Keystroke Template Update with Adapted Thresholds

Abir Mhenni, Christophe Rosenberger, Estelle Cherrier, Najoua Essoukri Ben Amara

Keystroke Template Update with Adapted Thresholds

Abir Mhenni
SAGE Research Unit
GREYC Research Lab
National Engineering School of Tunis
University of Tunis El Manar
BP 94, Rommana 1068 Tunis, Tunisia
Email: abirmhenni@gmail.com

Christophe Rosenberger
and Estelle Cherrier
Normandie Univ, UNICAEN,
ENSICAEN, CNRS, GREYC,
14000 Caen, France
Email:estelle.cherrier@ensicaen.fr
Email:christophe.rosenberger@ensicaen.fr

Najoua Essoukri Ben Amara
SAGE Research Unit
National Engineering
School of Sousse
University of Sousse
BP 526, 4002 Sousse, Tunisia
Email:najoua.benamara@eniso.rnu.tn

***Abstract*—The use of biometrics for authentication mechanisms is becoming more and more important for research as well as industry. Keystroke dynamics is a biometric authentication method that improves the security of password-based applications. The performance of biometric keystroke recognition is still an open research issue. In fact, the extracted features relevant to a personal way of typing become less representative over time. This can lead to a failure in the biometric verification task. Because of the changes of such features, the representative model has always to be updated. In this paper, we use growing and sliding windows as template update methods based on a statistical classifier. We also demonstrate that user-specific thresholds, varying from an update session to another, allows to reduce the error rates compared to the update with a fixed threshold.**

Keywords*—Keystroke dynamics, template update, variable thresholds.

I. Introduction

Keystroke dynamics is a behavioral biometric modality that authenticates individuals according to their way of typing on a keyboard. Therefore, it is interesting to enhance the security of logical access control based on the use of login and password. Such a security system is non expensive because it requires just a keyboard. It is also easily accepted by the users as they are accustomed to use passwords for authentication. For example, the creation of a customer account is requested by the majority of web applications (social networks, online banking, e-commerce, email ...). Unfortunately, the attacks to impersonate the account holder or to access other personal data has gone up. According to TeleSign, in 2014, personal information comprising accounts or passwords of 2 out of 5 people have been hacked. Moreover, 8 out of 10 people are worried about their online security and 7 out of 10 people no longer trust passwords to protect their online accounts. According to the same report, 68% of people want companies to provide an extra layer of internet security. Keystroke dynamics can offer a solution to the increasing security needs.

Preliminary research in the domain of biometric keystroke dynamics dates back to 1980, based on the Rand Corporations report [1]. The first related techniques emerged in this preliminary study. It demonstrated, that the latencies are good features to represent users keystroke dynamics. Latency correspond to the time between the pressure and release of successive keys. After that, various studies have been conducted to improve the performances of keystroke dynamics recognition approaches that are generally divided into two phases: an enrollment phase, where the biometric reference is computed and a verification one which consists in comparing a novel capture to the reference. Bleha et al. [2] demonstrated that the longer the password is, the lower identification error is. Also, the more numbers of typing samples are used for training, the smaller identification error is obtained.

A successful utilization of keystroke dynamics information for personal recognition requires to follow the changes on the typing manner over time. It is therefore usual that the behavioral biometric reference, created at a given time, becomes less representative or even obsolete as time elapses [3]. As a solution, a user can be re-enrolled when a system malfunction is noticed. Thus, we can reinsert novel captures of the user's dynamics as model. This operation reduces the impact of poor representativeness of the biometric reference, but it can be costly in terms of time and money. Moreover, adding auxiliary information like soft biometrics [4] or other biometric traits such as face information [5] can improve the performance of these biometric systems. However, it can complicate the system functioning and degrade user's experience. Another interesting solution to cope with keystroke reference aging is the template update method. It consists in adapting the biometric reference automatically while using the system. Thereby, the variation of the user's way of typing is taken into consideration progressively.

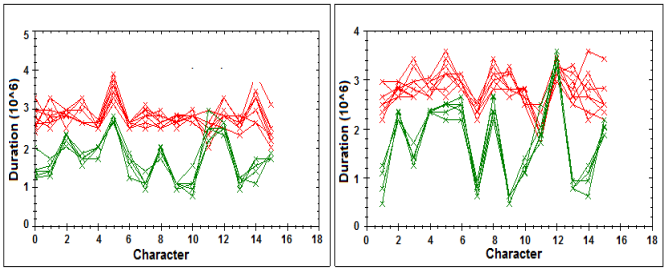
Considering the existing works about keystroke dynamics, the biometric systems update always consists in adapting the reference according to the variation of the person's typing rhythm. Nevertheless, the update decision threshold remains unchanged for all sessions. In this paper, we propose to adapt both the reference model and the threshold from one session to another. In our work, we capture more variable characteristics over time, so we get a better performance.

The remainder of this paper is structured as follows. Section II presents the keystroke template update. Section III

details the proposed approach. Section IV shows the obtained experimental results. Finally, the conclusions are presented in section V.

II. KEYSTROKE TEMPLATES UPDATE

Various biometric modalities suffer from aging problems (mainly face and fingerprint modalities). However, behavioral modalities are more relevant. For keystroke modality, the problem of template limited time has several causes. The most important one is the intra-class variability. As illustrated in Figure 1, one month later (green curves) the keystroke dynamics of the user is quite different from the initial one (red curves). This variability may be due to the acquired habit of typing the password of the password. In fact, through its frequent use, user's way of typing changes and causes the drift from the initial reference. In addition, emotions and state of mind have an impact on keystroke dynamics (stress, anger, happiness, beginning of the day, end of the day ...). Furthermore, changing the keyboard (azerty/qwerty) leads to altering the interactions with it.



(a) Time between two keys pressure

(b) Time between two keys release

(c) Time between one release and one pressure

(d) Time between one pressure and one release

Fig. 1: Intra-class variability of a user after one month

Several studies focus on using a template update process to solve intra-class variability. Four parameters are generally considered while updating the reference [6] :

- the choice of update decision criteria,
- the update periodicity (real-time or delayed),
- the update mode (supervised or semi-supervised),
- the adopted strategy (the used algorithm)

Different approaches have been developed depending on the application and its purpose.

Giot et al. [7] used a composite reference which contained various sub-references. Each sub-reference was modified with a different template update methods. Three update strategies were used. Both parallel sliding and parallel growing were utilized where one biometric sub-reference was never updated, but the other one was updated, with the sliding window or the growing window respectively. The third one was double parallel, where one biometric reference was updated using the sliding window and the other one was updated utilizing the growing window. The sliding window consists in adding the new biometric data (the accepted request) to the user's gallery, while deleting older data. Therefore, the number of examples in the gallery remains fixed. Besides, the growing window involves inserting a new accepted example to the gallery of the user. The number of examples in the gallery continues to increase. Nevertheless, the authentication was done with a unique biometric authentication method based on distances. This contribution was inspired by the co-update methods [8].

Serwadda et al. [3] contributed to the update decision criteria. In this study, it was proven that temporal error distributions could be useful information about the biometric systems performance. Indeed, a continuous sequence of false rejections means that the users template was aged and needed an update mechanism to capture variations in the users' features. Instead, a sudden false rejection could indicate that the system was disturbed, due to a change in the acquisition conditions, for example a change in users' keyboard.

Pisani et al. [9] used the parallel growing update method applied to a composite reference. They improved a statistical classifier to make the reference length constant through time. This method was named "improved double parallel".

Generally, template update can be performed in three ways : (1) adapt the system parameters depending on the user (or the type of users) [10] or the quality of the capture [11]; (2) adapt the decision frontier overtime [12]; and (3) update the biometric reference of the user while using the system [13]. For keystroke authentication, most studies focused on the third type of template update. In fact, the update threshold is generally fixed, even if some papers deal with individual ones [14], namely the threshold is different for each user, while remains unchanged during all update sessions.

Our contribution consists in adapting both the reference and the decision threshold from one update session to another. It is detailed in the next session.

III. PROPOSED APPROACH

For this study, we use the statistical recognition method [15], which is chosen because it is quicker to calculate and gives a good performance compared to other methods [16]. For this method, the reference is represented by both the mean μ and the standard deviation σ of the training samples. The comparison score between the query V and the reference is computed by equation (1):

$$Score = 1 - \sum_{i=1}^{n} e^{-\frac{|V_i - \mu_i|}{\sigma_i}} \quad (1)$$

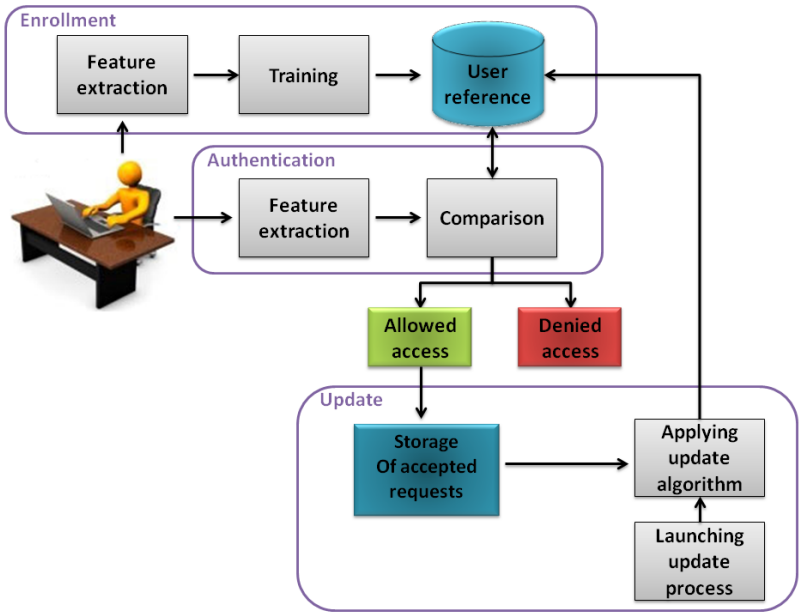


Fig. 2: Delayed template update process of keystroke recognition system

with :
V is the query,
μ is the mean of the training samples,
σ is the standard deviation of the training samples,
i = 1 : n where n is the length of the password.

Concerning the update, for each parameter of the proposed approach, we detail the chosen method. For the update decision criterion, we use the double thresholding method. The verification threshold first serves to accept or deny the new query. The update threshold is used to decide whether to use the query for the reference template update. This method avoids the inclusion of impostor samples in the biometric reference. Regarding the periodicity, the delayed method is used. For the mode, the semi-supervised approach is applied by the statistical classifier. Concerning the strategies, different methods have been adopted to modify the reference : selection, addition or replacement methods. The sliding window and the growing one are the most used ones. That is why they are adopted in the proposed approach. A scheme detailing the different steps of the proposed system for updating the keystroke recognition is illustrated in Figure 2.

The choice of the update threshold is very important. On one hand, a strict threshold does not manage intra-class variability. On the other hand, a very high threshold raises the possibility of including impostor information to the reference template. In the literature, the decision threshold is chosen using one of the following approaches: opting for the same threshold for all users; or using a specific threshold for each user [12]. It can be empirically or automatically defined, depending on the security level to reach.

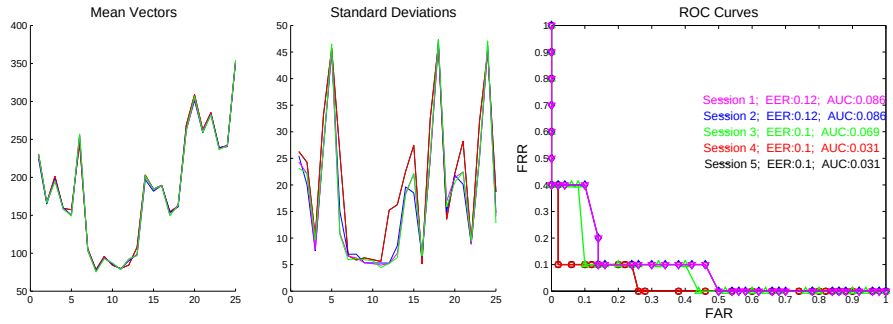
It is known that the measured system performance is different depending on the targeted choice [10] [17]. Moreover, it has been demonstrated that the individual threshold approach is more advantageous in terms of calculated error percentage [16]. Some studies [12] have analyzed the influence of age progression on the classifier thresholds and have proved that there is a conditional dependency between age progression and classifier scores. This motivates us to test if the classifier score distribution for the keystroke dynamics is dependent on the time progression for all individuals. The proposed approach consists in using a variable threshold. More precisely, we use an individual threshold that varies from one update session to another. In our test, it is demonstrated that by decreasing the score from one session to another, we can obtain a better performance. The threshold is modified according to equation (2):

$$T_{i+1} = T_i - e^{-\frac{\mu}{\sigma}} \tag{2}$$

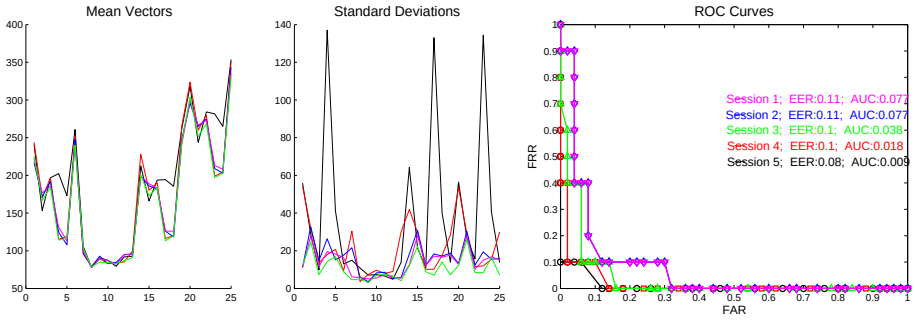
where μ is the average of the mean vector of the reference, σ is the standard deviation of the standard deviation vector and T_i is the threshold value for session i. Using this approach, it is demonstrated that more variability is captured in the user characteristics, hence getting a better performance. Figure 3 represents the vectors constituting the reference of the same user in each update session. It is clear that the mean and standard deviation vectors are quite different from session 1 to session 5. Moreover; the EER value and the ROC curves indicate that there are better performances with variable thresholds.

IV. Experiments and Results

Two different keystroke dynamics datasets are used in order to validate the proposed approach. The first one is the GREYC2009 [18], which contains data of a total of 133 users, captured during several one-week-spaced sessions. During each session, individuals were asked to type six times the same password. Some users did not participate in all sessions,



(a) User's characteristics variations in each update session using individual threasholds



(b) User's characteristics variations in each update session using variable threasholds

Fig. 3: Performances results (EER, AUC) and characteristics curves (mean and standard deviation) for 5 sessions. For the variable threshold case (b), characteristics are more dissimilar and performances are quite better than those of the fixed threshold case (a).

but 100 of them participated in five sessions and provided 60 patterns. These data are used in our experiments. The second database is the WEB-GREYC [19], which is the first public dataset where each user has its own password. 118 individuals participated in the creation of this database. Only 45 users participated in five sessions and provided 60 patterns.

We start by applying the sliding window algorithm. For each user, 10 captures are used for training, that is to calculate the mean and the standard deviation vectors. Initially, the threshold is set empirically. After that, five update sessions are conducted. For every update session, we use 10 captures of a genuine user and 10 others of different impostors. We make their recognition using the statistical keystroke distance computation function presented in equation (1). If a request is accepted, it is inserted into the reference of the user and the oldest data is deleted. The same process is repeated for all the database users. At the end of one update session, we calculate the system performances (ie, EER, ROC). Naturally, when the gallery is changed, it is necessary to recalculate the user model (ie, Mean, standard deviation). Before starting the new update session, we calculate the new threshold. It is obtained by changing the first threshold, according to equation (2). We first present the results obtained by applying this approach on the GREYC-Keystroke database. Figure 4 represents results validated with fixed, individual and variable thresholds. The basic scenario without template update is "None". The scenarios using a template update strategy are "Sliding" and "Growing".

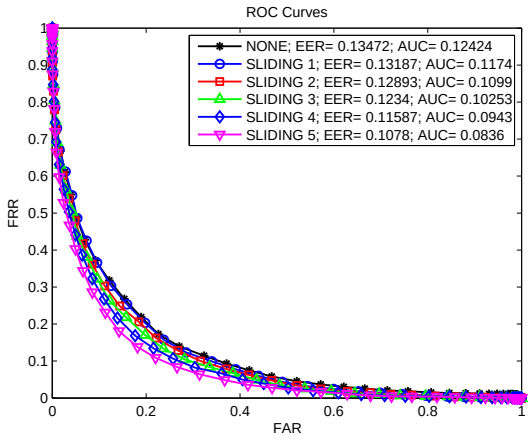
TABLE I: Experiment parameters

Parameter	**Value**
Modality	Keystroke dynamics
Authentication method	Statistical classifier
Update decision	Double threshold
Update mode	Semi-supervised
Update periodicity	Delayed
Update strategy	None, sliding window, growing window
Number of sessions	5 sessions
Respect to chronology	Yes
Enrollment samples	10 user's samples
Verification samples	10 user's samples, 10 impostor's samples
Evaluation metrics	EER , ROC

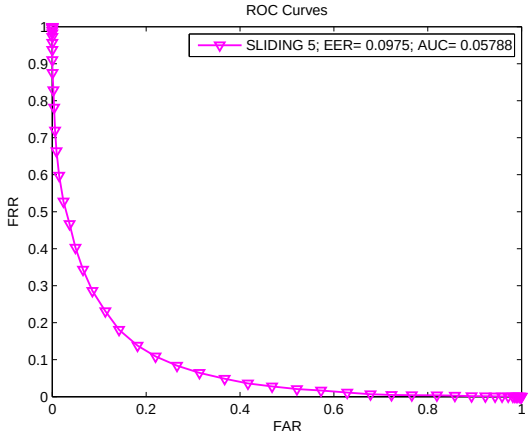
Comparing the obtained results to the existing work [16], it is noticed that the variable threshold gives a better performance compared with fixed and individual thresholds. The same process is repeated to the same database, but by using the growing window approach for the update strategy. Figure 5 illustrates the obtained results. For the Web-GREYC database, we have obtained the results shown in Figure 6.

Experimentally, taking into consideration the intra-class

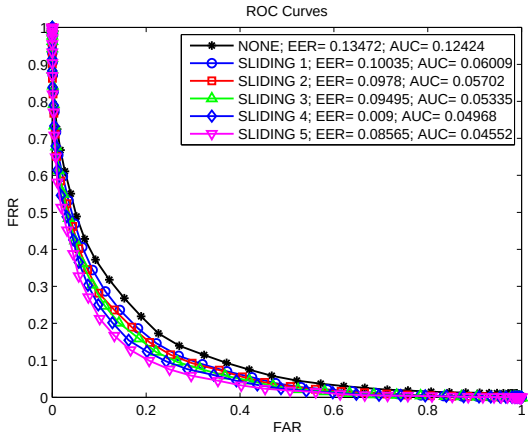
variation, the characteristics of the users change over time. Thus, by decreasing the threshold, we only pick the most



(a) Performances of update method with global threshold



(b) Performances of update method with individual threshold



(c) Performances of update method with variable thresholds

Fig. 4: Performances of sliding window update method applied on GREYC-Keystroke database.

TABLE II: Global results for all thresholds tested on GREYC-Keystroke database

Performances	**Previous work [16]**	**Our results**	
		Sliding window	**Growing window**
EER fixed threshold %	10.75	10.78	10.78
EER individual threshold %	9.22	9.75	9.92
EER variable threshold %	-	8.56	8.65
Gain %	1.53	2.22	2.13

similar characteristics to the reference. However, we do not suddenly reduce the threshold. Instead, we start with a high threshold (but which differentiates between samples of authentic user and those of impostors) in the first update session. Consequently, we introduce into the reference some new users' samples that are different. From one session to another, we slightly decrease the threshold so that we can capture less dissimilarity. Finally, samples similar to the modified reference (containing new samples added in last sessions) are captured. This is explained by the fact that by mastering the password, there is a noticeable stability in the typing manner.

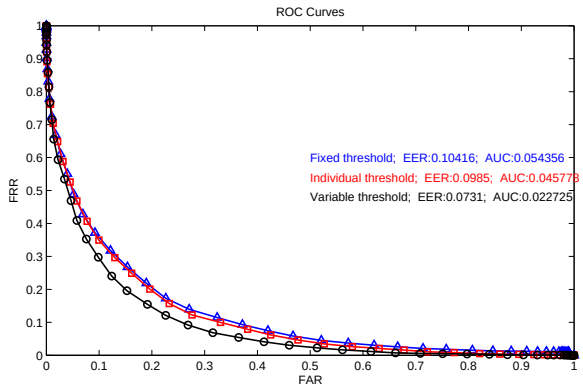


Fig. 5: Performances of growing window update with different thresholds

V. Conclusion

We have presented a template update method using variable thresholds. Varying the update thresholds from one session to another allows reducing the update error rates, so the performance gets better over time in comparison to using a fixed or individual threshold. The method has been validated on a template update system for keystroke dynamics on two datasets. We have shown that the proposed approach (based on sliding or growing windows) gives a better performance than the classical ones (EER 2% lower). Our implementation uses the delayed periodicity for the template update scenario; we are analyzing the online scenario too, to compare the performances. We are also interested to investigate the effect of the password length on the chosen thresholds.

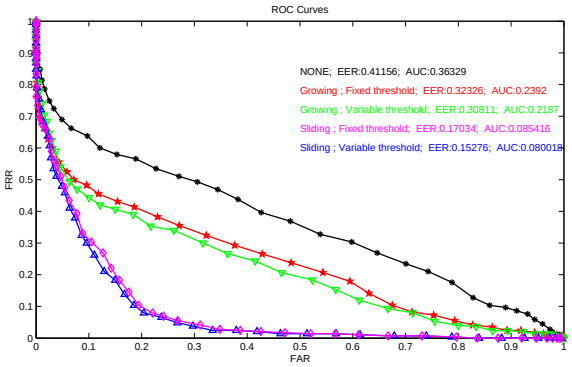


Fig. 6: Performances of growing and sliding window updates with different thresholds tested on Web-GREYC database

REFERENCES

[1] R. S. Gaines, W. Lisowski, S. J. Press, and N. Shapiro, "Authentication by keystroke timing: Some preliminary results," DTIC Document, Tech. Rep., 1980.

[2] S. Bleha, C. Slivinsky, and B. Hussien, "Computer-access security systems using keystroke dynamics," *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, vol. 12, no. 12, pp. 1217–1222, 1990.

[3] A. Serwadda, K. Balagani, Z. Wang, P. Koch, S. Govindarajan, R. Pokala, A. Goodkind, D.-G. Brizan, A. Rosenberg, and V. V. Phoha, "Scan-based evaluation of continuous keystroke authentication systems," *IT Professional, IEEE*, no. 4, pp. 20–23, 2013.

[4] S. Z. S. Idrus, E. Cherrier, C. Rosenberger, and P. Bours, "Soft biometrics for keystroke dynamics," in *Image analysis and recognition*. Springer, 2013, pp. 11–18.

[5] R. Giot, M. El-Abed, and C. Rosenberger, "Fast learning for multi-biometrics systems using genetic algorithms," in *High Performance Computing and Simulation (HPCS), 2010 International Conference on*. IEEE, 2010, pp. 266–273.

[6] A. Rattani, B. Freni, G. L. Marcialis, and F. Roli, "Template update methods in adaptive biometric systems: a critical review," in *Advances in Biometrics*. Springer, 2009, pp. 847–856.

[7] R. Giot, C. Rosenberger, and B. Dorizzi, "Hybrid template update system for unimodal biometric systems," in *Biometrics: Theory, Applications and Systems (BTAS), 2012 IEEE Fifth International Conference on*. IEEE, 2012, pp. 1–7.

[8] H. S. Bhatt, S. Bharadwaj, R. Singh, M. Vatsa, A. Noore, and A. Ross, "On co-training online biometric classifiers," in *Biometrics (IJCB), 2011 International Joint Conference on*. IEEE, 2011, pp. 1–7.

[9] P. H. Pisani, A. C. Lorena, and A. C. de Carvalho, "Adaptive approaches for keystroke dynamics," in *Neural Networks (IJCNN), 2015 International Joint Conference on*. IEEE, 2015, pp. 1–8.

[10] S. Hocquet, J.-Y. Ramel, and H. Cardot, "Estimation of user specific parameters in one-class problems," in *Pattern Recognition, 2006. ICPR 2006. 18th International Conference on*, vol. 4. IEEE, 2006, pp. 449–452.

[11] N. Poh, J. Kittler, S. Marcel, D. Matrouf, and J.-F. Bonastre, "Model and score adaptation for biometric systems: Coping with device interoperability and changing acquisition conditions," in *Pattern Recognition (ICPR), 2010 20th International Conference on*. IEEE, 2010, pp. 1229–1232.

[12] A. Drygajlo, W. Li, and K. Zhu, "Q-stack aging model for face verification," in *Signal Processing Conference, 2009 17th European*. IEEE, 2009, pp. 65–69.

[13] L. Didaci, G. L. Marcialis, and F. Roli, "Analysis of unsupervised template update in biometric recognition systems," *Pattern Recognition Letters, Elsevier*, vol. 37, pp. 151–160, 2014.

[14] M. Seeger and P. Bours, "How to comprehensively describe a biometric update mechanisms for keystroke dynamics," in *Security and Communication Networks (IWSCN), 2011 Third International Workshop on*, May 2011, pp. 59–65.

[15] S. Hocquet, J.-Y. Ramel, and H. Cardot, "User classification for keystroke dynamics authentication," in *Advances in biometrics*. Springer, 2007, pp. 531–539.

[16] R. Giot, "Contribution to keystroke dynamics: mulitbiometrics, soft biometrics and template update," Theses, Université de Caen, Oct. 2012. [Online]. Available: https://tel.archives-ouvertes.fr/tel-00748915

[17] D. Hosseinzadeh and S. Krishnan, "Gaussian mixture modeling of keystroke patterns for biometric applications," *Systems, Man, and Cybernetics, Part C: Applications and Reviews, IEEE Transactions on*, vol. 38, no. 6, pp. 816–826, 2008.

[18] R. Giot, M. El-Abed, and C. Rosenberger, "Greyc keystroke: a benchmark for keystroke dynamics biometric systems," in *Biometrics: Theory, Applications, and Systems, 2009. BTAS'09. IEEE 3rd International Conference on*. IEEE, 2009, pp. 1–6.

[19] R. Giot, M. El-Abed, and R. Christophe, "Web-based benchmark for keystroke dynamics biometric systems: A statistical analysis," in *Intelligent Information Hiding and Multimedia Signal Processing (IIH-MSP), 2012 Eighth International Conference on*. IEEE, 2012, pp. 11–15.