

Analysis of Template Update Strategies for Keystroke Dynamics

IEEE Workshop on Computational Intelligence in Biometrics and
Identity Management

Romain Giot* Bernadette Dorizzi† Christophe Rosenberger*

*GREYC - CNRS, ENSICAEN, Université de Caen

†Institut Télécom; Télécom SudParis

April 14, 2011

- 1 Introduction
- 2 Proposed Method
- 3 Results
- 4 Conclusion

Contents

- 1 Introduction
 - Keystroke Dynamics
 - Template Stabilisation
- 2 Proposed Method
- 3 Results
- 4 Conclusion

Behavioural modality

The aim is to recognize individuals on their way of using a keyboard

Two families

- Static: **fixed string**
- Dynamics: anything

Extracted features

- Duration of pressure of each key
- Interval between two keys



Recognition [de Magalhaes et al.2005]

Enrollment

- User provides several samples
- Compute the template with several samples
 - Multivariate modelisation of the distribution of the times
- Store the samples in the gallery
- Use the template to represent the user

Verification

- Claimant provides a query sample and a claimed identity
- Compute a distance with the query and the template
 - Verification if the query matches the model for most of its dimensions

Performance

- around 15% of EER on big datasets

Problem with keystroke dynamics

This modality is subject to an high variability

- It is a behavioral biometric, and way of typing change with time
- Various factors can greatly impact the typing
 - user happy
 - user stressed
 - user talking
 - user concentrated or not
- *We cannot track enough variability during enrolment*

Solutions

- Supervised template update works well
- Semi-supervised template has not been studied
 - *This is one aim of this paper*
- Capture more samples during enrollment

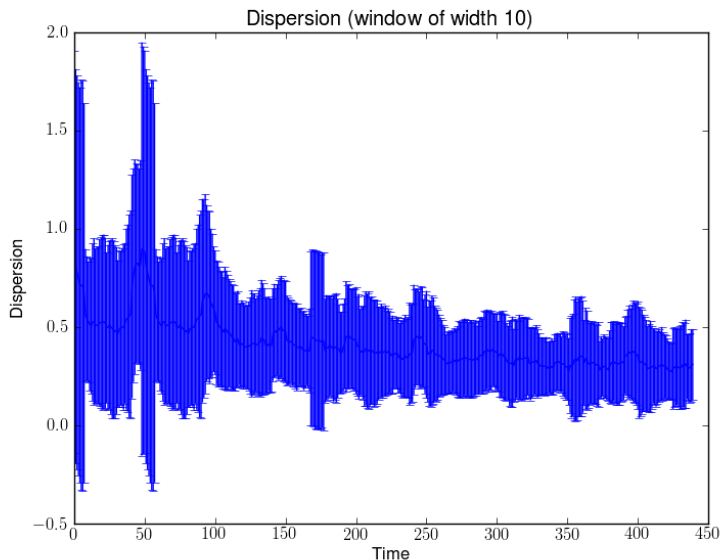
Template Behavior Through Time

To understand why supervised template works, we have computed the dispersion measure of the samples through time. Results are presented for a window of size 10.

Dispersion mesure

- We compute the mean vector of the samples in the window
- We compute the distance of each sample of the window against its mean vector
- We compute the mean value of these distances
- We slide the window of one simple (move in time)

Result



Interpretation

Dispersion's mean decreases with time

- ⇒ The dispersion value of each user decrease with time
- ⇒ Way of typing stabilize with time
 - Conclusion
 - As typing change regularly through time, until stabilisation, it is necessary to use template update schemes

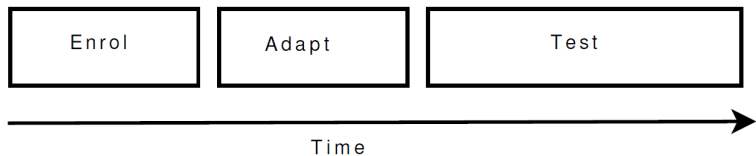
Dispersion's standard deviation decrease with time

- ⇒ Same trend for most users

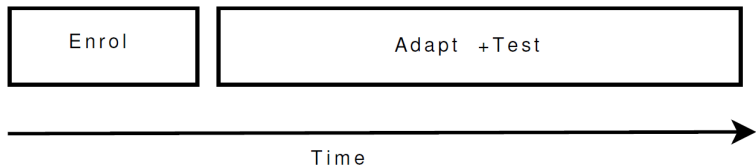
Contents

- 1 Introduction
- 2 Proposed Method
 - Évaluation
 - Dataset
 - Template Update
- 3 Results
- 4 Conclusion

Usually, datasets are separated in 2 or 3 parts [Poh et al.2009]



(a) Separate adapt-and-test strategy



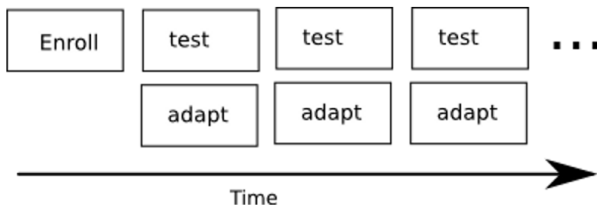
(b) Joint adapt-and-test strategy

Usually, datasets are separated in 2 or 3 parts [Poh et al.2009]

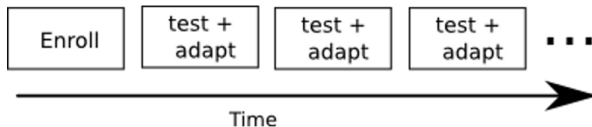
Observation

- Proposition interesting for main modalities
 - which have not a variability through time due to a cognitive learning
- But, it is inadequate for behavioral biometrics or datasets on several sessions
 - It does not take into account the various sessions
 - In the literature, the time chronology is almost never respected (shuffling of samples)
 - This point is very important for behavioral biometrics
- That is why we propose an adaptation of these two schemes to fit our requirements

Evolution of the scheme



(a) Separate adapt-and-test strategy



(b) Joint adapt-and-test strategy

Evaluation Procedure

- Template update is applied independently for each user
- We build a pool of queries for each user at each session
 - Samples are chronologically ordered
 - The frequency of impostor samples is chosen
 - Genuine or impostor samples are randomly chosen (based on the impostor frequency)

DSL2009 [Killourhy and Maxion2009]

- 51 users
- 8 sessions
- 50 samples per session \Rightarrow 400 samples per user
- 1 keyboard
- password: ".tie5Roanl"

Update Decision

Double thresholding

- **Acceptance threshold**
 - If distance lower to threshold, user is accepted
- **Update threshold**
 - If distance lower to threshold, new sample must be integrated in the template

Update Strategies

Growing window

- The new sample is added to the gallery
- The mathematical model is computed from scratch

Sliding window

- The oldest sample is removed from the gallery
- The mathematical model is computed from scratch

Fixed window

- The gallery is not updated
- The mathematical model does not change

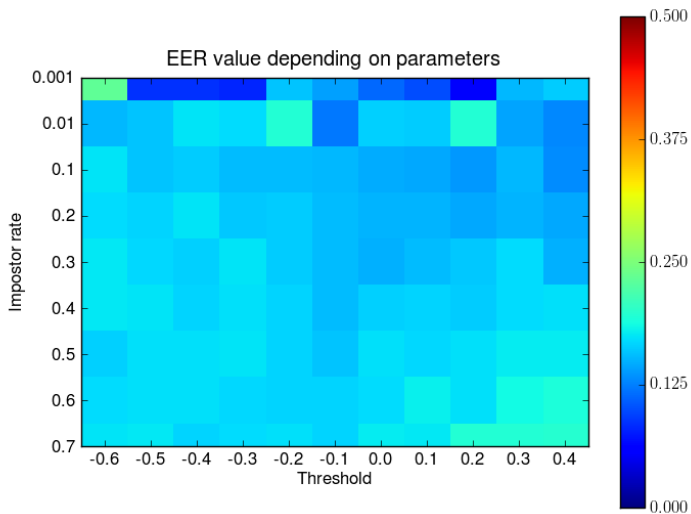
Contents

- 1 Introduction
- 2 Proposed Method
- 3 Results**
 - EER evolution through time
 - Performance evolution for one configuration
- 4 Conclusion

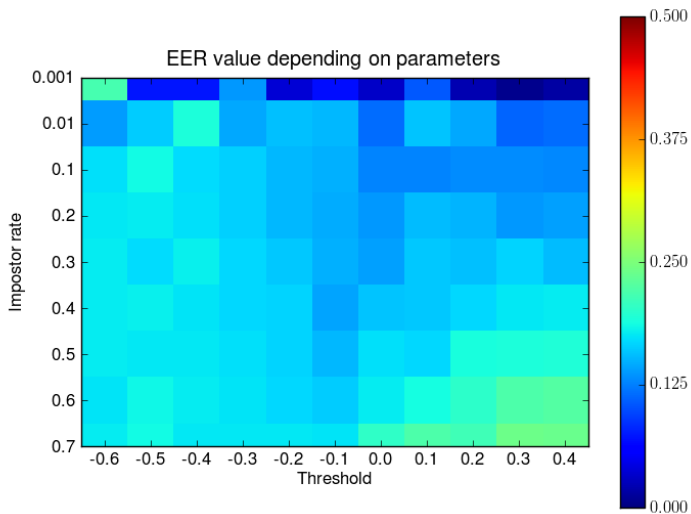
Configuration

- Variation of update threshold
- Variation of impostor rate
- Study of EER through sessions

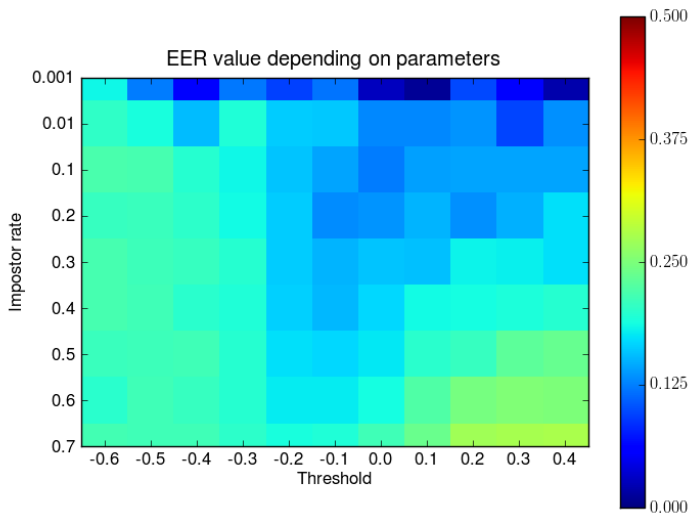
Session 1



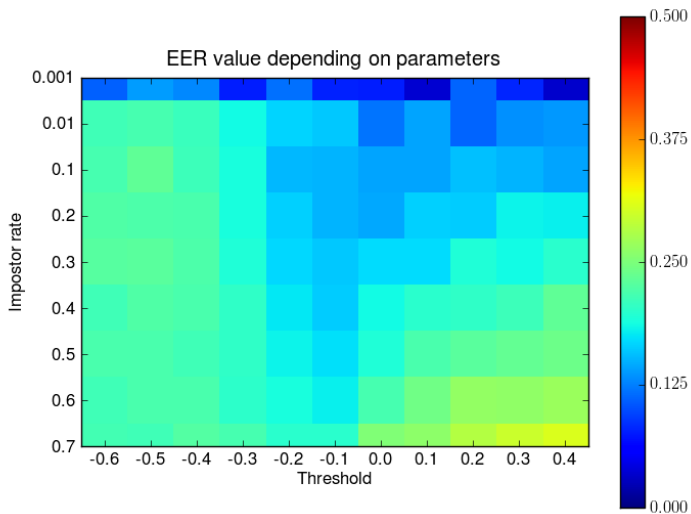
Session 2



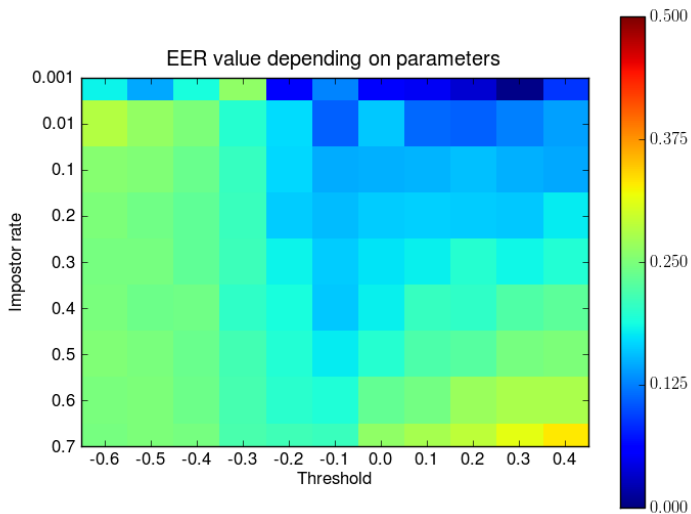
Session 3



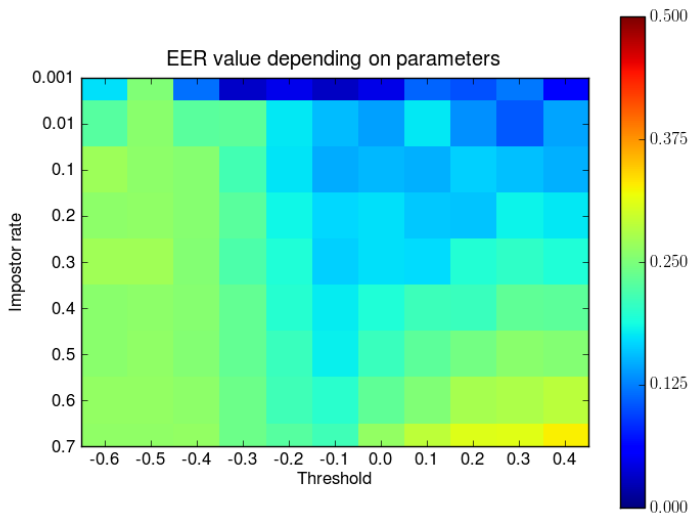
Session 4



Session 5



Session 6



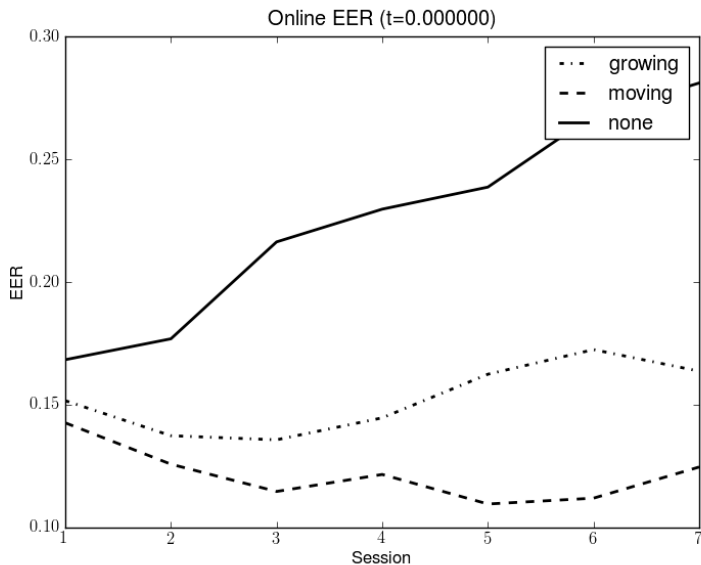
Interpretation

- The higher the percentage of attacks, the higher the EER increases (bottom of figures)
- A too high update threshold decreases the performance
 - ⇒ addition of impostors to the template
- A too low update threshold decreases the performance
 - ⇒ not enough template update done

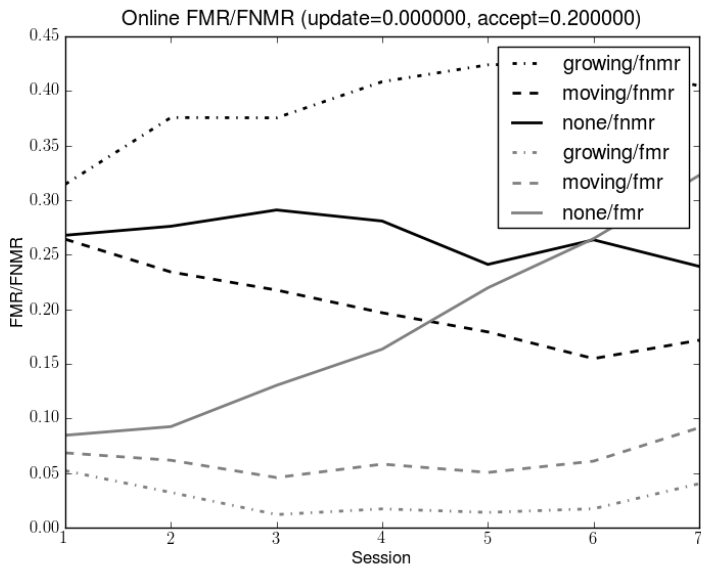
Interpretation of a particular configuration

- 20% of impostors
- Acceptation threshold at 0.2
- Update threshold at 0.0

EER



FAR/FRR



Interpretation

- False Reject Rate increases with growing window
 - False Rejection Rate decreases with sliding window
 - False Acceptance Rate increases with time when no update is done
- ⇒ These results depend on the configured thresholds

Contents

- 1 Introduction
- 2 Proposed Method
- 3 Results
- 4 Conclusion**
 - Discussion
 - Conclusion

Template update evaluation I

Which are the best metrics ?

- Error Equal Rate ?
 - It is hard (impossible ?) to set the right thresholds in order to obtain it
- False Acceptance Rate/False Rejection Rate ?
 - Which threshold configuration to use ?

How often to compute the Errors ?

- One time at the end of the evaluation ?
 - One time for each session ?
- ⇒ It seems that last presentation of the session will answer this question ...

Template update evaluation II

Which dataset to use ?

- Almost no dataset contains enough data on an enough period to really track variability though time

Conclusion

- We have shown the evolution of the keystroke biometric data through time
- We have demonstrated the possibility of using semi-supervised template update for keystroke dynamics
 - Littérature only treat
 - semi supervised techniques for morphological modalities
 - supervised techniques for keystroke dynamics
- We obtained interesting results
 - ⇒ Improvement of 50% with supervised techniques
 - ⇒ Improvement of 45% for semi supervised techniques




Thanks for your attention

Appendix

Contents

- 5 Bibliography
 - Bibliography

Bibliography

-  de Magalhaes, T., Revett, K., and Santos, H. 2005.
Password secured sites: stepping forward with keystroke dynamics.
In International Conference on Next Generation Web Services Practices.
-  Killourhy, K. S. and Maxion, R. A. 2009.
Comparing anomaly-detection algorithms for keystroke dynamics.
In 39th Annual International Conference on Dependable Systems and Networks (DSN-2009), pages 125–134.
-  Poh, N., Wrong, R., Kittler, J., and Roli, F. 2009.
Challenges and research directions for adaptive biometric recognition systems.
In Advances in Biometrics, pages 753–764.