I KNOW IT'S YOU: TOUCH BEHAVIORAL CHARACTERISTICS RECOGNITION ON SMARTPHONE BASED ON PATTERN PASSWORD

Yong Liu, School of Economic Information Engineering, Southwestern University of Finance and Economics, Chengdu, P. R. China, liuyong.swufe@qq.com
Jiafen Liu, School of Economic Information Engineering, Southwestern University of Finance and Economics, Chengdu, P. R. China, Corresponding Author: jfliu@swufe.edu.cn
Zhangxi Lin, School of Economic Information Engineering, Southwestern University of Finance and Economics, Chengdu, P. R. China, zhangxi.lin@gmail.com
Xubin Luo, School of Economic Information Engineering, Southwestern University of Finance and Economics, Chengdu, P. R. China, xubinluo@gmail.com
Jiang Duan, School of Economic Information Engineering, Southwestern University of Finance and Economics, Chengdu, P. R. China, duanj_t@swufe.edu.cn

Abstract

In recent years, pattern password has been widely used for user authentication on smartphones and other mobile devices in addition to the traditional password protection approach. However, pattern password authentication mechanism is incapable of protecting users from losses when a user's login credential information is stolen. We propose an identity verification scheme based on user's touching behaviors when inputting a pattern password on the smartphone screen. By exploiting the biometrical features, such as position, pressure, size, and time when a user inputs a pattern password to a smartphone, the proposed user verification mechanism can validate whether the user is the true owner of the smartphone. We adopted fuzzy logic, artificial neural network, and support vector machine, to build classifiers, using the behavioral data collected from 10 users. The experimental results show that all the three algorithms have significant recognition capacity, and the fuzzy logic algorithm is the best one with its false acceptance rate and false rejection rate as 4.7% and 4.468% respectively.

Keywords: Pattern Password, User Authentication, Biometric Recognition, Mobile Devices

1 INTRODUCTION

Smartphones and other mobile devices have been gradually replacing PCs as the preferred tool for people on the Internet. More and more people use mobile devices to search information, shop online, pay from their electronic bank account, store important information, send emails, and perform different kinds of social networking. According to statistics, in 2014, mobile online shopping transactions in China have reached 861.66 billion yuan, up 229.3% from a year ago. The third-party mobile payment reached 7.766 trillion yuan, an increase of nearly 500% over the same period. Meanwhile, the online attacks via mobile devices, such as fraud SMS, phone lost, become a serious threat to the security of mobile payments. QR code Trojan phishing and electronic password upgrading fraud are currently typical network tricks used in attacking mobile payments, which steal the credentials of payment account, intercept and modify messages sent to the bank, or directly empty your account. As a consequence, security authentication mechanism based on pattern password is widely used to deal with these smartphone security problems (Jermyn 1999). Pattern password in smartphone is widely used for locking screen or the smartphone application to prevent unauthorized users from operating the mobile phone device or mobile phone applications. With the input of specific Numbers, letters or other characters of the password, the pattern password mechanism combined with the touch function of the smartphone. One just simply uses a finger to connect given points on the screen as a password. It works not only very fast, but also is easy to remember. However, this mechanism is not necessary secure enough. For example, a recent study shows that an attacker can extract the oily residues left on the phone screen to analyze the password (Aviv 2010). An attacker could also use the accelerator and gyroscope to detect the position pressed on the phone screen and then extract the letters or numbers (Cai 2011; Miluzzo 2012; Owusu 2012; Xu 2012). Therefore, a more efficient user authentication mechanism for improving the security of smartphones becomes necessary. Such a mechanism must be able to verify whether the user who mastered the pattern password is the real owner of the smartphone.

This paper explores the feasibility of using touching behavioral characteristics extracted to authenticate the user. Our working principle is that each person has a unique pattern of touching the screen in input the pattern password. We used sensors of smartphone to collect user's touching features. When a user was drawing a pattern, we extracted behavioral characteristics in user's touching process, which include touch coordinates, pressure, contact area and time and so on. Based on these characteristics, we constructed an accurate classifier. In practice, our authentication mechanism can be seamlessly integrated into existing password system, transparent to users and with no additional cost. It. Experimental results based on 10 users' touching data show that our authentication mechanism can effectively improve the security of smartphone.

The remainder of this paper is as follows. The second part of the paper reviews the background and related work. The third section explains our data collection and processing, including analysis of the data. Section 4 describes the experimental research methods in details. Section 5 presents the experimental results and discussion. Section 6 concludes this paper and propose the ideas for future work.

2 BACKGROUND AND RELATED WORK

Advanced IT benefits our society with revolutionary changes but also brings up the new challenges in security. OpenSSL has been referred to as the best-known security vulnerabilities in 2014 (Durumeric 2014), because it let a hacker have the potential to access approximately 30% user's login account and password beginning with HTTPS URLs. In that case, security attacks may reach popular shopping webs, online banking, social networks, portals, Sina Weibo, WeChat, email and so on, affecting at least two hundred million Chinese Internet users. Thus static authentication, though it is the most widely used authentication mechanism, is unable to protect users given that password is stolen.

Authentication mechanism based on biometric characteristics is a good supplement to static authentication. Since signature has been used to verify one's identity in various scenarios of our daily life, such as goods receipting and banking transactions confirmation, scholars studied the feasibility of automatic verification by recognizing an individual's handwritten signatures (Abuhaiba 2007). A digitizer or tablet was used to obtain the digital image of handwritten signature. Static information, such as trajectory and direction, and dynamic information, such as speed and pressure of writing were collected. Jain describes an on-line handwritten signature verification scheme(Jain 2002). Signatures were acquired using a digitizing tablet, which captured both dynamic and static information of the writing. All strokes were combined into one long stroke in preprocessing. Both spatial and temporal features were extracted from the shape of the signature. It achieved optimal threshold after several approaches. It collected a database containing 1232 signatures of 102 individuals. The result yields a false accept rate of 1.6% and false reject rate of 2.8%. Nowadays, some high-end smartphones have been equipped with fingerprint identification. But both fingerprint and handwritten signature authentication need special hardware to work with.

Besides signature, there has been a long history for keystroke dynamics being used in user authentication. In 1895, through the observation of transmitter operators, Bryan and Harter found that each operator has its own unique keystroke pattern when sending the same period of the packet. According to the commas, periods and other characteristics, we can distinguish who sent the telegram. Their research in the field of behavior recognition opened the door of keystroke dynamics (Bryan 1897). In 1977, Forsen introduced a theory about whether users can be recognized through the unique typing pattern of entering their name(Forsen 1977). Since then a variety of algorithms for keystroke dynamics sprung up everywhere. In 1980, K Gaines discussed the possibility of using keystroke time to identify the user at the first time. They thought that using keystroke authentication method could effectively prevent imposters who obtained password illegally, thereby enhance the security of authentication (Gaines 1980). The characteristics of one's keystroke were considered to be unique, and hard to imitate (Bolle 2004; Jin 2004; Nguyen 2005). Since 4 PIN code for authentication on mobile devices is vulnerable to peeping or guessing attack, Seong-seob Hwang et al. conducted a study based on keystroke dynamics to improve the accuracy of user authentication on mobile devices, and claimed the use of artificial rhythm will improve user recognition accuracy (Hwang 2009). In their research, some special preprocessing for keystrokes rhythm of short passwords can improve the quality of data, such as different time pauses. We select keystroke interval time and duration, and adopt different lengths of passwords for experiments.

As the development of keystroke dynamics, it has been introduced to smartphones. Smartphones have more sensors that can easily record the user's behavior patterns. Mantyjarvi used the accelerometer

and orientation sensors to monitor the user's behavior pattern for the first time. They could reach an accurate rate of 60-85% (Mantyjarvi 2005). Conti also proposed a model based on the pattern of swing arm to authenticate users (Conti 2011). With acceleration sensor and direction sensor, devices recorded the user's pattern data when they waved the arm in a call. Their experiments improved the accuracy of user authentication greatly with false acceptance rate of 4.44%, false rejection rate of 9.33%. Lingjun Li et al. established a user authentication mechanism by monitoring finger real-time movement on the smartphone screen without affecting user(Lingjun 2013). Their system monitors five types of gestures: sliding up, sliding down, sliding left, sliding right, and tap. In other words, they extracted biometrical features in vertical and horizontal directions, and authenticated user by moving angle and interval time of gesture.

Pattern password, as a widely used authentication method, is applicable on almost all smartphones. However, there are only a few researches on pattern password. This paper throws a sight to pattern in pattern password to provide an extra layer of authentication, in case a pattern password was leaked. In this paper, we exploit the user's touching behaviors including position, pressure, contact area, time and other information when user draws a pattern password. We build three classifiers to re-authentication user besides pattern password itself. This authentication is transparent to users, and can be combined seamlessly with the original pattern password mechanism.

3 DATA SET

3.1 Data Collection

We used a smartphone LG-F320L equipped with Android 4.2 OS to collect the data from a total of 10 users. At first user should input his name to log in as shown in Figure 1, then there is a pattern password interface as in Figure 2. Each user was required to enter the same unlock pattern in Figure 3 for 51 times. That is to say, we collected up to 510 pieces of gesture data for the same pattern, and saved them in our database.

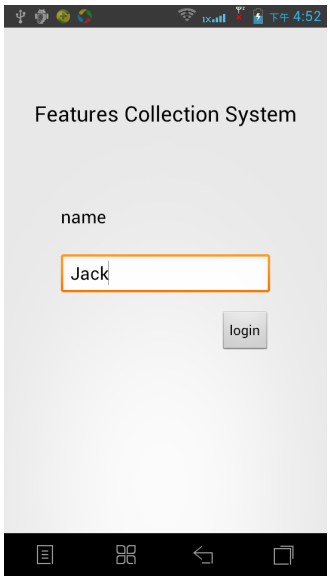

Figure 1. Landing Interface

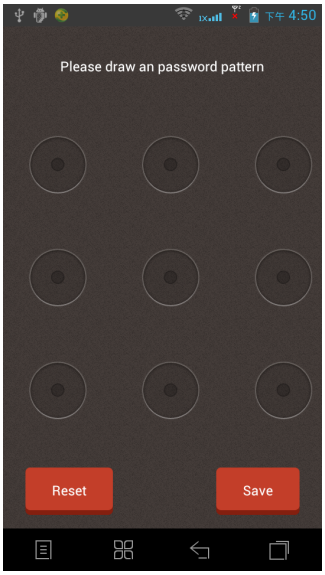

Figure 2. Initial Interface

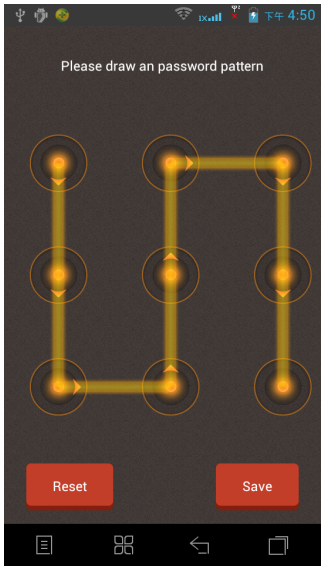

Figure 3. Data Collection Interface

3.2 Data Description

We adopted edge-sensing mechanism in our experiment, and nine spots in pattern password interface were expanded to nine circles as shown in Figure 4. These nine circles were numbered from left to right, top to bottom: 0, 1, 2, 3, 4, 5, 6, 7 and 8. Once user's finger touched edge of any circle, that circle was triggered and touch pattern at that time was recorded. We collected number of point, action type, touch point coordinates (x, y), touching area, start time, entry time, elapsed time, pressure, IMEI (smartphone unique identification) and smartphone model for future analysis.

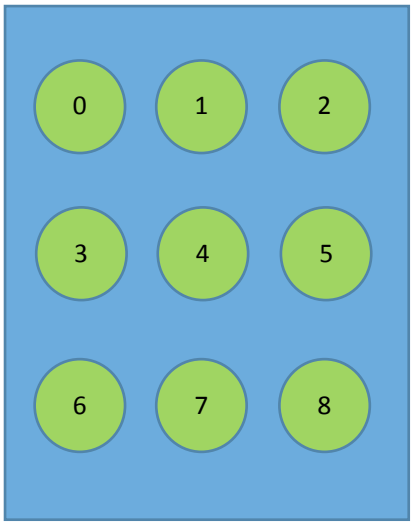

Figure 4. The Index of Point

Here is the explanation of each parameter in detail.

Number of point: the index number of the pattern password code, it could be 0, 1,2,3,4,5,6,7 and 8.

Action type: touch event type obtained by Android API MotionEvent.getAction(). It may have a value of 0, 1 or 2, which means press event, release event, and repeat touching event respectively.

Touch point coordinates: Once user's finger entered he circular edge of each pattern password, we can obtain the x and y coordinates of the contact point by an Android API with unit pixel (PX).

Touching area: denoted as S, means contact area of the finger on smartphone screen, obtained through Android API MotionEvent.getSize (), with value 0 or 1.

Start time: denoted as T_s, that is the absolute time of touch event from the program running time, in milliseconds. It can be obtained through Android API MotionEvent.getDownTime ().

Entry time: denoted as T_e, represents the absolute time of user's finger entering a circle from the program running time in milliseconds. It can be obtained through Android API MotionEvent. getEventTime ().

Elapsed time: denoted as T, represents the time consuming from start of stroke to entering the current point. That is to say, it is equal to enter time minus the start time. Its unit is millisecond.

Pressure: denoted as P, touching pressure obtained through Android API MotionEvent.getPressure (). Its value is a real number from 0 to 1.

IMEI: International Mobile Equipment Identity, unique sign of smartphones, used to identify

cellphones.

We adopted SQLite in Android to store our collected data, each piece of gesture data in a text field. Data of nine recording points was separated with semicolon, while metadata was separated with comma. Then we created two tables, user table and touching pattern table as shown in table 1 and 2, to store user's information and user's touching pattern separately.

Name	Declared Type	Type	Size
id	integer	integer	0
name	varchar(20)	varchar	20

Table 1. Structure of user table

Name	Declared Type	Type	Size
id	integer	integer	0
name	varchar(20)	varchar	20
track	text	text	0

Table 2. Structure of gesture record table

There are 510 pieces of data of 10 users, 51 unlock pattern data per person. Data fields such as number of point and IMEI data has no direct relation to one's touch pattern, so we removed them in our experiment. Since user should complete the pattern in one stroke, value of touch type for 9 points should always be 0. Start time denotes absolute time of the whole touch event, and it also remains the same for 9 points. We remove them from our parameter list as redundant data. Then we studied five features left in detail. Figure 5 shows coordinates distribution of the first touching point of all 510 pieces of data. The scatterplot indicates that one's first touching point has subtle difference with others, so it could be used to help distinguish smartphone users.

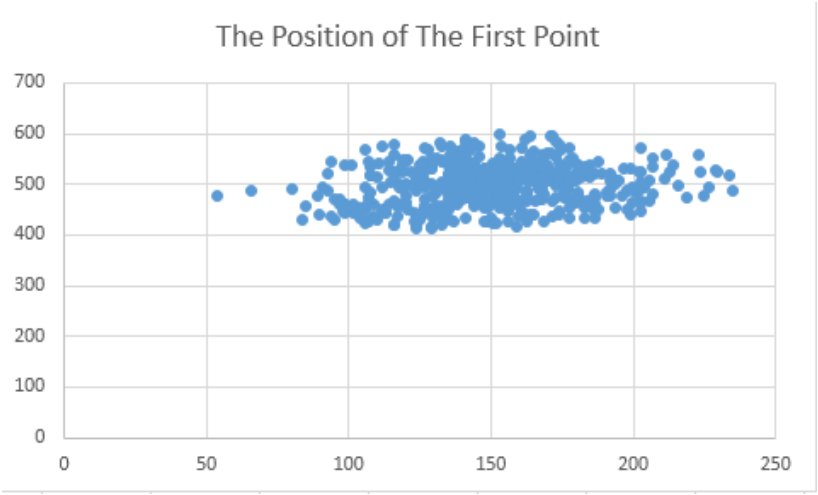

Figure 5. Coordinates of the first touching point

We computed the average touch area of the first touching point for ten users and compared them in Figure 6.This bar chart shows that touching area of different users has significant difference, and we should consider it in our experiment to distinguish users.

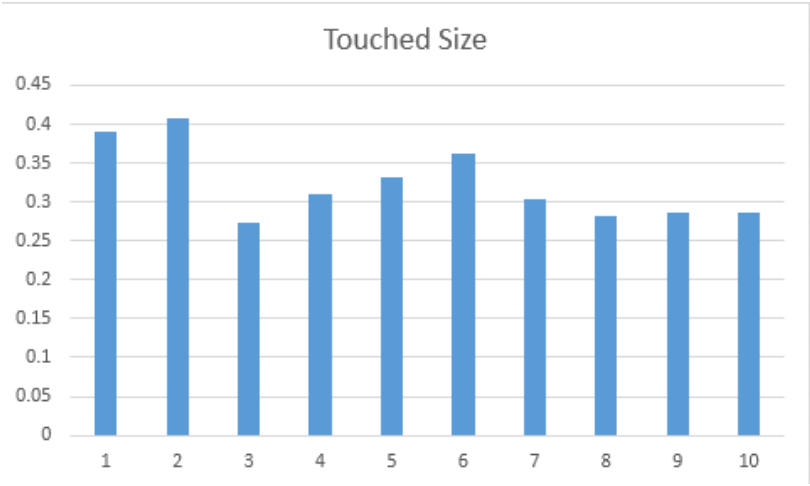

Figure 6. Average touching area of the first touching point

We could see from Figure 7, 8 and 9 that average elapsed time and average touching pressure of one user are different with others, so we use them in our experiment to distinguish users.

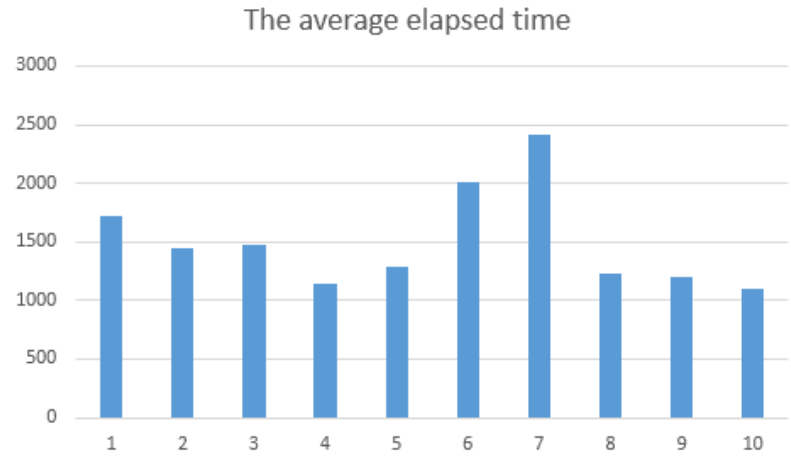

Figure 7. Average elapsed time of touch

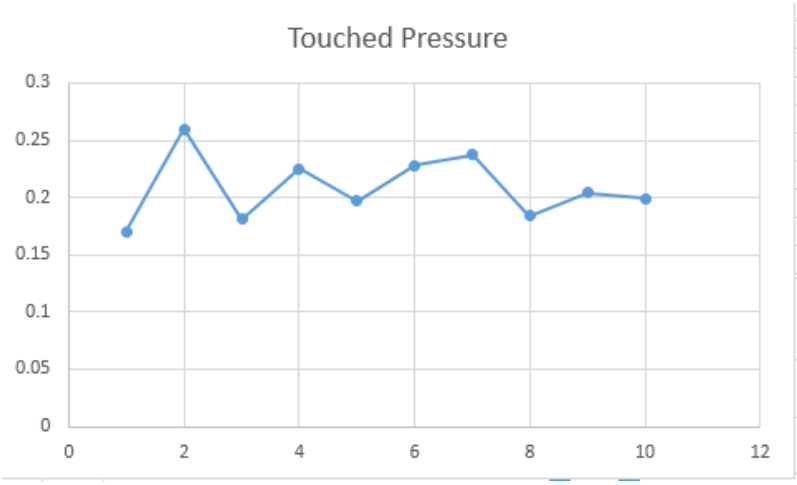

Figure 8. Average touching pressure of the last touching point

Based on the above analysis, we may use x coordinate, y coordinate, touching area, elapsed time, and touching pressure as input parameters for each pattern point. Each password consists of nine points, so

there are 45 features altogether. We could construct a vector of 45-dimension as our system input.

4 EMPIRICAL METHODS

The first step of our experiment is to extract touching features, and the second step is to propose a authentication scheme based on touching features by constructing a classifier. We used three classification methods including fuzzy logic, neural networks, and support vector machine to compare their experimental results.

Artificial neural networks (ANNs), mimic the structure and function of biological neural networks, interconnected together by a large number of neurons (Zeidenberg 1990). It can change the internal structure on the basis of the outside information. It is an adaptive system, with learning capacity. However, ANN requires a lot of parameters, such as network topology, the initial value of the weights and thresholds, etc. We can't observe the learning process, and the output is difficult to explain. This will affect the credibility and acceptability of the results. In addition, the learning time is too long, and ANN may even reach the purpose of the study. Support vector machine (SVM) was proposed by Corinna Cortes and Vapnik in 1995 (Cortes 1995). SVM can solve the problems of high-dimensional and nonlinear. However, it is sensitive to missing data, and there is no universal solution for nonlinear problem, so you must choose carefully the kernel function. Fuzzy logic is Boolean logic extensions which deal with part of the real concepts. As is grey between black and white, there are imprecise concepts in its linguistic form, such as "a little", "quite" or "very" (Klir 2003; George 2008). Fuzzy logic is good at processing the qualitative knowledge and experience. It distinguishes fuzzy set, and processes fuzzy relation by means of fuzzy membership function. Fuzzy logic simulates the human brain to conduct fuzzy comprehensive judgment, and it reasons to solve the classification problem of the fuzzy information, which is difficult to deal with common methods (Hwang 2009).

5 RESULTS AND DISCUSSION

5.1 Experimental Environment

In this paper, all data are collected on a LG brand smartphone equipped Android 4.2 operating system. Our computer for training and testing is an ASUS brand laptop, with Intel core i5 4 dual-core processor, 4g memory and 32-bit Windows 7 OS. Programs run in MATLAB R2010b and Eclipse, and our programming languages are Java and C.

5.2 Experimental Parameters

As described above, we gathered 51 unlock gesture for each user. Since all 10 users drew the same pattern, we used other users' data as negative samples to simulate the attacker. We selected 4 pieces of touching data of each other 9 users, and then we got 36 negative samples for each user. We use 31 positive samples out of 51 as training data, other 10 positive samples and 18 negative samples out of 36 as validation data, the other 10 positive samples and 18 negative samples as testing data. In order to eliminate the influence of data unit, all data was normalized.

We adopted FAR (False Accept Rate) and FRR (False Rejection Rate) to evaluate the experimental results. FAR denotes the probability of negative samples being considered as positive samples mistakenly, and FRR indicates the probability of positive samples being considered as negative samples mistakenly. We also used program-running time including training and testing time to

evaluate authentication efficiency.

5.3 Results Contrast

5.3.1 Artificial Neural Network

First we tried BP neural network in Matlab. Numbers of nodes in both the hidden layer and output layer are set to 45 as the dimension of feature vector, and times of maximum training is set to 2,000. Experimental result of BP neural network is summarized as following. The maximum value of FRR is 0.500000, and the minimum value is 0.000000, with mean value of 0.190000. The maximum of FAR is 0.555556, and the minimum value is 0.111111, with mean value of 0.277778. Running time of program is 552.4121 seconds.

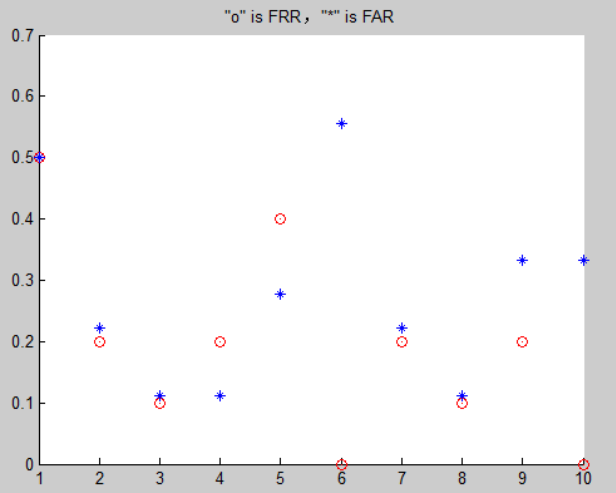

Figure 9. FAR and FRR using ANN

5.3.2 Support Vector Machine

We adopted LIBSVM developed by professor Chih-jen Lin as our support vector machine tool. We tried several loops to find the most proper parameters setting. Experimental result of SVM is summarized as following. The maximum value of FRR is 0.600000, and the minimum value is 0.000000, with mean value of 0.230000. The maximum of FAR is 0.222222, and the minimum value of 0.000000, with mean value of 0.055556. Running time of program is 99.296 seconds.

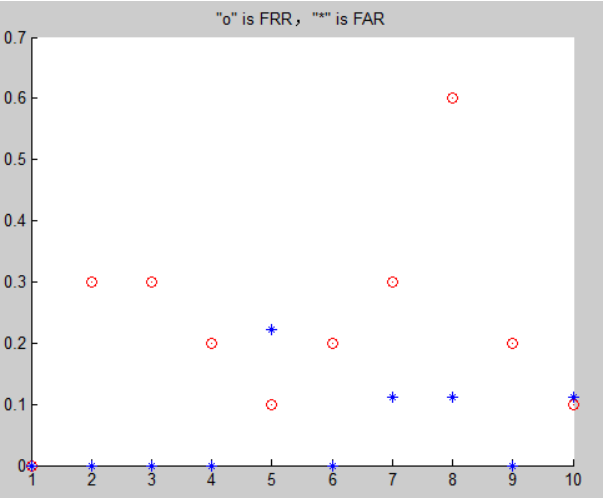

Figure 10. FAR and FRR using SVM

5.3.3 Fuzzy Logic

We calculated the membership value of each feature in training samples, and tried to find the optimal threshold by loops. The value of threshold is set from 0 to 45, adding 1 in each cycle. In order to eliminate contingency, we take 50 iterations in our experiment. The result of fuzzy logic is described as following. The maximum value of FRR is 0.080000, and the minimum value is 0.020000, with mean value of 0.049800. A maximum of FAR is 0.083333, the minimum value is 0.022222, with mean value of 0.047778. Running time of program is 22.9852 seconds.

We compared all three classifiers in Table 3. As we can see, result of fuzzy logic is better than that of artificial neural network and support vector machine in all parameters FRR, FAR and program running time. FAR of SVM is close to fuzzy logic, but running time and FRR are not so good. Neural network has the longest running time among three, FRR and FAR are also the worst. As to our experiment, Fuzzy logic is optimal in terms of efficiency and accuracy, which FRR and FAR were about 4.98% and 4.78%.

Classifiers	FRR（mean）	FAR（mean）	Runtime(seconds)
ANN	0.190000	0.277778	552.4121
SVM	0.230000	0.055556	99.296
Fuzzy Logic	0.049800	0. 047778	22.9852

Table 3. The comparative results of classification algorisms.

6 CONCLUSION AND FUTURE WORK

Smartphones bring great convenience to our daily life, but also bring risks of privacy losing. How to re-authentication users besides the original password system has been a hot topic. This paper proposed a novel authentication mechanism based on biometrics. With that authentication mechanism, mobile devices can tell whether the current user is the real owner by identifying user's touching pattern when inputting the pattern password, or just a penetrator who knows the pattern password by guessing, peeping or cheating. Our authentication mechanism can be seamlessly integrated with original authentication mechanism to enhance the security level without disturbing users. Experimental results show that fuzzy logic classification algorithm has the best performance on our experimental data, with mean FAR and FRR to be 4.98% and 4.78%.

However, our authentication mechanism is not applicable for all types of smartphones, because some of them do not support reading user's touching area and touching pressure. Secondly, we just use a little part of sensors to collect pattern data and it is not sufficient. For example acceleration sensor, gyroscope, direction sensor may be considered in our future work to improve accuracy rate of authentication.

7 Acknowledgement

This work was supported by National Natural Science Foundation of China [60903201, 91218301]; and the Fundamental Research Funds for the Central Universities [JBK120505, JBK140129].

References

Abuhaiba, I.S.I. (2007). Offline Signature Verification Using Graph Matching. Turkish Journal of Electrical Engineering & Computer Sciences, 15(1), 89-104.

Aviv, A. J., Gibson, K., Mossop, E. et al(2010). Smudge Attacks on Smartphone Touch Screens. In Proceedings of the 4th USENIX Workshop on Offensive Technologies(WOOT), 1-7, Washington DC.

Bolle, R. Connell, J. Pankanti, S. Ratha, N. Senior(2004). Guide to Biometrics. Springer-Verlag, Berlin Heidelberg, New York.

Bryan, W. L., Harter, N(1897). Studies in the Physiology and Psychology of the Telegraphic Language. Psychological Review, 4(1), 27.

Cai, L., Chen, H(2011). TouchLogger: Inferring Keystrokes on Touch Screen from Smartphone Motion. In Proceedings of the 6th USENIX Workshop on Hot Topics in Security.1-9, San Francisco.

Conti, M., Zachia-Zlatea, I., Crispo, B(2011). Mind How You Answer Me!: Transparently Authenticating the User of a Smartphone When Answering or Placing a Call. In Proceedings of the 6th ACM Symposium on Information, Computer and Communications Security. ACM,249-259,Hong Kong.

Cortes, C., Vapnik, V(1995). Support-vector Networks. Machine learning, 20(3), 273-297.

Durumeric, Z., Kasten, J., Adrian, D., et al(2014). The Matter of Heartbleed. In Proceedings of 2014 Conference on Internet Measurement Conference. ACM, 475-488,Vancouver.

Forsen, G. E., Nelson, M. R., Staron, Jr. R. J(1977). Personal Attributes Authentication Techniques. Pattern analysis and recognition. Crop Rome, NewYork.

Gaines, R. S., Lisowski, W., Press, S. J. et al(1980). Authentication by Keystroke Timing: Some Preliminary Results. Rand Corp. Santa Monica.

George, J. K., Bo, Y(2008). Fuzzy Sets and Fuzzy Logic, Theory and Applications. Prentice-Hall. New Jersey.

Hwang, S., Cho, S., Park, S(2009). Keystroke Dynamics-based Authentication for Mobile Devices[J]. Computers & Security, 28(1),85-93.

Jain, A. K., Griess F. D., Connell, S. D(2002). On-line signature verification. Pattern Recognition, 35(12), 2963-2972.

Jermyn, I., Mayer, A. J., Monrose, F. et al(1999). The Design and Analysis of Graphical Passwords. In Proceedings of Usenix Security Symposium.1-14. Washington DC.

Jin, L., Ke, X., Manuel, R. et al(2004). Keystroke Dynamics: A Software Based Biometric Solution. In Proceedings of the 13th USENIX Security Symposium.23-36. San Diego.

Klir, G. J., Folger, T. A(1988). Fuzzy Sets, Uncertainty, and Information. Prentice Hall, New Jersey.

Lingjun, L., Xinxin, Z. and Guoliang, X(2013). Unobservable Re-authentication for Smartphones. In Proceedings of the 20th Annual Network & Distributed System Security Symposium(NDSS). San Diego.

Mantyjarvi, J., Lindholm, M., Vildjiounaite, E., et al(2005). Identifying Users of Portable Devices from Gait Pattern with Accelerometers, In Proceedings of IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP), 973-976, Pennsylvania.

Miluzzo, E., Varshavsky, A., Balakrishnan, S. et al(2012). Tapprints: Your Finger Taps Have Fingerprints. In Proceedings of the 10th international conference on Mobile systems, applications and services(Mobisys). ACM, 323-336, Cumbria.

Nguyen, H. T., Walker, E. A(2005). A First Course in Fuzzy Logic. CRC press, Florida.

Owusu, E., Han, J., Das, S. et al(2012). ACCessory: Password Inference Using Accelerometers on Smartphones. In Proceedings of the 12th Workshop on Mobile Computing Systems & Applications. ACM, Phoenix.

Xu, Z., Bai, K., Zhu, S(2012). Taplogger: Inferring User Inputs on Smartphone Touchscreens Using On-board Motion Sensors. In Proceedings of the 15th ACM conference on Security and Privacy in Wireless and Mobile Networks. ACM, 113-124, Tucson.

Zeidenberg, M(1990). Neural Networks in Artificial Intelligence. Ellis Horwood Limited, Chichester.