- **Average Resultant Acceleration (1):** for each value in the segment of x, y, and z axes, the square roots of the sum of the values of each axis squared over the segment size (i.e., 10 seconds) are calculated.
- **Variance (3):** The second-order moment of the data.

3.2.2 Frequency domain features. The process of extracting frequency domain features is somewhat different from the time domain. Before extracting a frequency domain feature, a Fourier transform needs to be applied to the data. A set of frequency domain features are calculated which might be useful to create a discriminative feature vector for each individual. The extracted features are presented in Table 2. In the second and the fourth columns, NF stands for the number of generated features.

Table 2. Frequency domain features

Features	NF	Features	NF
Energy	3	Difference	3
Entropy	3	Zero crossing rate	3
Root Mean Square	3	Interquartile range	3
Maximum	3	Correlation Coefficients	3
Minimum	3	Percentiles25	3
Standard Deviation	3	Percentiles 50	3
Median	3	Skewness	3
Variance	3	Kurtosis	3
Average Absolute Difference	3	Average Resultant Acceleration	1

3.2.3 Validating features extracted from the smartwatch. In order to validate the effectiveness of the 143 generated features for a promising authentication technique, the data set was divided to form both reference and testing templates for all users in two scenarios (i.e., Same-Day and Cross-Day). The average Euclidean distance between the reference template and testing templates was calculated; this distance value represents the similarity between the two templates: the smaller the value, the more similarity between the reference and testing templates and vice versa. As a result, in order for this technique to work, a small distance value should be presented when the reference and testing templates are from the same user; while a large distance value should be expected when these templates are from different users – representing the intra and inter sample variances. The results of 36 users' movement data for the Same-Day and Cross-Day scenarios are presented in Tables 3 and 4 respectively.

Table 3 shows (for the Same-Day) the acceleration templates of the same user competitive average Euclidean distance scores, ranging from 0.55 (subject 17) to 1.41 (subject 34). When gyroscope data was used, the distance scores of the same subject were in the range of 1.41 (subject 35) – 4.61 (subject 14). In comparison, average Euclidean distance scores for reference and testing templates of different subjects that are extracted on the same day are much larger: 2.54 (subject 25) to 3.33 (subject 34) for acceleration and 3.57 (subject 4) - 5.85 (subject 20) for gyroscope.

The imposter distance scores from each genuine user are further analyzed separately (Figures 1 and 2 for acceleration and gyroscope data respectively). The given acceleration based- results in Figure 1 show that the user's arm movement is highly consistent and each subject has a distinctive arm pattern. Moreover, the majority of imposters are more likely to be rejected by the system as their distances scores were far away from the genuine user. In contrast, when gyroscope data was applied, the average Euclidian distance scores between imposters and a genuine user were greatly dependent on the subject (Figure 2). For example, subjects 3, 5, 7, 11, 13, 15, 17, 21, 27, and 33 had low inter-variance, which means the chance of accepting an imposter is high. One reason for this is that using a large number of features might influence the system performance.

Table 3. Results of Same-Day Scenario

ID	Dist to Self		Dist to Others		ID	Dist to Self		Dist to Others	
	Acc	Gyr	Acc	Gyr		Acc	Gyr	Acc	Gyr
1	0.89	2.08	2.76	3.82	**19**	0.97	1.96	3.08	4.45
2	0.72	1.82	3.3	3.72	**20**	1.58	2.2	3.13	5.85
3	0.77	3.21	3.18	4.03	**21**	0.83	2.45	2.62	3.51
4	1.11	2.2	3.07	3.57	**22**	0.74	1.53	2.82	3.74
5	0.89	2.35	2.69	3.75	**23**	1.32	2.96	2.68	5.13
6	1.01	1.73	2.65	3.84	**24**	1.04	2.35	2.72	3.73
7	1.15	2.57	2.78	3.7	**25**	1.01	1.76	2.54	3.74
8	1.02	2.56	2.67	4.14	**26**	0.91	1.91	3.13	3.72
9	0.84	1.78	2.7	3.61	**27**	1.17	3.05	3.69	3.97
10	1.18	2.21	2.84	3.9	**28**	1.12	2.23	2.58	3.61
11	1.19	4.94	2.93	5	**29**	1.2	2.18	2.89	3.8
12	0.76	2.35	2.57	4	**30**	1.02	2	2.8	3.59
13	1.02	3.9	2.71	4.82	**31**	1.01	1.9	2.85	3.72
14	1.23	4.61	3.17	5.24	**32**	0.89	2.33	3.06	3.49
15	0.98	2.44	2.83	3.72	**33**	0.86	2.95	2.73	3.79
16	1.4	1.87	3.23	4.78	**34**	1.41	3.14	3.33	4.64
17	0.55	3.59	2.91	4.5	**35**	0.91	1.41	2.62	3.93
18	0.97	2.39	2.92	4.72	**36**	0.85	1.51	2.57	4.07

A more realistic test for a behavioural based-biometric comes when the Cross-day scenario is applied to show the influence of the variation of human movement over time. Therefore, the Cross-day scenario was also evaluated and the results shown in Table 4.

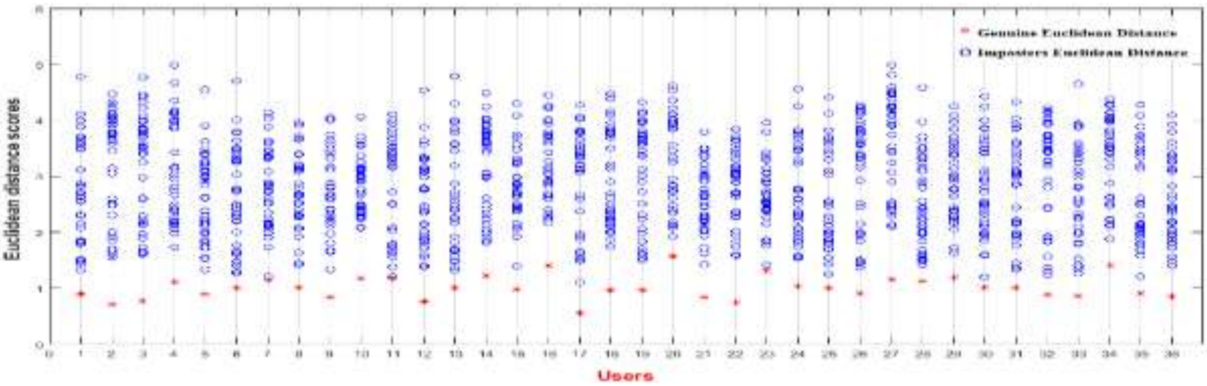

Figure 1. Acceleration Euclidean Distance Scores Using All Features

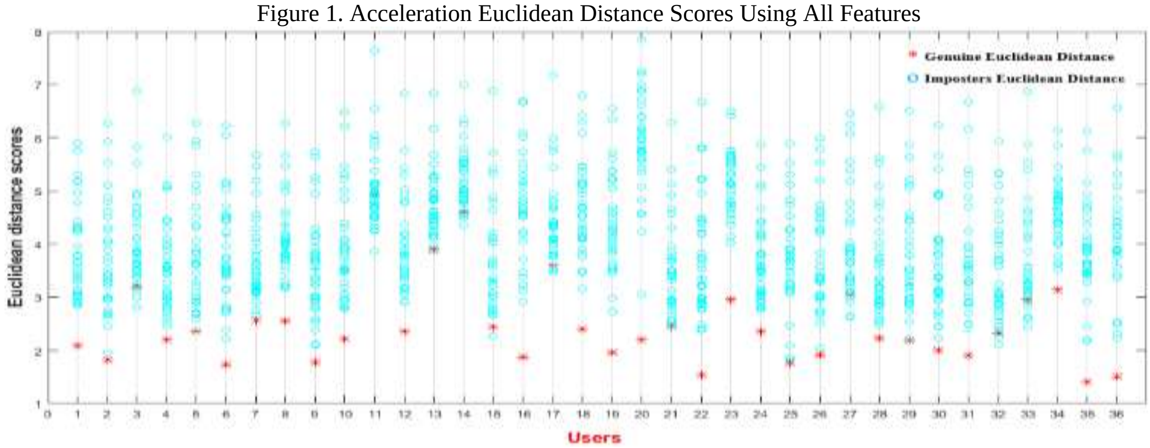

Figure 2. Gyroscope Euclidean Distance Scores Using All Features

While the distance scores under this more realistic evaluation scenario for acceleration and gyroscope templates of the genuine user were increased, they were still viable to be used for discriminating users: ranging from 0.58 (subject 17) to 2.11 (subject 10) for acceleration and from 1.52 (subject 22) to 4.51 (subject 13) for gyroscope. In comparison, the resulting distance scores for reference and probe templates of imposters were generally quite high: 2.47 (subject 25) to 3.39 (subject 27) for acceleration, which is an indication that imposters are more likely to be rejected by the system. In contrast, the distance scores for the gyroscope were slightly larger ranging from 3.6 (subject 9) to 6.21(subject 20); this could cause more imposters to be falsely accepted. The results also show the necessity of using a sensor fusion approach (i.e., combining the smartwatch sensors data) in order to have a balance between security and usability. In addition, an improved feature selection method to select a set of features that have low-intra and high inter-variance is definitely required.

Table 4. Results of Cross-Day Scenario

ID	Dist to Self		Dist to Others		ID	Dist to Self		Dist to Others	
	Acc	Gyr	Acc	Gyr		Acc	Gyr	Acc	Gyr
1	1	3.28	2.65	4.5	**15**	1.69	3.9	2.67	4.38
2	0.91	1.94	3.29	3.83	**16**	1.33	2.1	2.93	4.49
3	0.9	3.38	3.09	3.61	**17**	0.58	4.0	2.91	4.29
4	1.16	2.43	2.94	3.68	**18**	1.12	2.3	2.76	4.18
5	1.05	2.29	2.62	4.15	**19**	1.2	2.9	2.91	4.27
6	0.95	1.58	2.63	3.85	**20**	1.4	2.7	2.68	6.21
7	1.03	2.71	2.76	3.81	**21**	1.0	2.4	2.62	4.5
8	0.97	3.19	2.67	4.3	**22**	1.0	1.5	2.63	3.83
9	1.1	1.86	2.63	3.6	**23**	1.2	4.1	2.6	3.61
10	2.11	3.1	3	4.18	**24**	0.9	3.3	2.74	3.68
11	1.13	3.8	2.83	4.39	**25**	1.3	1.6	2.47	4.15
12	0.89	2.81	2.59	4.1	**26**	1.0	2	3.12	3.85
13	1.1	4.51	2.69	4.87	**27**	1.2	1.9	3.39	3.81
14	0.85	4.18	3.14	5.69	**28**	1.0	3.9	2.52	4.3
29	1.4	2.6	2.91	3.6	**33**	0.9	2.7	2.66	4.87
30	0.8	2.1	2.68	4.18	**34**	1.1	2.7	3.32	5.69
31	1.1	2.0	2.79	4.39	**35**	1.0	2.0	2.64	4.38
32	0.8	2.1	3.04	4.1	**36**	0.8	1.5	2.55	4.49

4. Feature selection approach

The feature selection step has become the focus of many research studies in the area of authentication in order to reduce potentially large dimensionality of input data and thus system performance could be enhanced by selecting the most optimal and unique features for individual. Furthermore, it will be easier to manipulate small feature subsets on digital devices (i.e., smartphones and smartwatches). The majority of activity recognition systems select common features (e.g., features that have the smallest standard deviation) for all the population.

This could be very useful if it is considered that the authentication system is based on identifying the genuine user only. However, a balance between security and usability needs to be taken for Transparent Authentication Systems (i.e., low false acceptance rate (FAR) and low false rejection rate (FRR)). FAR shows the percentage in which the system incorrectly accepts an imposter as the legitimate user while FRR displays the percentage in which the authorized user is wrongly rejected by the system.

The current study focused on creating a dynamic feature vector that contains unique features for each subject. This was achieved by measuring the standard deviation (STD) for each feature and, subsequently, selecting feature subsets that have the smallest STD for each user independently. Using this method 30 features were identified for each subject. For example, the reference template of subject 1 could be created by using features 1, 2, 3, and 7 (features with smallest STD) while features 3, 4, 5, and 7 might be used to form the reference template of subject 2. This could result in low FRR and FAR. Moreover, selecting small feature subsets will greatly reduce the complicated computations on smartphones, which limit processing resources as compared to standard computers.

To evaluate the effectiveness of the selected feature subsets (30 features) for classification, the Euclidean distance metric for both scenarios (i.e., Same-Day and Cross-Day) are calculated and the results were presented in Tables 5 and 6 accordingly. The results in Table 5 indicate that applying small feature subsets yields very small distance scores between the training and test of the genuine user ranging from 0.03 (subject 6) to 0.2 (subject 16) for acceleration and 0.19 (subject 22) to 0.39 (subject 18) for gyroscope (compared to 0.55 and 1.41 for acceleration and 1.41 to 3.59 for gyroscope when the entire feature sets are used). These results suggest that the chance of a genuine user being correctly authenticated by the system is high. Also, the system would be able to identify imposters as their Euclidean distance scores are large: 0. 57 (subject 28) to 1.65 (subject 27) and 0.48 (subject 26) to 1.1 (subject 15) for acceleration and gyroscope respectively. Interestingly, the results in Table 6 show that the selected feature subsets are more resistant to changes of the user's behavior as the Euclidean distance scores of Same and Cross-day scenarios for most subjects are nearly similar, apart from subjects 10, 15, 25, 27, 29, and 31 for acceleration and subjects 9, 10, 15, and 23 for gyroscope.

By using features associated with the acceleration data, Figure 3 shows that all imposters will be more likely to be rejected by the system (apart from subject 8 as one or two imposters might be able to deceived the system).

Table 5. Results of Same-Day Scenario by using 30 Features

ID	Dist to Self		Dist to Others		ID	Dist to Self		Dist to Others	
	Acc	Gyr	Acc	Gyr		Acc	Gyr	Acc	Gyr
1	0.08	0.32	0.99	1.06	**19**	0.14	0.25	0.78	0.51
2	0.06	0.23	1.05	0.5	**20**	0.11	0.36	1.02	1
3	0.07	0.23	1.18	0.84	**21**	0.06	0.26	0.82	0.56
4	0.08	0.29	0.91	0.56	**22**	0.04	0.19	0.84	0.95
5	0.05	0.26	0.9	0.52	**23**	0.1	0.29	0.9	0.64
6	0.03	0.25	0.7	1.03	**24**	0.08	0.3	0.71	0.51
7	0.09	0.3	0.83	0.58	**25**	0.06	0.24	0.88	0.55
8	0.1	0.3	0.69	0.53	**26**	0.06	0.3	0.89	0.48
9	0.07	0.2	0.7	0.5	**27**	0.11	0.29	1.65	0.94
10	0.15	0.2	1.02	0.95	**28**	0.07	0.21	0.57	0.8
11	0.1	0.3	1.06	0.59	**29**	0.08	0.23	0.58	1.05
12	0.04	0.21	0.97	0.89	**30**	0.08	0.25	0.76	0.91
13	0.08	0.33	0.65	0.59	**31**	0.06	0.24	0.65	0.53
14	0.08	0.27	1	0.62	**32**	0.05	0.21	1	0.84
15	0.09	0.23	0.71	1.1	**33**	0.05	0.25	0.77	1.07
16	0.2	0.33	1.16	0.68	**34**	0.2	0.31	1.12	0.61
17	0.05	0.31	0.9	0.56	**35**	0.1	0.24	0.85	0.58
18	0.13	0.39	1.1	0.99	**36**	0.07	0.25	0.86	0.89

Table 6. Results of Cross-Day Scenario by using 30 Features

ID	Dist to Self		Dist to Others		ID	Dist to Self		Dist to Others	
	Acc	Gyr	Acc	Gyr		Acc	Gyr	Acc	Gyr
1	0.09	0.31	0.9	1.08	**19**	0.17	0.31	0.74	0.51
2	0.11	0.26	1.15	0.53	**20**	0.18	0.35	0.64	0.93
3	0.12	0.25	1.19	0.55	**21**	0.06	0.27	0.81	0.62
4	0.1	0.28	0.69	0.59	**22**	0.05	0.22	0.95	1
5	0.12	0.31	1	0.48	**23**	0.17	0.48	0.74	0.83
6	0.05	0.23	0.69	0.99	**24**	0.11	0.38	0.88	0.57
7	0.07	0.31	0.83	0.61	**25**	0.21	0.25	0.46	0.59
8	0.09	0.31	0.75	0.5	**26**	0.12	0.29	0.56	0.51
9	0.12	0.3	0.79	0.52	**27**	0.2	0.26	0.67	1.19
10	0.28	0.31	1.1	0.99	**28**	0.08	0.23	0.54	0.97
11	0.13	0.29	1.02	0.61	**29**	0.17	0.27	0.95	1.07
12	0.05	0.22	0.99	1.01	**30**	0.1	0.3	0.86	0.48
13	0.08	0.31	0.67	0.66	**31**	0.12	0.23	0.5	0.57
14	0.09	0.31	0.69	0.65	**32**	0.05	0.25	0.98	0.96
15	0.27	0.39	0.78	1.15	**33**	0.1	0.3	0.77	1.03
16	0.18	0.3	1.01	0.62	**34**	0.18	0.29	1.09	0.65
17	0.06	0.27	0.88	0.59	**35**	0.12	0.27	0.86	0.86
18	0.17	0.3	1.04	0.77	**36**	0.08	0.26	0.85	0.8

When gyroscope features are used, Figure 4 reveals that the system was still able to identify the majority of imposters. While some of the gyroscope results may not seems that positive, they are actually quite impressive when one considers that they were produced from only 30 features. Compared to the previous experiment, which used the whole gyroscope feature set (143 features), it can be clearly noticed that the imposters overlapping with subjects 3, 5, 11, 15, 21, 27, and 33 are greatly reduced. This is due to the fact that selecting more discriminative feature sets could result in low intra-variance and high inter-variance. The results show that accelerometer features are unique and more distinctive than gyroscope features as the distance scores between the reference and test templates of the genuine user are small (i.e., low intra- variance), as well as provide a significant distinction between the genuine user and imposters (i.e., high inter-variance).

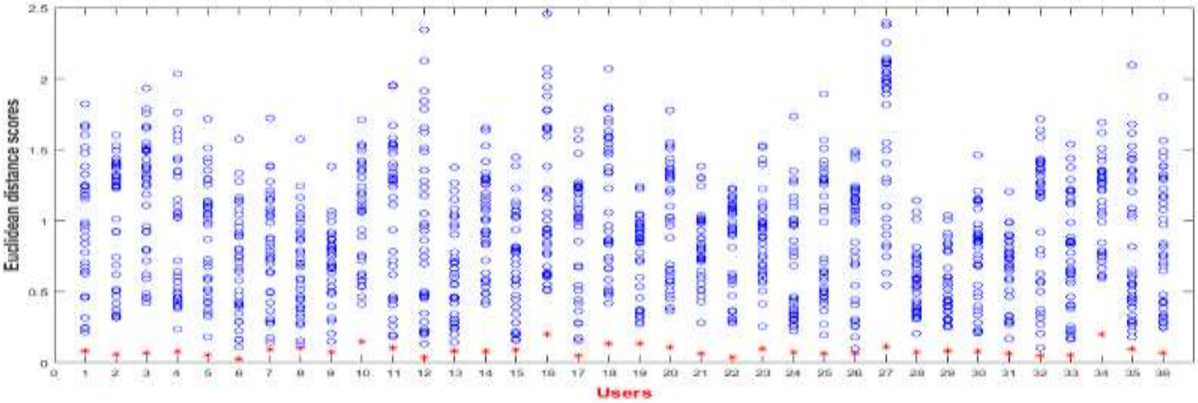

Figure 3. Acceleration Euclidean Distance Scores Using 30 Features

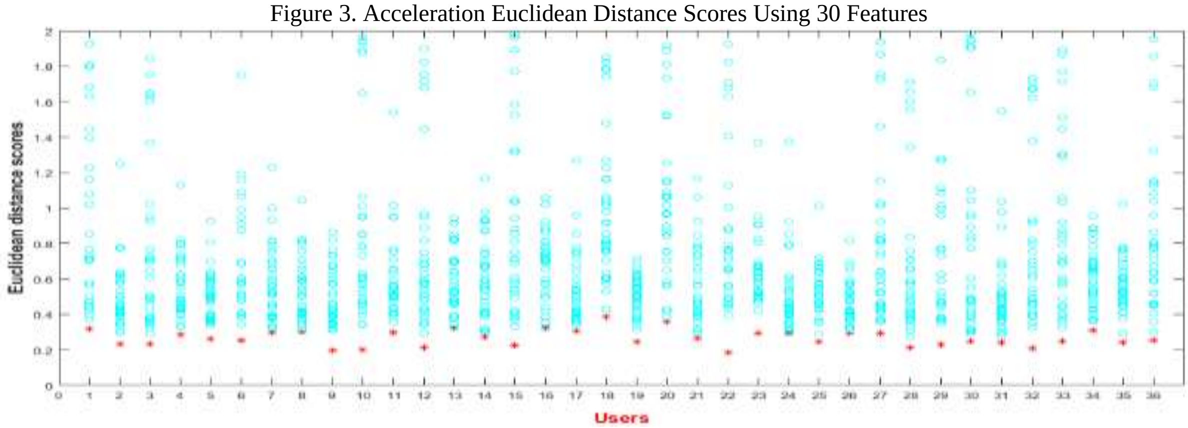

Figure 4. Gyroscope Euclidean Distance Scores Using 30 Features

5. Proposed Architecture to support Smartwatch-based Activity Recognition

A high-level architecture of the proposed system is presented in Figure 5. The prior art has established that managing the complex and varying signals of real-life use is a significant barrier. In order to overcome this, a context aware approach will be used in order to predict the user's activity at a specific point of time. This can be achieved by obtaining information from other smartwatch sensors (e.g., GPS) and using the information to create a multi-classifier approach that is trained to specific activities. This should result in a reduction in the variability in the feature set and provide better classification performance.

Unlike most of the prior studies that utilized information from a single sensor only (i.e., accelerometer or gyroscope), the proposed system aims to collect the movement data of both sensors as well as GPS information.

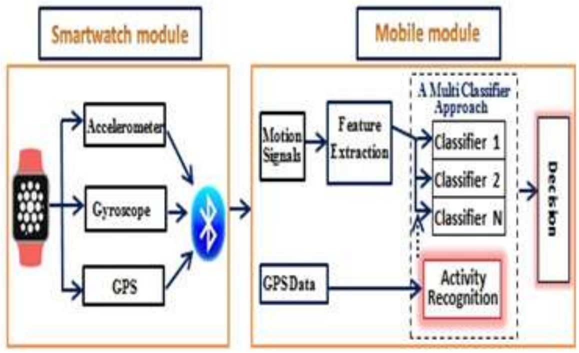

Figure 5. Proposed Architecture for Motion-based Activity Recognition

It is possible that the fusion of acceleration and gyroscope data would offer a greater level of accuracy than either sensor alone. Thereafter, feature selection needs to be sophisticated enough before the classification phase takes place. This can be achieved by selecting the features that are more resistant to changes of the user's behaviour. Finally, a set of classification methods will be evaluated to create a model for each individual activity

6. Conclusion and Future Work

In the experimental study, movement data was collected from 36 subjects and the feature set analysed to determine its uniqueness. The data collection process was more realistic than previous studies [20], as each subject was asked to walk a predefined route that included flat ground/ multiple turns and opening doors. The results of this paper show that smartwatch motion sensors (i.e., accelerometer and gyroscope) can be effectively used in a Transparent Authentication System and future work needs to focus upon developing appropriate classification strategies to maximise performance. The study also shows some good results using the more realistic Cross-day scenario by utilizing small feature subsets. Unlike most of the previous smartphone based activity recognition systems, the proposed feature selection method utilized a dynamic feature vector for each user in order to have a trade-off between FAR and FRR. This feature reduction will help to decrease the computation burden of creating the test template on smartphones and/or smartwatches. However, more experimental work is required to evaluate whether the selected features of this study are the most effective feature sets. This can be carried out by using advanced techniques (e.g., decision-tree based classifiers and neural networks).

Unlike most existing motion-based authentication studies implemented within a controlled environment (i.e., all participants were asked to perform specific activities in an indoor environment), future work will also aim to design a methodology in order to collect real life data (i.e., users would wear a smartwatch during their day-to-day activities). By collecting unconstrained data a richer user profile can be generated. This could be extended to include interacting and typing on the smartphone touch screen and collecting different walking paces. As the nature of the real life signals is likely to be noisy, data from other smartwatch sensors (e.g., GPS) could be used in order to develop a context-aware approach (which will be useful to predict the user's activity).

7. Acknowledgment

The authors would like to thank the test subjects who participated in this study. I would also like to express my sincere gratitude to the University of Kufa for their financial support of this research.

8. References

[1]THE RADICATI GROUP, INC., (2015). Mobile Statistics Report, 2015-2019. [online] LONDON, UK, pp.2-3. Available at: http://www.radicati.com/wp/wp-content/uplo ads/2015/02/Mobile-Statistics-Report-2015-2019-Executive-Summary.pdf [Accessed 30 Jan. 2017].

[2] Lifestylegroup, "Data stored on a phone more precious than the phone itself," 2011 .[Online]. Available: http://www.lifestylegro up.co.uk /content/Data-stored-on-a-phone-more-precious-than-the-phone-itself.html. [Accessed: 27-Jan-2017].

[3] N. L. Clarke and S. M. Furnell, "Advanced user authentication for mobile devices", Computer and Security, vol. 26, no. 2, pp. 109–119, 2007.

[4] M. Hamblen, "Mobile phone security no-brainer: Use a device passcode", 2013. [Online]. Available: http://www. computerworld.com/article/2497183/mobile-security/ mobile-phone-security-no-brainer--use-a-device-passcode. html. [Accessed:25-Jan-2017].

[5] N. Clarke, Transparent user authentication: biometrics, RFID and behavioural profiling. Springer London, ISBN: 978-0-85729-805-8, 2011.

[6] H. Saevanee, N. L. Clarke, and S. M. Furnell, "Multi-modal Behavioural Biometric Authentication for Mobile Devices", in IFIP Advances in Information and Communication Technology, vol. 376 AICT, 2012, pp. 465–474.

[7] PWC, (2016). The Wearable Future. [online] United States, pp.4-6. Available at: https://www.pwc.com/us/en/technology /publications/ assets/pwc-wearable-tech-design-oct-8th.pdf [Accessed 30 Jan. 2017].

[8] L. Rong, D. Zhiguo, Z. Jianzhong, and L. Ming, "Identification of Individual Walking Patterns Using Gait Acceleration", in 2007 1st International Conference on Bioinformatics and Biomedical Engineering, 2007, pp. 543–546.

[9] D. Gafurov, E. Snekkenes, and P. Bours, "Spoof Attacks on Gait Authentication System", IEEE Transactions on. Information. Forensics Security., vol. 2, no. 3, pp. 491–502, Sep. 2007.

[10] M. O. Derawi, P. Bours, and K. Holien, "Improved Cycle Detection for Accelerometer Based Gait Authentication", in 2010 Sixth International Conference on Intelligent Information Hiding and Multimedia Signal Processing, 2010, pp. 312–317.

[11] M. O. Derawi, C. Nickel, P. Bours, and C. Busch, "Unobtrusive User-Authentication on Mobile Phones Using Biometric Gait Recognition", in 2010 Sixth International Conference on Intelligent Information Hiding and Multimedia Signal Processing, 2010, pp. 306–311.

[12] J. R. Kwapisz, G. M. Weiss, and S. A. Moore, "Cell phone-based biometric identification", in 2010 Fourth IEEE International Conference on Biometrics: Theory, Applications and Systems (BTAS), 2010, pp. 1–7.

[13] C. Nickel, H. Brandt, and C. Busch, "Benchmarking the performance of SVMs and HMMs for accelerometer-based biometric gait recognition", in 2011 IEEE International Symposium on Signal Processing and Information Technology (ISSPIT), 2011, pp. 281–286.

[14] C. Nickel, M. O. Derawi, P. Bours, and C. Busch, “Scenario test of accelerometer-based biometric gait recognition”, in 2011 Third International Workshop on Security and Communication Networks (IWSCN), 2011, pp. 15–21.

[15]C. Nickel and C. Busch, “Classifying accelerometer data via Hidden Markov Models to authenticate people by the way they walk”, in 2011 Carnahan Conference on Security Technology, 2011, vol. 28, no. 10, pp. 1–6.

[16] M. Muaaz and R. Mayrhofer, “An Analysis of Different Approaches to Gait Recognition Using Cell Phone Based Accelerometers”, in Proceedings of International Conference on Advances in Mobile Computing & Multimedia - MoMM ’13, 2013, pp. 293–300.

[17] T. Hoang, T. Nguyen, C. Luong, S. Do, and D. Choi, “Adaptive Cross-Device Gait Recognition Using a Mobile Accelerometer”, Jornal of Information. Processing System., vol. 9, no. 2, pp. 333–348, Jun. 2013.

[18] A. H. Johnston and G. M. Weiss, “Smartwatch-based biometric gait recognition,” in 2015 IEEE 7th International Conference on Biometrics Theory, Applications and Systems (BTAS), 2015, pp. 1–6.

[19] G. M. Weiss, J. L. Timko, C. M. Gallagher, K. Yoneda, and A. J. Schreiber, “Smartwatch-based activity recognition: A machine learning approach”, in 2016 IEEE-EMBS International Conference on Biomedical and Health Informatics (BHI), 2016, pp. 426–429.

[20] N. Al-Naffakh, N. Clarke, P. Dowland, and F. Li. ACTIVITY RECOGNITION USING WEARABLE COMPUTING. Barcelona, Spain: ICITST-2016, pp.189-195. Available at: http://icitst.org/ICITST-2016-Proceedings /ICITST-2016-Proceedings.pdf.