Extended Sammon Projection and Wavelet Kernel Extreme Learning Machine for Gait-based Legitimate User Identification

Muhammad Ahmad*  
Innopolis University, Russia  
University of Messina, Italy  
mahmado0@gmail.com

Adil Mehmood Khan  
Innopolis University,  
Innopolis, Russia  
a.khan@innopolis.ru

Manuel Mazzara  
Innopolis University,  
Innopolis, Russia  
m.mazzara@innopolis.ru

Salvatore Distefano  
University of Messina, Italy  
sdistefano@unime.it

Amjad Ali  
MNC Lab, School of Electrical Engineering, Korea University, Seoul, South Korea  
amjad.khu@gmail.com

Ali Tufail  
Faculty of Computer and Information Systems, Islamic University of Medina, Saudi Arabia  
ali.tufail@iu.edu.sa

ABSTRACT

Smartphones have pervasively integrated into our home and work environments managing confidential information but their owners still rely on as explicit as inefficient and insecure identification processes. Therefore, if a device is stolen, a thief can have access to the owner’s personal information and services through the stored password/s. To avoid such situations, this work demonstrates the possibilities of legitimate user identification in a semi-controlled environment through the built-in smartphone motion dynamics captured by two different sensors. This is a two step process: sub-activity recognition followed by user/impostor identification. Prior to the identification, Extended Sammon Projection (ESP) method is used to reduce the redundancy among the features. To validate the proposed system, we first collected data from four users walking with their device freely placed in one of their pants pockets. Through extensive experimentation, we demonstrated that time and frequency domain features, optimized by ESP to train the wavelet kernel based extreme learning machine classifier, implement an effective system to identify the legitimate user or an impostor with 97% accuracy.

KEYWORDS

Smartphone; Sensor; Feature Extraction; Feature Selection; Legitimate User; Imposture.

ACM Reference Format:

Muhammad Ahmad*, Adil Mehmood Khan, Manuel Mazzara, Salvatore Distefano, Amjad Ali, and Ali Tufail. 2019. Extended Sammon Projection and Wavelet Kernel Extreme Learning Machine for Gait-based Legitimate User Identification. In Proceedings of April 8–12, 2019, Limassol, Cyprus (SAC’19). ACM, New York, NY, USA, Article 4, 4 pages. https://doi.org/https://doi.org/10.1145/3297280.3297579

1 INTRODUCTION

In the recent past, devices such as desktop computers and laptops were our best means for staying connected to the Internet community and to have access to online services. However, with enhanced capabilities, low cost, and user-friendly interfaces, smartphones (SP) have become people’s first choice to stay connected to the Internet. This is equally credited to the publicly available Internet which facilitates users to access their device contents regardless of their location. According to a recent study [1], over 80% of cell phone users globally use SPs. Furthermore, 2.56 billion people are estimated to own SPs by the end of 2018.

With this rapid growth trend in the usage of SPs, device control and data security have become extremely important. Not only personal and professional contact information is stored on SPs but users also store their sensitive and critical information on them [20]. If a SP is stolen, the stored information can be used to create problems not only to the owner but also to the owner’s contacts. To secure this access, it is important to develop fast and accurate methods for legitimate user identification and block-out impostors. Ideally, these methods should detect an impostor since when a device is stolen.

Current identification methods such as secret PIN number (SPN) or lock codes [19] are not only risky but also difficult to use. Another factor that restricts the security of SPNs-based mechanisms is the SPs are mostly used in public places with many other people around. This increases the chances for the codes to be found-out by potential attackers. To overcome these issues, fast and secure methods are required to intelligently verify the legitimate user and to block-out impostors.

In the literature several solutions were proposed addressing implicit user identification without involving the user, such as keystroke-based [20], touch screen [9], application set fingerprints [2], hybrid methods [5, 18], and gait-based identification [3]. However, these solutions only discuss either software or hardware aspects of user identification.

The first thing a thief might do after stealing a SP is to walk away from the owner. Considering this fact, the gait-based identification appears to be the most efficient, as it can instantly detect the impostor from the walking patterns and thus can trigger an alarm to inform the owner and/or relatives and friends about the theft.
This study focuses on the idea of identifying a SP user by employing different walking patterns, hereby referred to as sub-activities. Furthermore, it is assumed that the phone is freely placed without any particular orientation inside any of the user’s pants pockets.

In this regard, the aim of our present work is to investigate some research questions relevant to building a walking-based legitimate user identification system starting from Extreme Learning Machine: 1) Does the Extended Sammon Projection (ESP), a non-linear unsupervised feature selection method, improve the identification accuracy more than the other existing and well-studied unsupervised feature selection methods such as Principal Component Analysis (PCA)? 2) Is kernel-based Extreme Learning Machine (KELM) an effective classifier for non-linear signal-based user identification? 3) Does data variation affect the identification performance? Our research contributions are towards answering these questions using ESP and KELM methods.

2 SYSTEM DESIGN

Nowadays, SPs are equipped with a variety of motion sensors and two of these are the accelerometer (ACC) and the linear acceleration (LACC) sensors [4]. We have collected raw signals from these two sensors, as the user performed daily sub-activities i.e., walking while the SP was placed in one of the subject’s pants pockets, e.g. back right (BRP) and left pocket (BLP), front right (FRP) and left pocket (FLP). This dataset was gathered from four users each one performing individual sub-activities each day. Each sub-activity was performed at least twice a day, for an entire month. The dataset was gathered from all users at a constant rate of 50Hz. The dataset was pre-processed by applying noise filters, using 2.56 sec windows.

The ACC and LACC sensors generates time series signals which are highly fluctuating and oscillatory in nature [8, 16], thus making the user identification more difficult. Therefore, it is compulsory to gather the nontrivial signals from the raw data through the feature extraction process. Given a sampling rate of 50Hz, we chose a window size of 50 samples for both sensors. The selected window size provides enough data to be able to extract the quality features while ensuring a fast response [17].

We found that the time domain features, including the coefficients from the time series model, provided the same accuracy as the frequency domain features [3, 12]. Therefore, in this work, we have extracted both the time and frequency domain features as similar to our previous works [2, 3]. In total 72 features were extracted from each window. Prior to the feature extraction, a moving average filter of order three was employed for noise reduction purposes.

It is a known fact that the output of SP sensor depends on the position of the SP while walking. This could result in high within class variance [13]. Therefore, it is desirable to improve both the discriminatory power and achieve dimensionality reduction, by employing an optimum method. The advantages of the feature selection process are to avoid the curse of dimensionality [15], as well as to reduce the abundant, irrelevant, misleading and noisy features, but above all, to be able to reduce the system’s running cost pertaining to real-time applications [6]. In addition to the above, effective feature selection can increase the accuracy of the resulting model.

In this regard, the most commonly used method is the principal component analysis (PCA). However, we have used extended Sammon projection (ESP) for the first time for SP-based user identification due to its flexibility. The ESP is one of the most successful non-linear metric multidimensional scaling methods, indeed, and it projects the high dimensional space into a lower dimensional space while preserving the structure of inter-point distances. Let us assume $d_{i,j}$ be the Euclidean distance between two adjacent samples $x_i$ and $x_j$, $i \neq j$ in the original space and $d_{i,j}'$ be the distance between two samples $x_i'$ and $x_j'$ in the mapped space, the Sammon stress measure $E$ can be defined as:

$$E = \frac{1}{\sum_{i=1}^{n-1} \sum_{j=i+1}^{n} d_{i,j}} \sum_{i=1}^{n} \sum_{j=i+1}^{n} \frac{d_{i,j}' - d_{i,j}}{d_{i,j}}. \quad (1)$$

The error measurement is minimized by using the steepest and gradient decent methods, respectively, as $x_{ik}(t + 1) = x_{ik}(t) - \alpha \frac{\partial E(t)}{\partial x_{ik}(t)}$ and $x_{ik}(t + 1) = x_{ik}(t) - \alpha \frac{\partial^2 E(t)}{\partial x_{ik}(t)^2}$. In both cases, $x_{ik}$ is the $k$th coordinate of the position of $x_i$ in the low dimensional space. Since the steepest descent method has issues at the inflection points, where the second-order derivative appears to be quite small [14], we set $\alpha$ in the $[0.3, 0.4]$ range as the optimal value using a grid operation between $[0, 1]$, but there is no reason to expect this range to be optimal for all problems and datasets.

Neural networks (NN) have quite diverse applications and among their different types, extreme learning machines (ELM) have shown better generalization capabilities and fast learning speeds [11]. An ELM is a single hidden layer feed-forward NN which randomly determines the initial parameters of weights input and hidden biases with simple activation functions [7]. Moreover, ELMs with a tuneable activation function were proposed to handle the data dependency on hidden neurons. However, the selection of suitable combinations for activation functions is still an open research question. Kernelized ELMs are known to improve the generalization capabilities, when the feature mapping function of hidden neurons is unknown. However, the parameters for the kernel function must be selected carefully. In our work, the said parameters are optimized through a swarm optimization based method. Further information about ELM can be found in [10].

3 EXPERIMENTAL RESULTS

To validate our proposed user identification system, we conducted different experiments. In all experiments, the values used for the different parameters for ESP, PCA, and KELM methods were optimized using cross-validation. For the comparison of feature selection process, some numbers of PCs were selected as the number of features selected by the ESP. Prior to the classification, all obtained features were normalized between $[0, 1]$. Our first experiment details the process for analyzing the effect of using different number of samples per window (i.e. $\{25, 50, 75, 100, 125, 150, 175, \text{and} 200\}$), on the performance user identification system starting from Extreme Learning Machine: 1) Does the Extended Sammon Projection (ESP), a non-linear unsupervised feature selection method, improve the identification accuracy more than the other existing and well-studied unsupervised feature selection methods such as Principal Component Analysis (PCA)? 2) Is kernel-based Extreme Learning Machine (KELM) an effective classifier for non-linear signal-based user identification? 3) Does data variation affect the identification performance? We found that the time domain features, including the coefficients from the time series model, provided the same accuracy as the frequency domain features [3, 12]. Therefore, in this work, we have extracted both the time and frequency domain features as similar to our previous works [2, 3]. In total 72 features were extracted from each window. Prior to the feature extraction, a moving average filter of order three was employed for noise reduction purposes.

It is a known fact that the output of SP sensor depends on the position of the SP while walking. This could result in high within class variance [13]. Therefore, it is desirable to improve both the discriminatory power and achieve dimensionality reduction, by employing an optimum method. The advantages of the feature selection process are to avoid the curse of dimensionality [15], as well as to reduce the abundant, irrelevant, misleading and noisy features, but above all, to be able to reduce the system’s running cost pertaining to real-time applications [6]. In addition to the above, effective feature selection can increase the accuracy of the resulting model.
Table 1: Average Accuracy, Confidence Intervals and Time taken for Legitimate User Identification for 50 Sample Per Window With Different Feature Selection Methods and Different Number of Features on Both Sensors Data

| Features | Metric       | Back Left Pocket | Back Right Pocket | Front Left Pocket | Back Right Pocket |
|----------|--------------|------------------|-------------------|-------------------|-------------------|
|          | ACC          | LACC             | ACC               | LACC              | ACC               | LACC             |
|          | PCA          | ESP              | PCA               | ESP               | PCA               | ESP              |
| 5        | Accuracy     | 55.6±3.1         | 50.3±2.9          | 50.4±2.8          | 55.9±3.1          | 50.2±2.9         | 58.4±3.9         | 54.4±4.0         | 61.8±2.9         | 54.3±3.8         | 52.4±4.5         |
|          | Time         | 0.140±0.138      | 0.209±0.243       | 0.265±0.261       | 0.533±0.527       | 0.175±0.173      | 0.268±0.270      | 0.164±0.133      | 0.201±0.202      |                  |                  |
| 10       | Accuracy     | 74.6±3.3         | 52.9±1.1          | 74.6±3.2          | 52.5±2.1          | 74.6±3.2         | 52.5±3.2         | 74.6±3.2         | 52.5±3.2         | 74.6±3.2         | 52.5±3.2         |
|          | Time         | 0.140±0.142      | 0.209±0.214       | 0.255±0.254       | 0.535±0.523       | 0.177±0.178      | 0.270±0.271      | 0.131±0.131      | 0.194±0.188      |                  |                  |
| 15       | Accuracy     | 72.6±2.8         | 98.1±1.1          | 72.6±2.8          | 98.1±1.1          | 72.6±2.8         | 98.1±1.1          | 72.6±2.8         | 98.1±1.1          | 72.6±2.8         | 98.1±1.1          |
|          | Time         | 0.142±0.142      | 0.211±0.238       | 0.254±0.254       | 0.536±0.517       | 0.177±0.178      | 0.270±0.271      | 0.135±0.135      | 0.197±0.203      |                  |                  |
| 20       | Accuracy     | 73.6±1.1         | 97.1±1.4          | 73.6±2.0          | 97.1±1.4          | 73.6±1.1         | 97.1±1.4          | 73.6±1.1         | 97.1±1.4          | 73.6±1.1         | 97.1±1.4          |
|          | Time         | 0.142±0.142      | 0.216±0.219       | 0.259±0.256       | 0.538±0.529       | 0.178±0.189      | 0.269±0.281      | 0.132±0.132      | 0.194±0.212      |                  |                  |
| 25       | Accuracy     | 78.6±3.8         | 99.0±0.5          | 78.6±2.7          | 99.0±0.5          | 78.6±3.8         | 99.0±0.5          | 78.6±3.8         | 99.0±0.5          | 78.6±3.8         | 99.0±0.5          |
|          | Time         | 0.148±0.140      | 0.221±0.219       | 0.258±0.257       | 0.532±0.528       | 0.180±0.179      | 0.254±0.270      | 0.134±0.133      | 0.200±0.203      |                  |                  |
| 30       | Accuracy     | 78.6±3.8         | 97.0±0.4          | 78.6±1.8          | 97.0±0.4          | 78.6±3.8         | 97.0±0.4          | 78.6±3.8         | 97.0±0.4          | 78.6±3.8         | 97.0±0.4          |
|          | Time         | 0.148±0.143      | 0.216±0.217       | 0.260±0.258       | 0.539±0.532       | 0.179±0.179      | 0.268±0.284      | 0.132±0.149      | 0.197±0.181      |                  |                  |
| 35       | Accuracy     | 77.6±4.5         | 99.0±0.5          | 77.6±2.8          | 99.0±0.5          | 77.6±4.5         | 99.0±0.5          | 77.6±4.5         | 99.0±0.5          | 77.6±4.5         | 99.0±0.5          |
|          | Time         | 0.143±0.143      | 0.216±0.208       | 0.258±0.259       | 0.536±0.548       | 0.180±0.179      | 0.267±0.275      | 0.134±0.133      | 0.194±0.201      |                  |                  |
| 40       | Accuracy     | 74.5±5.9         | 98.6±0.6          | 74.5±2.8          | 98.6±0.6          | 74.5±5.9         | 98.6±0.6          | 74.5±5.9         | 98.6±0.6          | 74.5±5.9         | 98.6±0.6          |
|          | Time         | 0.144±0.142      | 0.216±0.212       | 0.262±0.259       | 0.537±0.532       | 0.180±0.172      | 0.276±0.278      | 0.133±0.148      | 0.196±0.196      |                  |                  |

In our second experiment, we investigated the average identification performance for all users and each sub-activity within a fixed size window (50 samples per window), for both feature selection methods with a different number of features. Table 1 lists the results for the case of the PCA-KELM and ESP-KELM. According to these results, the best performance was obtained with KELM when the normalized features were extracted by both sensors and processed with the ESP method. PCA performed slightly better than the case where no feature selection was used, but on average, there was no major difference with and without the feature selection process when using PCA.

Figure 2 shows the accuracy obtained through ESP using a different sized windows, as explained earlier, and 30 selected features, since 30 features provided the best average results in our previous experiments for all pockets data. By Figure 2, we can observe that the accuracy increased from 70% to 98% for 50 samples per window, which is a significant improvement for any user identification system. We thus conclude that the BLP and FRP have some variations according to the number of samples per window; however, this variation is not enough to exclude these sub-activities from our experimental setup, except the ones with 75 and 100 numbers of samples per window. This degradation happens due to sudden changes in users walking patterns. In future research, we will further investigate legitimate user behaviors, while performing the same sub-activity to minimize the ambiguity in the proposed legitimate user identification method.

Figures 1 and 2 present the 99% confidence intervals, pertaining to the average user identification by using the pairwise T-test between groups with and without feature selection data at the 99% confidence level. Looking at Figure 2, significant statistical results can be clearly seen, showing that the classifier performs much better when selected features of data are used in all cases: an 80% and 88% to 99% performance increase. This leads us to prefer the use of the feature selection method in future applications.

4 CONCLUSION AND FUTURE WORK

This study substantiates the idea to be able to detect legitimate SP users based on their walking patterns through various sub-activities in a semi-controlled environment. By using a commonly available

Figure 1: Average Accuracy for legitimate user identification obtained without selecting features.
mass-market consumer hardware in our experimental platform, we have demonstrated the global applicability of our proposed method with minimal accuracy variations. The KELM requires minimal battery consumption, and in order to run it as a practical application, one needs to limit the number of required samples.

There are many ways to extend our current work. This work demonstrated that ACC data is more useful than LACC, but it is possible that a fusion of ACC and LACC data will yield improved results. We have also been experimenting with more realistic features, which capture specific elements of a user’s gait, and plan to investigate if these features can yield additional improvements.

One of our goals for future work is to expand the evaluation of the proposed system so that it is applied to more real-life situations. Thus, we plan to expand our user base significantly, increasing the diversity of the users (especially with respect to gender and age), and evaluate how the system operates when the training and test samples are collected over lower window size. This is an important future direction for our work, in which we will split the size of the smallest windows (25 or 50 samples per window) in to even smaller windows within the classification process, moving to different challenges: how to maintain performance, within each sub-window, and how to control the computational complexity during real-time deployment.

ACKNOWLEDGMENTS

The authors would like to thank the three anonymous Referees for their valuable comments and helpful suggestions.

REFERENCES

[1] 2016. Smartphones in 2018. (2016). https://www.emarketer.com/Article/2-Billion-Consumers-Worldwide-Smartphones-by-2016/101694

[2] Muhammad Ahmad, Mohammed A. Alqarni, Khan Asad, and et al. 2018. Smartwatch-Based Legitimate User Identification for Cloud-Based Secure Services. Mobile Information Systems 2018, Article ID 5107024 (2018), 14 pages. https://doi.org/10.1186/s13670-018-07024

[3] Muhammad Ahmad, A. M. Khan, J. A. Brown, S. Protasov, and A. M. Khattak. 2016. Gait fingerprinting-based user identification on smartphones. In 2016 International Joint Conference on Neural Networks (IJCNN). 3060–3067. https://doi.org/10.1109/IJCNN.2016.7727588

[4] Muhammad Ahmad, Adil Mehmood Khan, Manuel Mazzara, and Salvatore DiStefano. 2019. Seeking Optimum System Settings for Physical Activity Recognition on Smartwatch. In Proceedings of the Computer Vision Conference (CVC). Springer-Verlag, Berlin, Heidelberg. https://arxiv.org/pdf/1706.01720.pdf

[5] Cheng Bo, Lan Zhang, Xiang-Yang Li, Qiujuan Huang, and Yu Wang. 2013. SilentSense: Silent User Identification via Touch and Movement Behavioral Biometrics. In Proceedings of the 19th Annual International Conference on Mobile Computing & Networking (MobiCom ’13). ACM, New York, NY, USA, 187–190. https://doi.org/10.1145/2400423.2400472

[6] Xue-wen Chen. 2003. An Improved Branch and Bound Algorithm for Feature Selection. Pattern Recogn. Lett. 24, 12 (Aug. 2003), 1925–1933. https://doi.org/10.1016/S0167-8655(03)00002-5

[7] Shiwei Ding, Han Zhao, Yanan Zhang, Xinzheng Xu, and Ru Nie. 2015. Extreme Learning Machine: Algorithm, Theory and Applications. Artif. Intell. Rev. 44, 1 (June 2015), 103–155. https://doi.org/10.1007/s10462-013-9405-z

[8] Muhammad Fahim, Iram Fatima, Sunyoung Lee, and Young-Tack Park. 2013. EFM: Evolutionary Fuzzy Model for Dynamic Activities Recognition Using a Smartphone Accelerometer. Applied Intelligence 39, 3 (Oct. 2013), 475–488. https://doi.org/10.1007/s10489-013-9427-7

[9] Christian Holz, Senaka Buthipitiya, and Marius Knaust. 2015. Bodyprint: Biometric User Identification on Mobile Devices Using the Capacitive Touchscreen to Scan Body Parts. In Proceedings of the 3rd Annual ACM Conference on Human Factors in Computing Systems (CHI ’15). ACM, New York, NY, USA, 3011–3014. https://doi.org/10.1145/2702123.2702518

[10] Guang-Bin Huang, Hongming Zhou, Xiaoqian Ding, and Rui Zhang. 2012. Extreme Learning Machine for Regression and Multiclass Classification. Trans. Sys. Man Cyber. Part B 42, 2 (April 2012), 513–529. https://doi.org/10.1109/TSMB.2011.2186604

[11] Y. A. Jeroudi, M. A. Ali, M. Latief, and R. Akmalawiawat, 2015. Online sequential extreme learning machine algorithm based human activity recognition using inertial data. In 2015 10th Asian Control Conference (ASCC). 1–6. https://doi.org/10.1109/ASCC.2015.7244597

[12] Adil Mehmood Khan, Muhammad Hameed Siddiqui, and Seok-Won Lee. 2013. Exploratory Data Analysis of Acceleration Signals to Select Light-Weight and Accurate Features for Real-Time Activity Recognition on Smartphones. Sensors 13, 10 (2013), 13099–13122. https://doi.org/10.3390/s131013099

[13] Adil Mehmood Khan, Ali Tofail, Asad Masood Khattak, and Teemu H. Laine. 2014. Activity Recognition on Smartphones via Sensor-Fusion and KDA-Based SVMs. International Journal of Distributed Sensor Networks 10, 5 (2014), 503291. https://doi.org/10.1155/2014/503291

[14] T. Kohonen, M. R. Schroeder, and T. S. Huang (Eds.). 2001. Self-Organizing Maps (3rd ed.). Springer-Verlag, Berlin, Heidelberg.

[15] Jundong Li, Kewei Cheng, Suhang Wang, Fred Norrander, Robert P. Trevino, Jiliang Tang, and Huan Liu. 2017. Feature Selection: A Data Perspective. ACM Comput. Surv. 50, 6, Article 104 (Dec. 2017), 45 pages. https://doi.org/10.1145/3136625

[16] Jorge-L. Reyes-Ortiz, Luca Oneto, Albert Samà, Xavier Parra, and Davide Angelucci. 2016. Transition-Aware Human Activity Recognition Using Smartphones. Neurocomput. 171, C (Jan. 2016), 754–767. https://doi.org/10.1016/j.neucom.2015.07.085

[17] Theresia Ratih Dewi Saputri, Adil Mehmood Khan, and Seok-Won Lee. 2014. User-Independent Activity Recognition via Three-Stage GA-Based Feature Selection. International Journal of Distributed Sensor Networks 10, 3 (2014), 706287. https://doi.org/10.1155/2014/706287 arXiv:https://doi.org/10.1155/2014/706287

[18] Weidong Shi, Jun Yang, Yifei Jiang, Feng Yang, and Yingen Xiong. 2011. SenGuard: Passive User Identification on Smartphones Using Multiple Sensors. In Proceedings of the 2011 IEEE 7th International Conference on Wireless and Mobile Computing, Networking and Communications (WiMOB ’11). IEEE Computer Society, Washington, DC, USA, 141–148. https://doi.org/10.1109/WiMOB.2011.6085412

[19] Xinhua Wang, M. Hossain Heydari, and Hua Lin. 2003. An Intrusion-Tolerant Password Authentication System. In Proceedings of the 12th International Symposium on Recent Advances in Intrusion Detection (RAID ’09). Springer-Verlag, Berlin, Heidelberg, 224–243. https://doi.org/10.1007/978-3-642-05450-0_12

Figure 2: Average Accuracy for legitimate user identification where the sensor signal processed through ESP-KELM.