A NEURAL NETWORK BASED COMPUTER ACCESS SECURITY SYSTEM
FOR MULTIPLE USERS

A.S. ANAGUN* and I. CIN**

* Osmangazi University, Department of Industrial Engineering, Bademlik 26030 Eskisehir, Turkey.
  sanagun@vm.baum.anadolu.edu.tr  ** Teknik A.S., Cerkezkoy, Tekirdag, Turkey.

ABSTRACT
A neural network architecture is proposed to deal with a situation of multiple users where each user has his/her own password with different length. Three experiments were conducted to find a better way for a computer access security system. The neural networks are trained using time intervals between successive keystrokes during password entry through keyboard. The performance of the neural networks designed for each of the experiments are evaluated in terms of recognition accuracy. Two major issues, preventive security and detection of violations, which may be in question for a security system, are examined. A procedure is proposed to make a security system more reliable when a time interval pattern close to the one produced by the group members during on-line entry. © 1998 Elsevier Science Ltd. All rights reserved.

KEYWORDS
Neural networks, computer access, security systems, pattern recognition, keystroke dynamics

INTRODUCTION
In today's competitive market, a knowledge, consisting of results of marketing research, customer complaints, and related issues should be established for being succeed and survive. The stored information is valuable, critical, and/or more sensitive for an organization. Therefore, a computer access security system, a reliable way of preventing unauthorized people for accessing, changing or deleting, and stealing the information, should be designed and implemented within the organization.

A computer security system should not only be able to identify a person and let him/her access to the system if he/she has a correct security code or deny the access otherwise (preventive security), but also be capable of identifying the person whether he/she is indeed the right person (detection of violations). In order to accomplish these goals, a software-based, a hardware-based, or a software and hardware-based security system may be used. In either case, although each person, who is eligible for accessing to the system, has his/her own security code, the code may be found with a trial-error process or stolen from the authorized person by someone else. When this occurs, the attempt made by a person may not be prevented. Due to the drawbacks of the common approaches, a different method in terms of computer access security, which may prevent copying or duplicating the security code issued, should be developed to differentiate an authorized person from the others such that a valuable and/or more sensitive information for an organization should be secured.

Recent studies, in the area of computer access security, have concentrated on user identification based on individual's typing pattern, considered as a special characteristic for each person, using classical pattern recognition techniques, fuzzy algorithms, and neural networks, powerful tools for pattern recognition and classification applications (Bleha et al., 1990; Bleha and Obaidat, 1991; Hussein et al., 1989; Obaidat et al., 1991; Bleha and Obaidat, 1993; Obaidat and Macchiarulo, 1993). In these studies, time intervals between successive keystrokes while entering a known and long password, an example of software-based security system, on a keyboard were considered as an alternative security code to prevent unauthorized person for accessing the system involved and
changing some information. Since the same password has been entered by a group of people, this situation may be classified as multiple users-one password. In addition, the studies mentioned have focused on preventive security which basically classifies people into two groups, people who know the correct password and who do not know without evaluating whether they are indeed authorized.

However, due to the developments in computer technology and the complexity of information systems which the organizations might have, there may be different situations which needed to considered such as multiple users-multiple passwords, a user-one password, and a user-multiple passwords. As discussed in the study of (Anagun and Cin, 1996), each of these situations may be applied to the computer access security systems considering passwords with different lengths based on his/her preferences or system's requirements, if applicable. Also, the computer access security system should be designed for not only the purpose of preventive security, but also the purpose of detection of violations to make the system more reliable. Here, an intelligent computer access security system using a multilayered feedforward neural network trained with a backpropagation learning algorithm is proposed for a situation of multiple users-multiple passwords with different lengths. In this study, a password was assigned to each user and users were tried to differentiate from each other according to the time intervals between successive keystrokes obtained based on their typing techniques.

DATA COLLECTION

For the application purpose, five faculty/staff members were randomly selected from the Department of Industrial Engineering and five passwords with different lengths, ranging from 7 to 14, were determined according to their preferences and assigned to the group members. Afterwards, each member of the group was asked to enter his/her own password and other passwords of the remaining members, which were represented by *"* during the typing process, along with a user identification number in a random order based on a proposed data collection structure. After each entry, a typed phrase via keyboard was displayed on the bottom of the screen followed by a return key.

The time intervals between successive keystrokes of the password which was typed correctly were computed and automatically recorded in a file according to the user and password identification numbers. During this process, for instance, if the password OSMANGAZI were entered by the first member of the group, the time intervals of (O,S), (S,M), (M,A), (A,N), (N,G), (G,A), (A,Z), (Z,I), would be computed and recorded in a file as U-P1,nnn using a program coded in Pascal 7.0. Such a file, for each entry, consists of the time intervals of the passwords typed, user identification number that represents who typed the password, and password identification number that represents which password typed as follows:

\[ T_1, T_2, T_3 \ldots T_n, U_1, U_2, U_3, U_4, U_5, P_1, P_2, P_3, P_4, P_5 \]

where,

- \( T_i \) is the \( i^{th} \) time interval for the \( P_j^{th} \) password typed by the \( U_k^{th} \) user, \( i = 1,2,3,\ldots,n \)
- \( P_j \) is the \( j^{th} \) password typed by the \( U_k^{th} \) user (0 or 1), \( j = 1,2,3,4,5 \)
- \( U_k \) is the \( k^{th} \) user (0 or 1), \( k = 1,2,3,4,5 \)

The data were collected from five users each of whom has different levels of computer skills over a five-week period. Each user was asked to enter each password at least 5 times a week in a random fashion in terms of order of password entry and the days of the week. After the entries completed, the typing patterns for each user were evaluated and additional entries were made until the necessary number of patterns has been reached statistically.

NEURAL NETWORK ARCHITECTURE

A layered neural network provides more potential alternatives than traditional pattern recognition techniques (Huang and Lippmann, 1987). On the other hand, a pattern recognition technique, defined as a classification model, is concerned with performing feature extraction, learning the transparent mapping, and classifying patterns (Pao, 1989). For the task of pattern recognition using layered neural networks, inputs correspond to features, connections between layers correspond to mapping, and outputs correspond to pattern classes. In addition, a layered neural network may contain one or more hidden layers, which represent the domain knowledge and help to perform feature extraction.
The neural network used in this study was made up of three layers; input, hidden, and output. Each layer was connected to the upper layer, inter-layer connections, via weights, randomly generated real values. A sigmoidal function was used to determine the new activation values of the neurons. The number or neurons in the network architecture was varied depending on the experiments. The input layer was composed of 8-13 neurons, represented time intervals between successive keystrokes obtained from the passwords selected. The output layer was made up of 4-8 neurons for the desired output values of each pattern; for instance, 10001000 represented that the first user entered the first password.

As indicated in (Klimasauskas, 1991), more neurons in the hidden layer reduces total error in training; however, fewer neurons increases the network performance for generalization, meaning that the ability to correlate a pattern with previously used patterns. For that reason, the number of neurons in the hidden layer(s), as needed, which yield to extract features between the input and the corresponding output pattern, were varied between 4 and 8 depending on the experiments. The learning rate was arbitrarily assigned to 0.15, and momentum term to 0.4.

EXPERIMENTAL RESULTS AND DISCUSSIONS

The collected data from password entry process were normalized time intervals based on the fraction of the largest element in the data set before presenting them to the neural network. In regard to the computer access security system, three experiments were conducted. First, the time intervals obtained from the users were fed to a neural network consisting of 13 input and 4 output neurons to verify whether the users may be classified into appropriate groups assuming that 4 users may have 4 passwords which could be used by each of the users. This experiment was conducted for the purpose of preventive security.

In the second experiment, each password was assigned to each user only based on his/her preferences. Then, the data consists of time intervals belong to a specific password selected by a user were introduced to a neural network, which had 13 input and 8 output neurons, to initiate a multiple users-multiple passwords situation. This experiment was performed for both preventive security and detection of violation purposes. Since each user has his/her own password to access to a part of or complete system, this situation may be considered as password-dependent recognition. In the last experiment where a neural network with 8 neurons for the input and output layers, respectively, was used. The length of the patterns, time intervals obtained from the password entered, were formed, such that each pattern has an equal number of time intervals regardless of the number of characters they may have, to discuss an issue of password-independent recognition in computer access security.

The multilayered neural networks were trained using the proper data prepared for each of the experiments. In testing phase, the patterns which were not included in the training set, were fed to the designed neural networks and the performance of the each neural network was evaluated according to the correct/wrong classifications (Type I error). Also, time intervals obtained from an unauthorized person for the system involved and not included in training data were tested to verify whether the person may be considered as an authorized one by checking the similarity of the time intervals with the ones produced by the group members (Type II error).

In the first experiment, the patterns were recognized with the accuracy of 96.3% in training and 94.2% in testing. On the other hand, when the patterns of the unauthorized person were tested, the person was recognized as authorized one with the error level of 26.9% because of the similarity in the patterns. A recognition accuracy of 98.7% was obtained for training phase in the second experiment, and Type I and II errors were decreased down to 1.4% and 4.2%, respectively, in testing since both user identification number and the pattern code of the entered password were examined together at each query. In the last experiment, a value of 16.3% was obtained for the Type I error, and 8.2% for the Type II error. The results concluded that when the user identification number and a password for that user were questioned sequentially, a better performance in computer access security system may be obtained.

In addition, for the purposes of preventive security and detection of violations, a procedure was employed to evaluate an attempt by a person who is unauthorized for accessing to the system involved. The procedure may be used to make the system more reliable when a time interval pattern close to the one produced by the group members during on-line entry. As in real applications, a user was asked to enter his/her identification number and password. Then the
absolute difference between the desired and the actual outputs was calculated. When the calculated difference was higher than a predetermined value, the attempt was denied and user was asked to enter the necessary information until the number of trials reached to three.

As it is expected, the time intervals between keystrokes are shortened as more entry made. To increase system’s reliability, when the absolute differences for the entries made are between a lower limit and the predetermined value, such entries could be added to the training data files and the designed neural network may be retrained. By changing the limits appropriately after every retraining, the user may be recognized more precisely and the reliability of the system may be increased.

CONCLUSION

Multilayered neural networks were applied to computer access security system. Time intervals between successive keystrokes during password entry, which are considered special characteristics of the users, were used to differentiate users, prevent unauthorized person to access the system, and try to detect the violators. Three experiments were conducted via neural networks having different architectures. The experimental results showed that the neural network trained using time intervals of a specific password along with the user identification number provided better performance than the others. A procedure was employed to be able to increase the reliability of a security system.

Other issues, such as determining the optimal values of parameters of the neural network used to improve recognition accuracy, comparing the results of this research with other pattern recognition techniques, seeking better security code alternatives to differentiate users more precisely, dealing with the effect of computer configurations in terms of speed during password entry, and modifying the neural network architecture or integrating the neural network with an expert system, if applicable, to be able to implement this approach in an on-line mode are still remain for further investigation.

REFERENCES


