3. The proposed approach for user action authentication for mobile applications

A framework based on Clarke and Furnell (2007) was used to address the above concept, as shown in Figure 1. The proposed framework consists of a number of key components, including a Data Collection Engine, a Biometric Profile Engine, and an Authentication Engine. These engines perform various tasks, such as collecting biometric data, generating user profiles, and verifying the user's identity, respectively. There are two main system components. The first is the Authentication Manager, which controls the three engines referred to previously, sets the confidence level, observes the current security level and makes authentication decisions if the user requests access to a service within the application (intra-process). The Authentication Manager achieves this by comparing the risk level value for this intra process, which is retrieved from the Risk Database, with the confidence level value, which is calculated by the Authentication Engine. If the process risk value exceeds the threshold (confidence level), the user will be allowed access. However, if the process risk value is less than the threshold, the user will be denied access to the service. The second main system component is the Intra-Process Determination System, which observes the user's action on a specific application. This value is passed to the Authentication Manager to compare with the risk value for the process (a predefined value). The risk value is based on this new component. The novel elements are the ability to determine and identify the current user action on the application (intra-process), which is the key task of the Intra-Process Determination System. The outputs from this component are the application name and the intra-process name within this application, both of which are sent to the Authentication Manager in order to decide the legitimacy of the user to accomplish the action or not. This research work is an extension of prior published work [17].In this work, the research focus on user action only within each application without counting the application access.

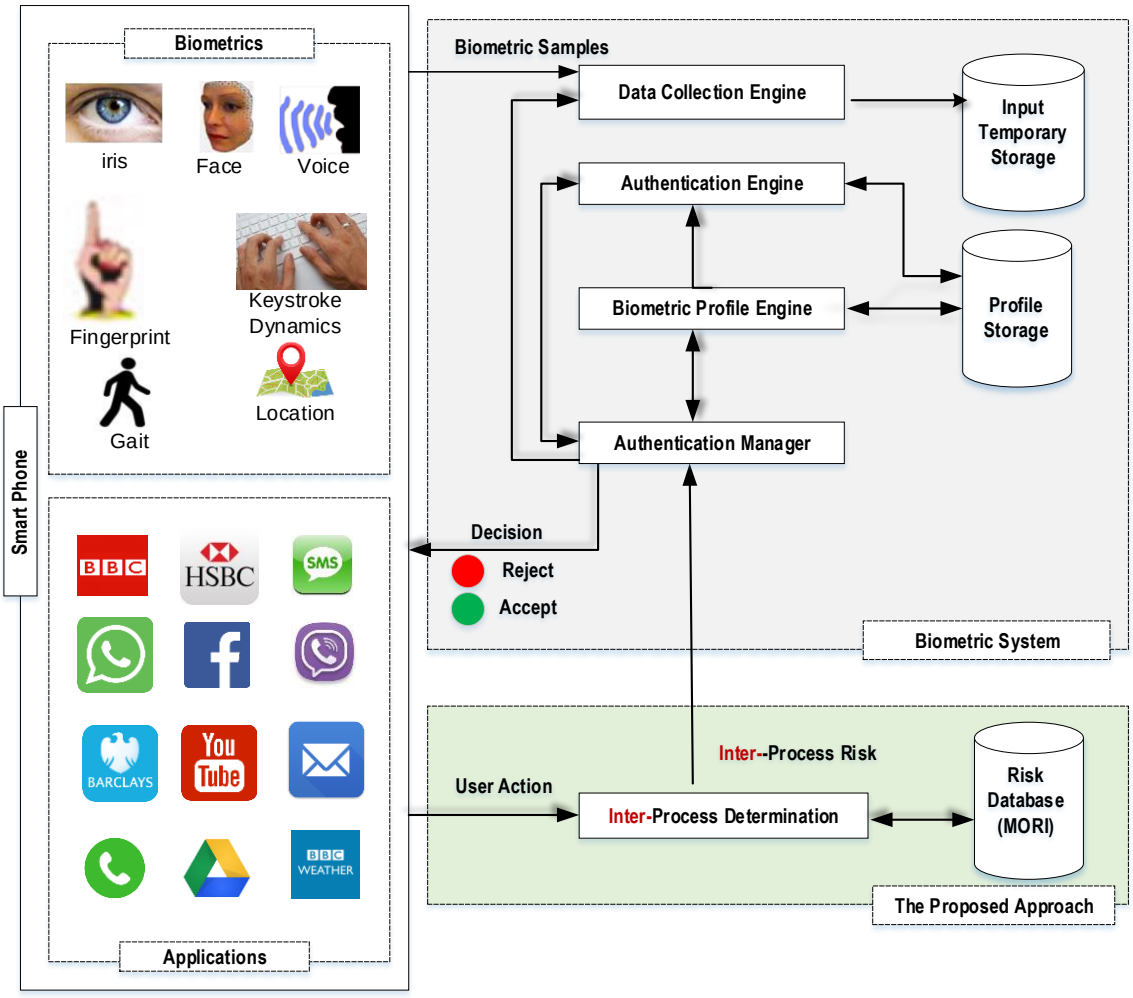

Figure 1. Framework for mobile application security (adapted from Clarke and Furnell, 2007)

It is first necessary to collect samples of genuine user interactions with their mobile devices/apps based upon a substantive period of real-world use [17]. At the end of the collection process, data had been collected from 76 users and they were ready for the analysis phase. Each user's data were stored in an individual text file and each record contained the following fields: a date in two formats (human time and a timestamp e.g., 2016-06-28 20:22:30, 1467141750071), application name, action type, extra information, such as message/email length, and call duration.

The proposed approach was based upon assessing intra-process (within the application) user interactions and testing the impact of an intra-process approach on the overall transparent user authentication for mobile applications by including application access with other actions within the application for 76 participants. Before starting the data analysis, a risk model (MORI) (Alotaibi et al., 2016a) was used to calculate the risk level for each action within each application. A wide range of biometrics were used in this research: facial, voice and iris recognition, keystrokes, behavioural and linguistic profiling, and fingerprint recognition, due to the ability of smartphones to capture multiple biometric modalities. Moreover, EERs published in prior studies in this domain were also used in this study [17].

In order to compute an identity confidence level based on the simulated biometric scenario, a weighted majority voting (WMV) formula [18] was utilized. In this approach, for each individual biometric technique, weights are assigned inversely proportionate to their EERs. More specifically, based on the WMV, a lower EER corresponds to a higher weighting than a high EER [18]. NICA was selected to analyse the data and compute the identity confidence level [3]. The NICA framework is designed to be a mobile-based solution by utilizing a combination of secret knowledge authentication and a number of biometric techniques in order to provide transparent and thus continuous authentication while the user interacts with the mobile device, despite an intrusive request at the beginning of the session [3], [18].In addition, the main aim of this framework is to observe the level of trust of the user in order to allow or restrict access to applications or services. Furthermore, based upon the biometric samples captured, the level of confidence fluctuates in a continuous manner [3], which has an effect on permissions to access applications. More specifically, if no biometric samples are captured to cause the confidence level to exceed the threshold value, the device will be locked. To provide effective security in a NICA system, two security mechanisms are considered imperative and define the core operation of the framework: Alert Level (AL) and Integrity Level (IL). These two levels are mapped to confidence levels to maintain security within the system, as well as its usability [3], [18]. NICA has a function that is defined as a degradation function, to decrease the value of the Integrity Level (-0.5) periodically every 30 minutes for frequent users and 50 minutes for infrequent ones, as defined by NICA (2007), when the device is inactive.

During a specific time window, the AL process seeks valid samples. If there are no samples, the identity confidence level will be periodically reduced by a degradation function that is 10% of current confidence in order to protect the mobile device the mobile while it remains inactive. In the case of the mobile user requesting to perform a task, the IL is applied to check the legitimacy of the mobile user. If the identity confidence is greater than or equal to the specified risk action level, transparent access is allowed. Otherwise, an intrusive authentication request is required in order to proceed with the service. In summary, each user file from the dataset was produced to generate different files. The first file was produced after applying the risk model [19] and the second after generating possible biometric samples and then computing the identity confidence value. Finally, the two files were compared and matched at a specific time. If the confidence level is more than the threshold (action risk level), the user can access the service (non-intrusive authentication request); otherwise, the mobile device is locked (intrusive authentication request).

This methodology was applied to each user file in order to compute the number of the intrusive authentication requests made during the intra-process (within the application) access to evaluate the average for the intrusive authentication requests for all 76 users. To do this, a number of scripts were generated and run with the participants' data for a combination of time windows: AL 2 min and IL 5 min; AL 5 min and IL 5 min; AL 5 min and IL 10 min; AL 10 min and IL 10 min; AL 20 min and IL 10 min; and AL 20 min and IL 20 min. The reason for changing the time window each time was to provide further insight into whether this would affect the number of intrusive authentication requests for each user.

4. Experimental Results and Discussion

In this study, the main aim is to compute the number of intrusive authentication requests (i.e. entering PIN or username and password): the higher the percentage of intrusive requests becomes, the less usable the system. It should be noted that there is no need to calculate biometric accuracy such as false positive and false negative as the biometric was simulated.

To provide further insight into whether applying a transparent authentication system at the action level would enhance security and usability, this experiment was applied to each user file to compute the average intrusive authentication requests for all 76 users. This second experiment differs from the first by focusing on user action access only (intra-process access) and not application access (inter-process access). To do this, after applying the risk model, the code was run with the participants' data to generate biometric samples (based on [18]) and then calculate the confidence level and intrusive authentication requests for each user for each user action by utilizing NICA across various ALs and ILs with the actions (within application only). The reason for trying different combinations of time windows was to investigate their effect on the system performance. As demonstrated in Figure 2, the distribution of user intrusive requests for 76 participants on an intra-process level based on minimum, median, and maximum values over the different time windows was considered. In this figure, and as mentioned in Table 5-10, the majority of user intrusive requests for the AL = 2 min / IL = 5 min time window were between 15% and 20% for 26 users. For instance, participant 46 had the highest intrusive requests at 33%, whereas participant 71 had 4% intrusive requests. It can be interpreted from these results that the total usage of these participants played a significant role. In this context, the total usage for participant 46 was 27,576 over 592 days, which, in turn, means one action per hour approximately. This low usage could have led to the poor performance and is likely to lead to a large number of intrusive requests.

On the other hand, the highest usage might be the cause of the fewest intrusive requests, such as participant 71 with a usage of 13,702 over 51 days, which, in turn, means three actions per hour approximately. In contrast, the vast majority of user intrusive requests for the AL = 20 min / IL = 20 min time window were less than 10% (73 participants) which was envisaged to be the case given the longer length of time to collect biometric samples or a longer time in which to recall the degradation function to reduce the user identity level. Another observation regarding this figure is that the result was mostly identical if there was no change in the AL value, such as AL = 5 min / IL = 5 min and AL = 5 min / IL = 10 min, which could suggest that AL is important.

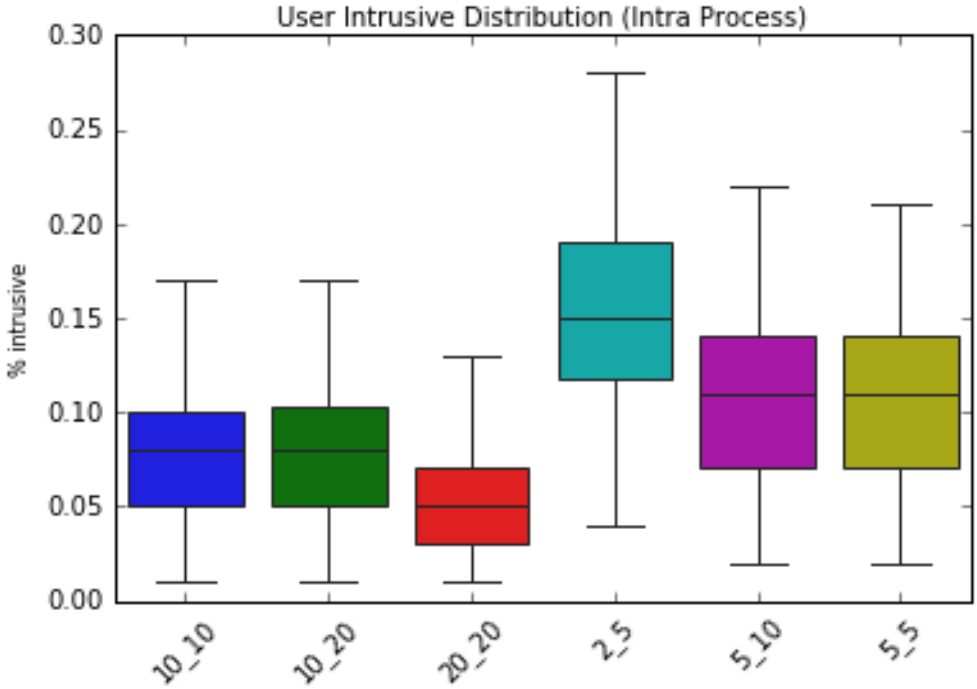

Figure 2. Average user intrusive requests distribution for intra-process access

As depicted in Table 1, the performance results for experiment across various ALs and ILs were promising for the intra-process level (actions within application only). The experimental results range from 15% average intrusive authentication requests at AL = 2 min / IL = 5 min to 5% at AL = 20 min / IL = 20 min for the same total of requests (2,561k). Accordingly, it is clear from Table 5-10 that the more substantial the AL and IL values, the fewer intrusive authentication requests. This is logical, as in cases in which the biometric samples were

insufficient or not available for capture, the user identity was reduced by the degradation function and resulted in a high FRR for the smaller time windows. For instance, the percentage of average intrusive authentication requests gradually reduced by approximately 50% for the AL = 10 min / IL = 10 min window to 7% from 15% for AL = 2 min / IL = 5 min. As a result, the shorter time windows could have the effect of raising the security level in relation to users' convenience, which was the opposite case for the larger time windows. The larger time windows might also lead to preserving a high level of identity confidence even though no biometric samples could be captured, which means there is an opportunity for misuse of the mobile device by an unauthorised user 2.

Table 1: Percentages of intrusive authentication requests for intra-process access

		Time Window					
		AL = 2	AL = 5	AL = 5	AL = 10	AL = 10	AL = 20
		IL = 5	IL = 5	IL = 10	IL = 10	IL = 20	IL = 20
Intra	% Average Intrusive Requests	15	10	11	7	8	5
	Total Requests	2,561 k					
	Intrusive ≤ 10% (# users)	20	37	37	58	57	73
	10% < Intrusive ≤ 15%	18	34	29	17	17	2
	15% < Intrusive ≤ 20%	26	4	6	1	2	1
	Intrusive > 20%	12	1	4	0	0	0

As previously mentioned, in the data collection stage, 47 actions were collected with the following distribution of risk types: 36% were high risk, 47% were medium risk, 13% were low risk, and 4% were no risk. One possible reason for the high percentage of intrusive authentication requests for some participants is that the majority of these actions are considered high and medium risk (83%), so the threshold (i.e., risk level) would require a greater confidence value to access the service.

In this context, Figures 3 and 4 show the intrusive/non-intrusive request results for the types of risk for the AL = 2 min / IL = 5 min and AL = 10 min / IL = 10 min time windows, respectively, for intra-process access. In both figures, the majority of intrusive requests come from high-risk actions, leading to an increase in the average intrusive authentication requests. Only 3% of the total requests come from medium-risk actions for the AL = 2 min / IL = 5 min time window.

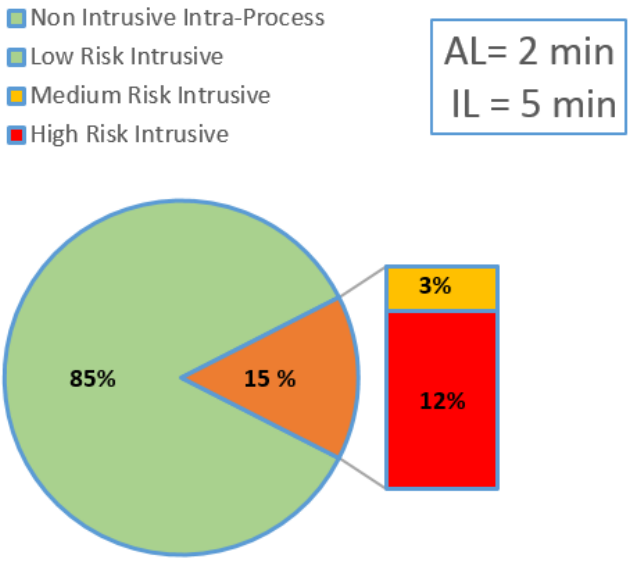

Figure 3. Intrusive/mon-intrusive request results for intra-process access at AL = 2 min / IL = 5 min

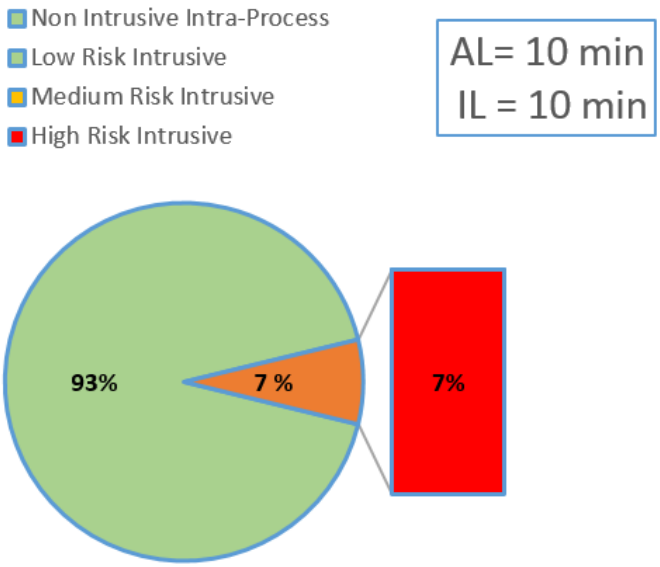

Figure 4. Intrusive/non-intrusive results for intra-process access at AL = 10 min / IL = 10 min

The experimental results clearly demonstrate that the proposed framework is able to provide a transparent authentication system for intra-process security. In addition, paying closer attention to the intrusive request results for different types of usage might lead to reducing the total average intrusive requests. For instance, participants 46, 71 and 57 received intrusive requests of 33%, 4% and 6%, respectively, for the shortest time window (AL = 2 min / IL = 5 min). To assess this, the 76 participants were categorized into three usage groups based on the user actions per hour. The primary aim of the participant categories was to gain greater insight into how low usage would affect the total average intrusive authentication requests for the entire dataset. The categorization was also aimed at testing whether all the time windows considered were reasonable and would tend to be more suitable for different types of users and thereby affect the intrusive authentication requests.

The experimental results for the 76 participants were categorized into three groups of usage (27 users had high usage, 24 users had medium usage, and 25 users had low usage). Accordingly, it can be seen that the results significantly improved following this classification and could lead to gradually reduced intrusive authentication requests. For instance, participants 36, 67, and 15 attained the highest average intrusive authentication requests at 18%, 17%, and 15%, respectively, for the shortest time window (AL = 2 min / IL = 5 min), whereas they

achieved 4%, 2%, and 3%, respectively, with the largest time window (AL = 20 min / IL = 20 min). A possible reason for this is that there is sufficient time to find and capture biometric samples, thereby raising the user identity level with enough time to reduce the confidence level (IL = 20 min).

For the same group of usage, however, participants 71, 4, and 60 obtained the lowest average intrusive authentication requests of 4%, 5%, and 5%, respectively, with the shortest time window (AL = 2 min / IL = 5 min). Similarly, they achieved 3%, 2%, and 3%, respectively, with the largest time window (AL = 20 min / IL = 20 min), which was expected to have fewer intrusive authentication requests. What can also be noticed in Table 2 is that the vast majority of participants achieved less than 10% intrusive authentication requests across all the different time windows (ranging from 15 participants at AL = 2 min / IL = 5 min to 27 participants at AL = 10 min / IL = 10 min).

Table 2. Average percentages of intrusive authentication requests for intra-process (usage)

		Time Window Intra-process					
		AL = 2	AL = 5	AL = 5	AL= 10	AL = 10	AL = 20
		IL = 5	IL = 5	IL= 10	IL = 10	IL = 20	IL = 20
High Usage	% Average Intrusive Requests	10	6	6	4	4	2
	Total Requests	1,772 k					
	Intrusive ≤ 10% (# users)	15	26	26	27	27	27
	10% < Intrusive ≤ 15%	9	1	1	0	0	0
	15% < Intrusive ≤ 20%	3	0	0	0	0	0
	Intrusive > 20%	0	0	0	0	0	0
Medium Usage	% Average Intrusive Requests	18	12	13	9	9	6
	Total Requests	396,640					
	Intrusive ≤ 10% (# users)	1	5	5	19	19	24
	10% < Intrusive ≤ 15%	4	17	14	5	5	0
	15% < Intrusive ≤ 20%	12	2	4	0	0	0
	Intrusive > 20%	7	0	1	0	0	0
Low Usage	% Average Intrusive Requests	18	13	13	10	10	8
	Total Requests	392,795					
	Intrusive ≤ 10% (# users)	4	7	7	13	13	23
	10% < Intrusive ≤ 15%	6	15	15	11	11	2
	15% < Intrusive ≤ 20%	9	2	2	1	1	0
	Intrusive > 20%	6	1	1	0	0	0

On the other hand, for the medium and low usage groups, a further interesting point to be noticed in these results is that the average intrusive authentication requests increased compared with the entire dataset for the same time windows (15% vs 18%). In addition, the vast majority of participants achieved around 15% intrusive authentication requests across the shorter time windows. For instance, at medium usage, participant 21 has the highest percentage of intrusive requests (25%) due to 21,880 actions being produced over 443 days, which means two actions per hour. In contrast, participant 65 has the lowest intrusive requests of 6%. These results support the conclusion that a short time window might mean the required service is protected by intrusive requests if no interaction is performed between the mobile user and his/her device and biometric samples are not available. Although the short time windows prompted a high degree of protection and intrusive authentication, this intrusiveness might lead to exaggerated re-authentication of the original user. As a result, short time windows appear to work well for security but are not quite sufficient for usability.

With regard to the low usage group results, approximately 56% of user intrusive requests were more than 15% for the shortest time window. For instance, participants 46 and 58 achieved 33% and

28%, respectively, which are the highest percentages of intrusive requests, whereas participant 44 achieved a much lower rate of intrusive requests (6%). In addition, the intrusive requests for this participant improved to 2% for the longest time window (AL = 20 min / IL = 20 min). One of the reasons for this could be that the degradation function was recalled very few times due to the AL taking a long time to collect biometric samples, thereby increasing the probability of raising the user identity level. Therefore, a larger time window can be considered to perform well with the majority of low user usage.

7. Discussion

In this research, two pervious works [17], [20] were compared with this study. In this study, only 11 applications were selected for consideration with a limited number of user actions, which would be highly likely to lose interactions, causing the loss of many biometric samples. In addition, 47 actions were collected and categorized as high risk (35%), medium risk (47%), low risk (13%) and no risk (4%). With this in mind, the majority of these actions were considered high and medium risk (83%), which, in turn, means identity confidence should be higher in order to exceed the threshold and access the required service. In addition, the experimental results showed that the majority of intrusive requests came from high-risk actions. Despite previous challenges, the experimental results for the intra-/inter-process and intra-process only for the 76 participants were promising across the various ALs and ILs considered, as demonstrated in Table 3, together with the worst and best performing time windows for each access level. It is clear from the table that the larger AL/IL time windows led to fewer intrusive authentication requests. The reason for the larger time windows outperforming the shorter time windows could be that a high number of user interactions with a mobile phone leads to the collection of many more biometric samples, thereby raising the identity confidence level.

Furthermore, this study highlights the clear effect of AL value on the average intrusive authentication. Likewise, the degradation function was significantly affected in terms of the total confidence level, as this automatically dropped. This is logical if there were no biometric samples collected or the quality of the modality was poor, especially with the shorter time windows. A further point to be noticed in these results is that the vast majority of intrusive requests came from high-risk actions and very few from medium-risk actions, while there was full transparency for low-risk actions. With regard to the system's robustness and users' convenience, a short time window is likely to lead to a large percentage of intrusive authentication requests, which could become a problem, thereby disturbing legitimate mobile users. As a result, short time windows would lower the security of the system, which might, in turn, allow an imposter to access a service.

Table 3: Average percentages of intrusive authentication requests

	Intra + Inter [17]	**Intra**	**Inter [20]**
Total Requests	3,006 k	2,561 k	1,364 k
Least Effective Time Window	AL = 2 min / IL = 5 min	AL = 2 min / IL = 5 min	AL = 2 min / IL = 5 min
% Intrusive Requests	18	15	27
Most Effective Time Window	AL = 20 min / IL = 20 min	AL = 20 min / IL = 20 min	AL = 20 min / IL = 20 min
% Intrusive	6	5	13

To consider this in more detail, further investigation was undertaken in order to explore how low usage would affect the total percentage of users' intrusive authentication requests. This was achieved by classifying the 76 participants into different types of users to gain greater insight into optimising the performance results and determining whether a particular grouping of time windows would perform better with a particular type of usage. Classifying participants into three groups of usage indicated a notable improvement and achieved promising experimental results with regard to intrusive authentication requests compared with those previously reported in the first experiment for all differing AL/IL timings, from the shortest time window (AL = 2 min / IL = 5 min) to the longest time window (AL = 20 min / IL = 20 min). The results for the three usage groups underline the evidence for the effect of low user usage on the total average intrusive authentication requests for the time window selected. One possible reason for this could

be that there is a suitable time window for each group of usage and, therefore, a high probability of gathering biometric samples when the user interacts with his/her mobile device and the degradation function is not recalled to reduce the identity confidence level when the device is inactive for very short intervals.

To conclude, the experimental results highlight that the proposed approach achieved a desirable level in terms of applying a transparent authentication system to intra-process security. As a result, this system would, in turn, enable control of the overall authentication process, thereby enabling a continuous and non-intrusive authentication approach.

Table 1. Average percentages of intrusive authentication requests by usage type

Usage Type	Comparison	Intra + Inter[17]	Intra	Inter [20]
High	Total Requests	2,045 k	1,772 k	833,679
	Least Effective Time Window	AL = 2 min / IL = 5 min	AL = 2 min / IL = 5 min	AL = 2 min / IL = 5 min
	% Intrusive	12	10	22
	Most Effective Time Window	AL = 20 min / IL = 20 min	AL = 20 min / IL = 20 min	AL = 20 min / IL = 20 min
	% Intrusive	3	2	12
Medium	Total Requests	464,869	396,640	260,468
	Least Effective Time Window	AL = 2 min / IL = 5 min	AL= 2 min / IL = 5 min	AL = 2 min / IL = 5 min
	% Intrusive	21	18	29
	Most Effective Time Window	AL = 20 min / IL = 20 min	AL = 20 min / IL = 20 min	AL = 20 min / IL = 20 min
	% Intrusive	7	6	12
Low	Total Requests	496,096	392,795	270,532
	Least Effective Time Window	AL = 2 min / IL = 5 min	AL = 2 min / IL = 5 min	AL = 2 min / IL = 5 min
	% Intrusive	22	18	31
	Most Effective Time Window	AL = 20 min / IL = 20 min	AL = 20 min / IL = 20 min	AL = 20 min / IL = 20 min
	% Intrusive	9	8	15

8. Conclusions

This paper presented and evaluated a novel framework for transparent user authentication for mobile applications. An experiment was devised to explore the intra-process (within the application) access levels across different time windows. In summary, the experimental results demonstrate that this approach achieved results that would fulfil security obligations and a desirable level of results for applying a transparent user action authentication level. The shortest time window (AL=2/IL=5 min) produced an average of 15% intrusive authentication requests, whereas the largest time window (AL=20/IL=20 min) generated 5%.

9. References

[1] K. Tam, S. J. Khan, A. Fattori, and L. Cavallaro, "CopperDroid: Automatic Reconstruction of Android Malware Behaviors," in *Proceedings 2015 Network and Distributed System Security Symposium*, 2015.

[2] V. M. Patel, R. Chellappa, D. Chandra, and B. Barbello, "Continuous User Authentication on Mobile Devices: Recent progress and remaining challenges," *IEEE Signal Process. Mag.*, vol. 33, no. 4, pp. 49–61, Jul. 2016.

[3] N. Clarke, S. Karatzouni, and S. Furnell, "Flexible and transparent user authentication for mobile devices," in *IFIP Advances in*

Information and Communication Technology, 2009, vol. 297, pp. 1–12.

[4] T. Ledermüller and N. L. Clarke, "Risk assessment for mobile devices," 2011.

[5] Y. H. Chuang, N. W. Lo, C. Y. Yang, and S. W. Tang, "A lightweight continuous authentication protocol for the Internet of Things," *Sensors (Switzerland)*, vol. 18, no. 4, Apr. 2018.

[6] N. Clarke, *Transparent User Authentication: Biometrics, RFID and Behavioural Profiling*. London: Springer Science & Business Media, 2011.

[7] J. Hatin, E. Cherrier, J.-J. Schwartzmann, and C. Rosenberger, "Privacy Preserving Transparent Mobile Authentication," in *Proceedings of the 3rd International Conference on Information Systems Security and Privacy*, 2017, pp. 354–361.

[8] H. Khan and U. Hengartner, "Towards application-centric implicit authentication on smartphones," in *Proceedings of the 15th Workshop on Mobile Computing Systems and Applications - HotMobile '14*, 2014, pp. 1–6.

[9] D. Yousefpor, M., Bussat, J. M., Lyon, B. B., Gozzini, G., Hotelling, S. P., and Setlak, "Fingerprint Sensor in an Electronic Device.U.S. Patent Application 14/451,076," *US Pat. App. 14/ 451,076*, 2014.

[10] P. Tresadern *et al.*, "Mobile biometrics: Combined face and voice verification for a mobile platform," *IEEE Pervasive Comput.*, vol. 12, no. 1, pp. 79–87, Jan. 2013.

[11] T. Feng *et al.*, "Continuous mobile authentication using touchscreen gestures," in *2012 IEEE International Conference on Technologies for Homeland Security, HST 2012*, 2012, pp. 451–456.

[12] P. Koundinya, S. Theril, T. Feng, V. Prakash, J. Bao, and W. Shi, "Multi resolution touch panel with built-in fingerprint sensing support," in *Design, Automation & Test in Europe Conference & Exhibition (DATE), 2014*, 2015, pp. 1–6.

[13] S. Alotaibi, S. Furnell, and N. Clarke, "Transparent authentication systems for mobile device security: A review," in *2015 10th International Conference for Internet Technology and Secured Transactions (ICITST)*, 2015, pp. 406–413.

[14] E. Hayashi, O. Riva, K. Strauss, A. J. B. Brush, and S. Schechter, "Goldilocks and the Two Mobile Devices: Going Beyond All-Or-Nothing Access to a Device's Applications," in *Proceedings of the Eighth Symposium on Usable Privacy and Security*, 2012.

[15] O. Riva, C. Qin, and K. Strauss, "Progressive authentication: deciding when to authenticate on mobile phones," in *Proceedings of the 21 st ...*, 2011, pp. 1–16.

[16] F. Li, N. Clarke, M. Papadaki, and P. Dowland, "Misuse Detection for Mobile Devices Using Behaviour Profiling," *Int. J. Cyber Warf. Terror.*, vol. 1, no. 1, pp. 41–53, Jan. 2011.

[17] S. N. Alotaibi, S. Furnell, and N. Clarke, "A novel transparent user authentication approach for mobile applications," *Inf. Secur. J.*, vol. 27, no. 5–6, pp. 292–305, Nov. 2018.

[18] A. Abdulwahid and A. Abdullah, "Federated Authentication using the Cloud (Cloud Aura)," Plymouth University, 2017.

[19] S. Alotaibi, "A Novel Taxonomy for Mobile Applications Data," 2016.

[20] S. N. Alotaibi, S. Furnell, and N. Clarke, "Transparent and Continuous Identity Verification for Mobile Applications Security," *In Proceedings of the Annual Information Institute Conference, Eds. G. Dhillon and S. Samonas, April 29 – May 1, 2019. Las Vegas, NV. USA. ISBN: 978-1-935160-20-5*