International Journal of Applied Engineering Research ISSN 0973-4562 Volume 12, Number 19 (2017) pp. 8727-8731

Customized Data Filtering For Mobile Signature Verification

Seungsoo Nam[1], Hosung Park[2], Changho Seo[1] and Daeseon Choi[2]

[1]Department of Conversions Science, Kongju National University, Korea.
[2]Department of Medical Informatics Management, Kongju National University, Korea.

Orcid ID: 0000-0001-8849-345X

Abstract

Many studies for signature verification focus on the improvement of verification accuracy, the most significant issue. They makes subject models with sufficient sample signatures since the verification accuracy is generally proportional to the number of sample data. However, in the mobile environment, excessive requests of signatures could exhaust users to abandon the corresponding services. This paper therefore proposes customized signature filtering scheme that reduces the number of signature requests with preserving the verification accuracy. On the basis of the observation that large variations between sample signatures disturb fast and precise learning of a classifier, the proposed scheme filters out unnecessary samples. In other words, it compares sample signatures, eliminates out-of-range signatures, and uses only others as the sample data. The experiment results show that the proposed scheme reduces 51% signature requests without decrease of the verification accuracy.

Keywords: signature, AutoEncoder, Customized filtering, Mobile biometric

INTRODUCTION

According to the increase of mobile devices, many studies try to provide higher security using the biometric authentication not the mere password authentication [1][2][3]. Signature verification is one of the most focused authentication scheme in the mobile environment due to its convenience, familiarity, and versatility. It does not need to memorize identification numbers or passwords and to possess stamps, keys, or cards. It is also safe from loss, leakage, and theft like other biometric authentication. Moreover, users are less resistant to using signatures than other biometrics such as fingerprints and iris. Signature verification is already utilized in a wide range of applications such as credit card, bank account, immigration control, electronic document, and electronic commerce in both online and offline. Generalization of mobile signature verification may lead to various services by working along with the existing applications.

The most significant issue for signature verification is to provide high accuracy since the miss-classification could cause critical damages like personal information leakage and financial loss. Therefore, many studies focus on the improvement of verification accuracy. Studies for dynamic signature verification [4][5] use, as features for verification, not only the signature shape but also behavior characteristics of signing such as accelerometer, pressure, velocity, pressure derivative and intersecting points. Other studies try to increase verification accuracy by improving classifiers. Fischer et al. [6] refines DTW (Dynamic Time Warping) to enhance the throughput and the accuracy. Antal et al. [7] separately reconstitute the neural network for each user.

Existing studies for signature verification assume sufficient training data, i.e. many signatures of users. However, excessive requests of signatures could exhaust users to abandon the corresponding services. In deep learning, the performance (verification accuracy) is generally proportional to the number of sample data (user signatures). In other words, this issue has tradeoff between decreasing the number of signature requests and increasing the verification accuracy. Mobile signature verification therefore encounters the following question: “Can the proper verification accuracy be maintained with relatively few signature requests?” That question is also the goal of this paper.

This paper proposes customized signature filtering scheme that reduces the number of signature requests with preserving the verification accuracy. Signatures have some differences with each other even though they are made by a user. The large variations between sample signatures could disturb fast and precise learning of a classifier. It consequentially makes it difficult to distinguish original signatures from others, especially skilled forgeries. Therefore, the proposed scheme compares sample signatures, eliminates out-of-range signatures, and uses only others as input data of the classifier. A filtering criterion is calculated based on the basis of each sample’s standard deviation with a median and a weighted value. Each user has own filtering criterion since the sizes of signatures and behavior characteristics of signing are depend on the user. Experiment al results show that customized signature filtering could reduce signature requests with maintaining the verification accuracy in spite of their tradeoff.

This paper is organized as follows. The proposed scheme is presented in section 2 and evaluated based on experimental results in section 3. The paper ends with concluding remarks in section 4.

THE PROPOSED SCHEME

The mobile signature verification consists of two main phases: signature training and signature verification. Figure. 1 shows the procedure of the proposed mobile signature verification scheme.

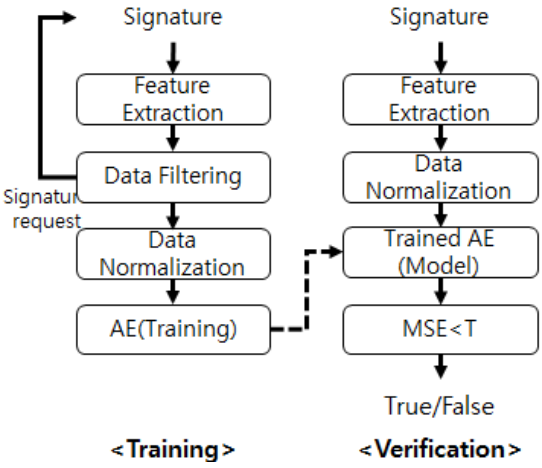

Figure 1: Overall structure block diagram

In the training phase, a user is initially required to sign several times on a smartphone screen. After feature extraction from the signatures, the proposed scheme executes data filtering that compares sample signatures and eliminates out-of-range signatures with the certain filtering rule presented in section 2-2. At least three sample signatures are needed for comparing each other. If a signature is eliminated, the proposed scheme requires a new signature. This process is repeatable. If all sample signatures satisfy the filtering rule, the process is transited to the next step. After data filtering, the features are normalized and then used as the input to an AE (Auto-Encoder). AE is trained with the sample data and the trained AE becomes the user's signature model to be used as a classifier in the verification phase.

In the verification phase, a new input signature to be verified, the test signature is entered into the trained AE after feature extraction and data normalization. The test signature is compared with the output, test signature, by MSE (Mean Square Error). If the difference is less than a pre-defined threshold T, the signature is accepted as a valid one. The threshold is affected by the characteristics of the user's signature and the result of data filtering. In the proposed scheme, data filtering in the training phase could lead to the stricter threshold which means the higher verification accuracy.

Feature Extraction

When a user signs on a smartphone screen, the coordinates of points (Px , Py), the sensed value of the accelerometer (Ax, Ay, Az), and distance from (Px_{i-1}, Py_{i-1}) to (Px_i Py_i) dis are sampled every 32ms. For comparing with other signature, the coordinates are aligned to set the starting point to (0, 0). The features of a signature s are defined as follows:

$$s = (Ax_{i,} - Ax_0, Ay_{i,} - Ay_0, Az_{i,} - Az_0, Px_{i,} - Px_0, Py_{i,} - Py_0 dis_{i,} - dis_0) \quad i = 0,,\ldots,n \tag{1}$$

Data Filtering

To compare signatures, the proposed scheme utilizes the median of signatures' vector values. The vector value of a signature V is calculated by Eq. 2.

$$V = \sqrt{\sum_{i=0}^{n} Ax_i^2 + Ay_i^2 + Azi^2 + x_i^2 + y_i^2 + disi^2} \tag{2}$$

With three sample signatures, the three vector values are sorted in ascending order. The median, i.e. the second vector value, not a mean value becomes a reference value $\overline{V}$. If the mean value is used as the reference value, data filtering could be critically affected by an extreme vector. For example, when a vector has wide gap from other vectors, the mean value is unfair and ineffective for data filtering. Note that the purpose of data filtering is to eliminate dissimilar signatures disturbing AE's training.

With the reference value, the proposed scheme utilizes standard deviation for the filtering rule since the sample signatures follow a normal distribution. The standard deviation σ is calculated as follows:

$$\sigma = \sqrt{\frac{(V_1 - \overline{V} + V_2 - \overline{V} + V_3 - \overline{V})}{n-1}} \tag{3}$$

where n is the number of sample data, 3. The filtering rule is defined as follows:

$$if \begin{cases} True & \overline{V} - \sigma k < V_n \le \overline{V} + \sigma k \\ False & else \end{cases} \tag{4}$$

where k is a constant for adjustment of filtering range. The filtering rule means that a sample is accepted if its vector value V_n is not out of bound σk from the reference value $\overline{V}$. A signature which does not satisfy Eq. 4 is eliminated and the proposed scheme requires a new sample signature. If all sample signatures satisfy the filtering rule, the process is transited to the next step. For our goals, k is important constant that affects the overall performance. Too low k leads to excessive signature requests and too high k is worthless for the accuracy in the verification phase. Initial k is heuristically configured to 0.7. On the normal distribution, $1*\sigma$ typically covers 68% samples [8] and $0.7*\sigma$ permits sample signatures included within 26%. Initial k value is adjusted on the basis of

the σ size as defined in Eq. 5. The value of k is proportional to the size of σ.

$$\begin{aligned} &\sigma \le 60, && k = 0.7 \\ &60 + 20(i-1) < \sigma \le 60 + 20i && i = 1,\ldots,n \\ &k = 0.7 + 0.2i \end{aligned} \quad (5)$$

Data Normalization

In signature verification, data normalization is essential since the sizes of signatures and value range of each feature are different. All features are normalized into values range between 0 and 1 [9]. Every signature length is also normalized to a mean length of the reference signature.

AE (Auto-Encoder) for Training and Verification

AE is a type of FNN (Feedforward Neural Network) that learns inherent characteristics of sample data and has the same dimensions for input and output [10]. AE is trained by putting the same data (sample signatures) as input and output but, in the test phase, trained AE generates an output corresponding to an input (test signature). AE could generate highly similar output for trained data pattern whereas it does not for others. AE therefore is useful for authentication.

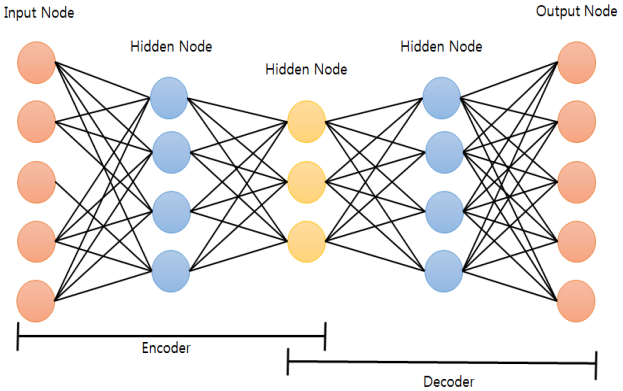

Figure 2: Structure of Auto-Encoder

The proposed scheme uses a 5-dimensional AE described in Figure. 2. AE itself consists of two blocks: encoder and decoder. The outputs of encoder and decoder are denoted as z and s` respectively and defined as Eq. 6 and 7. The encoder signature data $s$ as the input and produces $z$ that is implemented as a hidden node in the neural network. The decoder accepts the $z$ and produces regenerated signature $s$`. σ_1 and σ_2 are activation functions.

$$z = \sigma_1(Ws + b) \quad (6)$$

$$s' = \sigma_2(W'z + b') \quad (7)$$

$$L(s, s') = \left\| x - \sigma_2(W'(Ws + b)) + b \right\| \quad (8)$$

The loss function L defined in Eq. 8 is the objective for training a neural network [11]. The training of AE is a procedure involving the modification of weights W and W' for minimizing L with s as both input and output.

In the test phase, i.e. the verification phase, a new input signature s to be verified is entered into trained AE and compared with the output s` of the AE. The proposed scheme applies MSE (Mean Square Error) for the similarity comparison [12]. MSE means the difference between $s$ and $s$` and is defined as Eq. 9. With MSE, the verification rule is defined as Eq. 10, where k is the constant used in data filtering in Eq. 4 and 5.

$$Diff = MSE(s, s') \quad (9)$$

$$Diff < t \times k, \; t : threshold \quad (10)$$

Threshold t is adjusted with k used in data filtering process. If the test signature satisfies the verification rule, i.e. MSE is less than the adjusted threshold, the signature is accepted as a valid one. In general cases, adjusted threshold are less than the predefined threshold t and it could improve the verification accuracy. Note that decreasing threshold without data filtering may degrade the accuracy due to the increase of false rejection rate.

EXPERIMENTS

Experimental Settings

We conducted a test set by ourselves since no public test set of dynamic signatures has been made on smartphones. On a smartphone Galaxy S3, we gather 20 subjects' 20 signatures (400 original signatures). The subjects' signatures are shown to 5 forgers and they then make 4 imitations per 7 original signatures (140 forged signatures). In the verification phase, we conduct two types of verification tests and table 1 shows the test data configuration for a subject.

Table 1: Test data configuration for each subject

Verification test	Subject	Others
Original / Others	3EA (Original)	19 users X 20EA (Others)
Skilled forgery	3EA (Original)	20EA (Skilled forgery)

Experiments are implemented on the Android Studio 2.1.3. The AE for subject modeling consists of the input layer, three hidden layers (50, 20, and 50 nodes), and the output layer. The number of nodes of input and output layers depends on the subject.

EER (Equal Error Rate) is used as an evaluation metric for signature verification [13]. FAR (False Acceptance Rate) means the rate of other signatures misclassified as subject's ones. FRR (False Rejection Rate) means the rate of Subject's signatures misclassified as other's one. EER is the rate when FAR and FRR are same and we measure the lowest EER.

Experimental results

Table 2 shows the signature verification accuracy (by EER) and the number of signature requests for 5 types of experiment cases. In the experiments, the proposed scheme requests average 4.9 signatures for training AE with 3 signature samples.

Table 2: Comparison of signature verification results

Experiment case	# of signature requests	# of training signatures	EER (original / others)	EER (skilled forgery)
AE (10 samples) [8]	10	10	2.8%	18.4%
AE (3 samples)	3	3	7.2	25.5%
Filtering	4.9	3	6.3%	14%
Adjusted threshold	3	3	5.7%	15.2%
Filtering + Adjusted threshold	4.9	3	2.99%	7.5%

In other words, the data filtering process eliminates and redemands average 1.9 signatures. The EERs of three experiment cases, 1) AE trained with 3 samples without filtering, 2) AE trained with 3 filtered samples without adjusted threshold, and 3) AE trained with 3 samples with adjusted threshold without filtering, are inferior to well-trained AE with 10 samples. In the proposed scheme (filtering + adjusted threshold), the EER for original / others test is a little higher (0.19%) than the AE with 10 samples but the EER for skilled forgery test is much lower (10.9%), i.e. better. Consequently, the proposed scheme reduces 5.1 signature requests (51%) without decrease of the verification accuracy. Figures 3 and 4 show the results in details by ROC(Receiver Operating characteristic) curves.

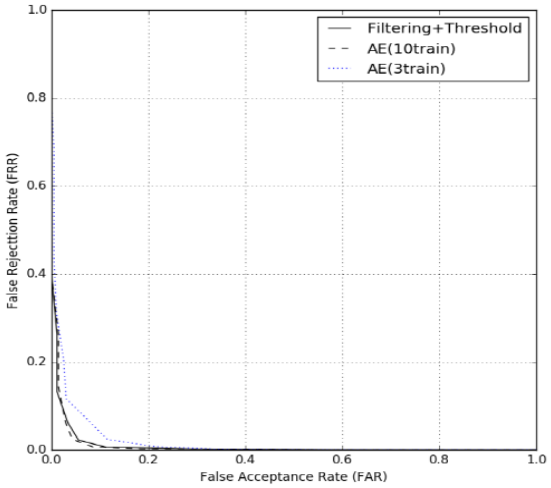

Figure 3: ROC curve (original / others)

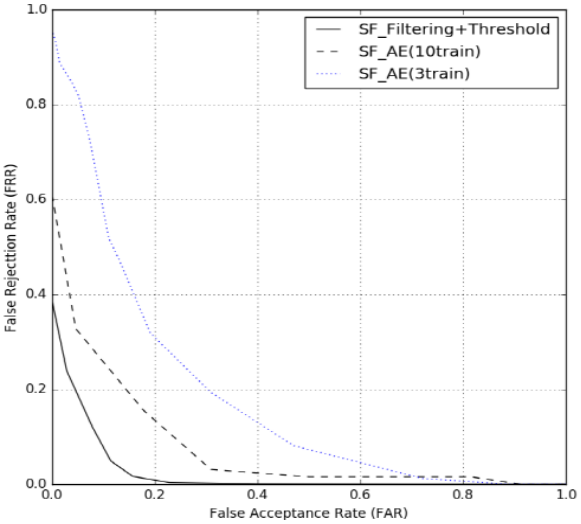

Figure 4: ROC curve (skilled forgery)

CONCLUSION

On the basis of our observation that the large variations between sample signatures could disturb fast and precise learning of a classifier, we propose customized signature filtering scheme for mobile signature verification. The proposed scheme reduces the number of signature requests with maintaining the verification accuracy. We expect that the results could contribute to mobile users in terms of security as well as convenience.

ACKNOWLEDGMENTS

This work was supported by Institute for Information & communications Technology Promotion (IITP) grant funded by the Korea government (MSIT) (No.B0717-16-0084, Development of Information security core technology)

REFERENCES

[1] H. Crawford, K. Renaud, and T. Storer, "A framework for continuous, transparent mobile device authentication," Computers and Security, vol. 39, pp. 127–136, 2013.

[2] W. Meng, D. S. Wong, S. Furnell, and J. Zhou, "Surveying the development of biometric user authentication on mobile phones," IEEE Communications Surveys and Tutorials, vol. 17, no. 3, pp. 1268–1293, 2015.

[3] J. Bonneau, C. Herley, P. C. Van Oorschot, and F. Stajano, "The quest to replace passwords: A framework for comparative evaluation of web authentication schemes," Proceedings - IEEE Symposium on Security and Privacy, pp. 553–567, may 2012.

[4] T. Van Nguyen, N. Sae-Bae, and N. Memon, "Finger-drawn pin authentication on touch devices," in Proc. IEEE Int. Conf. Image Process. (ICIP), pp. 5002–5006.

[5] Yang, Junshuang, Yanyan Li, and Mengjun Xie. "MotionAuth: Motion-based authentication for wrist worn smart devices." Pervasive Computing and Communication Workshops (PerCom Workshops), 2015.

[6] Fischer, Andreas, et al. "Robust score normalization for DTW-based on-line signature verification." 13th International Conference on Document Analysis and Recognition (ICDAR), 2015.

[7] Antal, Margit, and László Zsolt Szabó. "Some remarks on a set of information theory features used for on-line signature verification." 5th International Symposium Digital Forensic and Security (ISDFS), 2017.

[8] B. Ferris, D. Hahnel, and D. Fox, "Gaussian processes for signal strength-based location estimation," in Proc. Robotics Sci. Syst., 2006.

[9] S. S. Nam, "Mobile Finger Signature Verification Robust to Skilled Forgery", Journal of The Korea Institute of Information Security & Cryptology, Vol. 26, No 5, Oct. 2016 (in Korean).

[10] L. Badino, C. Canevari, L. Fadiga, and G. Metta, "An autoencoder based approach to unsupervised learning of subword units," in Proc. ICASSP, 2014.

[11] https://en.wikipedia.org/wiki/Autoencoder

[12] W. Dennis, W. Mendenhall, and R. Scheaffer. "Mathematical statistics with applications," Nelson Education, 2007.

[13] https://en.wikipedia.org/wiki/Biometrics