以類神經網路探討按鍵輸入型樣認證問題
Neural Network Technique in Authentication Using Key Stroke Pattern Recognition

汪仲甫
C.F.Uang
Department of Computer
Engineering and Science
Yuan-Ze University
No. 135 Yuan-Tung Rd.,Chung-Li,
Taiwan.R.O.C.

黃朝章
C.J.Hwang
Department of Computer Engineering and
Science
Yuan-Ze University
No. 135 Yuan-Tung Rd.,Chung-Li,
Taiwan.R.O.C.
cschwang@cs.yzu.edu.tw

摘要

電腦與通訊技術日新月異，電腦網路已成為人們相互交換信息的重要管道。然而其中安全上潛藏的問題與危機是必需考量的，如何確保通訊內容不會被他人卻竊取，或惡意串竄改，如何確認雙方身份而不被有心假冒，特別是在金融機構的應用上，保密安全上的考量是必需注意的。

藉由身份認證與金匙分配密碼協定(Authentication and Key Distribution Protocal)可以防止資料在傳遞時被截取、竄改及有效確對方身份，已有效應用於電腦網路中。不過在電腦或終端機如何認證使用者身份，最常用還是人們鍵入通行碼或ID。

由使用者在電鬧腦鍵盤上鍵入 ID 與通行碼，是最常見鑑別身份方式，不過也是最危險。如何預防使用者鍵入個人password，而不被他人截取，是非常重要課題。本論文即探討如何在使用者鍵入密碼時，防止他人的截取.

關鍵字:認證,類神經網路,型樣辨識

Abstract

The technique of Computer and Communications are renovated daily, the Computer Network role in interchanging information. Therefore to concern the security of data transfer and ensure the data assurance become more important than ever. For example, how to protect the data not been stolen or mischief while connection? how to confirm the personal identification from both end? Specially in the usage of Finance Institution, the security system should be tremendous considerate.

The "Authentication and Key Distribution Protocol"[1]-[5] has proved his efficiency in preventing the break-in , damage or tamper data, and to confirm the identification code for the Computer Network Application package. But, most people omitted one important point: how to identify the user authorization through the terminal or monitor? The most common situation is Key in the user ID or Password from keyboard. Allowed the user to key in their own ID and Password in the normal but also the dangerous way to enter the Application system . This paper is to discuss how to prevent the mischief personal password.

Keywords : Authentication, Neural Network, Pattern Recognitio

1. Introduction

As a user keys in Password, added with the consideration of personal keying in habit to strengthen the protection of attack by intentional offenders shall be the subject of research of this paper. Because it is believed that each person shall have his own fixed mode in keying in, and the modes differ from one from another which cannot be intercepted, learned or simulated by others.

In 1987, S. Bleha and C. Silvinsky [6], in 1980, B. Hussien, R. McClaren and S. Bleha [7], in 1991, S. Bleha and M.S. Obaidat [8], in 1992, Mohammad S. Obaidat and David T. Macchiarolo [9] [10] all proposed related studies. Among them, except for Mohammad S. Obaidat and

David Macchiarolo use the Artificial Neural Network as method of problem solution to attain larger improvement, the formerly mentioned other studies tend to use the method of acquiring the minimum distance or only base on the improvement of this type to make the users to be required to key in at least more than 30 words in order to attain the less error.

The study of this paper, similarly, is made on Artificial Neural Network as nucleus with reinforcement on personal keying in behavior habit consideration:

1. It shall be able to record the time interval of continued keying and also the time spent for each making and breaking of each key. In this way, one can further differentiate the key making behavioral mode of each person.
2. The time consideration shall be made on unit of 10-3 second but not the basic unit second of computer time or 1/18.2 second.
3. It is required to analyze the length of keyed in character string in order to effectively differentiate the user.

Based on the above considerations, this paper will focus on particular analysis concerning:

1. In keying in pure character keys, the individual identification rate of each of the 3 kinds of keying in condition.
2. While keying in pure number keys, can it also be able to be successfully identified.
3. How long the keyed-in word string is can still possess lower error ratio and better safety.
4. Analyze whether the user, under condition of keying in with different computers, will cause different influences.

2. System Structure

The integral key-in type identification structure can be divided into two major parts: training stage, classification stage. Besides, 10-3 second serves as the basic time measurement unit. The time interval between each of the continued key making, the idling time for each make and brake of each key, and synthetic of the former two shall be considered separately.

In training stage, the users may key in the same word string for several times to allow the system to be able to learn the key-in behavioral habit of a user by means of the learning and practice method of reverse artificial neural network [11] [12].

The classification stage is the recall stage of artificial neural network, used in normal verification process, through the personal behavioral mode data obtained in training stage, to judge each individual user.

3. Results Analysis

3.1 Key-in Condition Identification Ratio of 3 Kinds of Pure Text Input

An analysis is made on the 40 batches of word strings keyed in by 6 people with word string of 15 bytes " wonderful world ". Among them, the first 20 batches shall serve as test data, the latter 20 batches shall serve as training data.

The following shows the correct identification analysis of continued key-press time interval, make/break idling time of each single key, and synthetic consideration time sampling:

	continued key-press time interval	make/break idling time of each single key	synthetic consideration
Correct Ratio	85.8%	91.7%	99.2%

Table 1: Comparison of Identification Correct Ratio of 3 Kinds of Key-in Condition

If only the continued key-in time interval is considered, the result differs not much from the experiment conducted by former people. But in case consideration is made on the single-key make/break time sampling proposed by the paper, then the identification ratio will not be worse than the one with mere consideration on continued key-in time interval. It is obvious that this parameter can certainly be in use. Besides, if both are to be considered, 99.2% correct ratio can be obtained.

3.2 An analysis is made on the 40 batches of strings

keyed in by 6 people, with word string of 15 bytes "046388003913455", if test is in concurrent consideration of the make/break time condition of each key, the identification ratio shall be:

	synthetic consideration
Correct Ratio	80.8%

Table 2: Identification Ratio of Keying in Pure Numbers

Since the habit of keying in number key by most people tends to use a single finger and the number key is not so familiar as that for word key; the memory of long number string to partial people is very difficult. Therefore, even if two time sampling of two kinds is considered concurrently, only 80.8% of correct identification ratio can be obtained.

3.3 Analysis of Key-in Making Length and Identification

3.3.1. Time Interval of Continuous Keying

If consideration shall be made on the continuous keying time interval, the identification ratio change can be observed from the word string " wonderful world" at different key-in length.

	1 key	2 keys	3 keys	4 keys	5 keys	6 keys	7 keys	8 keys
Correct Ratio	45.8%	55.8%	64.2%	70.0%	77.5%	74.2%	81.7%	87.5%

	9 keys	10 keys	11 keys	12 keys	13 keys	14 keys	15 keys	
Correct Ratio	85.8%	87.5%	88.3%	87.5%	87.5%	86.7%	85.8%	

Table 3: Identification Ratio of Each Kind of Key Making on Continuous Key-in Time Interval

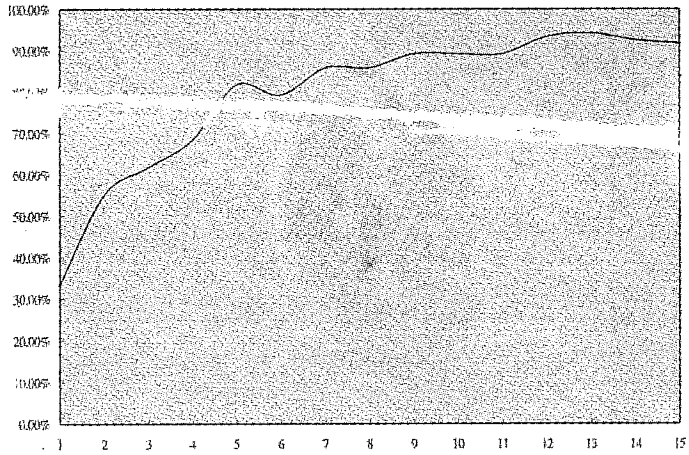

Diagram 1: Analysis Diagram for Identification Ratio of Each Kind of Key Making on Continuous Key-in Time Interval

From the diagrams and charts, one may discover, under most conditions, the longer the key-in length is, the higher is the correct identification ratio. From the start of 7-key making, there is more than 80% correct identification ratio.

3.3.2 Make/Break Idling Time for Each Single Key

If consideration can be made on the make/break idling time for each single key observation can be made on the identification ratio of the word string "wonderful world" in case of different lengths:

	1 key	2 keys	3 keys	4 keys	5 keys	6 keys	7 keys	8 keys	9 keys	10 keys	11 keys	12 keys	13 keys	14 keys	15 keys	
Correct Ratio	33.3%	55.0%	61.7%	68.3%	81.7%	79.2%	85.8%	85.8%	89.2%	89.2%	89.2%	93.3%	94.2%	92.5%	91.7%	

Table 4: Identification of Various Key-in Length Concerning the Make/Break Idling Time for Each Single Key

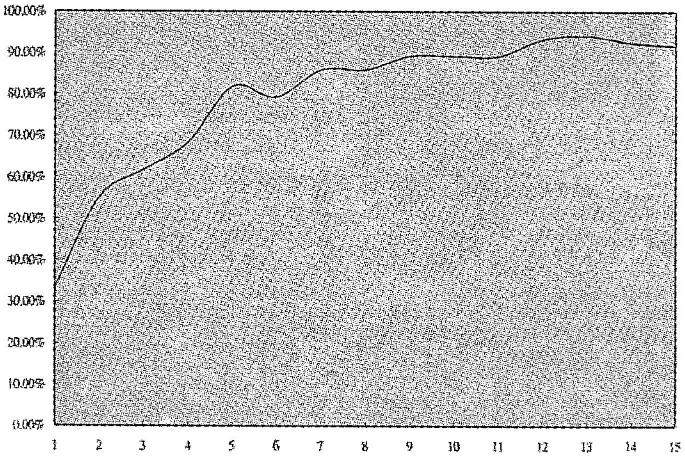

Diagram 2: Analysis Diagram for Identification of Various Key-in Length Concerning the Make/Break Idling Time for Each Single Key

From the diagrams and charts, one may discover, under most conditions, the longer the key-in length is, the higher is the correct identification ratio. From the start of 6 or 7-key making, there is commendable correct identification ratio.

4. Conclusions and Perspective

In concluding the above discussion, if concurrent consideration is made on continuous key-in time interval, and the make/break idling time of each single key, individual key making behavioral habit can be further identified. That is to say, the safety system, while the user is keying in Password, besides inspecting the user's

ID, and Password, personal keying behavioral habit will be further inspected and it is feasible.

In the future, continued effort can be made toward the following research directions:

1. Solve the influence on an unfamiliar keyboard by the user.
2. In practical application, attack protection consideration shall be strengthened. For example, attack with a keyboard simulator.
3. Apply the same method to th- identification of personal facial type, palm print and finger print analysis.
4. Apply the proposed methods of the paper to differentiate document key-in person to attain signature effect.
5. Apply the proposed methods of the paper to Password safety and develop a practical product.

References

[1] R.M. Needham and M.D. Schroeder, "Using Encryption for Authentication in Large Networks of Computers," Communications of the ACM, Vol. 21, No. 12, December 1978, pp.993-999.

[2] J.G. Steiner, C. Neuman, and J.I. Sehiller, "Kerberos: An Authentication Service for Open Network Systems," Proceedings of the Winter 1988 USENIX Conference, 1988, pp.191-202.

[3] L. Gong, "A Security Risk of Depending on Synchronized Clocks," ACM Operating Systems Review, Vol. 26, No. 1, January 1992, pp.49-53.

[4] A. Kehne, J. Schonwalder, and H. Langendorfer, "A Nonce-Based Protocol for Multiple Authentications," ACM Operating Systems Review, Vol. 26, No. 4, October 1992, pp.84-89.

[5] B.C. Neuman and S.G. Stubblebine, "A Note on the Use of Timestamps as Nonces," ACM Operating Systems Review, Vol. 27, No. 2, April 1993, pp.10-14.

[6] S. Bleha, and C. Slivinsky, "Computer access security system based on time dutations betteen termial keystrokes," in IEEE Proc. Miami Technicon '87 Conf. Miami, FL, pp. 217-220, Oct. 1987.

[7] B. Hussien, R. McClaren, and S. Bleha, "An application of fuzzy algorithm in a computer access security system," Pattern Recognition. Letter., vol. 9, no. 1, pp.39-43, Jan. 1989.

[8] S. Bleha and M. S. Obaidat, "Dimensionality reduction and feature extraction applications in identifying computer users," IEEE TRANSCATIONS ON SYSTEM, MAN, AND CYBERNETICS, vol. 21, pp. 452-456, Mar./Apr. 1991.

[9] Mohammad S. Obaidat and David T. Macchiarolo, "A Multilyer Network System for Computer Access Security," IEEE TRANSCATIONS ON SYSTEM, MAN, AND CYBERNETICS, Vol. 24, No. 5, MAY 1994, pp.806-813.

[10] Mohammad S. Obaidat and David T. Macchiarolo, "An On-Line Neural Network System for Computer Access Security," IEEE TRANSCATIONS ON INDUSTRIAL ELECTRONICS, Vol. 40, No. 2, APRIL 1993, pp.235-242.

[11] D. E. Rumelhart and J. L. McClelland, eds., Parallel Distributed Processing：Explorations in the Microstructure of Cognition, vol. 1, Cambridge, MA:MIT Press, 1986.

[12] D. E. Rumelhart, G.E. Hinton and R. J. Willians, "Learning representations by back-propagating error," Nature, vol. 323, pp.533-536,1986.