Soft Biometrics and Its Implementation in Keystroke Dynamics

Mohd Noorulfakhri Yaacob¹, Syed Zulkarnain Syed Idrus¹,², Wan Nor Ashiqin Wan Ali¹,², Wan Azani Mustafa³, Mohd Aminudin Jamlos³ and Mohd Helmy Abd Wahab⁴

¹School of Human Development and Technocommunication, Universiti Malaysia Perlis, Perlis, 01000 Perlis, Malaysia.
²Center of Excellence Geopolymer and Green Technology, Universiti Malaysia Perlis, 01000 Perlis, Malaysia.
³Faculty of Engineering Technology, Universiti Malaysia Perlis, 01000 Perlis, Malaysia.
⁴Faculty of Electrical and Electronic Engineering, Universiti Tun Hussein Onn Malaysia, Malaysia.

Abstract. Biometrics is a unique art that exists within a person and allows it to be used to differentiate between one another. These biometrics can be divided into two categories namely behaviour and physical such as face, fingerprint, hand, voice and gait. There are previous studies which examine the personal’s characteristic or personality as gender, age, cultural, weight, height, colour of hair etc. This personal’s characteristic or personality also known as soft biometric. Several previous studies have shown that the use of soft biometric element (one or combination of elements) in the process of identifying individuals can improve the performance of individual recognition. This paper will elaborate and concludes past studies related to the use of soft biometric elements in KD. Several soft biometric elements applied in various methods of recognition on previous studies have been listed and the results of the studies are compared.

1. Introduction

Nowadays, most application systems have used biometric as a security feature before a person can use the system. The use of fingerprint, iris and face recognition is often used in systems with high confidentiality information. Most of these systems use one or a combination of two different biometric elements for identification purposes. For example, using fingerprint and iris simultaneously before entering the vault. However, the use or combination of biometric soft elements found in individuals is still lacking.

Most application systems use the username and password as the main tool to use the system. However, if the username and password are known to other users, then the system may be easily breached. To reduce and prevent password intrusion problem, the use of keystroke dynamic and soft biometric can be used as a second layer of verification after username and password. For example, the recorded typing of username and password during the enrolment process will be compared against the correct username and password entered. This writing presents and summarizes the results of past
studies which incorporated soft biometric dan Keystroke Dynamics. This will further help researchers continue their studies on KD in the future.

This writing can be divided into two categories, the first category is related to soft biometric that used in other area of study. The seconds category is related to the research of KD that apply the soft biometric criteria.

2. Soft Biometrics

Soft biometrics is a technique that can be used to recognize an individual. Examples of soft biometrics criteria on a person which can be differentiate from one another are height, weight, gender and eye colour. Generally soft biometrics criteria can be classified into two categories; continuous (height and weight) and real time (gender and eye colour).

Soft biometrics method implementation to identify users in criminology was pioneered by GALTON [1] in the beginning of the 18th century. This study used 3 categories of identification methods: (i) anthropometric measurements (arm length), (ii). Scar effect and Mole, (iii). Body shape. Later in 2001, Heckathorn, Broadhead [2] conducted research to identify people by combining some unique or general identities of a person such as gender, ethnicity, eye colour and height. These features require combination with other physical biometrics method to produce satisfying in identifying a person [2].

Besides that, soft biometrics was also implemented to filter or classify large biometrics data by adding gender features and age [3]. Research results found that this record classification could speed up search process and identification of a person. For example, if the user is a woman or man, database search can be specified right away to related category.

Application of soft biometrics in KD was commenced by Jain in his research regarding recognition using KD [4]. This study incorporated the usage of gender classification, ethnic, and height as additional features in physical biometrics system using fingerprint. The research concluded that user recognition process in database was accelerated.

To further justify research results by Jain, Dass [4], a research undertaken by [5] also ruled that soft biometrics method application does accelerate user recognition performance. Soft biometrics approach applied in this research are gender, ethnic, age, hair colour and some others.

Aside from that, in 2006, research conducted by Ailisto, Vildjionautaite [6] used BMI features or body weight to enhance finger scan biometric recognition. He managed to reduce EER rates to 1.5% from 3.9%. Later in 2008, a person’s walking style caught Ran, Rosenbush [7] interest to conduct a research on identifying people from walking style. Soft biometric features used in his research includes length, height and gender derived from video sequence.

The use of hair colour and ethnic recognition was also researched where ethnic recognition yield higher EER compared to hair colour with the result at 1.5% [8].

In 2010, a research utilizing soft biometrics was applied into recognition via face recognition which taken into account the features of gender, ethnic, scars, tattoo or mole available on top of a person’s face [9]. Meanwhile, recognition research using eyebrow based on gender was conducted with the results amounting to between 89% and 97%.

Tiwari, Singh [10] introduced a recognition method against new born babies using baby ear shape and general other soft biometric features including gender, height, weight and blood type. They took about 2100 images from 210 babies where each baby taken 10 pictures. Research result amounts to 90.7% precision in recognition.

Soft biometrics was also utilized by Koga, Yamazaki [11] to identify user using camera surveillance system, This research yielded 71.43% when using height criteria as identifier and using walking speed as criteria also yielded 71.43%. Almost similar research was conducted by Moctezuma, Conde [12] by identifying people using a series of video recordings. They managed to produce 97.90% in the process of identifying people in a series of video recordings. In 2014, recognition research based on soft biometrics was also conducted based on image distance taken by surveillance camera [13].

User recognition using smart phone was also conducted by utilizing soft biometric features such as heart beat rate and phone conversation style [14]. Heart beat rate were taken using a few
sensors available on smart phone and conversation style was used to identify user emotion when using smart phone at a particular time. Research results in 84.7% precision in user emotional recognition.

Apart from that, soft biometrics application in research was also conducted by Yang, Yang [15]. This research emphasized the new features in soft biometrics namely the finger vein image and ‘the width of phalangeal joint’. This leads to the result of EER between 5.53% and 8.08% for open sourced database, whereas EER between 3.4% and 5% for data gathered themselves in the research.

Research on combination of soft biometrics and hard biometrics in continuous authentication using computer was conducted by Prakash and Mukesh [16]. They have identified user features based on shirt colour and user facial skin. Registration was being done automatically by the system each time a user uses the system.

Most research on the field of soft biometrics at present is being conducted for forensics purposes and recognition such as identity recognition of a person through video images or static images. [17, 18]. Also, in the year 2016, a research pertaining soft biometrics was conducted using the method extracted from the body and the correlation between [19]. Research was conducted on 5 people instructed to pose some action such as walking and sitting. Research result gained raised min up to 4.39% after these soft biometrics was incorporated into the recognition criteria. Recognition process using tattoo on the body was also being conducted in 2016 [20]. Difference between original tattoo and other images was identified in this research.

Moving on to 2017, research on the use of soft biometrics in the investigations using video recordings was conducted by analysing unique features main in the way a person walks and facial figure [21].

2.1. Summary of Soft Biometric

Studies regarding soft biometric have begun from 1896 to present. Numerous soft biometric elements have been used in previous studies in various aspects. Table 1 shows the summary of soft biometric elements application in previous studies. The results of this study have shown that soft biometric elements can be used as an additional element in identifying a person.

| Year  | Mole | Scar Effect Tattoo | Body Figure | Gender | Cultural | Eye Colour | Height | Weight | Hair Colour | Age | BMI | Walking/sitting Style | Blood Type | Heart Beat | Talking Style | Vein Image | Facial Skin / Figure | Ear Shape |
|-------|------|-------------------|-------------|--------|----------|------------|--------|--------|-------------|-----|-----|----------------------|------------|-----------|---------------|------------|---------------------|----------|
| 1896  | [1]  | √                 | √           |        |          |            |        |        |             |     |     |                       |            |           |               |            |                     |          |
| 1997  | [3]  |                   |             | √      |          |            |        |        |             |     |     |                       |            |           |               |            |                     |          |
| 2001  | [2]  |                   |             | √      | √        |            |        |        |             |     |     |                       |            |           |               |            |                     |          |
| 2004  | [4]  |                   |             | √      | √        |            |        |        |             |     |     |                       |            |           |               |            |                     |          |
| 2005  | [5]  |                   |             | √      | √        |            |        |        |             |     |     |                       |            |           |               |            |                     |          |
| 2006  | [6]  |                   |             |        |          | √          |        |        |             |     |     |                       |            |           |               |            |                     |          |
| 2008  | [7]  |                   |             |        |          |            |        |        |             |     |     |                       |            |           |               |            |                     |          |
| 2009  | [8]  |                   |             |        |          |            |        |        |             |     |     |                       |            |           |               |            |                     |          |
| 2010  | [9]  |                   |             | √      | √        |            |        |        |             |     |     | √                     |            |           |               |            |                     |          |
| 2012  | [10] |                   |             |        |          |            |        |        |             |     |     |                       |            |           |               |            |                     |          |
| 2013  | [11] |                   |             |        |          |            |        |        |             |     |     |                       |            |           |               |            |                     |          |
| 2013  | [12] |                   |             |        |          |            |        |        |             |     |     |                       |            |           |               |            |                     |          |
| 2013  | [14] |                   |             |        |          |            |        |        |             |     |     | √                     |            |           |               |            |                     |          |
| 2014  | [13] |                   |             |        |          |            |        |        |             |     |     |                       |            |           |               |            |                     |          |
3. Soft biometrics application for keystroke dynamics

Research on use of soft biometrics in KD to identify a user’s emotion while using a keyboard was conducted [22]. Research resulted in accuracy between 77.4% and 87.8 % for emotion category of confidence, hesitant, nervous, relax, sad, and fatigue. Aside from that, for emotion category of anger and joy, results yielded up to 84% accuracy.

Gender recognition method caught the interest of Fairhurst and Da Costa-Abreu [23] to identify users of social media using GREYC dataset. They utilized the method of 10-fold cross-validation to analyze obtained data and resulted with 95% accuracy.

Next, Giot and Rosenberger [24] researched the capability to identify the gender of a user based on typing style. This research produced accuracy in identifying gender between 87.32% and 91.63% based on testing of data retrieved from GREYC dataset. Reduction in 20% of EER measurement aspect was also achieved compared to earlier methods.

Gunawardhane, De Silva [25] researched KD from the aspect of user’s stress. Results from the research allows to differentiate whether a user is currently under stress or not. Additionally, research by Nahin, Alam [26] had proven that a person’s emotion could be identified based on his/her typing style. That research resulted with more than 80% accuracy by classifying emotion into 7 categories namely anger, disgust, guilt, fear, joy and shame.

Research regarding user emotion were also conducted by other researchers in the same year by studying user utilizing keyboards, mouse and touch screen [27]. This research compared other normal methods regularly used to identify emotion such as Electroencephalography (EEG) machines, facial expression, voice and body language. Accuracy percentage increased approximately 5% totalling to 93.2% compared to other methods.

Combination of soft biometrics and keystroke dynamics inspired Idrus et. al. [28-30] to conduct research to identify users based on gender classification, age, and left or right hand handedness. EER produced is a total of 5.41%.

Research to identify user gender based on KD and using touch screen input was conducted in 2016 [31]. Research result established recognition accuracy obtained totalling to 64.76% against dataset keystroke records and 57.16% against touch screen usage.

The latest KD research relating to soft biometric is performed by [32]. She move forward the continuity of study made by Nahin [26] to identify people’s emotion while typing. There were 5 types of emotions that are attempted to be identified in her study and those emotions are Happiness, Boredom, Fear, Anger and Sadness. She concluded that the user’s emotion does influence the way a person types, however a person’s strength also affects the typing activity.

Also in this 2018, Katerina and Nicolaos [33] studied user’s behavior based on the usage of mouse and typing style while using End-User Development (UED) software tool. Users are classified into 4 different levels that is knowledgeable with using computer, knowledgeable in using web, knowledgeable in handling database and knowledgeable in programming.

As a result of earlier researches, we can conclude that there a handful of researches regarding soft biometrics using KD being conducted. Albeit the numbers of research, there is still no research available that focused on the scope and objective of this research.

3.1. Summary of implementation Soft Biometric in Keystroke Dynamic

The usage of soft biometric elements in KD studies has also been carried out by some earlier researchers beginning in 2011. This paper will summarize the results of previous studies using soft biometric elements in KD studies. The result of previous studies can be concluded that the addition of
new elements in the identification process of using KD can improve the level of individual’s recognition. This writing will provide a significant inspiration for KD researchers to merge some of the other soft biometric elements in the next study. Table 2 shows the summary of the relevant past studies for soft biometric on KD.

| Year  | Confidence | Hesitance | Nervous | Relax | Sad | Fatigue | Anger | Joy/Happiness | Stress | Disgust | Guilt | Fear | Shame | Boredom | Gender | Age | Handedness | Mouse usage |
|-------|------------|-----------|---------|-------|-----|---------|-------|---------------|--------|---------|------|------|-------|---------|--------|-----|-------------|-------------|
| 2011  | √          | √         | √       | √     | √   | √       | √     | √             |        |         |      |      |       |         |        |    |             |             |
| 2011  | [22]       |           |         |       |     |         |       |               |        |         |      |      |       |         |        |    |             |             |
| 2012  | [23]       |           |         |       |     |         |       |               |        |         |      |      |       |         |        |    |             |             |
| 2013  | [24]       |           |         |       |     |         |       |               |        |         |      |      |       |         |        |    |             |             |
| 2014  | [25]       |           |         |       |     |         |       |               |        |         |      |      |       |         |        |    |             |             |
| 2014  | [26]       |           |         |       |     |         |       |               |        |         |      |      |       |         |        |    |             |             |
| 2014  | [27]       |           |         |       |     |         |       |               |        |         |      |      |       |         |        |    |             |             |
| 2014  | [28]       |           |         |       |     |         |       |               |        |         |      |      |       |         |        |    |             |             |
| 2016  | [31]       |           |         |       |     |         |       |               |        |         |      |      |       |         |        |    |             |             |
| 2018  | [32]       |           |         |       |     |         |       |               |        |         |      |      |       |         |        |    |             |             |
| 2018  | [33]       |           |         |       |     |         |       |               |        |         |      |      |       |         |        |    |             |             |

4. Conclusion
The used of ancillary user information or soft biometrics such as cultural, gender, age, hair colour and skin colour can be used for individual identification process. The combination of soft biometric elements in KD can help classify and distinguish patterns and how someone is typing based on soft biometric categories. Although some of these soft biometrics data may change from time to time, but they can be used for item classification. This writing has summarized the results of previous studies that have incorporated soft biometrics elements together with KD studies. Generally, based on previous studies, the use of biometric soft elements in KD has improved performance in individual identification. Based on Table 1, it can be concluded that there are other soft biometric elements that can be studied together with KD. Therefore, based on the previous study, a conclusion can be made that studies in the field of KD are still to be continued by combining several elements of soft biometric to enhance recognition using KD. In addition, this writing is also expected to assist researchers in the same field to obtain additional inspiration for enhancement of KD for the next study.

5. References
1. GALTON, F., Signaletic Instructions, including the Theory and Practice of Anthropometrical Identification. 1896.
2. Heckathorn, D.D., R.S. Broadhead, and B. Sergeyev, *A methodology for reducing respondent duplication and impersonation in samples of hidden populations.* Journal of Drug Issues, 2001. **31**(2): p. 543-564.

3. Wayman, J. *Large-scale civilian biometric systems-issues and feasibility.* in *Proceedings of Card Tech/Secur Tech ID.* 1997.

4. Jain, A.K., S.C. Dass, and K. Nandakumar, *Soft biometric traits for personal recognition systems,* in *Biometric Authentication.* 2004, Springer. p. 731-738.

5. Ambalakat, P. *Security of biometric authentication systems.* in proceedings of 21st Computer Science Seminar. 2005.

6. Alliston, H., et al., *Soft biometrics—combining body weight and fat measurements with fingerprint biometrics.* Pattern Recognition Letters, 2006. **27**(5): p. 325-334.

7. Ran, Y., G. Rosenbush, and Q. Zheng. *Computational approaches for real-time extraction of soft biometrics.* in *Pattern Recognition, 2008. ICPR 2008. 19th International Conference on.* 2008. IEEE.

8. Marcialis, G.L., F. Roli, and D. Muntoni, *Group-specific face verification using soft biometrics.* Journal of Visual Languages & Computing, 2009. **20**(2): p. 101-109.

9. Park, U. and A.K. Jain, *Face matching and retrieval using soft biometrics.* IEEE Transactions on Information Forensics and Security, 2010. **5**(3): p. 406-415.

10. Tiwari, S., A. Singh, and S.K. Singh. *Fusion of ear and soft-biometrics for recognition of newborn.* Signal and Image Processing: An International Journal, 2012. **3**(3): p. 103-116.

11. Koga, Y., Y. Yamazaki, and M. Ichino. *A study on the surveillance system using soft biometric information.* in *Consumer Electronics (GCCE), 2013 IEEE 2nd Global Conference on.* 2013. IEEE.

12. Moctezuma, D., et al. *Incremental Learning with Soft-Biometric Features for People Re-Identification in Multi-Camera Environments.* in *Digital Image Computing: Techniques and Applications (DICTA), 2013 International Conference on.* 2013. IEEE.

13. Tome, P., et al., *Soft Biometrics and Their Application in Person Recognition at a Distance.* IEEE Transactions on Information Forensics and Security, 2014. **9**(3): p. 464-475.

14. Zhang, W., et al. *A hybrid emotion recognition on android smart phones.* in *Green Computing and Communications (GreenCom), 2013 IEEE International Conference on and IEEE Cyber, Physical and Social Computing.* 2013. IEEE.

15. Yang, L., et al., *Exploring soft biometric trait with finger vein recognition.* Neurocomputing, 2014. **135**: p. 218-228.

16. Prakash, A. and R. Mukesh, *A Biometric Approach for Continuous User Authentication by Fusing Hard and Soft Traits.* IJ Network Security, 2014. **16**(1): p. 65-70.

17. Azam, S. and M. Gavrila, *Soft biometric: Give me your favorite images and I will tell your gender.* in *Cognitive Informatics & Cognitive Computing (ICCI* *CC), 2016 IEEE 15th International Conference on.* 2016. IEEE.

18. Tome, P., et al., *Facial soft biometric features for forensic face recognition.* Forensic science international, 2015. **257**: p. 271-284.

19. Gharaghabi, S., et al. *Improving person recognition by weight adaptation of soft biometrics.* in *Computer and Knowledge Engineering (ICCKE), 2016 6th International Conference on.* 2016. IEEE.

20. Hrkac, T., K. Brkic, and Z. Kalafatic. *Tattoo Detection for Soft Biometric De-Identification Based on Convolutional Neural Networks.* in *1st OeAGM-ARW Joint Workshop.* 2016.

21. Correia, P.L., et al, *On Using Soft Biometrics in Forensic Investigation,* in *Handbook of Biometrics for Forensic Science,* M. Tistarelli and C. Champod, Editors. 2017, Springer International Publishing: Cham. p. 241-266.

22. Epp, C., M. Lippold, and R.L. Mandryk. *Identifying emotional states using keystroke dynamics.* in *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems.* 2011. ACM.
23. Fairhurst, M. and M. Da Costa-Abreu. *Using keystroke dynamics for gender identification in social network environment*. in *Imaging for Crime Detection and Prevention 2011 (ICDP 2011), 4th International Conference on*. 2011. IET.

24. Giot, R. and C. Rosenberger, *A new soft biometric approach for keystroke dynamics based on gender recognition*. International Journal of Information Technology and Management, 2012. 11(1-2): p. 35-49.

25. Gunawardhane, S.D., et al. *Non invasive human stress detection using keystroke dynamics and pattern variations.* in *Advances in ICT for Emerging Regions (ICTer), 2013 International Conference on*. 2013. IEEE.

26. Nahin, A.N.H., et al., *Identifying emotion by keystroke dynamics and text pattern analysis*. Behaviour & Information Technology, 2014. 33(9): p. 987-996.

27. Bakhtiyari, K., M. Taghavi, and H. Husain. *Implementation of emotional-aware computer systems using typical input devices*. in *Asian Conference on Intelligent Information and Database Systems*. 2014. Springer.

28. Idrus, S.Z.S., et al., *Soft biometrics for keystroke dynamics: Profiling individuals while typing passwords*. Computers & Security, 2014. 45: p. 147-155.

29. Idrus, S. Z. S. (2014). *Soft Biometrics for Keystroke Dynamics (Biométrie douce pour la dynamique de frappe au clavier)* (Doctoral dissertation, University of Caen Normandy, France).

30. S Syed Idrus, E Cherrier, C Rosenberger (2016). Fusion et biométrie douce pour la dynamique de frappe au clavier. *Colloque COMpression et REprésentation des Signaux Audiovisuels (CORESA)*.

31. Antal, M. and G. Nemes. *Gender recognition from mobile biometric data*. in *Applied Computational Intelligence and Informatics (SACI), 2016 IEEE 11th International Symposium on*. 2016. IEEE.

32. Kołakowska, A., *Usefulness of Keystroke Dynamics Features in User Authentication and Emotion Recognition*, in *Human-Computer Systems Interaction: Backgrounds and Applications 4*, Z.S. Hippe, J.L. Kulikowski, and T. Mroczek, Editors. 2018, Springer International Publishing: Cham. p. 42-52.

33. Katerina, T. and P. Nicolaos, *Mouse behavioral patterns and keystroke dynamics in End-User Development: What can they tell us about users’ behavioral attributes?* Computers in Human Behavior, 2018. 83: p. 288-305.