A MULTI-EXPERT BASED APPROACH TO CONTINUOUS AUTHENTICATION OF MOBILE-DEVICE USERS

Oleksiy Mazhelis, Alexandr Seleznyov, Seppo Puuronen
Computer Science and Information Systems Department
University of Jyväskylä
P.O. Box35, FIN-40351, Jyväskylä, Finland
{*mazhelis, alexandr, sepi*} *@it.jyu.fi*

Abstract

Currently used in mobile devices PIN-based user authentication cannot provide a sufficient security level. Methods based on multi-modal user authentication involving biometrics (i.e. physical and behavioral characteristics of a person) may be employed to cope with this problem. However, dealing with physical characteristics only, these methods are either unable to provide continuous and user-friendly identity verification, or are resource consuming.

In this paper, we aim at the provision of continuous, user-friendly and accurate verification of the user identity while preserving scarce resources of a mobile device. Rather than physical, behavioral characteristics are analyzed. The normal behavior of the user is modeled by a set of complementary behavioral aspects that can be used to uniquely identify the user. We develop an approach, where these aspects are separately monitored by dedicated software-based experts. By analyzing the deviations of the current behavior from the modeled one, each expert infers separately its decision about the user identity. The final decision is derived from these multiple expert decisions by applying a decision fusion technique. The monitoring of multiple behavioral aspects helps to improve the authentication accuracy and enables the continuous verification of the user identity. The user-friendliness is supported by the use of transparent authentication methods that do not require direct user participation. Finally, the analysis of behavioral characteristics and the decision fusion process do not involve complicated computational steps and, therefore, are conservative in resource consumption.

Keywords: mobile device security, authentication, continuous identity verification, user profiling, expert decision fusion

---

The original version of this chapter was revised: The copyright line was incorrect. This has been corrected. The Erratum to this chapter is available at DOI: 10.1007/978-0-387-35696-9_19

E. Nardelli et al. (eds.), *Certification and Security in E-Services*

1. INTRODUCTION

Nowadays mobile devices have become a conventional element of our everyday life. Their computational power and functionality evolve constantly. More and more crucial personal and corporate data is kept on such devices. These devices (so called personal trusted devices, or PTDs) are often used for carrying out mobile e-transactions. Thus, more strict requirements emerge regarding security level provided [13].

There are four main security services: confidentiality, integrity, availability, and accountability. In order to guarantee these services a set of security mechanisms should be employed such as authentication, authorization, and audit. In this paper, we focus on the user authentication, whereby the device verifies that a person is the one who is eligible to use it. A number of methods can be used to identify a user. They can be based on i) something one knows (password, PIN, etc.), ii) something one possess (e.g. smart-card), and iii) something one is i.e. on biometric characteristics of the user.

In current mobile devices, the authentication is usually based on the user's knowledge and implemented by Personal Identification Numbers (PINs). It is assumed that the valid user is the only person who knows the PIN, and by showing the knowledge of the PIN the user identity can be proven. However, due to their small size the mobile devices can be easily lost or stolen. According to the estimation of F-Secure Corporation [4], more than ten mobile devices are lost or stolen in the world every minute. Should it happen, the PIN-based authentication might be unable to protect the device since the device can be "on" and unlocked, or the PIN can be compromised. Therefore, additional authentication mechanisms are needed to assist the PIN-based authentication in protecting the lost/stolen devices from the illegal use.

What requirements should an authentication mechanism meet? Evidently, it must guarantee the negligible probability of an impostor being granted the access to the device. Meanwhile, it should work continuously verifying that the user substitution has not occurred. Besides, the mechanism should be user-friendly and should not consume much computational resources. Unfortunately, the use of PINs fails to provide continuous and user-friendly authentication. As indicated by the survey of Clarke et al. [3], 41% of respondents do not consider the usage of PIN as user-friendly. It is not continuous either because the identity verification is performed only at some points in time.

A possible way to strengthen PIN/password-based authentication is the incorporation of biometrics into authentication process. As opposed to the PINs and passwords, the biometrics cannot be easily compromised or forgotten. Besides, some biometrics measures could provide continuous and user-friendly identity verification. All this makes them especially useful in user authentication.

Biometric authentication can be based on physical characteristics of a person (fingerprints, voice, face shape, etc.) and/or on behavioral rhythms (mobility, keystroke dynamics, activities, etc.). Authentication methods based on mobility monitoring construct a model of user's movements. Displacements that do not fit to the model signalize the substitution of the user. Keystroke based methods analyze characteristics of user's typing rhythms. Using similarity measures these characteristics are compared with patterns stored in a reference profile. Significant differences between the characteristics and the profile are assumed to indicate the user substitution. In a similar manner, activity monitoring involves building user behavior profiles that contain information about user actions (calls, applications run, etc.) including their temporal characteristics and ordering in time. By comparing current user actions with the profile it is possible to reveal anomalous actions, which are caused by an impostor's activity.

The problem is that biometric measures vary in time for the same person. Consequently, the authentication based on such biometric measures may result in a poor accuracy. One possible solution is to analyze multiple biometrics simultaneously. A number of studies was devoted to the problem of the user authentication based on multiple biometrics. This form of authentication is referred to as multi-modal user authentication. Most of these studies address the problem of the combination of visual and acoustic features for the identity verification purposes (see, for example, [2], [1], [11], [14]). Combinations of face, fingerprint, hand geometry, and speech-based biometrics were investigated at Michigan State University. The reported studies dealt with integrating face and fingerprints [6], fingerprints, face, and speech [7], face, fingerprint, and hand geometry [9] within a single authentication approach. However, some of the modalities involved do not support continuous and user-friendly authentication (as fingerprints and hand geometry) while the others (visual and voice based authentication) require considerable computational resources and therefore are not practical given restricted computational power of modern mobile devices.

In this paper, we describe our approach to user authentication aimed at the provision of continuous, user-friendly, and accurate verification of the user identity while preserving scarce resources of a mobile device. The problem of user authentication is formulated as an anomaly detection problem, where the anomalies in a user behavior indicate the substitution of the user by an impostor. The normal behavior of a user is modeled by a set of complementary behavioral aspects each of which is assigned to a dedicated software-implemented expert. Every expert constructs the model of normal user behavior from the perspective of the assigned behavioral aspect, and matches the current user behavior against this model. Based on the result of this matching, the expert provides its individual opinion regarding the user identity. The

decisions of multiple experts are integrated by applying a decision fusion technique. The continuous monitoring of multiple aspects enables the continuous verification of the user identity. The user-friendliness is supported by the use of transparent authentication methods that do not require direct user participation. The fusion of the opinions of multiple experts taking into account the information about their competence and confidence helps to improve the authentication accuracy.

The rest of the paper is organized as follows. Section 2 introduces our approach to continuous user authentication. The process of the user identity verification is considered in section 3 separately for local experts and for the decision fusion center. In section 4, the characteristics of the approach are discussed and possible directions for future work are outlined.

2. MULTI-EXPERT BASED APPROACH TO USER AUTHENTICATION

Throughout the paper, the continuous user authentication is considered as a constant process of collecting arguments in support of the hypothesis that the substitution of the user by an impostor has taken place. Once a number of these arguments overpasses a certain critical limit, the decision about the user substitution is made followed by adequate response (as, for example, blocking the device). A *legitimate user*, or *user* is considered to be a person authorized to use the device, while an *impostor* represents a malicious person who claims to be the legitimate one. It is assumed that only one user is allowed to interact with the device.

Two kinds of errors can take place in the authentication process: false acceptance (FA) error and false rejection (FR) error. A *probability of false rejection*, or a *false rejection error rate* P_{fr} is the probability of the legitimate user being classified as an impostor. A *probability of false acceptance*, or a *false acceptance error rate* P_{fa} corresponds to the probability of the situation when the impostor is erroneously considered to be the legitimate user and therefore is granted her privileges. The FA and FR probabilities characterize the accuracy of an authentication method. Another related characteristic is a *probability of detection* P_d that is the probability of the impostor being detected. This value complements the value of the probability of the FA error to unity.

In the context of the mobile-device user authentication, it may be reasonable to strive for low value of the FR error to the detriment of the FA error. Indeed, large values of the FA probability reduce significantly the security level provided by the system. However, if the values of the FR probability are not negligible, the user is disturbed by often false alarms and, therefore, it is likely for such authentication facility to be disabled by the user. As a result, the security level will evidently suffer much more. Thus, we argue that it is highly

desirable to limit the FR error rate by a certain level accepted by the user. In turn, the system should maximize the detection probability for the accepted FR error probability.

2.1. MONITORING OF USER BEHAVIOR

In the course of authentication, characteristics of the user behavior are analyzed. These characteristics are expressed by a set of features that are monitored. A *user behavior profile* aggregates the normal values of the features, which can be used for distinguishing the legitimate user and impostors. To verify the user identity, the instant values of the features are matched against those stored in the profile.

With the aim to provide continuous, user-friendly and resource preserving identity verification, we propose to monitor the following three characteristics of the user behavior: i) typing rhythms, ii) mobility patterns, and iii) activity patterns. This kind of monitoring does not involve the user participation and, therefore, is not likely to disturb her. At the same time, if implemented efficiently, an authentication system based on the monitoring of these characteristics will produce a little computational overhead.

The involved characteristics represent three different aspects of the user behavior: typing, mobility, and application usage. To improve the authentication accuracy these behavioral aspects are monitored in parallel. Several features to be monitored describe each behavioral aspect involved. Below these features are considered.

- *Typing rhythms.* The features to be monitored are the keystroke duration and keystroke latency time. Keystroke duration is the time of a key being pressed; the keystroke latency time corresponds to the time interval between two subsequent keystrokes (Figure 1). The statistical parameters of these features (means and standard deviations) are stored in the user behavior profile [8].

- *Mobility pattern.* Mobility monitoring deals with routes a user takes. The aim of this monitoring is to detect a situation when a user is faraway from his/her usual location [10]. According to this, the cells a user visits along with the timestamps of handovers are being monitored. In turn, the user behavior profile keeps the probabilities of a user traversed from one cell to another neighboring one, and mean cell residence times.

- *Activity patterns.* The actions a user takes are the subject of interest in this case. The set of possible user actions is classified into several action classes, as e.g. incoming calls, writing an email, etc. Similarly to the keystroke dynamics based features, the action time and tempo-

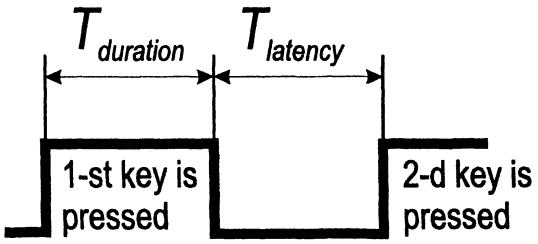

Figure 1 Keystroke dynamics monitoring

ral distance between subsequent actions are being monitored [12]; their statistical parameters are stored in the user behavior profile.

2.2. DETECTION OF IMPOSTORS BY LOCAL EXPERTS

As discussed above, for the authentication mechanism to be accepted by the users it is highly desirable to limit the FR error by a certain level. At the same time, the FA error probability should be minimized. This argumentation motivated us to employ the Neyman-Pearson test in our approach. According to Neyman-Pearson lemma, given the two-category classification problem, the Neyman-Pearson test is guaranteed to minimize one type of misclassification error subject to the constraint that the other type of misclassification error is no larger than a certain fixed level.

To implement the Neyman-Pearson test, the probability distribution of the feature values have to be obtained. This is performed at the learning stage whereon the user behavior profiles are created. A difficulty is that the size of data needed for learning grows exponentially with the number of used features due to the course of dimensionality. Moreover, should a feature be added or excluded, the entire system would have to be relearned. In order to eliminate these disadvantages, we consider the learning and authentication processes for each aspect of the user behavior separately, i.e. assign every behavioral aspect to a dedicated *local expert*. If R-aspects are being monitored for user authentication purposes, then R local experts should be employed. Each expert, in turn, is assigned to a set of related features to be monitored.

In order to leverage the distinction between the behavior of the user and the impostor, a so-called *punishment function* is implemented by each local expert. This is a single-output function, whose arguments are the values of the features analyzed by the expert. It compares the current, measured values of features with their normal values described in the user behavior profile, and penalizes

(or "punishes") the feature values that are distinct from normal. We will refer to the output of the punishment function as a *penalty*. Consequently, as the feature values for impostors are assumed to differ significantly from those for the legitimate user, this function is supposed to penalize the impostors to a greater extent than the user.

The mean value of the penalty for the legitimate user should be less than for impostors. The distributions of the penalty values for the user and for an impostor should be linearly separable. It means that the space of admitted values of the penalty F can be divided into two subspaces by the line $F = F_c$ such that within the first subspace the probability density is greater for the user, and within the second one — for the impostor. Let us denote the probability density function of the penalty values for the legitimate user and the impostor respectively as $f^U(F)$ and $f^I(F)$. Allow the penalty to have mean value F_μ^U for the user and F_μ^I for the impostor. Then F_c should be the only value such that $f^U(F_c) = f^I(F_c)$ and the following conditions should be met:

$$\begin{aligned} &F_\mu^U < F_c < F_\mu^I \\ &f^U(F) > f^I(F), \ F < F_c \\ &f^U(F) < f^I(F), \ F > F_c \end{aligned} \tag{1}$$

An example of the probability density function of penalty values is given in Figure 2.

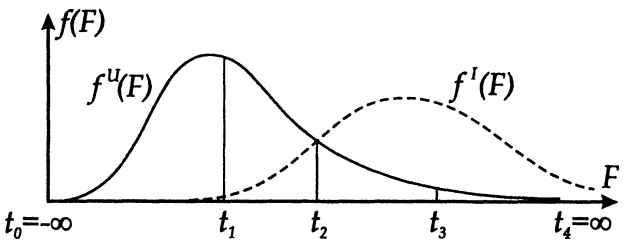

Figure 2 An example of the probability distribution of penalty values for the user and impostors

The exact form of the punishment functions depends on the features involved and, therefore, is different for every expert. For the three experts based on the features above the punishment functions are as follows.

Typing based expert. Denoting keystroke duration and keystroke latency time as T^{dur} and T^{lat}, and corresponding values of the means and

standard deviations as μ^{dur}, σ^{dur}, μ^{lat}, and σ^{lat}, the punishment function is expressed as:

$$F_{pun}(T_i^{dur}, T_i^{lat}) = \frac{1}{w_{KD}} \sum_{j=i-w_{KD}+1}^{i} \left(\frac{|T_j^{dur} - \mu_j^{dur}|}{\sigma_j^{dur}} + \frac{|T_j^{lat} - \mu_j^{lat}|}{\sigma_j^{lat}} \right) \quad (2)$$

where the index i corresponds to the i-th step of the work of the authentication system. In order to reduce the noise, a trailing window mean value filter is employed; w_{KD} denotes the window length.

Mobility based expert. Let the user have traversed from cell c_{i-w_M} through cells $c_{i-w_M+1}, \ldots, c_{i-1}$ to cell c_i, where w_M is the length of the trailing window. Then the penalty value can be calculated as:

$$F_{pun}(c_{i-1}, c_i) = \frac{1}{w_M} \sum_{j=i-w_M+1}^{i} \left[\left(\frac{1}{n_j^{cell} P_{j-1,j}} - 1 \right) \frac{|T_{j-1}^{res} - \mu_{j-1}^{res}|}{\sigma_{j-1}^{res}} \right]. \quad (3)$$

Here, $P_{j-1,j}$ is the probability that the user being situated in cell c_{j-1} will move to cell c_j, which is one of the n_{j-1}^{cell} cells surrounding cell c_{j-1}. T_{j-1}^{res}, μ_{j-1}^{res}, and σ_{j-1}^{res} denote respectively the residence time in previous cell c_{j-1} and its mean and standard deviation values.

Activity based expert. Similarly to the typing based expert, this expert for two subsequent actions implements the punishment function in a form:

$$F_{pun}(T_i^{act}, T_i^{tr}) = \frac{1}{w_M} \sum_{j=i-w_A+1}^{i} \left(\frac{|T_j^{act} - \mu_j^{act}|}{\sigma_j^{act}} + \frac{|T_j^{td} - \mu_j^{td}|}{\sigma_j^{td}} \right), \quad (4)$$

where T^{act} and T^{td} are the action time and temporal interval between two subsequent actions. Their means and standard deviations are denoted correspondingly as μ^{act}, σ^{act}, μ^{td}, and σ^{td}.

2.3. LEARNING AND DETECTING PHASES

The proposed authentication comprises learning and impostor detection activities. The learning consists of creation and updating of the user behavior profile, which is used by the detection process for the continuous verification of the user identity.

During profile creation, the probability distributions of the penalty values for the user are obtained based on the feature values. The distributions of the penalty values for the impostors are assumed to have been obtained during the

extensive tests conducted by an expert supplier. Based on the above distributions, the thresholds for every local expert are set according to the FR error rate required.

At the detection (verification) stage, when a new event is registered (for example, new call was made or the location of the device has been changed), the corresponding expert measures the values of the relevant features and calculates the penalty value. By comparing this value with the local thresholds, the expert makes its local decision regarding the user identity. A *local decision* can be defined as a decision provided by a local expert independently from other experts.

All the local decisions are transmitted to the decision fusion center (DFC) inferring a central decision from the partial decisions of several local experts [5]. The basic goal of such fusion is improvement of the authentication accuracy. Besides, the consideration of multiple clues collected from different sources makes it possible to cover most part of the user-device interaction process and, therefore, supports truly continuous authentication.

Provided the local decisions of the experts, the Neyman-Pearson test (NP-test) is implemented at the DFC. As a result of the test the central decision regarding the presence of the impostor is made. Should this decision be positive, the immediate measures have to be taken. For example, the device can be blocked and the user can be asked to enter a PIN unlock key (PUK) to unblock it. Alternatively, a notification can be sent to a user's contact address as e.g. email address or another mobile number.

In this section, we introduced the approach to the continuous authentication that analyses multiple aspects of the user behavior. These aspects are analyzed separately by dedicated experts each of which provides its confidence in the legitimacy of the user. These confidences are taken into account by the decision center in order to infer the final decision about user's authenticity. Both the local experts and the decision center compare the values of the parameters based on current user behavior with the thresholds obtained during the learning process. In next section, the thresholds selection process at the local experts and at the decision fusion center is described.

3. FUSION OF EXPERT DECISIONS

In this section, the procedure of threshold selection is considered in details. The local experts and the decision fusion center are discussed separately in the next two subsections. Firstly, we describe the threshold selection process at the local experts. Secondly, the inference of the central decision at the fusion center is explained.

3.1. LOCAL EXPERTS

Local experts match current values of features against their usual values, described by the profile, calculating penalty values according to deviations between them. After this, comparison of the obtained penalty values with a set of the local thresholds derives the local decisions with respect to the user identity.

Instead of hard decisions, N-bit soft local decisions u_i are to be provided by each expert i, i.e. $u_i \in [0, 2^N-1]$. Zero corresponds to the greatest expert confidence that the user is the legitimate one, while 2^N-1 corresponds to the greatest confidence that an impostor is interacting with the device[1]. To provide such soft decisions, the area of the penalty values is divided into 2^N *confidence regions*. These regions are bounded by thresholds $t_0, t_1, ..., t_{2^N}$, where $t_0 = -\infty$ and $t_{2^N} = \infty$ are the auxiliary thresholds. For every confidence region C_j, $j \in [0, 2^N-1]$ bounded by the thresholds t_j and t_{j+1}, the probability of FR $\alpha_i(j)$ and probability of detection $\beta_i(j)$ can be calculated as follow:

$$\alpha(j) = \int_{C_j} f^U(F)\, dF, \tag{5}$$

$$\beta(j) = \int_{C_j} f^I(F)\, dF, \tag{6}$$

where the index i was omitted for the sake of simplicity.

The local decision submitted by the expert to the DFC is the index of the confidence region to which its current penalty value belongs. Along with its local decision, the expert delivers the values of the α and β parameters for the corresponding confidence region. Together, α and β values provide the information about expert competence and confidence in the legitimacy of the user. The expert confidence is described by the value of α, while the ratio α/β characterizes the expert competence.

The local-threshold selection process described below is aimed at providing additional information about expert confidence. The expert uncertainty is insignificant when the penalty values are either very small or very large, and is high for the penalty values situated around F_c, which is the threshold minimizing Bayes risk. Accordingly, we try to construct more fine-grained confidence regions in the proximity to F_c, in the area bounded by F_μ^U and F_μ^I.

The first and the middle thresholds are assigned to F_μ^U and F_c respectively:

$$t_1 := F_\mu^U, \tag{7}$$

$$t_{2^N-1} := F_c, \tag{8}$$

[1]Currently, the decisions of all the experts are assumed to have the same number N of bits. Although, the approach can be adopted to support the decisions with various values of N.

and the last threshold t_{2^N-1} is selected to be such that the probability of FR for the last confidence region would be equal to the admissible level of the probability of FR P^*_{FR}:

$$t_{2^N-1}: \qquad \alpha(2^N-1) = \int_{t_{2^N-1}}^{\infty} f^U(F)\,dF = P^*_{FR}. \tag{9}$$

As will be explained in next subsection, this ensures that the admissible level of FR error probability will be achieved at the fusion center. It may happen that $t_{2^{N-1}} > t_{2^N-1}$. Should it take place, these two thresholds are swapped so that the numeration remains consistent.

For $N = 2$ the process of the threshold selection is illustrated in Figure 2. For the penalty values greater than threshold t_{2^N-1} the expert is considered as absolutely confident that the user was substituted. Consequently, these values are united within a single confidence region. In a similar way, the penalty values to the left of $t_1 = F^U_\mu$ are considered to reflect the absolute expert confidence in the validity of the user. These values also are united within a single confidence region. In turn, the area between t_1 and t_{2^N-1} refers to expert's uncertainty and is divided into several confidence regions (if $N > 2$). The regions to the left and to the right of the middle threshold correspond to the equal probabilities of the user being legitimate. Thus, the respective thresholds are selected (after having selected the above three thresholds) to satisfy:

$$\begin{aligned} t_2, \ldots, t_{2^{N-1}-1}: &\qquad \alpha(1) = \ldots = \alpha(2^{N-1}-1), \\ t_{2^{N-1}+1}, \ldots, t_{2^N-2}: &\qquad \alpha(2^{N-1}) = \ldots = \alpha(2^N-2). \end{aligned} \tag{10}$$

It is necessary to note that the described method of the local threshold selection is suboptimal. The optimal choice of the local thresholds would provide a superior accuracy. However, the evaluation of such threshold values would be computationally expensive and, hence, is impractical given restricted power and computational capability of modern mobile devices.

3.2. DECISION FUSION CENTER

At the fusion center, the NP-test is performed in order to obtain the central decision, i.e. the likelihood ratio $\Lambda(u_1, u_2, \ldots, u_R)$ is calculated:

$$\Lambda(u_1, u_2, \ldots, u_R) = \frac{P(u_1, u_2, \ldots, u_R|H_I)}{P(u_1, u_2, \ldots, u_R|H_U)} \gtrless_{H_U}^{H_I} t_f \tag{11}$$

where H_I and H_U denote the hypotheses that respectively the impostor and the legitimate user are interacting with the device, u_r denotes the local decision made by expert $r, r \in [0, R]$, and t_f is the central threshold, selected at the

learning stage according to the desirable probability of the FR errors:

$$t_f : \sum_{\Lambda(u_1,u_2,\ldots,u_R)>t_f} P(\Lambda(u_1,u_2,\ldots,u_R|H_U)) \leq P^*_{FR}. \qquad (12)$$

Here, $P(\Lambda(u_1,u_2,\ldots,u_R|H_U))$ is the probability of the likelihood ratio value under the hypotheses that the user is valid.

Different aspects of the user behavior are analyzed by different experts. It is assumed that the features describing one aspect are independent from the features describing the others. Consequently, we assume the independence of the expert decisions from each other. Based on this assumption, it follows that:

$$\Lambda(u_1,u_2,\ldots,u_R) = \prod_{r=1}^{R} \frac{P(u_r|H_I)}{P(u_r|H_U)} = \prod_{r=1}^{R} \frac{\beta_r(u_r)}{\alpha_r(u_r)}, \qquad (13)$$

and

$$P(\Lambda(u_1,u_2,\ldots,u_R|H_U)) = \prod_{r=1}^{R} P(\Lambda(u_r|H_U)) = \prod_{r=1}^{R} \alpha_r(u_r), \qquad (14)$$

$$P(\Lambda(u_1,u_2,\ldots,u_R|H_I)) = \prod_{r=1}^{R} P(\Lambda(u_r|H_I)) = \prod_{r=1}^{R} \beta_r(u_r). \qquad (15)$$

where $\alpha_r(u_r)$ and $\beta_r(u_r)$ are calculated using (5) and (6) depending on the decision value of local expert r.

In order to select t_f, the conditional distributions (14) and (15) are obtained. Based on the conditional distribution $P(\Lambda(u_1,u_2,\ldots,u_R)|H_U)$ the values of t_f satisfying (12) are selected. For every value of t_f obtained the corresponding value of the probability of the false acceptance is calculated according to

$$P_{FA} = 1 - \sum_{\Lambda(u_1,u_2,\ldots,u_R)>t_f} P(\Lambda(u_1,u_2,\ldots,u_R|H_I)) \qquad (16)$$

using the conditional distribution $P(\Lambda(u_1,u_2,\ldots,u_R)|H_I)$. After that, the minimum of P_{FA} is selected, and the corresponding value of t_f is chosen as the central threshold.

In detection process, the central decision regarding the user's authenticity is derived by calculating the likelihood ratio (according to (11)) and comparing it with the central threshold. Once the value of this ratio exceeds the threshold, the decision is made that the user has been substituted.

The least value of t_f satisfying (12) can be calculated as $\prod_{r=1}^{R}\alpha_r(2^N\text{–}1)$. As was described in previous subsection, the last local thresholds are selected such

that $\alpha_r(2^N - 1) = P^*_{RF}$. From $P^*_{RF} < 1$ it follows that $\prod_{r=1}^{R} \alpha_r(2^N - 1) \leq P^*_{FR}$ for $R = 1, 2, \ldots, \infty$. The least value is therefore less than or equal P^*_{FR} ensuring that at least one value of t_f satisfying (12) exists and, therefore, that the central threshold can be selected that provides the admissible P^*_{FR} level at the DFC.

In this section, the decision-making process at the experts and at the DFC was described. The information about relative strength of the experts along with the information about expert confidences in their decisions is taken into account by this process in order to improve the authentication accuracy. In next section, the abilities of the approach are analyzed, and possible directions for further work are considered.

4. DISCUSSION AND CONCLUDING REMARKS

In previous sections we introduced the approach to the continuous user authentication based on the monitoring of the user behavior and aimed at strengthening PIN-based authentication. The employment of the approach is expected to raise significantly the security level provided by the device since it provides continuous and transparent user identity verification. The verification is continuous because the behavior of a user is being monitored continuously, and it is transparent since no direct user participation is required.

In the course of continuous identity verification multiple behavioral aspects are analyzed by dedicated experts. They are capable of making their own decisions that are subsequently combined by applied decision fusion technique. As a result, the probability of detection is improved as compared with a single expert used alone.

Let us consider how the authentication accuracy can improve due to fusing decisions from several experts, depending on the quality of the individual experts. The accuracy criteria are two related characteristics — the probability of detection and the probability of FA error. The quality of a local expert is characterized by the classification error, which in case of two-bit local decisions can be calculated as:

$$E = \beta(0) + \beta(1) + \alpha(2) + \alpha(3). \tag{17}$$

The decision fusion of three local experts is considered. The admissible FR error level F^*_{fr} is assigned to 0.01; all three experts are assumed to have the same characteristics. These characteristics are represented by the probability distributions of the penalty values, which in turn are described by the values of $\alpha(j)$, $\beta(j)$, $j = 0, \ldots, 3$.

For the sake of simplicity, we approximate the probability density functions of penalty for the user and the impostors by normal distributions, which

have the same standard deviations, but different means. Given the normal distribution of the penalty values and the classification error, the local thresholds are calculated using (9), (7) and (8). At the same time, $\alpha(j)$ and $\beta(j)$ are derived according to (5) and (6). After that, the conditional distributions $P(\Lambda(u_1, u_2, \ldots, u_R)|H_U)$ and $P(\Lambda(u_1, u_2, \ldots, u_R)|H_I)$ are obtained using (14) and (15). Finally, we select the central threshold that minimizes the probability of FA (16) and is subject to constraint (12).

The improvement in accuracy depending on the quality of the individual experts is shown in Figure 3, where the probability of detection at the DFC is compared with the probability of detection of a single expert working alone for a range of classification error values. As could be seen, for the whole

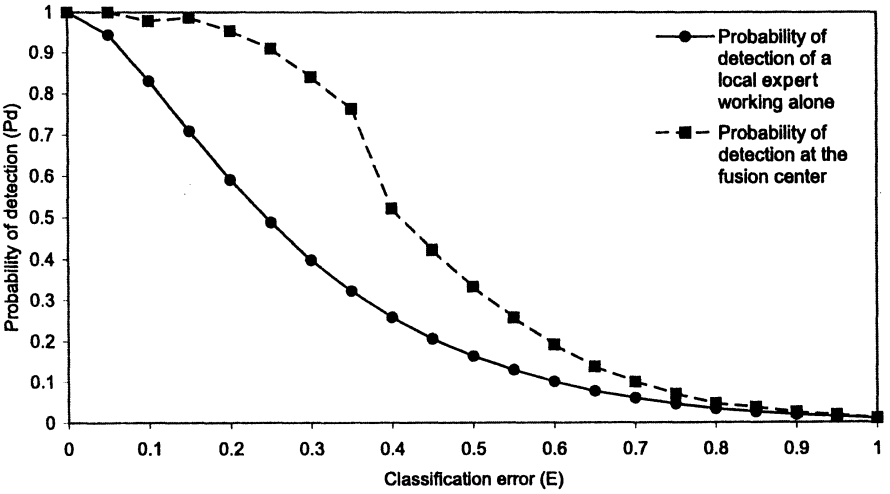

Figure 3 Probability of the impostor detection at the local experts and at the DFC

range of classification error values, describing the quality of the local experts, the fusion of their decisions results in an improvement of the authentication accuracy. Meanwhile, the magnitude of the improvement varies reaching its maximum when the classification error equals 0.35. Towards the boundaries of the classification error values, this magnitude declines and tends to zero. The accuracy of authentication, therefore, cannot be improved significantly by the proposed approach if the classification error of individual experts is close to unity. Nevertheless, even for poorly performing local experts the fusion of their decisions may significantly improve the overall accuracy. For instance, given

local experts with $E = 0.6$, the decision fusion provides 92% of improvement in accuracy.

From the resource consumption perspective, it is important to note that the proposed approach does not involve any sophisticated computations, neither it requires a large amount of memory to keep the user behavior profiles. As a result, the level of the resource consumption is expected to be conservative. The estimation of the real (exact) values of the computational overhead and required memory size is, however, for a further study.

Additionally, the approach supports easy expert management (addition or removal of the local experts) if it is required due to changes in the user behavior. For instance, if a user stops using a keyboard and adopts a stylus for text input, then the expert based on keystroke dynamics should be removed, followed by the potential inclusion of an expert monitoring the use of the stylus. In such situation, only the recalculation of the central threshold is required.

There is a potential issue of the correlation between experts decisions that has to be addressed. While herein the expert decisions are assumed to be independent, it is reasonable to expect that some degree of the correlation will exist. On one hand, should we ignore this correlation, an approximation error will be inherited in the decisions of the DFC. On the other hand, to take this correlation into account, the set of correlation coefficients has to be estimated. Consequently, as a training set available at the learning stage is restricted, the estimation of these coefficients will bring an evaluation error. Thus, to reveal the degree of the correlation between experts, compare the estimation and the approximation errors, and obtain the real values of the detection probability, the experiments in real environment are required.

Future study is required to address a number of other issues. Privacy risks due to the storage of sensitive data in the user profile are to be investigated. Synchronization of local experts should also be addressed. As well, a mechanism controlling the addition and removal of the experts should be elaborated.

Acknowledgments

This work was partly supported by the COMAS Graduate School of the University of Jyväskylä. The authors would like to thank the anonymous reviewers for their valuable comments and suggestions.

References

[1] S. Ben-Yacoub, Y. Abdeljaoued, and E. Mayoraz. Fusion of face and speech data for person identity verification. *IEEE Transactions on Neural Networks*, 10(05):1065–1074, 1999.

[2] T. Choudhury, B. Clarkson, T. Jebara, and A. Pentland. Multimodal person recognition using unconstrained audio and video. In *The 2nd International Conference on Audio-Visual Biometric Person Authentication*, pages 176–181, 1999.

[3] N.L. Clarke, S.M. Furnell, P.M. Rodwell, and P.L. Reynolds. Acceptance of subscriber authentication methods for mobile telephony devices. *Computers & Security*, 21(3):220–228, 2002.

[4] F-Secure Corporation. Content security at hand. A white paper on handheld device security, February 2002. Available from http://www.europe.f-secure.com/products/white-papers.

[5] B.V. Dasarathy. *Decision Fusion*. IEEE Computer Society Press, 1994.

[6] L. Hong, and A.K. Jain. Integrating faces and fingerprints for personal identification. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 20(12):1295–1307, 1998.

[7] A.K. Jain, L. Hong, and Y. Kulkarni. A multimodal biometric system using fingerprint, face, and speech. In *The 2nd International Conference on Audio-Visual Biometric Person Authentication*, pages 182–187, March 1999.

[8] F. Monrose and A.D. Rubin. Keystroke dynamics as a biometric for authentication. *Future Generation Computing Systems (FGCS) Journal: Security on the Web (special issue)*, March 2000.

[9] A. Ross, A.K. Jain, and Jian-Zhong Qian. Information fusion in biometrics. In *3rd International Conference on Audio- and Video-Based Person Authentication*, pages 354–359, Sweden, June 2001.

[10] D. Samfat and R. Molva. IDAMN: An intrusion detection architecture for mobile networks. *IEEE Journal on Selected Areas in Communications*, 7(15):1373–1380, 1997.

[11] C. Sanderson and Kuldip K. Paliwal. Adaptive multi-modal person verification system. In *First IEEE Pacific-Rim Conference on Multimedia*, Sydney, Australia, 2000.

[12] A. Seleznyov. A methodology to detect anomalies in user behavior basing on its temporal regularities. In *IFIP/SEC2001: 16th International Conference on Information Security*, Paris, France, June 2001.

[13] J. Veijalainen. Transactions in mobile electronic commerce. *Lecture Notes in Computer Science*, 1773:208–229, December 1999.

[14] P. Verlinde, G. Chollet, and M. Acheroy. Multi-modal identity verification using expert fusion. *Information Fusion*, 1(1):17–33, 2000.