자기연상 다층 퍼셉트론을 이용한 키 스트로크 기반 사용자 인증

조성준(*), 한대희(**)
포항공대 전자계산학과(*) SK 텔레컴(**)

Keystroke Dynamics based User Authentication with Autoassociative MLP

Sungzoon Cho(*) and Daehee Han(**)
POSTECH(*) and SK Telecom (**)

요 약

Password checking is the most popular user authentication method. The keystroke dynamics can be combined to result in a more secure system. We propose an autoassociator multilayer perceptron which is trained with the timing vectors of the owner's keystroke dynamics and then used to discriminate between the owner and an imposter. An imposter typing the correct password can be detected with a very high accuracy using the proposed approach. The approach can also be used over the internet such as World Wide Web when implemented using a Java applet.

제 1 절 Introduction

Password is the most widely used tool for computer access security. More often than desirable, however, easy-to-guess words such as a family member name, a birthday, a phone number, an address, etc. are chosen for password, which results in a security failure. Some other means should be devised to replace or be combined with the password.

In general, there are three different approaches to system security, possession-based (key or card), memory-based (password), and biometrics-based (fingerprints or keyboard dynamics) (Davis and Price, 1989). They are evaluated in such criteria as the error rate, cost, user discomfort and environmental requirement. Each approach has pros and cons. Possession based methods are very cheap and simple, but could be too simple in that they allow access for anyone with the key or card. More sophisticated methods which employ fingerprints and retinal patterns involve an extra hardware device and increased user discomfort. Since no approach is perfect, it is usually recommended to combine measures.

One approach that is both inexpensive and simple is the keyboard dynamics. When a user types a word, for instance one's password, the keyboard dynamics can be characterized by the "timing vector" consisting of the duration time of key strokes and the interval time between them. A

word of n characters and "Return" results in a timing vector of dimension $2n+1$. The owner's timing vectors are collected and used to build a model which is to discriminate between the owner and imposters. The approach has many advantages. It is low cost and causes no user discomfort. It can also fit with the internet environment (i.e., World Wide Web) A web server can dynamically send a user a Java applet which measures a timing vector using language construct java.awt.Event. The approach can also be naturally combined with the password, providing twofold security. The only disadvantage has been its relative low accuracy. Previous studies reported error rates much larger than 10% which is practically unacceptable.

In this paper, we propose an autoassociator neural network model which reduces the error rate significantly. Timing vectors from an owner were collected and used to build a neural network model which outperformed a more conventional Nearest Neighbor (k-NN) approach. Although experiments involving many more owners are required for practical use of the approach, the preliminary results are the best ever reported to authors' knowledge.

The next section briefly describes the results of the previous studies concerned with the keystroke dynamics based security. Then, we propose the neural network novelty detector. Data collection and experimental results are presented, followed by a short summary and discussion of ongoing and future research issues.

제 2 절 Previous Studies

Biometrics-based approaches have two types of errors. The false accept rate (FAR) denotes the ratio that an imposter passes while the false reject rate (FRR) denotes the ratio that the owner fails. One type of error can be reduced at the expense of the other. An appropriate middle point is usually used as a threshold based on the relative cost of the errors. Another widely used error measure is FRR when FAR is reduced to zero.

In the past, a short character string such as a password was regarded inadequate to be used for user authentication (Nelson, Forsen and Staron, 1977). A long string of 537 characters for example had to be employed to achieve 5.0% FAR and 5.5% FRR (Williams, Leggett and Usnick, 1991). Only recently through the use of neural networks, a comparable performance of 12% to 21% was achieved with short strings such as real life names (Brown and Rogers, 1993). These error rates are still too high to be practically acceptable. In addition, they trained the neural network not only with the owner's timing vectors but all those of the imposters in advance. In real

life situations, this is unacceptable because the owner's password has to be revealed to the network users at large. In the late 80's, two US patents were granted to the statistical approaches, but their performance is not available (Garcia, 1986; Young and Hammon, 1986).

A lower error rate of 2.5% was obtained when the user identification problem was solved (Obaidat and Macchairolo, 1994). The problem is to find who typed the password among several candidates instead of checking if the timing vector is from the owner. The network had to be trained with the timing vectors from all candidates. Unfortunately, the result can not be applicable to user authentication problem. Also recently 0% error rate was reported for user verification using 7 character-long login names (Obaidat and Sadoun, 1997). However, negative examples (i.e., intruder's typing patterns) as well as positive examples (i.e., owners' patterns) were used for training. Also the training data set was much larger (6,300 positive and 112 negatives). Also, the training and test patterns were not chronologically separated. These factors lead into a less practical scenario.

제 3 절 Autoassociative MLP Novelty Detector

User authentication is challenging from a pattern classification point of view. It is a two class (owner vs imposters) problem, yet the patterns from only one class, the owner's, are available in advance. Since there are millions of potential imposters, it is not practical to obtain enough patterns from all kinds of imposters. Also it is not desirable to publicize one's password in order to collect potential imposters' timing vectors. The only solution is to build a model of the owner's keystroke dynamics and use it to detect imposters using some sort of a similarity measure. This type of problem has been known as "partially exposed environment" (Dasarathy, 1980) or "novelty detection." Another important area of applications is in fault diagnosis where most of the time things are in a normal condition while information on abnormal conditions is necessary. Usually, a model of normal conditions is built and then used to detect abnormality or novelty.

A multilayer perceptron can be used to detect novelty (Frosini, Gori and Priami, 1997). Owner's patterns are used to train the network to be an autoassociator, i.e., by using a timing vector as both an input and a target output. The MLP is trained to learn to encode certain properties only present in the owner's timing vectors at the hidden layer. When a previously unseen timing vector of the owner arrives, the network is expected to output a vector that is reasonably close to the input. When an imposter's pattern

arrives, on the other hand, the network is expected to output a vector that is far from the input. That is, a timing vector X is classified as owner if and only if

$$||X - M(X)|| < \epsilon$$

where $M(X)$ and ϵ represent the MLP's output for X and a threshold. An issue of interest is whether the MLP based autoassociators can learn closed separation surfaces which contain the owner's training patterns. Frosini et. al provided the theoretical results supporting it when they applied the network to paper currency verification (Frosini, Gori and Priami, 1997).

The proposed autoassociative MLP approach is compared with a more conventional Nearest Neighbor (NN) approach. When a new timing vector arrives, the average "distance" to the k closest training patterns is computed. If the average distance is smaller than a predetermined threshold, the timing vector is classified as from the owner. Otherwise, it is classified as from an imposter. The distance between two vectors is defined as $(\vec{x}-\vec{y})^T M(\vec{x}-\vec{y})$. In order to give more weights to those elements with a smaller variance, the inverse of a covariance matrix Σ can be used for M which results in so called Mahalanobis distance.

제 4 절 Experimental Results

A program was developed to measure the key stroke duration times and interval times in X window environment on a Sun Sparcstation. A PC version was also developed but was not used in the experiment reported here. A password of 7 characters long results in a timing vector of dimension 15 since the duration of "Enter" key is included. An example of a timing vector is [120,60,120,90,120,60,150,-60,120,-30,120,-60,120,120,90,60,150] where each element was measured in miliseconds. A negative interval time results from a situation where a next key is stroked before a previous key is released.

A total of 25 subjects were asked to come up with one's new password. Each subject or owner typed one's password 150 to 400 times during a period of several days. The 75 timing vectors collected last were set aside for testing. The remaining timing vectors were used for training the network. If any of its element is larger than the upper 10 percent, however, the vector was classified as an outlier and discarded. Depending on the owner, 6 to 50% of the training vectors were discarded. There were four owners whose discard rates were higher than a one third. A high discard rate implies that the owner did not become comfortable with the new password. Since

Owner ID	Password	No Tr Ptn	Discard rate	FRR when FAR = 0	
				1-NN	MLP
1	loveis.	207	0.21	22.7	2.7
2	i love 3	330	0.15	30.7	0.0
3	autumnman	111	0.10	0.0	0.0
4	90200jdg	164	0.10	5.3	0.0
5	rla sua	101	0.18	8.0	1.3
6	dhfpql.	232	0.08	17.3	2.7
7	love wjd	101	0.19	54.7	0.0
8	dltjdgml	151	0.14	0.0	0.0
9	dusru427	365	0.27	0.0	0.0
10	manseiii	86	0.25	60.0	1.3
11	rhkdwo	205	0.20	18.7	0.0
12	beaupowe	76	0.24	9.3	4.0
13	tmdwnsl1	108	0.18	17.3	4.0
14	yuhwa1kk	388	0.12	0.0	0.0
15	anehwksu	319	0.10	10.7	0.0
16	tjddmswjd	337	0.10	33.3	0.0
17	drizzle	299	0.10	9.3	1.3
18	dlfjs wp	342	0.06	1.3	0.0
19	c.s.93/ksy	200	0.22	17.3	2.7
20	dirdhfmw	309	0.33	89.3	0.0
21	ahrfus88	260	0.20	5.3	0.0
Avg.		223	0.17	19.5	1.0
Min.		76	0.06	0.0	0.0
Max.		388	0.33	89.3	4.0

표 1: Passwords characteristics and error measures from respective models

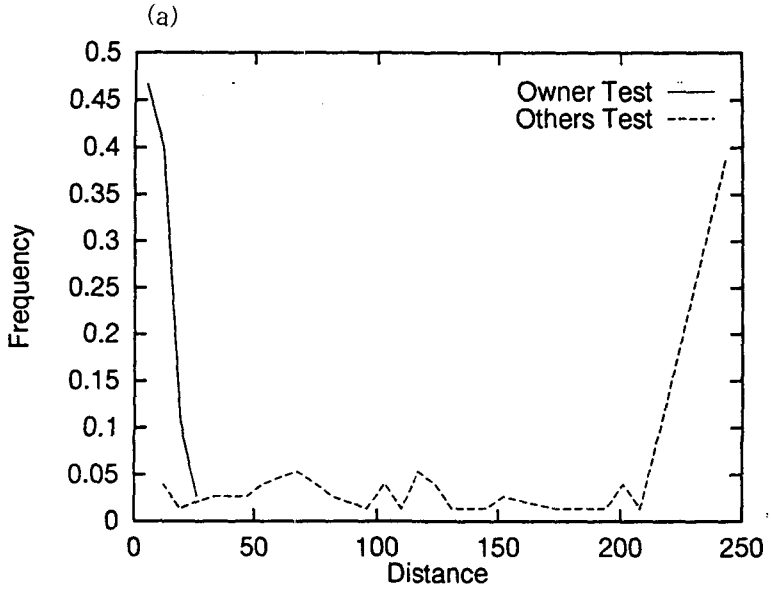

그림 1: Histograms of average distance measure (1-NN). The better separated the owner's and imposters's populations are, the better the classification is.

we only consider experienced owners, we removed those four owners from the experiment. A total of 15 imposters were given all the 21 passwords and asked to type each password five times, resulting in 75 imposter test vectors for each password. Combined with the owner's 75 test vectors set aside before, a total of 150 test vectors were obtained. Table 1 shows, for all 21 owners, the respective password, the number of training patterns and the discard rate. Some passwords, such as 5, 6 and 8, are words in Hangul, Korean alphabet. We simply show the corresponding English alphabets.

Also shown in Table 1 are the error rates for k-NN and MLP approaches. The error is the False Reject Rate (FRR) when False Accept Rate (FAR) was reduced to zero. For k-NN, we tried 1, 2, and 3 for k values and obtained the best result when $k = 1$ shown here. Each MLP contains the same number of hidden units as input units. All 21 MLPs were trained with a standard backpropagation algorithm with a learning rate of 0.1 and a momentum term of 0.3 for 500 epochs. The proposed MLP approach clearly outperformed the k-NN. A perfect authentication was achieved for 13 owners. The worst performance was from owners 12 and 13 with the error rate of 4.0%. The average error rate was 1.0%.

The MLP approach's performance advantage is clearly visualized in

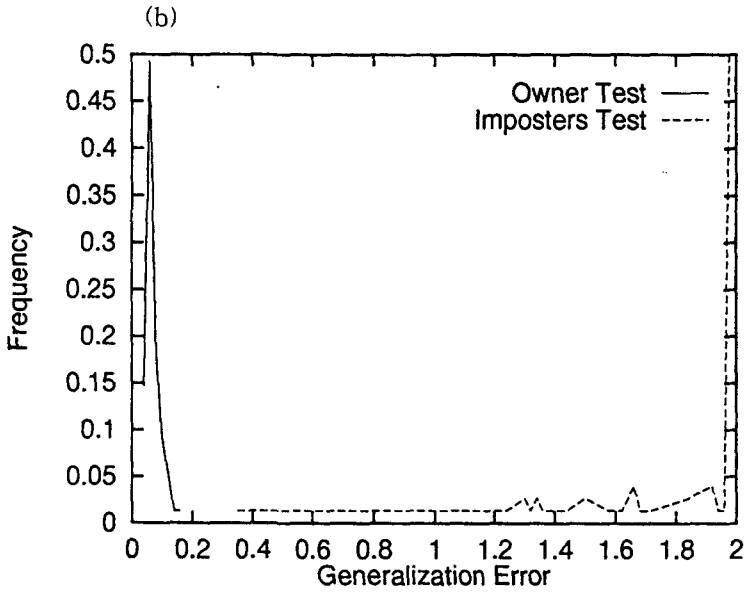

그림 2: Histograms of generalization error (MLP). The better separated the owner's and imposters's populations are, the better the classification is.

Fig.s 1 and 2. The histogram distributions are from owner 2. A total of 150 test patterns, one half from the owner and the other from 15 different imposters, were presented to two respective models. For 1-NN, the average distance from the nearest neighbor was computed. For MLP, the generalization error was computed. The resulting histograms show why the MLP gives perfect authentication (no overlap between owner's and imposter's test vector histograms) while the 1-NN gives 30% error (significant overlap).

제 5 절 Conclusions

An MLP-based novelty detector was proposed for user authentication using keystroke dynamics. An autoassociative MLP is built from a set of previously collected timing vectors from the owner. When a new timing vector arrives, it is presented to the MLP, the output is computed and compared with the input. If it is close enough to the input, the input timing vector is classified as from the owner. If not, it is classified as from an imposter. The experimental results involving 21 skilled users show that the proposed approach is significantly more effective than k-NN approach. For 13 owners, the MLP approach achieved perfect authentication. Among the rest,

the worst performance was just 4% error rate. The overall average error rate was 1%. The preliminary result reported here is quite promising. The proposed approach can be implemented in any password typing situations including the network environment, World Wide Web, for instance.

Further investigation is planned in the following areas. First, a much larger number of experiments involving human subjects are to be done. Also considered are those issues on how to deal with the inexperienced users as well as learning effects and fatigue effects. Second, the issue of how to make the system more practical. It is essential that the number of necessary training patterns should be minimized. Due to the nature of a neural network, however, a smaller training set makes it hard to learn the function appropriately. This is one reason why a three hidden layer autoassociator which does nonlinear encoding is not employed even though such a model could be more powerful. Finally, we are working on ways to implement Java applets which measure the timing vectors and send them back to a Web server so that the scheme could be used over the internet.

참고 서적

[1] M. Brown and S. J. Rogers. User identification via keystroke characteristics of typed names using neural networks. *International Journal of Man-Machine Studies*, 39:999–1014, 1993.

[2] B. V. Dasarathy. An alien identification approach to recognition in partially exposed environments. In *Proceedings of the 5th International Conference on Pattern Recognition*, pages 91–93, December 1980.

[3] D. Davis and W. Price. *Security for Computer Networks.* John Wiley & Sons, Inc, 1989.

[4] A. Frosini, M. Gori, and P. Priami. A neural network-based model for paper currency recognition and verification. *IEEE Transactions on Neural Networks*, 7(6):1482–1490, 1997.

[5] M. Nelson G. Forsen and R. Staron. Personal attributes authentication techniques. In A. Griffs, editor, *Rome Air Development Center Report RADC-TR-77-1033*, 1977. New York:RADC.

[6] J. Garcia. Personal identification apparatus. Patent No. 4,621,334. U.S. Patent and Trademark Office, Washington D.C. 20231, 1986.

[7] G. Williams J. Leggett, M. Usnick, and M. Longnecker. Dynamic identity verification via keystroke characteristics. *International Journal of Man-Machine Studies*, 35:859–870, 1991.

[8] M. Obaidat and S. Sadoun. Verification of computer users using keystroke dynamics. *IEEE Transactions on Systems, Man and Cybernetics, part B: Cybernetics*, 27(2):261–269, 1997.

[9] M. S. Obaidat and D. T. Macchairolo. A multilayer neural system for computer access security. *IEEE Transactions on Systems, Man, and Cybernetics*, 24(5):803–816, May 1994.

[10] J. Young and R. Hammon. Method and apparatus for verifying an individual's identity. Patent No. 4,805,222. U.S. Patent and Trademark Office, Washington D.C. 20231, 1989.