# Context-Aware Mobile Biometric Authentication based on Support Vector Machines

H. Witte, C. Rathgeb and C. Busch
da/sec – Biometrics and Internet Security Research Group
Hochschule Darmstadt, Germany
Email: {heiko.witte,christian.rathgeb,christoph.busch}@h-da.de

***Abstract*—The ubiquitous use of smartphones raises the need for stronger device protection. Traditional authentication methods on mobile devices are still knowledge-based, exhibiting well-known drawbacks. In addition, requests for passwords, PINs, or screen lock patterns represent an interruption of the device usage. In this paper a context-aware mobile biometric system is proposed. Modern smartphone devices comprise a multitude of sensors which can be utilized to measure a variety of environmental aspects, e.g. noise level or location. Based on this contextual information subject-specific context models are constructed in order to train SVMs, providing an alternative user-friendly authentication mechanism. In experiments a self-acquired database is employed where obtained results confirm the feasibility of the proposed system.**

## I. Introduction

Based on a study on mobile phone security in 2010 [1] only 13% (!) of mobile devices are secured with PINs or screen-lock patterns, where the main reason (74% of the cases) for this lack of security is a demand on fast access, i.e. security does not coincide with usability when it comes to mobile devices. While biometric recognition [2] provides an increased level of security, different studies [3], [4] confirm that people identify improved usability as the major benefit of biometric systems, compared to conventional authentication mechanisms. Improving the usability of computer applications and systems represents a core goal of the emerging research field "Human Computer Interaction" (HCI). With respect to mobile applications emphasis is put on letting technology fade into the background. One way to achieve such adaptable computing platforms is the application of contextual information in order to provide services, that do not require user interaction. Context-aware authentication systems (also referred to as implicit authentication systems) call for numerous applications: detection of anomalous interaction with mobiles devices can be observed in order to prevent from theft, or modality-dependent application of biometric authentication based on environmental conditions can improve recognition accuracy.

The contribution of this work is the proposal of a context-aware mobile biometric system based on support vector machines (SVMs) which enables the collection, pre-processing, aggregation, and evaluation of contextual information. To this end the context is captured by observing behavioral characteristics of an individual, i.e. a subset of collected data can be seen as behavioral biometric features, e.g. device usage patterns. Based on the construction of subject-specific context models, SVMs are trained and applied to derive a class probability, which indicates whether authentication is necessary at all, or determines the number of required authentication factors. The idea of combining decisions derived from the subject models with conventional authentication mechanisms or additional unobtrusive biometrics in order to compute a final confidence score is depicted in Fig. 1. Within this work emphasis is put on the authentication of subjects based on contextual features only. Additionally, the modular structure of the system encourages the development of sophisticated feature extractors as well as additional sensors.

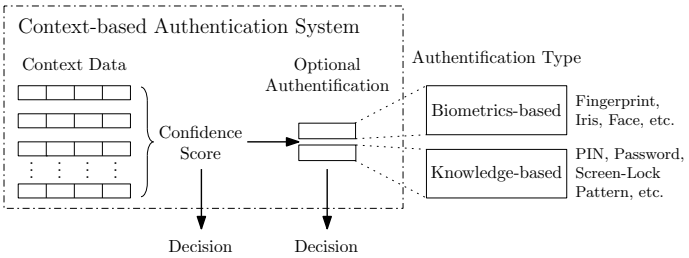


Fig. 1. Context-aware mobile system: confidence scores obtained from (biometric) contextual data are directly used for decision or to parameterize further authentication systems.

This paper is organized as follows: Sect. II introduces the term 'context' and reviews related works. In Sect. III key components of the presented system are described in detail and the proposed context-modeling technique is summarized. Subsequently, experimental evaluations are presented and discussed in Sect. IV. Finally, conclusions are drawn in Sect. V.

## II. Related Work

Schilit *et al.* [5] were the first to provide a definition for the terms 'context' and 'context-aware computing applications'. Factors such as lighting conditions, noise levels, communication bandwidth, or social situation are mentioned as additional features of a context description. In further work [6] the authors emphasize, that a context is characterized by more information than just the physical location. Dey and Abowd [7] claim that these definitions contain serious flaws since these are based on examples or define context as 'situation' or 'environment', which represent synonyms. Alternatively, the authors define context as the information that can be used to describe a situation: *"Context is any information that can be used to characterize the situation of an entity. An entity is a person, place, or object that is considered relevant to the interaction between a user and an application, including*

*the user and applications themselves"*. Based on this abstract definition of context, they define the term 'context-aware' as follows: *"A system is context-aware if it uses context to provide relevant information and/or services to the user, where relevancy depends on the user's task"*. In contrast to definitions that are centered around devices or environments [8], [5], this definition puts the subject into the center of contextual considerations, i.e. this definition fits with the goal of the presented paper.

### A. Context-Data vs. Biometrics

Behavioral biometric characteristics, e.g. speech patterns, on-line signatures, or keystroke dynamics, are understood as physiological or behavioral processes created by the body. This implies that, as mentioned earlier, distinct contextual data can be interpreted as biometric data per se, e.g. accelerometer-based context data directly relates to a subject's figure or sound levels during phone calls relate to a subject's voice patterns.

On the other hand, contextual data can be exploited in order to operate biometric authentication systems based on environmental conditions, e.g. a rather dark environment may reduce the reliability of face recognition. That is, context-awareness can be employed in order to improve the recognition performance of biometric systems by adapting authentication methods to a situation [9], i.e. parameterizing biometric authentication. In addition, the usage of contextual information enables implicit authentication. For instance, Shi *et al.* proposed a system that uses phone, SMS, GPS, and browser data to implicitly authenticate subjects [10]. Combining context-awareness to determine a security level and unobtrusive biometric authentication methods (e.g. gait recognition [11]) improves the usability of an authentication system as wel as social acceptance of biometrics.

## III. Context-Aware Mobile Biometric System

In the following subsections (1) the architecture of the proposed context-aware mobile biometric system and (2) the applied context modeling are described in detail:

### A. System Architecture

The BioAPI (ISO/IEC 19784-4:2006) compliant system comprises three components:

1) **Context Subsystem**: within the context subsystem contextual information is collected and processed in order to build a user model. After the model is derived from training data, a classification process computes the probability of the association between the current context and the genuine subject.
2) **Authentication Subsystem**: the authentication subsystem provides an interface to optionally integrate BioAPI compliant biometric service providers, utilized by the context subsystem.
3) **Result Action Subsystem:** this subsystem uses information on (un-)successful authentication attempts in order to execute according actions (e.g. combination of biometrics and token-based authentication [12]).

The Android API[1] is utilized to collect contextual data offering access to a variety of sensors. Raw data is pre-processed ($\simeq$normalized) and features are extracted. Subsequently, the resulting features are sent to the context component which aggregates the pre-processed data and combines it into a context frame. Context frames can then be written to a persistent store for later reference, implemented on the device or remotely on a server. Collected context frames and a set of training examples are employed, in order to generate a context model of the user. Finally, the current context is evaluated, i.e. the probability of the current context frame being associated with the enrolled subject is computed.

### B. Context Modeling

A context model is defined as a set of contextual information, that describes the situation of an entity and contains behavioral biometric traits. Since the system is designed as a mobile application, data acquired from all available sensors, as well as information related to user-smartphone-interaction is included. In addition to generic sensors that measure physical magnitudes, quantities such as the current time, network traffic, battery percentage, etc. are analyzed. The former type of sensors is referred to as physical sensor, the latter as logical sensor. Furthermore, a discrimination between active and passive sensors is necessary. Active sensors can be activated and deliver data in predefined intervals, while passive sensors only deliver data in case a subject interacts with the device. An example for a passive-logical sensor is the measurement of activations and deactivations of a device's screen. Sensors send the acquired data (together with a timestamp) to the pre-processor which performs normalization and/or filtering, prior to feature extraction.

Pre-processing of sensor data in the presented configuration of the system is limited to a reduction of the data by computing the mean and median of a set of raw sensor values. This simplification was imposed to keep emphasis on system design, development, and evaluation. Extracted features and associated data types, units, or ranges are summarized in Table I. After the data is collected and pre-processed, it is adapted for user modeling. For this purpose, aggregated data is stored in a set of vectors, which is referred to as a context data frame. Depending on the learning algorithm, data must be transformed into numerical values. Since all sensor values include a timestamp of the measurement, this information is used to group data into sections. We divide each day into eight sections of three hours duration, such that measurements fit into the cells of a seven-by-eight matrix, i.e. the temporal granularity of this specific context model is rather coarse. If a system shall be able to provide real time feedback, narrower intervals are required. A context data frame $C_{\pi,\rho}$ represents an instance of the context model and is constructed by combining the day index $\pi \in [1,7]$ and day section index $\rho \in [1,8]$ with matching values of all sensors, $C_{\pi,\rho} = \{(s_1,...,s_n)|s_i \in S_{\pi,\rho}\}$ where $S$ is the set containing pre-processed data of all sensors, that match the weekday index $\pi$, and day section index $\rho$.

## IV. Experimental Evaluations

We developed a mobile application referred to as *ContextCollector* for a data collection and published it in the

[1] Android, http://www.android.com/

TABLE I. TYPES OF EMPLOYED CONTEXT SENSORS, FEATURES, AND UNITS (P=PHYSICAL, L=LOGICAL, A=ACTIVE, PS=PASSIVE).

| Sensor | Type | Feature | Description |
|---|---|---|---|
| Location | P, A | latitude | float |
| | | longitude | float |
| | | accuracy | m |
| | | altitude | m |
| | | speed | m/sec |
| Accelerometer | P, A | median $X$,$Y$,$Z$ | m/sec$^2$ |
| | | mean $X$,$Y$,$Z$ | m/sec$^2$ |
| Magnetic Field | P, A | median $X$,$Y$,$Z$ | m/sec$^2$ |
| | | mean $X$,$Y$,$Z$ | m/sec$^2$ |
| Microphone | P, A | median amplitude | $[0, 32767]$ |
| | | mean amplitude | $[0, 32767]$ |
| Light | P, A | median brightness | lux |
| | | mean brightness | lux |
| Battery | L, A | percentage | float |
| | | charging status | boolean |
| ScreenState | L, PS | percentage | float |
| | | charging status | boolean |
| ShutdownBoot | L, PS | boot time | timestamp |
| | | shutdown time | timestamp |
| Call | L, PS | duration | Seconds |
| | | direction | boolean |
| | | numberKnown | boolean |

Google Play Store[2]. In experiments the application had to run on a variety of Android devices (v2.3.3 - v4.1.2). Participants had to install the application on their personal device and enroll in the data collection by providing demographic information, as well as accepting the terms and conditions. The following subsections provide detailed information on the setup, acquisition, and modeling process, as well as an evaluation of the classification performance.

### A. Experimental Setup

In total 25 subjects participated in the data collection where each participant had to provide information about his/ her gender, age group and profession as part of the enrollment process of the data collection. The majority of the participants are male students in the age of 26-35. After successful enrollment, the application moves into the background. Subsequently, the application is invoked every 20 minutes, and the context sensors start to measure for the configured period of 20 seconds. Data of some of the sensors is then pre-processed as summarized in Table I. A critical issue of such a data collection is the support of a wide range of Android devices, for 25 participants 16 different devices had to be supported, where some devices did not comprise the entire set of the desired sensors. In addition, some subjects did not provide enough contextual information, i.e. the entire database had to be reduced to 15 subjects from which data units of at least three days were extracted.

### B. Context-based User Modeling

The goal of building a subject-specific model is to perform a classification of context data frames in order to determine if the current context is associated with the enrolled subject. Classification is performed based on SVMs where a predefined set containing negative examples is labeled with class 0, while the subjects' context data frames are labeled with class 1. The complete set of context data frames is divided into a training set of size 4/5 and a test set of size 1/5. The libsvm[3] implementation of a support vector machine is utilized for cross-validation on the training data.

[2]Google Play Store, https://play.google.com/store/

TABLE II. CLASSIFICATION RESULTS FOR A DECISION THRESHOLD OF 0.4, KERNEL PARAMETER $\gamma = 0.1$, COST PARAMETER $C = 10$.

| ID | Precision | Recall | $F_1$-score | Set size |
|---|---|---|---|---|
| 1 | 0.85 | 0.32 | 0.46 | 2364 |
| 2 | 0.78 | 0.43 | 0.56 | 2268 |
| 3 | 0.74 | 0.15 | 0.25 | 865 |
| 4 | 0.82 | 0.42 | 0.56 | 383 |
| 5 | 0.95 | 0.82 | 0.88 | 5536 |
| 6 | 0.97 | 0.84 | 0.90 | 2023 |
| 7 | 0.95 | 0.86 | 0.90 | 2782 |
| 8 | 0.76 | 0.21 | 0.33 | 2440 |
| 9 | 0.77 | 0.34 | 0.47 | 3190 |
| 10 | 0.78 | 0.18 | 0.30 | 1333 |
| 11 | 0.78 | 0.18 | 0.30 | 1106 |
| 12 | 0.83 | 0.43 | 0.56 | 3349 |
| 13 | 0.87 | 0.62 | 0.72 | 3938 |
| 14 | 0.80 | 0.41 | 0.55 | 1099 |
| 15 | 0.92 | 0.76 | 0.83 | 715 |

### C. Classification Performance

In order to evaluate the classification performance, we compute the common classification magnitudes, precision $P$, recall $R$ and the resulting $F_1$-scores (harmonic mean). Let $T_p$, $F_p$, $F_n$ denote the amount of true-positives, false-positives, and false-negatives, then the applied metrics are defined as:

$$P = \frac{T_p}{T_p + F_p} \ , \ R = \frac{T_p}{T_p + F_n} \ , \ F_1 = 2 * \frac{P * R}{P + R}. \quad (1)$$

We do not use the accuracy as a performance measure, because of the unbalanced classes in the training set. Precision and recall are directly related to False Match Rate and inverse False Non-Match Rate (ISO/IEC 19795-1:2006), respectively. It is important to note that we do *not* include geographic position as a contextual feature, i.e. in contrast to related works, e.g. in [13], GPS coordinates have not been exploited. This setting is motivated by the fact that any deployment of a context-aware system will always have to deal with geographically close entities. While including users which live in different geographical areas into a test set may improve the overall performance (if GPS coordinates are interpreted as contextual information), realistic use-cases, e.g. anomaly detection based on GPS coordinates in case of theft, will not be feasible. That is, the use of position coordinates clouds the picture of the underlying context-aware system, while excluding this information yields an even more challenging classification scenario. The best results of all evaluations are given in Table II. Varying the decision threshold for the class 1 probability leads to better precision or recall scores, but cannot improve the $F_1$-score. As previously mentioned, sizes of training sets play an important role. Fig. 2 shows

[3]libsvm, http://www.csie.ntu.edu.tw/~cjlin/libsvm/

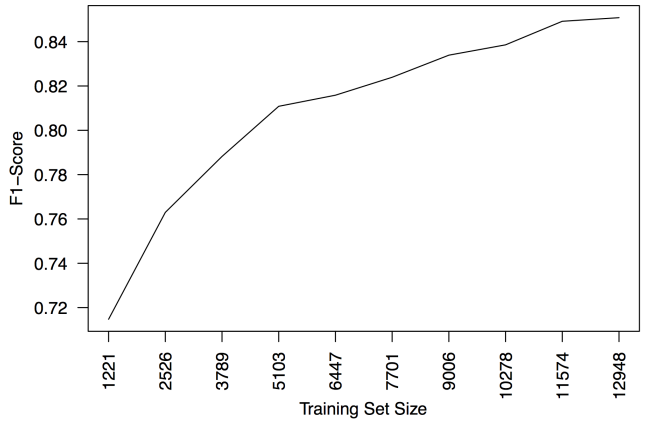


Fig. 2. $F_1$-scores for the model with ID 5 significantly improve for gradually increasing training set sizes.

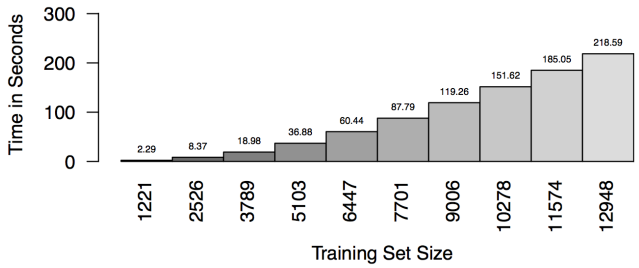


Fig. 3. Time (in seconds) for the training of the SVM for model with ID 5 with gradually increasing training set sizes.

an example how the $F_1$-score improves for a single subject with an increasing amount of training samples. Fig. 3 depicts the required time for the training of the SVM with respect to increasing training set sizes, respectively. Without a doubt, sufficiently large training sets are required in order to construct representative context models. In addition, it is interesting to examine which contextual features turn out to be the most reliable ones. Fig. 4 depicts the sums of all feature weights across all different models. It becomes clear that (comparable to conventional biometric features) distinct features turn out to be more reliable than others. For instance, battery charging behavior turned out to be a rather unreliable feature, such as day or day section. In contrast, call duration and brightness represent the most reliable features. Due to the fact that the collected data set comprises a variety of different devices (like in reality) feature weights may also turn out to be device-dependent. As expected, a small amount of distinct contextual features correlation was observed, e.g. accelerations or magnetic fields in different directions, highly correlated while the vast majority of features does not. Still, the great variety of un-correlated features provides high entropy of the extracted contextual information and confirms the soundness of the presented system.

## V. Conclusion

In the presented paper we proposed an architecture for a context-aware mobile biometric system and described constituent components in detail. The feasibility of the proposed architecture has been demonstrated by developing a mobile application which has been applied for data collection purposes. Furthermore, the results of the data collection were

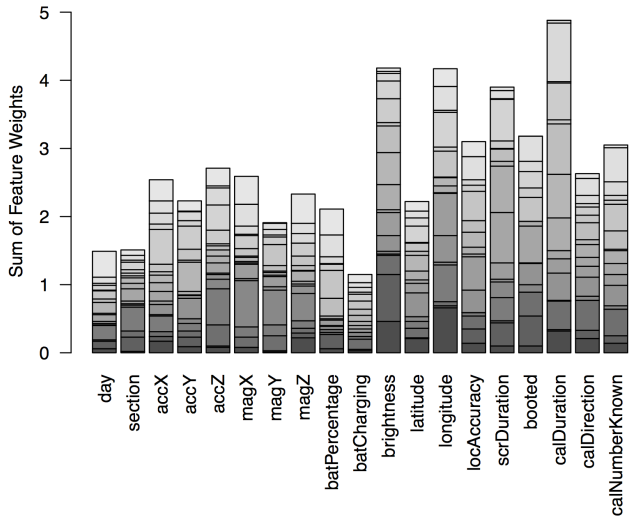


Fig. 4. Sum over the normalized feature weights across all features and for all different models.

utilized to train an adequate classifier in order to obtain models of according subjects. The models were evaluated by applying standard statistical classifier evaluation metrics. The results demonstrate the feasibility of the proposed system and classification approach.

## References

[1] F. Breitinger and C. Nickel, "User Survey on Phone Security and Usage," in *In Proc. of the Int'l Conf. of the Biometrics Special Interest Group (BIOSIG'10)*, pp. 139–144, 2010.

[2] A. K. Jain, A. Ross, and S. Prabhakar, "An introduction to biometric recognition," *IEEE Trans. on Circuits and Systems for Video Technology*, vol. 14, pp. 4–20, 2004.

[3] R. R. Heckle, A. S. Patrick, and A. Ozok, "Perception and acceptance of fingerprint biometric technology," in *Proc. of the 3rd Symposium on Usable Privacy and Security (SOUPS'07)*, pp. 153–154, 2007.

[4] M. El-Abed, R. Giot, B. Hemery, and C. Rosenberger, "A study of users' acceptance and satisfaction of biometric systems," in *IEEE Int'l Carnahan Conf. on Security Technology (ICCST'10)*, pp. 170 –178, 2010.

[5] B. Schilit, N. Adams, and R. Want, "Context-Aware Computing Applications," *1st Workshop on Mobile Computing Systems and Applications, WMCSA'94*, pp. 85–90, 1994.

[6] A. Schmidt, M. Beigl, and H. W. Gellersen, "There is more to context than location," *Computers & Graphics*, vol. 23, no. 6, pp. 893–901, 1999.

[7] G. Abowd and A. Dey, "Towards a better understanding of context and context-awareness," *Handheld and ubiquitous computing*, pp. 304–307, 1999.

[8] J. Pascoe, "Adding generic contextual capabilities to wearable computers," *Wearable Computers, 1998. Digest of Papers. Second International Symposium on*, pp. 92–99, 1998.

[9] X. Geng, K. Smith-Miles, L. Wang, M. Li, and Q. Wu, "Context-aware fusion: A case study on fusion of gait and face for human identification in video," *Pattern Recognition*, vol. 43, pp. 3660–3673, Oct. 2010.

[10] E. Shi, Y. Niu, M. Jakobsson, and R. Chow, "Implicit authentication through learning user behavior," *Information Security*, pp. 99–113, 2011.

[11] C. Nickel, H. Brandt, and C. Busch, "Classification of acceleration data for biometric gait recognition on mobile devices," *In Proc. of the Int'l Conf. of the Biometrics Special Interest Group (BIOSIG'11)*, 2011.

[12] M. Derawi, S. McCallum, and H. Witte, "Biometric access control using Near Field Communication and smart phones," *Proc. of Int'l Conf. on Biometrics (ICB)*, 2012.

[13] R. J. Hulsebosch, M. S. Bargh, G. Lenzini, P. W. G. Ebben, and S. M. Iacob, "Context sensitive adaptive authentication," in *EuroSSC'07: Proc. of the 2nd European Conf. on Smart sensing and context*, Springer-Verlag, Oct. 2007.