Precision	Accuracy
BayesNet	50.5	50.5
NaiveBayes	50.8	50.7
Kstar	46.8	47.2
Id3	35.15	35.2
J48	39.7	40.12
Nnge	40.7	40.99
Decision Table	31.41	27.5
Conjunctive Rule	2.1	7.51
Enhanced NM-W Alg.	90.3	80

IV. CONTRIBUTIONS

Figure 4 depicts this study's framework for keystroke recognition system. Applying Artificial Neural Network Model - multilayer perceptron.

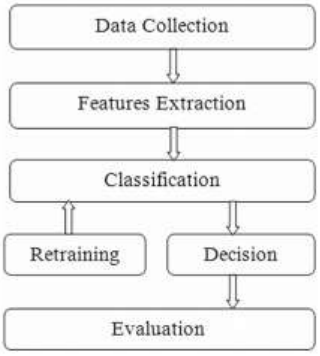

Fig. 4 Keystroke Recognition System

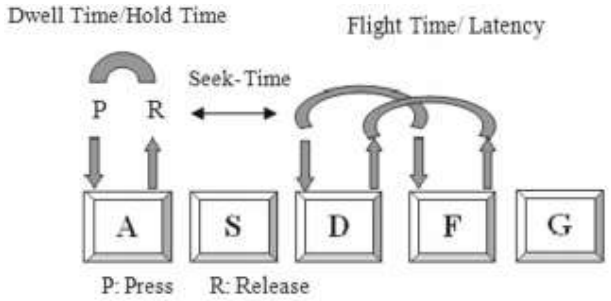

Fig. 5 Dwell time and flight time

Keystroke data sets available will be used to test accuracy of the developed add-on agent for LMS. The available Blackboard enrolled users' data sets will also be used in training valid and recognized users and learners.

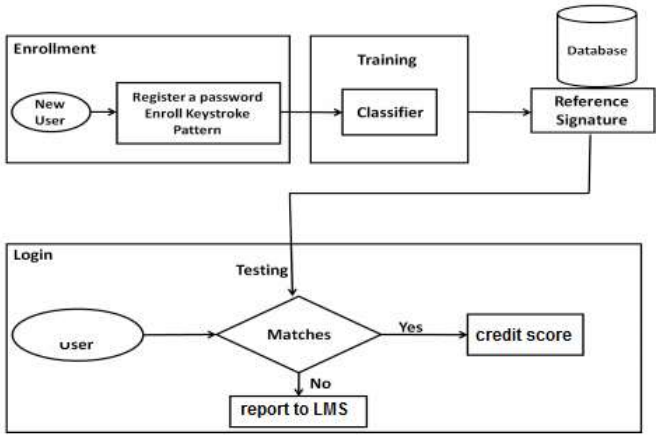

Fig. 6 User recognition process using keystroke dynamics method

The positive recognition of each user will merit exam score after the examination is taken, as reflected in Figure 6. False identifications will send a report to the LMS of fraudulent account taking the exam. Measures to address this shall be based on the protocol when this framework is in place

V. EVALUATION

Typing patterns are believed to be unique to an individual. KB systems generally have lower implementation and deployment costs compare to other biometric authentication systems. The typically KB system can be fully implemented by software and has low dependency on specialized hardware. From the user's perspective, keystroke analysis systems are transparent and non-invasive. They offer increased password strength and lifespan and continuous monitoring. In this study, the keystroke analysis is used to provide recognition of online exam takes in addition to user authentication. The development of the keystroke recognition agent is sought to become a novel framework for administering online examination.

VI. CONCLUSION

Current online courses have been prevalent today. Coursera and Edx are two of the top providers for a student's access to online course content and certificate. Locally, the use of blended learning is implemented with the use of LMS. In these online environments, detection and recognition of enrolled students as they take examinations is an issue that must be addressed. The framework provides student authentication and recognition for the online environment.

VII. FUTURE WORK

Further tests and algorithm refinement must be done to increase accuracy of keystroke analysis agent for online examinations in LMS.

ACKNOWLEDGMENT

This work was supported in part by the Commission on Higher Education Faculty Development Program Phase II. The support is gratefully acknowledged.

REFERENCES

[1] M. Corporation, "Online proctored exam delivery," [Online]. Available: https://www.microsoft.com/en-ph/learning/online-proctored-exams.aspx. [Accessed 20 May 2016].

[2] G. R. C. Jr., "Thwarting online exam cheating without proctor supervision," *Journal of Academic and Business Ethics*, 2015.

[3] M. A. Sarrayrih, "Challenges of Online Exam, Performances and problems for Online University Exam," *International Journal of Computer Science Issues*, vol. 10, no. 1, pp. 439-443, 2013.

[4] S. B. Wankhede, "Keystroke Dynamics Authentication System Using Neural Network," *International Journal of Innovative Research and Development*, vol. 3, no. 1, pp. 158-164, 2014.

[5] M. L. Ali, "Keystroke Biometric Systems for User Authentication," 2016.

[6] F. Monrose, "Keystroke dynamics as a biometric for authentication," *Future Generation Computer Systems.*

[7] S. Bamatraf, "Keystroke Authentication on Enhanced Needleman Alignment Algorithm," *Intelligent Information Management*, pp. 211-221, 2014.

ARNEL C. FAJARDO finished his Doctor of Philosophy (Ph.D.), Computer Engineering Hanbat National University in 2014. His activities and societies include Research Assistant in IISPL (Intelligent Information Signal Processing Laboratory) Teaching Assistant for Seoul Accord Project in the Department of Computer Engineering.

MIDETH B. ABISADO received the Master of Science degree in Computer Science from the Mapua Institute of Technology and the Master in Information Technology degree from the Technological University of the Philippines in 2016 and 2004, respectively. She performs her research works in the field of adaptive learning, data mining and information security at the Technological Institute of the Philippines and University of Sto. Tomas.

BOBBY D. GERARDO finished his BS in Electrical Engineering in 1994, with High Distinction from Western Institute of Technology and his Master of Arts in Education Major in Mathematics from University of the Philippines in Diliman Quezon City in 2000 being the grantee of DOST-SEI scholarship for Math and science Faculty. He pursued his Ph.D. in Information and Telecommunications with major in Distributed Systems at Kunsan National University, Korea in 2007 being the grantee of Korean Scholarship for Brain Korea (BK-21) project.