Master Thesis

Gesture-based User Authentication on Mobile Devices using Accelerometer and Gyroscope

Dennis Guse

Supervisors:

Prof. Dr.-Ing Sebastian Möller
(Technische Universität Berlin)

Prof. Dr. Michael Rohs
(Ludwigs-Maximilians-Universität München)

Advisors:

Niklas Kirschnick
Sven Kratz

31 May 2011 (Submission)
08 August 2011 (Corrected Version)

Abstract

Mobile devices offer their users lots of possibilities and a feeling of freedom. However, this freedom comes along with new security threats. Sensitive data might be stolen and abused, if an unauthorized person gets unrestricted access to such devices. Therefore, user authentication mechanisms are required. So far authentication mechanisms like PINs and passwords do not take into account the limited capabilities of user interfaces of mobile devices. So, it is necessary to create and develop specially adapted mechanisms, which are designed to be usable under these restrictions.

Prior work has shown that gestural interfaces have potential as gestural interaction is natural to humans. In this thesis a behavioral biometric user authentication mechanism for mobile devices based upon gestures is presented using an embedded 3-dimensional accelerometer and a 3-dimensional gyroscope. As recognition algorithms *Dynamic Time Warping* and *Hidden Markov Models* were studied. The designed mechanism of this thesis was evaluated in a user study including a realistic attack. The conducted user study confirmed that gesture-based authentication is feasible, usable and promising for mobile devices.

Zusammenfassung

Mobile Endgeräte bieten ihren Nutzern unzählige Möglichkeiten und somit ein Gefühl von Freiheit. Doch diese Freiheit birgt auch Gefahren. Zusammen mit den Geräten können sensible Daten entwendet werden, die von unautorisierten Personen missbraucht werden können. Daher sind Authentifizierungsmechanismen essentiell für Sicherheit an mobilen Endgeräten. Bisherige Mechanismen basierend auf PINs oder Passwörter sind den begrenzten Kapazitäten und Nutzeroberflächen dieser Geräte jedoch unzureichend angepasst. Es ist notwendig speziell auf mobile Endgeräte angepasste Authentifizierungsmechanismen zu entwickeln und zu erproben.

In der vorliegenden Arbeit wird ein Authentifizierungsmechanismus für mobile Endgeräte präsentiert, der den Nutzer anhand einer Geste authentifiziert. Das mobile Endgerät nutzt einen eingebetteten Beschleunigungssensor sowie ein Gyroskop, um die Bewegung zu messen. Für die Verifizierung einer Geste werden die maschinellen Lernalgorithmen *Dynamic Time Warping* und *Hidden Markov Model* untersucht. Der für diese Arbeit entwickelte Mechanismus wurde in einer Nutzerstudie getestet, in der auch ein realistischer Angriff simuliert wurde. Die durchgeführte Studie bestätigt, dass gestenbasierte Authentifizierung machbar und für mobile Endgeräte sehr viel versprechend ist.

I Contents

II List of Figures

III List of Tables

IV Abbreviations

DTW	Dynamic Time Warping
EM-Algorithm	Expectation-Maximization-Algorithm
EER	Equal-Error-Rate
FAR	False-Acceptance-Rate
FTA	Failure-to-Acquire
FTE	Failure-to-Enroll
FRR	False-Rejection-Rate
GSM	Global System for Mobile Communication
GPS	Global Positioning System
HCI	Human-Computer Interaction
HMM	Hidden Markov Model
NFC	Near Field Communication
PIN	Personal Identification Number
PMTK3	Probabilistic Modeling Toolkit 3
UMTS	Universal Mobile Telecommunication System
WLAN	Wireless Local Area Network

1 Introduction

Mobile devices support us in our everyday life. The main advantage of these devices is that we can take them with us and use them almost everywhere and at any time: we can check our e-mails, read online news, communicate via social networks and do many other things on the go. To support their owners mobile devices create and store a lot of sensitive personal data. Recently, two things became important for the success of mobile devices: capable hardware and mobile internet access. Today, mobile devices are ubiquitous, but are potentially accessible to unauthorized persons. To avoid misusage it is important to implement and use reliable authentication mechanisms. Widely used knowledge-based mechanisms like PINs and passwords are not well suited for mobile devices as the capabilities of user interfaces are very limited.

It is important to develop new authentication mechanisms specially adapted to the limitations and options of mobile devices. It is necessary to understand the capabilities and limitations of mobile devices beforehand. Also is it necessary to analyze and develop effective countermeasures to new upcoming threats. This chapter gives a short introduction into mobile devices and is presenting possible threats. Afterwards, currently widespread user authentication mechanisms for mobile devices are analyzed and their deficiencies are discussed. All this is important to understand the need for alternative user authentication mechanisms especially designed for mobile devices. At the end of this chapter the objectives and the outline of this thesis are presented.

1.1 Mobile Devices

Having the freedom to move is one of the concepts described by the adjective *mobile* (Authorless, 2004). Therefore, mobile devices like mobile phones, smartphones and PDAs have to cope with certain limitations to achieve this type of freedom. They need to be small and lightweight, so they can be carried around easily. They need to be always operational, even when the user is on the move to allow user interaction whenever desired (Schmidt, et al., 1999). In certain situations the user may not be able to give full attention to device (Hinckley, et al., 2005), e.g. because he likes to interact while walking. In contrary to traditional desktop computers the user interaction with mobile devices is usually very short but happens frequently (Falaki, et al., 2010). These limitations demand specially adapted user interfaces, so mobile devices are usable in a broad range of situations. Currently, available user interfaces consist typically of a display and a keyboard or a touch screen. Additional sensors like motion sensors, GPS, proximity sensors, camera, microphone, compass etc. can provide contextual information and can be used as alternative input devices (Hinckley, et al., 2000).

Most mobile devices have wireless communication capabilities with long range like GSM/UMTS, medium range like WLAN and short range like Bluetooth, ZigBee and NFC. This enables the device, and therefore the user, to communicate anywhere at almost any time. This includes telephony and mobile internet access.

A mobile device with communication capabilities can either wait for inbound communication or start interaction with other reachable devices. To be available for inbound communication a mobile device needs to be switched on permanently. Clearly, mobile devices need to be battery-powered (Jansen, 2003), which limits their operation time and makes energy saving essential. If the battery is empty, the device is almost useless until power supply is available again. Furthermore, the form factor restricts the processing power embeddable due to size and cooling capacity.

Currently, mobile devices are personalized single user systems (Karlson, et al., 2009; Eren, et al., 2006), designed to support the user to solve specific tasks (Schmidt, et al., 1999). User specific tasks are related to storage and retrieval of information, communication, social interaction, entertainment etc. (Karlson, et al., 2009). This requires that such a device needs to store user specific data like contact information, messages etc. Furthermore, build-in functionality can be extended by installing small applications. These so called *Apps* allow the user to adapt the device to his own individual needs.

The ability to use a mobile device almost everywhere includes private and public places. This implies that other persons might be around during usage. Therefore, usage and interaction with mobile devices also depend on the social context (Rico, et al., 2010). In some situations it might be inappropriate to use some functionality such as taking a call during a lecture at university. Also, some features of the user interface may be regarded as inappropriate in some situations.

1.2 Data on Mobile Devices

Mobile devices offer their users lots of possibilities. Considering these, it is clear that mobile devices need store a lot of sensitive user data. Therefore, only the genuine user should be able to read, modify or delete this data. Obviously, the most sensitive data is generated by using the device as a telephone book, calendar or other applications managing personal data. If the device is also used as interface to services like telephony, e-mail, internet, mobile banking (Chong, et al., 2009) and also m-payment (Schwiderski-Grosche, et al., 2002) even more sensitive data might be stored. For many services, like websites and e-mail, password-based authentication mechanisms are used. Usually, the username and the password for a service are stored on the device, so the user is not required to re-enter his credentials every time he uses the service.

Mobile devices also generate user specific data. For instance how often and who the user calls. A device can collect and store data about the individual user interaction and may use this data to support the user, e.g. providing a record of recent calls. Also, mobile devices can collect data using embedded sensors. For example a built-in GPS allows determining the current position and allows tracking the path of the device. If embedded motion sensors are available, it can be determined whether the user is moving and gather information about the current context. Mobile devices need to store a large amount of data to be useful. Furthermore, these devices are able to collect a large amount of sensible user specific data, which may reveal very private individual features of the owner/user.

1.3 Threats to Mobile Devices

As the data on a mobile device is very private and sensitive, it might be highly interesting to other persons, i.e. potential attackers. In contrast to desktop computers mobile devices are designed to be movable and thus not restricted to a specific location. In many cases a mobile device is possibly accessible by unauthorized persons. For example if it is unattended by its owner. Also, an attacker may steal the device or a stranger may find the device after the owner lost it. Therefore, it cannot be safely assumed that the current user is authorized to use the device. In fact, this would be required to guarantee security. The process of identity verification of the current user is called authentication. Without an active authentication mechanism every person with physical access to the device could use it as if he would be the genuine user. This includes access to stored data and also services, which do not require additional authentication or if the credentials are stored on the device.

An attacker might also modify hardware and software of a device, so it functions as intended by the attacker and hand it back to the owner (Baumgarten, et al., 2001). For example a device could be modified, so it sends its current location on a regular basis (Dworschak, 2011). Using this data an attacker would be able to track the owner of the device. Alternatively, an attacker may change the genuine device for a similar one including stored data of the genuine device.[1] So, it is theoretically necessary to verify that the current device is genuine. This is usually not done for practical reasons.

Mobile devices are designed to be single user devices, but sometimes the owner shares the device with other persons freely (Karlson, et al., 2009). In this case, an enabled authentication mechanism would be useless as the owner would unlock the device and then hand it over.

For completeness, an attacker does not necessarily require physical access to a mobile device. An attack may also be possible via enabled communication interfaces. If the communication is unprotected an attacker may be able to eavesdrop, modify, retain and also fake communication (Bishop, 2002 p. 7f.). This class of threats is not new to mobile devices and exists since the emerging of computer networks. However, those threats are much more problematic, because mobile devices are able to communicate with nearby electronic equipment, i.e. the network structure is dynamic (Bichler, et al., 2005). An attacker may be able to attack a device, if he gets close enough and communication is therefore possible. This class of attacks is omitted in this thesis, because they are not related to user authentication. Nevertheless, they are real security threats.

[1] An example is given in the movie "Enemy of the State": the mobile phone of the protagonist Robert Dean (Will Smith) is exchanged against a modified similar looking version to spy on him.

1.4 Available User Authentication Mechanisms

> "The mobile […] device must be able to protect its information and control the processes that can access and use it. It may be assumed that the powerful, static server can fairly easily look after itself using established technologies and procedures, but the designer of a trusted mobile device must be very conscious of the practical concerns of costs and power consumption, portability and usability in addition to basic technical feasibility." (Hulsebosch, et al., 2003 p. 385)

For mobile devices diverse mechanisms for user authentication are available. Broadly in use are knowledge-based mechanisms like PINs and passwords. Those mechanisms are easy to implement, computational inexpensive and often use already available capabilities of the user interface. Nevertheless, they have several drawbacks regarding usability. The genuine user needs to memorize an arbitrary and complex secret, which he can recall precisely. It is important that a secret is not guessable by other persons and therefore might be hard to remember. Knowledge-based mechanisms have another drawback on mobile devices, which is related to the user interface. The user needs to enter the secret during the authentication process, but the capabilities of the user interfaces are limited. One example are touch screens, which use the same space to allow visual output and touch input, but do not provide haptic feedback. Therefore, the user needs to look at the screen, while he enters his secret on a virtual keyboard, to see if he hit the intended key. The input might be more time-consuming and exhausting to the user. This is problematic, because the user needs to be authenticated every time he engages interaction with the device to achieve security.

A drawback directly related to knowledge-based mechanisms is that every person who knows the secret is able to authenticate successfully. This is indeed problematic, because an attacker may be able to observe the genuine user during authentication and able to acquire the secret.

Alternative approaches to user authentication are stroke-based mechanisms using a touch screen (Weiss, et al., 2008). This approach may be regarded as an extension to PIN-based authentication with the difference that the user remembers a shape of strokes than a number. The user enters the shape by applying a sequence of strokes using his fingers. An attack to such authentication mechanisms is possible by analyzing smudge residues on the touch screen (Aviv, et al., 2010). The residues might be a result of a successful authentication and may give an attacker hints about the genuine shape.

Another alternative are graphical passwords. A graphical password does not consist of digits and letters, but of one or more images. One approach is to select a correct sequence of presented images or alternatively to locate special points in a presented image (Jansen, 2004). The usage of images as a secret allows the genuine user to attach meaning more easily and further might be a cue that supports the user to recall. Graphical passwords can be used on mobile devices, but require a certain screen size and resolution, and a sufficient method to select the images or points. One major drawback is that the user needs to give full visual attention to the device for authentication. Another problem is that images might give hints about the secret to an attacker and therefore should be protected (Hayashi, et al., 2008). One approach is to present images blurred (Hayashi, et al., 2008).

Another alternative are mechanisms based upon biometric features. Proposed for mobile devices were fingerprint, face and voice identification (Haze, et al., 2007). Mobile devices with built-in fingerprint scanner are already available on the market.[2] However, for diverse reasons like costs and size, biometric authentication mechanisms are currently not very common.

An approach to overcome the single user limitation of mobile devices is presented in Ni et al. (Ni, et al., 2009). They propose to implement complex user access controls, so functionality and access can be restricted depending on the rights of the current user.

The general problem of user authentication is that it is not the goal of the user to authenticate, when using a mobile device. The primary goal of the user is interaction with the device to carry out his tasks.

1.5 Objectives

In this thesis a gesture-based authentication mechanism for handheld mobile devices is developed and evaluated. The user authenticates by doing a hand gesture while holding the device one-handed. Prior work (presented in detail in section 2.3) has shown that gesture-based user authentication is feasible and has some advantages over widespread knowledge-based mechanisms. In prior work with handheld devices 3-dimensional accelerometers were used.

[2] The Motorola Atrix is equipped with a fingerprint scanner.

The gesture-based authentication mechanism implemented in this thesis uses a 3-dimensional accelerometer and a 3-dimensional gyroscope embedded into a mobile device.[3] The combination of those sensors should lead to more precise measurements and is likely to increase the performance. The measurements of those sensors are evaluated by using machine learning algorithms namely *Dynamic Time Warping* (DTW) and *Hidden Markov Models* (HMM).

This thesis has two goals. First of all, to prove gesture-based authentication on mobile devices is feasible and is more usable than available knowledge-based mechanism on mobile devices. Second, to prove that gesture-based authentication is possible using hardware, which is available on the end-user market.

1.6 Outline

This work is structured as follows. After a general introduction about mobile devices and the objectives in this chapter, basic knowledge and related work are presented in chapter 2. This chapter starts with an introduction into gestural interfaces and gesture recognition techniques. Afterwards, an introduction to biometric user authentication is given. Chapter 2 closes with a presentation of related and prior work on gesture-based authentication. In chapter 3 the requirements of user authentication mechanisms especially designed for mobile devices are analyzed and the implemented gesture-based authentication mechanism is presented. At the end of chapter 3 potential attacks to gesture-based authentication are presented. The applied variants of machine learning algorithms are presented in chapter 4.

The implemented mechanism was evaluated in user study, which is presented in chapter 5. The user study allows drawing conclusions about the feasibility and about the performance of the used algorithms. The results of the user study, including user survey and machine learning algorithms, are presented and discussed in chapter 6. Chapter 7 closes this thesis with a summary and an outlook on future work.

[3] Farella et al. present a gesture-based authentication mechanism using a 3-dimensional accelerometer and point out that additional sensors like a gyroscope might increase the performance (Farella, et al., 2006).

2 Related Work

In this chapter an introduction into gestural interfaces and gesture-recognition techniques for human-computer interaction is given. Afterwards, biometric user authentication and prior work on movement-based authentication mechanisms are presented. This section also includes prior work on gesture-based authentication with embedded motion sensors.

2.1 Gestural Interfaces

For human-to-human interaction gestures provide an additional and alternative channel to spoken communication. This type of interaction is natural to humans and therefore very promising for human-computer interaction.

2.1.1 Gesture Definition

Gestures are used for interaction (Mitra, et al., 2007). Interlocutors do not communicate by speech only, but also by expressive body movements including arm-, hand- and facial expressions etc. This adds additional information on the topic and may point out emotional aspects. In fact, even the exclusive use of gestures can make a communication. The interpretation of gestures is ambiguous (Mitra, et al., 2007). It depends on the social, cultural and emotional context of the communication partners. For gesture interpretation the spatial, pathic, symbolic and affective information are considered useful (Mitra, et al., 2007).

Gestural interfaces for HCI can either use personalized or discrete gestures. In the first case, the interface adopts the gestures of the users. In the second case, the gestures are defined by the user interface and the user needs to adapt them. In contrary to knowledge, movements and also gestures are not learned but rather trained. To adapt a gesture the user trains the movement by repetition. In the training process the way to move is stored implicitly in the motor cortex (Klemmer, et al., 2006). If the user likes to execute a trained gesture, the stored information is used and he does not need to remember how to move. The motor cortex does not tend to forget and is not prone to information overload as the memory. The learned movement is bound to the individual and can only be picked up by another person by training. Nevertheless, a user needs to remember, which action will be initiated by each gesture for successful interaction using a gestural interface.

Gestures can be classified as either static or dynamic. Gestures implied in a specific posture are called static gestures. Dynamic gestures consist of a movement instead. These can be divided in pre-stroke, stroke and post-stroke phase (Mitra, et al., 2007).

2.1.2 Gesture Recognition Techniques

Gestural user interfaces need to measure the user movements as results of gestures for interaction. This can either be done by capturing the gestures from distance or by attaching sensors directly at the user. Both approaches have their advantages and drawbacks.

To measure a movement from distance usually video recordings techniques are used, which allow tracking the absolute path of a movement. This approach does not require the user to wear special equipment and is therefore less intrusive. But it requires preparing the environment in which the gestural interaction should take place. Thus, video recording techniques for gesture recognition are not well-suited for mobile devices, because in most cases they cannot be used spontaneously everywhere. Furthermore, the evaluation of video recordings is computational complex.

Attaching sensors directly to the user is more intrusive, but avoids preparing the environment. Common examples of such motion sensors are accelerometers and gyroscopes. Today, these sensors are small, cheap, and lightweight and consume only little energy. They can be embedded into a device to measure the movement of the device and use the measurements to deduce the user movements. Of course, both approaches can be combined. The user wears special equipment, which allows measuring the movement from distance. This can be done for examples by using an embedded compass. The compass is used to measure the position and the path of a magnet attached to the hand of the user (Ketabdar, et al., 2010). The user interacts with the device by moving his hand around the device, so the compass can measure the induced changes of the magnetic field.

However, measuring a gesture is not sufficient, because the measurements need to be interpreted. One task is to decide from a known set of gestures the one, which fits best to the measurements. An alternative task is to extract the embedded information of a gesture from the measurements (Wilson, et al., 1999). There are several obstacles to gesture recognition and interpretation. Sensors cannot perfectly measure a movement as result of a gesture, because they are limited by their accuracy and resolution (Mitra, et al., 2007). They are also prone to noise (Niezen, 2008 p. 18). Furthermore, users introduce variance, because it is very unlikely that they can repeat exactly the same movements. Differences in shape and most notably in timing will occur. The current movement of a gesture depends on the mental and physical state of the user. The movement also differs from one person to another, which is problematic for gestural interfaces using discrete gestures. Another difficulty is the segmentation. To interpret a continuous stream of measurements, it is necessary to determine the start and end points of the measured gestures. This can be done automatically, but is a computational complex task, or manually by the user.

To interpret the varying measurement of gestures usually machine learning algorithms are applied. Usually, the task is to learn automatically, how to distinguish different gestures using a provided set of training samples. This is a classification task, where for each class samples are provided and the learned characteristics of the samples are used to assign the correct class to unknown samples. However, not all algorithms are equally suited for gesture recognition, because they differ in their approach to find and learn the characteristics of the provided training samples, but also in their computational complexity (Marsland, 2009 p. 5ff.).

2.1.3 Gesture Interaction on Mobile Devices

Gestural interfaces are well suited for mobile devices, because they allow interaction in a natural manner and allow overcoming some limitations of the user interface. A gestural interface for mobile devices enables the user to input commands eyes-free, i.e. without requiring that he needs to look at the device.

One way of gestural interaction is presented by Williamson et al. (Williamson, et al., 2007). In order to get information about the battery level the user shakes the device. Instead of a visual feedback, the device emits a sound comparable to water in a container. The level of water depends on the available energy. A more general interaction scheme for handheld mobile devices is presented by Kallio et al. (Kallio, et al., 2009). The user interacts with the device by turn-based hand gestures, which use the ability of humans to rotate the hand with the wrist and the forearm. The usage of rotation has the advantage that the gestures do not require a large space. This makes this approach very well suited for mobile devices, because in crowed places space is a limited resource. A gesture-based interface using a mobile device as a remote control for ambient services is presented by Westermann (Westermann, 2010). The user enters commands via hand gestures on a mobile device, which interprets the gestures and relays the commands to ambient services.

Gestural interfaces for user authentication, i.e. prior work on gesture-based user authentication using embedded motion sensors, are presented in section 2.3.4.

2.2 Biometric User Authentication

In the following, an introduction to biometric user authentication is given. This section starts with a short introduction into security and user authentication in general. Afterwards, the overall performance of authentication mechanisms is discussed and a classification scheme of attacks to biometric authentication is presented.

2.2.1 Security and User Authentication

Security is about assumptions and trust (Bishop, 2002 p. 10). A system can be depicted as a state-machine with secure states, insecure states and state transitions (Bishop, 2002 p. 10). The system definition includes all potential interacting subjects - including users and machines - and communication channels. If no state transition from any secure state to any insecure state exists, a system is called secure (Bishop, 2002 p. 10f.). The goal of security is to prevent and detect insecure transitions and further to recover from transitions to insecure states (Bishop, 2002 p. 10f.). A system needs to be known completely to guarantee this, which in realistic use-cases almost impossible. Therefore, security policies and mechanisms are applied. Security policies describe what is allowed and forbidden in a system (Bishop, 2002 p. 11) and are enforced by security mechanisms (Bishop, 2002 p. 9). Policies and mechanisms are assumed to be precise and correct. In fact, if this assumption does not hold, the system is insecure.

With regard to data, security leads to integrity, confidentiality, availability, authenticity and non-repudiation (Bishop, 2002 p. 4ff.). *Integrity* is given, if only authorized parties are allowed to modify data. *Confidentiality* means that only authorized parties are allowed to access data. In fact, confidentiality is related to privacy. *Availability* describes that the data should be available, if authorized parties try to access it. Authenticity and non-repudiation are closely related. While *authenticity* means that the identity of a party is provable, *non-repudiation* describes that it is provable which party is accountable for an action (Baumgarten, et al., 2001).

Security also includes that the current user can only access data and functionality to which he is authorized. At first usage, an identity for a genuine user is created. The identity describes what the user is allowed to do and how he can prove that he is indeed the one, who he claims to be. The creation of a relationship of trust is called *enrollment process* (Stallings, et al., 2007 p. 75f.). The assumption is made that only the genuine user is able to provide the proof for the identity during authentication (Stallings, et al., 2007 p. 75f.). The *authentication process* consists of two steps. In the identification step the current user has to communicate the identity he intends to use. In the following, verification step he needs to provide the proof for the selected identity. If the proof is valid, the identity is accepted. The identification step can be omitted, if the proof is unique in the system. In this case, the authentication process is called identification. In fact, the authentication process is done more often than the enrollment process.

Authentication can either be discrete or continuous. It is called discrete, if after a successful authentication the access is allowed for certain period or until the user logs out. Continuous authentication mechanisms verify the user throughout the complete interaction. This is usually done in background, because otherwise it would require too much attention of the user and distract him from his tasks. In fact, continuous mechanisms do not need to assume that the user stays the same after successful authentication as for discrete mechanisms.

The proof for an identity can be based upon what a user knows[4], possesses, what he is or how he behaves (Stallings, et al., 2007 p. 75f.; Renaud, 2005). Knowledge-based approaches use a secret known by the genuine user(s) and the authentication mechanism only. Authentication mechanisms based upon possession assume that only the genuine user can provide a unique artificial token. If it can be safely assumed that the token cannot be copied and the genuine user possesses the token, then only he is able to authenticate successfully. Therefore, those mechanisms provide verifiable security.

Biometric features measure the characteristics of the user. These allow verifying the user with regard to what he is or how he behaves. Features that use measurable attributes of the user's body are called *physiological*. In fact, parts of the body are used as biological tokens like fingerprints or the iris. Usually, physiological biometric features cannot be changed easily. Once, the specific characteristics of used physiological features of a user are exposed to an attacker, he might able to create forgeries. In this case alternative attributes need to be used to achieve security. Furthermore, physiological features might reveal information about the user's body that may be regarded as private. *Behavioral* biometric features measure how the user behaves. These features depend on the physiological attributes, upon the knowledge and habits of the user (Yampolskiy, et al., 2008). Mechanisms based upon biometric features try to proof that the genuine user is present, whereas knowledge- and token-based mechanisms need to assume it.

A biometric authentication mechanism uses one or more sensors to acquire the used features of the current user. For behavioral biometric sensors must not be physical, but can also be virtual like issued commands on a command shell interface (Yampolskiy, et al., 2008).

[4] Renaud differentiates between knowledge-based authentication mechanisms by their type of recall or recognition (Renaud, 2005).

In contrast to knowledge-based authentication mechanisms the measurement and the evaluation of biometric features is difficult, because people are human beings, who are not static, i.e. are changing over time. In addition, the authentication might be required in different situations. Therefore, different measurements of the used features are very unlikely to be exactly similar. Thus, variance needs to be taken into consideration, when comparing the current measurement with the enrollment measurements to decide if the current user is genuine.

2.2.2 Requirements on Biometric Features

Not all possible biometric features are useful. A useful biometric feature should have a low intra-class variation, whereas the inter-class distance should be large. This means that samples of one person should differ only slightly whereas samples from different persons should be very different. As the intra-class variation increases the probability of accepting false samples as genuine increases as well and forgeries become easier (Fierrez, et al., 2007). For identification the characteristics of used feature(s) need to be unique for each genuine user (Yampolskiy, et al., 2008). This introduces a far greater need for a large intra-class distance, because otherwise the chance of misidentification increases.

In general, a biometric feature should be more or less permanent, so it does only change slowly over time (Yampolskiy, et al., 2008). Otherwise the genuine user would be required to repeat the enrollment process every time, when his characteristics changed to far. A further requirement for biometric features is that all potential users should possess the feature. This quality is called universality (Yampolskiy, et al., 2008). In addition, it is important that a biometric feature can be reliable measured (Yampolskiy, et al., 2008) and this process is fast and convenient for the user (Ross, et al., 2007 p. 7; Yampolskiy, et al., 2008).

2.2.3 Risks to User Authentication Mechanisms

User authentication mechanisms are used to verify the identity statement of the current user, but they are not perfect. In the following, risks to user authentication mechanisms are presented.

An authentication mechanism, which requires revealing the proof during the authentication process, may enable an attacker to acquire knowledge about the proof (Renaud, 2005 p. 105f.). This knowledge might increase his ability to attack the identity successfully. This is especially problematic for knowledge-based mechanisms, because a successful attack only requires knowledge of the secret and access to the mechanism. To guarantee a certain level of security it is either required that the proof is concealed during input or that creating a successful forgery is not achievable by average attackers.

Another problem is that most user authentication mechanisms do only prove the validity of the user's identity statement, but cannot verify the motivation of the user. The successful authentication of a genuine user does not guarantee that he will not break system security. For example a user may be forced or tricked to authenticate by an attacker, so that the attacker gains access to the system (Yampolskiy, 2008) or the genuine user attacks the system from inside (Bishop, 2002 p. 21).

Furthermore, the goals of security and usability are conflictive. The goal of security from the usability perspective is "making undesirable actions more difficult" (Kainda, et al., 2010 p. 1) while the goal of usability is to make things easier (Kainda, et al., 2010). From the user perspective authentication mechanisms are often seen as annoying avoidable obstacles, because usually the user needs to actively interact with the mechanism, which interrupts his workflow.

> "If security and/or privacy and usability collide, then usability always wins!" (Mayrhofer, 2007)

Users may disable annoying mechanisms, if they are allowed to (Mayrhofer, 2007 p. 7). They may also not activate mechanisms, which are off by default as they are not aware of them. It is often not sufficient to simply add an authentication layer to the user interface, but rather integrate it thoroughly and consistently during system design (Renaud, 2005). In fact, users are often not aware or underestimate the risks, which user authentication mechanisms should counter. Those users also tend to disable these mechanisms (Adams, et al., 1999). Also weak user authentication may degrade system security. The users may perceive a false feeling of security, although it is not guaranteed. A user who feels secure will interact with the system in a way, which demand high security standards. Furthermore, users can also undermine security by insecure behavior like writing passwords down (Adams, et al., 1999). Security is not only a technical but also an organizational issue, because users are part of the system (Bishop, 2002 p. 21).

> "The end-user plays a vital role in achieving system security. If security systems are designed to accommodate the average user's needs and limitations, it is more likely that the security system will succeed." (Renaud, 2005 p. 103)

2.2.4 Performance of User Authentication Mechanisms

The performance of user authentication mechanisms is manifold and depends on the situation. To assess the characteristics and infer the performance Renaud developed four categories: accessibility, memorability, security and costs (Renaud, 2005). They are also called deficiencies as they are composed of potential deficits.

Accessibility comprises requirements of the mechanisms, convenience of usage and inclusivity of users. Requirements are special hardware, software and technical knowledge of potential users. Convenience of usage is related to usability of the mechanism including the enrollment, authentication and proof replacement. Inclusivity of users describes the requirements of a mechanism that potential users need to fulfill. A mechanism can only be used by users, who have all physical, cognitive, and sensory capabilities.

Memorability includes the depth of processing at enrollment, retrieval strategy at authentication and meaningfulness of the information to memorize. Recall without cues, recall with cues or recognition can be used as retrieval strategies. From the usability perspective it is preferable to either use recall with cues, recognition or support the user to attach meaning to the information he needs to memorize.

Security is composed of disclosure, predictability and confidentiality, breakability and abundance, and privacy. Disclosure means that the genuine proof can be willingly or unintentionally disclosed to other persons. This also includes disclosure during the authentication process. Predictability is given, if an attacker can acquire enough knowledge about the genuine user, so he is able to guess the genuine proof. Confidentiality describes how the proof is protected during the authentication process. Breakability designates the possibility to break a mechanism without any prior knowledge. Abundance describes how much time is necessary to break the system. From the security perspective those two deficiencies are often regarded as important only. Privacy is about the information that a genuine user needs to provide about him, so he is able to use a mechanism. This includes any personal information about the genuine user, which needs to be available during authentication or is saved by the mechanism.

In practice not only the presented deficiencies are important, but also the *costs* (Renaud, 2005). The overall costs include development, deployment and also maintenance of a mechanism. In addition, the costs include further expenses due to broken system security.

The above presented performance metric is very general and can be used to evaluate and compare different authentication mechanisms, so the best fitting can be chosen for a certain situation. However, the initial performance evaluation of a biometric authentication mechanism is usually done more practical by only regarding feasibility, i.e. the security perspective.

The widely used metrics are presented in the following: the Failure-to-Enroll (FTE), Failure-to-Acquire (FTA), False-Rejection-Rate (FRR) and False-Acceptance-Rate (FAR) (Yampolskiy, et al., 2008). The FTE measures the rate of users, who cannot use a mechanism, because the users do not possess required feature. The FTA measures the times users could not authenticate, because the system could not acquire the features temporarily. The FRR measures how often genuine proofs are rejected as false and thus authentication of a genuine user fails. The FAR measures how often a mechanism accepts false proofs as genuine. In this case, the authentication process succeeds though the current user is not genuine.

The FAR and the FRR are closely related, because both depend on the allowed variation. To visualize the trade-off between FRR and FAR the Receiver operating characteristic (ROC) is used. To select the parameters for a mechanism, i.e. the allowed variation, usually the Equal-Error-Rate (EER), i.e. FAR and FRR are equal, is used.

The above presented metrics for biometric user authentication can be approximated in user studies only, because this type of information is not available in real situations. For this reason, the conclusions about an authentication mechanism depend severely on the setup and the circumstances of a conducted user study. As performance indicators for biometric mechanisms the FAR and FRR are often used only without regarding sophisticated and realistic forgery attempts (Zoebisch, et al., 2003). Omitting those in the evaluation will lead to better results, but will lead to false conclusions about the security performance. To evaluate the security performance thoroughly it is required to assess the resources, capabilities and skills of potential attackers and study open-minded possible attacks and use this findings in user studies.

2.2.5 Classes of Attacks to Behavioral Biometric User Authentication

Authentication mechanisms should protect a system from potential attackers. In the following, a classification scheme of attacks is presented depending on the knowledge available to the attacker. For a successful attack on a biometric authentication mechanism, an attacker needs to be able to create a forgery, so the sensors provide measurements similar to the proof under attack. The intensity of possible threats differs depending on the skills, motivation and time available for preparation of a potential attacker. A threat intensifies the more an attacker knows about the function of the mechanism and the more resources are available to him. Motivation could be sufficient to acquire skills and knowledge, but time and resources are always limiting factors.

Any knowledge about the genuine proof for the identity under attack increases the threat. If all these are available to the attacker, he can successfully overcome the authentication mechanism and it does not provide security against this attacker (Ballard, et al., 2007).

Attacks are called *naïve*[5], if an attacker has a certain but very limited motivation to attack a system and thus needs to break the authentication mechanism in place, but has no information about the genuine proof to forge and is not able to acquire this knowledge. This type of attack is often used in proof-of-concept evaluation of authentication mechanisms, where all participants are regarded as genuine users. All non-genuine samples are simply used as forgeries. However, these forgeries were created without the intention to attack (Ballard, et al., 2007). In fact, Ballard et al. stress that this verification cannot be used to estimate the realistic security performance, because a real attacker would be more likely to acquire knowledge about the genuine proof. Therefore, naïve forgery only allows a very limited evaluation of real world attacks. A forgery from an attacker, who has information, resources and motivation to attack, is called *skilled forgery* (Ballard, et al., 2007). A skilled forgery should always be more likely to be accepted in the authentication process than a naïve forgery. One class of attack cannot be prevented by biometric authentication: *coercive attack*. The attacker threatens the genuine user to authenticate and grant access to the attacker (Yampolskiy, 2008). One case of coercive attack is blackmailing (Baumgarten, et al., 2001).

To evaluate realistic achievable security of a biometric authentication mechanism, it is first of all necessary to identify main usage scenarios and then identify potential classes of attackers and their abilities. Using this knowledge it is possible to evaluate the resistance of the mechanism to attacks under more realistic conditions. The FAR and FRR of different authentication mechanisms are only comparable, if the studied attacks are comparable as well (Zoebisch, et al., 2003).

2.3 Movement-based User Authentication

In the following, movement-based user authentication and identification techniques are presented. The genuine user proves his identity by a measurable movement assuming that only the genuine is able to provide samples of the genuine movement. In the following, only movement-based approaches are presented, which require the user memorize a movement that he can repeat similar later on and the user is able to decide, when to do the movement. This excludes mechanisms like gait recognition as presented by (Mäntyjärvi, et al., 2004).

[5] The term *naïve forgery* is introduced in (Ballard, et al., 2007) and also called *random forgery* or *zero-effort forgery*. (Yampolskiy, 2008) calls it *blind forgery* and (Zoebisch, et al., 2003) *accidental forgery*.

Before presenting movement-based user authentication mechanisms, mechanisms for movement-based device pairing are presented. In section 2.3.4 prior work on gesture-based user authentication with personalized hand gestures is presented that is a subclass of movement-based authentication.

2.3.1 Excurse: Movement-based Device Pairing

Electronic devices often communicate via anonymous channels and therefore the source of a message is indistinguishable (Castelluccia, et al., 2005 p. 6). To integrate the goals of information security, it is necessary to establish a relationship of trust between two or more devices at first encounter. This is called pairing. It is equivalent to the enrollment process of user authentication. First of all, the user needs to select the devices, which he likes to pair. Second, he embodies a secure communication channel between the devices to protect the initial communication used to exchanges device dependent secrets or he makes one secret available to all devices. Using the secret(s) the relationship of trust is established. An example for pairing is the Bluetooth pairing process. This approach requires that the devices have a user interface with sufficient interaction capabilities to input a secret. For some classes of devices this might be not adequate or even possible. For example Bluetooth headsets do often use a fixed secret - usually 0000 -, because they have only a very limited user interface. However, a fixed secret is not secure and therefore not adequate for a pairing process.

In the following, alternative approaches using user movements as secrets for pairing are presented. Holmquist et al. present a pairing technique using one movement as a secret (Holmquist, et al., 2001). To pair two devices the user picks them up and moves them simultaneously in similar manner, e.g. the user holds both device in one hand and shakes them. Assuming that only those devices are moved, which should be paired, the movement is regarded as a secret. In fact, the devices need to be able to measure the movement and produce comparable readings. Holmquist et al. used accelerometers embedded in each device. The selection procedure of the devices to pair is quite natural, because the user simply picks up the devices. The user does not need to memorize the movement, because it is used as one-time secret and using similar movements for multiple pairing processes might make them attackable. A very similar approach is presented by (Mayrhofer, et al., 2009).

Patel et al. created a movement-based pairing technique to pair a mobile device with a public terminal for a limited time (Patel, et al., 2004). This kind of pairing might be necessary, if the terminal should be used as input device for the mobile device or to grant the terminal access to data stored on the mobile device. This approach requires that the terminal and the mobile device can communicate directly with each other. The mobile device announces its presence to the terminal and the user chooses his mobile device on the public terminal, which requests the user to move the device in a specific way. The terminal then connects to the mobile device and requests the device to send the measurements of the movement. The temporal pairing process succeeds, if the measurements are accepted as similar to the requested movement.

An attack to both paring approaches is outlined in (Mayrhofer, et al., 2009). An attacker needs to measure the movement, which is used as secret. Then he might be able to use the secret to interfere with the pairing process and establish a false relationship of trust or acquire other sensitive information. Nevertheless, it is a promising approach for pairing.

2.3.2 User Movements for Machine Interaction

In the following, authentication mechanisms based upon user movements for machine interaction are presented. All established user interfaces use user movements as input method. Classic examples are buttons, keyboards, pointing devices and touch screens. Usually, only a small portion of available information is used for human-computer interaction like which button was pressed or where the user clicked. In fact, the unused portion is interesting for authentication as how the user uses the input devices.

Wobbrock proposes an authentication mechanism called TapSong using a single button (Wobbrock, 2009). The user authenticates by pressing the button rhythmically similar to a chosen rhythm in the enrollment process. Rhythms are used, which are based upon songs, to support the user in the memorization tasks. This approach is well suited for mobile devices, because a single button is required only and the user can authenticate eyes-free. Furthermore, the movement for authentication can be very small and only poorly observable by potential spectators. Also, the user can authenticate while the device in his pocket and thus the proof can be concealed completely.

Alternative approaches considered useful are mechanisms using keystroke dynamics (Clarke, et al., 2002). Those mechanisms do not use the content that the user types, but rather verify an identity by the way a user interacts with the keyboard. Examples of useful features are the time of key presses and the inter-key time. Keystroke dynamics features can either be measured by letting the user enter a static text or use input, which the user provides during normal interaction. In the second case the authentication can be done continuously in background. Keystroke dynamics can be used on all devices, which use keyboards or keypads as input devices. In fact, it is required that the user types and therefore limits the usefulness for mobile devices.

An approach to extend PIN-based authentication for mobile devices by using a pressure sensitive touch screen as input device is presented by Saevanee et al. (Saevanee, et al., 2008). Usually, a button provides information about his current state only, i.e. if it is pressed or not. However, the button concept can be extended, so the applied pressure is taken into account. In a user study an EER of 1% was achieved without regard to skilled forgery. Another approach using pressure sensitive buttons was studied by Igarashi et al. for car driver identification using the driving style (Igarashi, et al., 2004). The driving style is measured using the interaction with the acceleration and brake pedal in background.

2.3.3 Written Signature

A hand written signature of a person is one common way to proof the identity. Hand writing and written signatures are regarded as individual to a person. In practice the written signature is used to sign contracts etc. A signing person states that he knows, understands and accepts the content of the signing paper. The signature is legally binding and forgery prohibited by law.

A signature can be verified using offline or online information. Offline verification mechanisms use the resulting image of the signature only. This image can be compared to a genuine template to measure the similarity. Online verification mechanisms do also use information on how the image was created. This information might include velocity, movement changes and pressure. Such information can be acquired by using graphic tablets etc. as input devices rather than a pencil and paper. Automatic approaches for online signature verification are presented in (Alpcan, et al., 2008), (Fierrez, et al., 2007) and also (Fang, et al., 2005). Online information adds complexity to forgery, because not only the resulting image needs to be similar, but also the process of creation. A generative approach for automatic forgery of hand writing using written examples is presented in (Lopresti, et al., 2005).

2.3.4 Hand Gestures

In the following, mechanisms, which use one-handed gestures for authentication measured with embedded motion sensors in handheld devices, are presented. This is prior work on gesture-based user authentication.

A mechanism using discrete gestures mapped to numbers for PIN input on mobile devices is presented by Chong (Chong, 2009; Chong, et al., 2010). A user enters each digit of the PIN by applying the corresponding gesture to the mobile device. The used gestures are short one-handed movements, which start and end all in the same position. The user needs to input each gesture accurately, so recognition is successful. In fact, during authentication the user needs to disclose is secret PIN gesture and it can be assumed that an attacker gets to know the discrete set of gestures, because they are defined by the mechanism. This may enable an attacker learn the PIN by observing the gestures easily. He can use this knowledge to mimic the genuine order of gestures as he is not required to forge the movements of the genuine user exactly, but only to perform the discrete sub-gestures. Therefore, the authentication mechanism should only be used in secure locations.

In the following, authentication mechanisms are presented, which use personalized gestures and therefore overcome the drawback of discrete gestures. A written signature may be regarded as a personalized hand gesture, which is projected on a 2-dimensional space. In fact, gesture-based user authentication is sometimes referred as *gesture signature* (Farella, et al., 2006).

Okumura et al. proposed a user authentication mechanism by shaking the device one-handed (Okumura, et al., 2006). The mechanism requires that the device is shaken along one dimension of the used 3-dimensional accelerometer. In a user study with 22 candidates an EER of 5% was achieved. The approach was refined by adding an update procedure (Matsuo, et al., 2007) to the enrollment samples to increase the long term performance. This is necessary, because the user refines the trained movements over time. Liu et al. created a gesture-based authentication mechanism using a Nintendo Wii Controller (Liu, et al., 2009). The mechanism distinguishes between non-critical and critical authentication. Critical authentication is necessary, if the mechanism is used to protect a critical system. In fact, this imposes higher requirements on security and thus the complexity of the gestures is limited by a lower bound (Liu, et al., 2009). To study skilled forgery for critical authentication video recordings were used. Potential attackers may consider this approach promising, because the gesture seems easy to observe. Overall, a FRR of 5% against a FAR of 10% was achieved for critical authentication. For non-critical authentication, i.e. without skilled forgery, a FRR between 1% and 11% depending on the genuine user was achieved. Guerra Casanova et al. developed a mechanism using an iPhone 3GS (Guerra Casanova, et al., 2010). They also simulated attacks via video recordings. Overall, an EER of 2.5% was achieved. So far, all presented mechanisms were evaluated with user studies, where the participants were allowed to choose their own gestures, and mechanism used DTW. Farella et al. present a mechanism for user identification using feature extracting techniques (Farella, et al., 2006). They achieved an accuracy of 95% using prior defined gestures, which the users needed to interpret individually.

A very different approach to gesture-based user authentication and identification is presented in (Ketabdar, et al., 2010). It is named MagiSign and does not use built-in motions sensors for gestural input, but a compass and a magnet. MagiSign uses an integrated 3-dimensional compass in an iPhone 3GS to measure magnetic changes generated by a small magnetic rod, which the user holds like a pen. With this pen the user authenticates by "writing" his proof in the air. This approach is metaphorical similar to the written signature, but 3-dimensional movements are possible. In a user study a FRR of 4.8% against a FAR of 0.3% was achieved.

All the presented mechanisms with embedded motion sensors used a 3-dimensional accelerometer. Therefore, a mechanism using the combination of accelerometer and gyroscope to measure the movement should be performing even better as presented in this thesis.

3 Requirement Analysis and Approach

Prior work has shown that gesture-based authentication with hand gestures and embedded accelerometers is feasible. Gesture-based authentication is especially promising for mobile devices, because the capabilities of the user interfaces are limited. Personalized hand gestures are behavioral biometric features and have therefore some promising characteristics. The genuine user is not required to memorize an arbitrary secret. Instead, he trains a movement, which is stored implicitly in the motor cortex. Therefore, the proof cannot be handed over to other persons easily. Gesture-based authentication mechanism might be a useful alternative to available knowledge-based mechanisms for mobile devices.

For the initial evaluation of the gesture-based authentication mechanism four questions are considered to be important and need to be studied in detail:

1. Are the measurable biometric features user-specific enough for authentication?
2. What does a gesture-based mechanism demand from potential users?
3. How usable is a gesture-based mechanism from the user perspective?
4. Which kinds of attacks are possible and realistic?

These questions were used as guidelines in this thesis. Question 1 is about the general feasibility at all. Hints to this question were given by prior work on gesture-based user authentication using accelerometers as presented in section 2.3.4. This thesis aims to prove that gesture-based authentication is possible on mobile devices using embedded motion sensors and that a gyroscope improves the security performance. Before discussing question 2 and 3 in detail, the requirements for authentication mechanisms on mobile devices need to be analyzed, which is done in the following section. This is of great importance as the user perspective may reveal requirements, which are not directly visible but very relevant. Afterwards, the gesture-based user authentication mechanism of this thesis is presented in general. This section also answers question 2 and 3. At the end of this chapter potential attacks are discussed.

3.1 Requirements for User Authentication Mechanisms on Mobile Devices

The usability of an authentication mechanism can either *"make or break system security"* (Renaud, 2005 p. 108). Therefore, it is not sufficient to analyze the security perspective only, but also the usability perspective. To study the latter it is necessary to focus on the user, i.e. identify main users and their main use-cases. However, this is nearly impossible for mobile devices, as they should be usable by everyone and in almost every situation. In the following, requirements induced by the characteristic of mobile devices are presented, which seem to be important for the usability of authentication mechanisms on this class of devices.

Mobile devices are movable and can be used almost everywhere. This implies that usage of mobile devices is not restricted to private and/or secure places. To be useful interaction needs to be possible in unknown and/or public places. Most notably public places can be considered as possibly insecure, because other people may be around. A widely useful authentication mechanism for mobile devices should not make any assumption and/or requirement about the location, where authentication might be required. Nor should an authentication mechanism require actions of the user, which may be regarded as insulting or aggressive by other persons. Every action required should be absolutely appropriate.

From the user perspective, an authentication mechanism needs to be fast and easy-to-use. When a user starts interacting with the device, he aims at fulfilling a certain task. His goal is not to prove his identity to the device. A mechanism needs to be discrete enough, so it is an acceptable interruption before the intended interaction can take place. Furthermore, an authentication mechanism should be designed in a way that the limitations and capabilities of the user interface are taken into account. Otherwise, the authentication process might be annoying and exhausting. Also, an authentication mechanism should not require full attention during authentication, because in some situations this might be not possible. In addition, authentication should also be possible while the user is moving. For example a user may be required to authenticate while he is walking.

Another limitation is introduced by the limited capacity of built-in batteries. An authentication mechanism is probably used very frequently and should consume only minimal energy per authentication. Otherwise, the usage of the mechanism would decrease the runtime of the device and might therefore be annoying to the user.

From the user perspective an authentication mechanism needs to be reliable. A user will only use a mechanism, if the verification process accepts his genuine proofs reliably. Also, from the security perspective a mechanism needs to be reliable, but in a different manner. The security perspective requires that the mechanism denies access, if a false proof is encountered. However, a user might only be aware of the first meaning of reliability and overlook the security perspective and its implications. The acceptance for non-reliability, i.e. rejection of genuine proofs, may increase, if users know and understand the concepts of the authentication mechanism in use. This is an organizational issue and not directly related to a mechanism or mobile devices, but important for the user acceptance especially for innovative approaches to user authentication.

3.2 Gesture-based User Authentication Mechanism

In the following, the gesture-based user authentication mechanism of this thesis is presented. The used machine learning algorithms are omitted here and presented in detail in chapter 4. After presenting the mechanism in general including advantages and premises, potential attacks are presented at the end of this chapter.

3.2.1 Implementation

The presented gesture-based authentication mechanism is designed for handheld mobile devices with embedded motion sensors, i.e. a 3-dimensional accelerometer and a 3-dimensional gyroscope. Prior work on gesture-based authentication - presented in section 2.3.4 - used 3-dimensional accelerometers only. Therefore, it is assumed that using both sensors will increase the security performance with regard to prior work. To authenticate the user enters his gesture by moving the device one-handed. Therefore, gestures are basically restricted to hand and arm movements. A handheld device is only suited for this mechanism, if it is small and lightweight enough, so moving the device is not exhaustive to the user.

A 3-dimensional accelerometer measures the overall applied acceleration including gravity, user applied acceleration and all other applied accelerations. An accelerometer cannot precisely distinguish different sources of acceleration. Nevertheless, an accelerometer can be used to extrapolate the attitude to a certain precision, if the direction of the gravitational force is distinguishable. This is possible, if the gravity is by far the greatest force. A 3-dimensional gyroscope measures the rotational speed around each axis. The combination of both types of sensors allows estimating the device movement more precisely as different properties are measured. However, these sensors measure the movement of the device, which means they measure a user movement only indirectly. Furthermore, the exact path cannot be extracted as those sensors are dedicated to noise and they only measure the first and second derivative of properties of the path. Nevertheless, similar movements should lead to similar measurements.

An automatic segmentation approach was discarded, because it is a computational complex task. It would consume energy and slow down the authentication process. Instead manual segmentation with a push-to-gesture-button was chosen. To enter a gesture, the user presses the button continuously during the movement. Therefore, the start and the endpoint of the gestural input are entered by the user to a certain precision. Authentication is done by regarding the marked part of the measurements only. Furthermore, the user interaction with the button becomes part of the gesture. This introduces additional input and makes forging more complex, because an attacker needs to interact with the button in a similar manner.

The enrollment process of the gesture-based mechanism is straight forward. At first, the user chooses or creates one gesture, he wants to use for authentication. After sufficient training on his own, the user provides genuine samples to the device. This allows estimating the variance individually for each user. The samples are used by the mechanism to "learn" the gesture using machine learning algorithms, which create a model under the assumption that the provided samples are representative. In the authentication step this model is used to decide, if a currently provided unknown sample might be genuine and authentication succeeds. The enrollment and authentication process of the gesture-based authentication mechanism are presented in detail from the machine learning perspective in chapter 4.

3.2.2 Advantages and Premises

Gesture-based user authentication has some advantages over available knowledge-based mechanisms for mobile devices. Knowledge-based mechanisms require that users memorize and protect an arbitrary secret. To guarantee security it is required that a secret is complex enough, so it is not guessable by potential attackers, which makes memorization difficult. However, people memorize movements by training in their motor cortex. This property avoids forcing the user to memorize a complex secret. In addition, a movement cannot be simply written down and the training property prohibits passing on a gesture to another person easily. Another advantage is that the input can be done eyes-free, because the user is not required to look at the device.

The gesture-based authentication mechanism uses behavioral biometric features for user authentication. An authentication mechanism based upon biometric features needs to store the characteristics of the genuine proof, which are used in the authentication process. However, information about the genuine user may be very private and disclosure therefore problematic. The features used by the implemented gesture-based mechanism are only related to short hand movements. This information should not allow an attacker to draw specific conclusions about the genuine user. Furthermore, the proof of the gesture-based authentication mechanism is biometric, but can be changed easily by simply training another gesture and repeat the enrollment step. This might be necessary.

The gesture-based authentication has two major requirements, which answers question 2. First of all, it requires that a user can perform at least one complex and individual gesture. The gesture needs to be complex enough, so it cannot be mimicked easily by other persons. Therefore, people, who are either restricted in their ability to make gestures at all or in their ability to make complex gestures, will not be able to use gesture-based authentication. This potentially excludes disabled and elderly people. Also, sick people are possibly excluded for their time of recovery. Secondly, gesture-based authentication requires that a genuine user is able to repeat the gesture in a similar manner on demand. However, it is very unlikely that users can reproduce a movement exactly. Different movements based upon the same gesture may vary in timing and path.

In addition, a gesture needs to be complex, so it can be assumed as secure. Complexity is not restricted to the path only, but also includes speed changes and rotation. This makes it even harder for the user to repeat a movement perfectly. In fact, the more precisely a user can repeat the gesture the less variance need to be accepted in order to achieve a low FRR. Low variance makes successful forgery more difficult.

Question 3 is omitted here, because it cannot be answered thoroughly from a theoretical perspective. This question is further studied in the user study, which is presented in chapter 5.

3.2.3 Potential Attacks

To evaluate the implemented gesture-based authentication mechanism, it is not sufficient to prove feasibility and usability. It is also essential to demonstrate that security can be enhanced by the mechanism. An authentication mechanism is only useful, if it cannot be broken easily. The security requirements depend on the capabilities, knowledge, resources and motivation of potential attackers. In the following, attack scenarios to gesture-based authentication on mobile devices are analyzed to answer question 4.

System security may be broken by circumventing an authentication mechanism at all. In the following, the assumption that potential attackers are not able to elude the authentication mechanism is made and therefore system security can only be broken by creating forgeries, which are similar enough to the genuine proofs and therefore accepted. If an attacker has no knowledge about the genuine proof, he can attack the mechanism only by brute force. This can be considered as almost impossible, because the search space is very large. In fact, the search space decreases, if an attacker is able to acquire knowledge about the genuine proof.

For authentication the presented gesture-based authentication mechanism requires that the genuine user enters his gesture completely. Therefore, the full proof may be observable by spectators. This is called *visual disclosure* (Liu, et al., 2009). However, the genuine proof is only revealed, when the user authenticates. The user can protect his proof by authenticating only, if he assumes that the situation is secure.

In contrast to knowledge-based mechanisms knowledge about the genuine proof is not sufficient for successful authentication, because the genuine gesture needs to be repeated similar. The difficulty of forgery depends on the complexity of the genuine gesture and the allowed variance. It might be safe to assume that for successful forgery an attacker requires training depending on the complexity of gesture to be forged. For easier training an attacker might to attempt to record a video of a successful authentication. However, a gesture is a trajectory in a 3-dimensional space and therefore a single camera might not be sufficient enough to capture the gesture completely, because depth information may be hidden from one perspective. To acquire the gesture completely more video recordings from different angles are required, which makes video recording a complex tasks.

Using video recordings from more than one perspective also complicates training, because the attacker needs to extract the embedded information from multiple videos until he knows the gesture exactly enough.

Training may be improved by extracting the information and using visualization techniques, which are able to project the acquired information into a 3-dimensional space. A promising technique for training might be augmented reality techniques, because they allow visualizing the trajectory as overlay to the reality and the attacker can see and follow the trajectory. This approach provides qualitative feedback to an attacker, because he directly sees the differences in his movement and the gesture. The extracted information can also be used in a different way. The mechanism does not require that the gesture is done by a person. It is only required that the device is moved in a similar manner including interaction with the push-to-gesture-button. Forgery is also possible using a robotic arm, which is programmed using the extracted information. This avoids training a person, which requires skills and time for training.

Regarding gesture-based authentication mechanism for mobile devices, attacks based upon visual disclosure seem to be realistic, if the attacker manually forges the gesture himself. The augmented reality and the robotic attack may be possible, but seem very unrealistic, because they require technologic knowledge and many resources. It is assumed that eluding the mechanism would be easier and therefore both attacks are discarded in this thesis.

For all attacks mentioned above, the attacker needs to acquire information about the genuine proof, so he is able to forge successfully. Nevertheless, this is not sufficient as an attacker still needs to get access to the device.

4 Machine Learning Algorithms

In this chapter the used machine learning algorithms of the implemented gesture-based authentication mechanism are presented. The goal of machine learning algorithms applied in biometric user authentication mechanisms is to automatically learn the characteristic(s) of the measured feature(s) of a genuine user, so this knowledge can be used to verify the user later. To learn the characteristic(s) each feature is usually measured several times in the enrollment process, because a single measurement might not be representative and does not allow estimating the variation. One measurement of the used feature(s) is in the following denoted as *sample*. The extracted characteristics of the measured features of one genuine user are denoted as *model*. A sample is genuine if the genuine user was measured otherwise it is a forgery. This decision must be made for unknown samples in the authentication process. For this decision the similarity of a unknown sample and the enrollment samples or the derived model needs to be computed. If the given sample is similar enough, it is assumed that it was created by the genuine user and authentication succeeds. User authentication may be regarded as a special classification problem as only one class, i.e. the class of the genuine proof, is known. More precisely it is a one-class classification problem with novelty detection. An unknown sample is rejected, if it is likely that a new class was encountered.

The measurements of dynamic gestures are sequential, i.e. a time-series. This means that in addition to single measurements, the order of the measurements is informative. For the implemented gesture-based authentication mechanism it is assumed that the complete sequence is equally informative. As machine learning algorithms DTW and HMM are evaluated. They are suited to learn time-series and regard each part of the sequence as equally important (Sakoe, et al., 1978 p. 44) and (Rabiner, 1989; Bishop, 2007 p. 605ff.). Furthermore, DTW and HMM allow differences in timing and are therefore suited for dynamic gesture recognition.

In the following, the enrollment and authentication procedure are presented. Then the preprocessing procedures, which are necessary before training, are described. Afterwards, the variants of DTW and then HMM are presented.

4.1 Enrollment and Authentication Process

The enrollment process of the implemented gesture-based authentication mechanism consists of 6 steps as visualized in Figure 1. First of all, the user choses a gesture, he likes to use for authentication. The user can either be allowed to choose the gesture own his own or interpret a prepared gesture individually. Then he trains his gesture until he is able to repeat the gesture similar enough. In the third step, the user provides a sufficient number of genuine enrollment samples. These samples are in the following denoted as T. The raw samples are preprocessed as described in section 4.2 and then used to create a model.

The model is later used to calculate the similarity of a yet unknown sample using the learned characteristics of the enrollment samples. In the last step, the threshold for acceptance is chosen or calculated. The threshold defines the lower bound of similarity, which needs to be achieved by unknown samples to be accepted as genuine. It can be chosen a priori or calculated using the enrollment samples. The preprocessing of the samples, training of the model and calculation of the threshold is done automatically by the mechanism without user interaction.

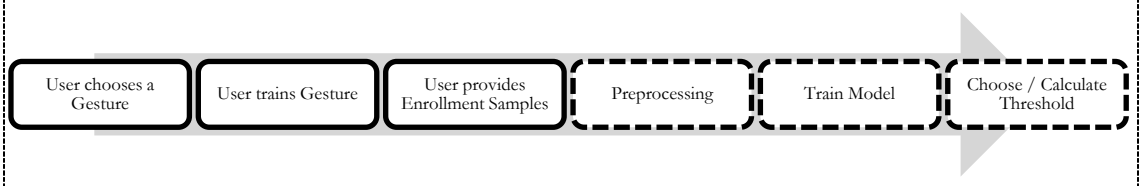


Figure 1: Enrollment procedure of a gesture-based authentication mechanism.
The solid bordered steps require user interaction and the dashed steps are done automatically.

In the authentication process the trained model and the calculated threshold are used to decide, if the current sample might be genuine. An unknown sample is accepted, if the *Acceptance Function* G evaluates to true. This function combines the result of the algorithm specific *Model Acceptance Function* M and the *Length Constraint*. The Model Acceptance Function for DTW is described in section 4.3.1 and for HMM in section 4.4.1. Introducing the Length Constraint is necessary as DTW and HMM are unfair with regard to variations in the length. The Acceptance Function is described in detail at the end of this chapter in section 4.5. It evaluates to true only and therefore an unknown samples is accepted, if the Length Constraint and the Model Acceptance Function are satisfied.

4.2 Preprocessing

Before presenting both algorithms in detail the required preprocessing steps are described at first. DTW and HMM require that the sampling frequency is constant, i.e. that the time difference between two successive measurements is equal for the complete sequence. Furthermore, it is required *that all samples* have an equal sampling frequency. During the first evaluation of the recorded measurements, it was found that the frequency varies.[6]

[6] The tool to record the movements is called Gesture Recorder and is described briefly in 5.1. The technical important details are described in the Appendix ii.i.

To overcome this problem, *linear interpolation* was applied to normalize the sampling frequency. The linear interpolation formula is shown in (1). For each timestamp t, where a measurement d^t should exist, d^t is approximated using the predecessor d_p^t and successor d_s^t. The timestamp of the predecessor is denoted as t_p and of the successor as t_s:

$$d^t = \frac{(t_s - t) * d_s^t + (t - t_p) * d_p^t}{t_s - t_p}. \quad (1)$$

Smoothing for HMM for gesture recognition using accelerometers was found useful in (Prekopcsák, 2008). Indeed, for HMM it was found that smoothing improves the performance of a model, which is discussed in section 4.6. All samples were smoothed for HMM with the *Simple Moving Average* algorithm. This algorithm calculates the value of the current point by computing the mean of the last n - denoted as span - points including the current.

4.3 Dynamic Time Warping

DTW is a dynamic-programming technique, which computes the difference between two sequences by allowing local variations in timing. DTW assumes that the start and the end points of the compared sequences are similar in their meaning. It was used in the prior work on gesture-based authentication as presented in 2.3.4 and also for verification of written signatures (Fang, et al., 2005). The derived variants of DTW for the implemented gesture-based authentication mechanism are based upon (Sakoe, et al., 1978).

DTW finds the cheapest temporal alignment between two sequences of feature vectors:

$$A \coloneqq a_1, a_2 \dots, a_i \dots, a_I \quad (2)$$

$$B \coloneqq b_1, b_2 \dots, b_j \dots, b_J \quad (3)$$

by calculating the I-by-J sized cost matrix D using the cost function $d(i,j)$. The cost function calculates the difference between two feature vectors a_i and b_j, where $1 \leq i \leq I$ and $1 \leq j \leq J$. The cheapest temporal alignment is found with respect to $d(i,j)$ by computing the path that minimizes the summed costs of all single steps. A possible path needs to fulfill the *boundary condition*, i.e. it starts at $i = 1, j = 1$ and ends at $i = I, j = J$, the *monotonic condition*, i.e. no steps backward in time are possible, and the *continuity condition*, i.e. no jumps in the alignment are allowed.

To avoid overheat by calculating the cost of each possible path separately the cost of the cheapest path is calculated recursively by storing intermediate results in D starting at $D(1,1)$. For each entry $D(i,j)$ the cost function is evaluated and the result is added to the summed up cost of the cheapest sub-path that allows reaching this entry. The new summed up partial cost is stored in $D(i,j)$. After evaluating each entry of D the total cost of the cheapest path is stored in $D(I,J)$. The cheapest path can then be calculated via backtracking from $D(I,J)$ to $D(1,1)$.

The set of possible path can be restricted, so paths are excluded that would be absurd temporal alignments. The set can be restricted globally by introducing the *adjustment window condition* R, which defines the maximal difference of the indices i and j: $i + j \leq R$. As alternative possibility the set of paths can also be restricted locally by constraining the slope. The *slope constraint* P limits the sub-paths depending on the slope from where an entry is reachable. Every P diagonal steps one horizontal or vertical step is allowed. A greater slope constraint restricts the ability for local temporal alignment. The slope constraint leads to a parallelogram of entries to calculate in the cost matrix known as the Itakura parallelogram, if for the complete cost matrix the same slope constraint is applied (Itakura, 1975). The slope constraint and the adjustment window condition are visualized in Figure 2. The two constraints do both restrict the allowed difference in length of comparable samples.

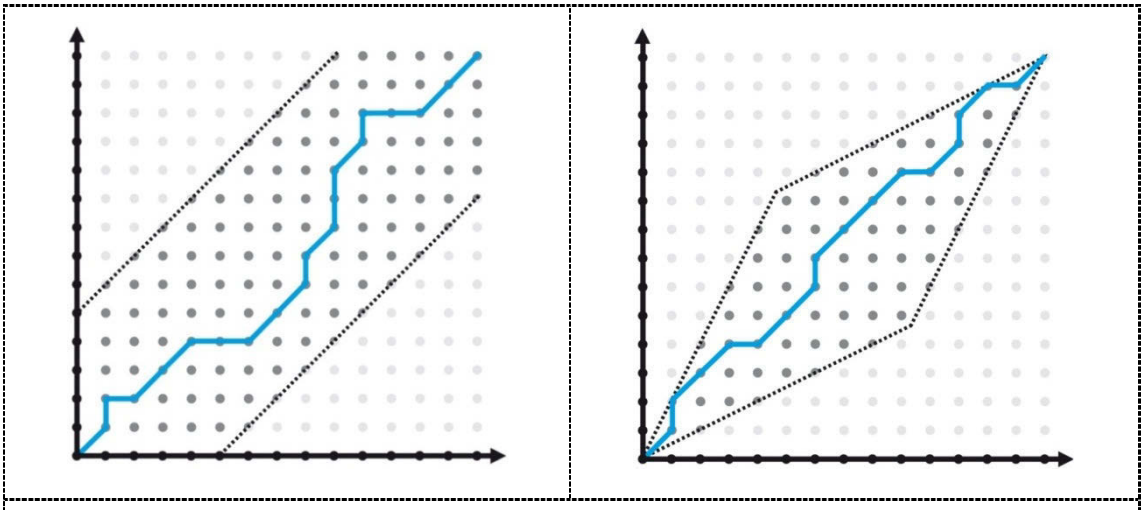


Figure 2: Visualization of the DTW algorithm with R=5; P=0 (left) and R=Inf; P=1 (right).

DTW as described above penalizes non-diagonal steps, because one horizontal step makes a vertical step necessary and vice versa. The combination of one vertical and one horizontal step adds two evaluations of $d(i,j)$ to the cost of a sub-path, whereas one diagonal step adds one evaluation only. To overcome this inequity the cost of diagonal steps is doubled. Nevertheless, non-diagonal steps should not be overused for temporal alignment. If two sequences are similar with minimal differences in the length of corresponding parts, then non-diagonal steps should be seldom necessary.

A large number of non-diagonal steps in the cheapest path between two sequences is an indicator that two those are unequal. To avoid overusing non-diagonal steps, the *non-diagonal alignment penalty* H is added to the cost of non-diagonal steps. Therefore, diagonal steps should be preferred, because they do not lead to additional cost due to penalty.[7]

The initial expression of the recursive function to calculate the entries of the cost matrix D is:

$$D(1,1) = 2 * d(1,1) \tag{4}$$

where the first step is regarded as diagonal. The recursive function depending on different slopes is shown in (5-7). The corresponding sub-paths are shown in Table 1.

$$P = 0:$$

$$D(i,j) = min\begin{cases} D(i-1,j) + d(i,j) + H \\ D(i-1,j-1) + 2 * d(i,j) \\ D(i,j-1) + d(i,j) + H \end{cases} \tag{5}$$

$$P = 1:$$

$$D(i,j) = min\begin{cases} D(i-1,j-2) + 2 * d(i,j-1) + d(i,j) + H \\ D(i-1,j-1) + 2 * d(i,j) \\ D(i-2,j-1) + 2 * d(i-1,j) + d(i,j) + H \end{cases} \tag{6}$$

$$P = 2:$$

$$D(i,j) = min\begin{cases} D(i-2,\mathrm{j}-3) + 2 * d(i-1,j-2) + 2 * d(i,j-1) + d(i,j) + H \\ D(i-1,j-1) + 2 * d(i,j) \\ D(i-3,j-2) + 2 * d(i-2,j-1) + 2 * d(i-1,j) + d(i,j) + H \end{cases} \tag{7}$$

If one of the sub-functions of the cost function is invalid, i.e. out of the valid range, then the expression is evaluated to infinity and therefore this sub-path discarded.

[7] Non-diagonal alignment penalty was used by (Guerra Casanova, et al., 2010). A discussion about penalty for directed graph alignment can be found in (Jones, et al., 2004 p. 184).

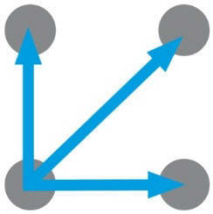
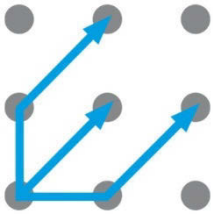
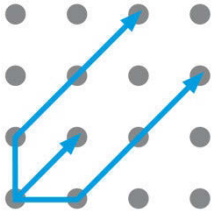
P = 0:	$P = 1$:	$P = 2$:
		

Table 1: DTW allowed sub-paths depending on the slope constraint.
Based on (Sakoe, et al., 1978).

On the one hand, the weighting of diagonal steps does achieve fairness with regard to the directions of the steps. On the other hand, the weighting introduces unfairness, because the implications of the cost of the cheapest path depend on the length of two compared sequences. Comparing a sequence with shorter or longer sequences will prefer shorter ones, because they require fewer steps with doubled costs, i.e. regarding combined horizontal and vertical steps as a single step, than longer sequences. Thus, the resulting overall costs are not similar in their implication and not useful as similarity metric. To overcome this, the overall costs are normalized (Sakoe, et al., 1978):

$$N(I,J) = \frac{D(I,J)}{I+J}. \tag{8}$$

Unrestricted DTW for two sequences of length I and J requires $I * J$ evaluations of the recursive function, i.e. it belongs to $O(N^2)$. If the slope and/or adjustment window is constrained, the number of evaluations usually decreases. Nevertheless, the length of sequences to compare is therefore limited due to practical reasons. An approach to approximate unrestricted DTW in linear time is presented in (Salvador, et al., 2007).

4.3.1 Model Acceptance Function

For DTW a model M is trained by choosing an enrollment sample or computing a sequence using the enrollment samples. The *Model Acceptance Function* decides if an unknown sample is accepted by the trained model as genuine. For DTW the cost of the cheapest path is used as metric for similarity and the threshold is chosen using the enrollment samples. The threshold for DTW is computed using M and the enrollment samples by calculating the costs of all enrollment samples against M. These costs are denoted as C_T . The threshold is calculated using the mean and the standard deviation. The influence of the both functions is selected and weighted with the positive two-dimensional row vector Q.

The equation used to calculate the threshold is:

$$Tr_Q(C_T) \coloneqq Q * \begin{bmatrix} \emptyset(C_T) \\ \sigma(C_T) \end{bmatrix}. \tag{9}$$

An unknown sample S is accepted by the model, if the normalized cost of the cheapest path against M is smaller than the threshold:

$$MA_{M,T}(S) \coloneqq N(M,S) < Tr_Q(C_T). \tag{10}$$

An optimal value for Q cannot be determined theoretically and therefore needs to be estimated practically in the user study.

4.3.2 Cost Functions

So far the DTW algorithm and the used Model Acceptance Function were presented. In the following, the cost function $d(i,j)$ is presented. For the gesture-based authentication mechanism each feature vector is 6-dimensional consisting of 3-dimensions for acceleration and 3-dimension for rotational speed. This requires that the cost function must be applicable to multi-dimensional feature vectors. In addition, the influence of each dimension should be equal, because no a priori knowledge about the importance for the authentication task of the different dimensions was available. As cost functions the *Manhattan Norm* (L_1) and *Euclidian Norm* (L_2) were chosen, which are instances of the Minkowski Norm (Marsland, 2009 p. 191):

$$L_k\left(a_i^V, b_j^V\right) = \left(\sum_{v=1}^{V} w(v)\left|a_i^v - b_j^v\right|^k\right)^{1/k}. \tag{11}$$

V denotes the dimensions of the feature vectors. One problem is that the accelerometer and gyroscope measurements are unequal scaled, because they differ in their range and also in their unit. This problem is addressed by integrating the Weighting Function $w(v)$ into the Minkowski Norm.[8] After calculating the difference between each dimension of the two feature vectors, the intermediate results are rescaled before summing them up. Nevertheless, this solution does not address the problem that the dimensions also differ in their meaning, because different properties, i.e. acceleration and rotational speed, are measured by the sensors.

[8] The idea of the Weighting Function is based upon (Redžić, et al., 2010).

In prior work on gesture-based authentication with DTW as presented in section 2.3.4 mostly the Euclidian Norm was used (Guerra Casanova, et al., 2010; Liu, et al., 2009; Okumura, et al., 2006).

One alternative was studied by Okumura et al. and in the follow up work of Matsuo et al. (Matsuo, et al., 2007; Okumura, et al., 2006). Instead of the Euclidian Norm the *Error of Angle* was found to be useful. The Error of Angle uses the difference of the angles only, rather than the combination of differences in angle and distance approach of the Euclidian Norm. However, the Error of Angle was applied to 3-dimensional acceleration feature vectors only and is therefore discarded, because in the gesture-based authentication mechanism of this thesis a gyroscope is additionally used.

4.3.3 Training Modes

DTW can only use one sequence as model M. This requires that the sequence used for M is a representative of the class of sequences that should be matched. Thus, it needs to be chosen or computed carefully. In the following, the three considered approaches are presented.

4.3.3.1 Choosing an appropriate Sample

Choosing an appropriate sample from the set of enrollment samples is a straightforward approach to choose M. It is assumed that all enrollment samples are representative samples of the genuine proof, but one sample is most representative. In fact, only the enrollment samples are available and therefore only the most representative sample with regard to the enrollment samples can be chosen. For the implemented gesture-based authentication mechanism it is chosen by calculating the cost of the cheapest path using DTW for all pairwise combinations of enrollment samples. The sample that minimizes the summed up costs is chosen as M.

The drawback of this approach is that only one enrollment sample is used and the variance of the other enrollment samples is only integrated into the threshold.

4.3.3.2 Computing the Model

An alternative to choosing the most representative enrollment sample is to integrate the enrollment sample into a combined sequence. To achieve this, the algorithm developed by Abdulla et al. was adapted (Abdulla, et al., 2003).

The algorithm of Abdulla et al. works as follows. First of all, the mean length of the enrollment samples is computed and the closest sample chosen as initial template. Then, the remaining enrollment samples are locally compressed and expanded to length of the initial template. This is done for each sample individually by calculating the cheapest path between the initial template and the sample. For all steps, where the slope of the cheapest path is not equal to 1, the current sample is locally *compressed* or *expanded*. After this procedure the length of the sample is equal to the length of the initial template. The model M is computed by the calculating the mean for each corresponding feature vector, i.e. similar index, of the initial template and all aligned enrollment samples.

Abdulla et al. used duplication of a feature vector for expansion and averaging multiple feature vectors into one for compression (Abdulla, et al., 2003). This approach changes the local sum of the measurements. This is adequate, if the sum is not important like in speech data or in position data. In such cases only the timing is changed, but not the "path". However, the sum should be important for acceleration and rotational measurements as small changes would lead to different trajectories. Therefore, an alternative approach that maintains the local sum is studied as alternative. For compression the sum of the feature vectors is calculated and expansion is done by distributing the value of the feature vector equally over the resulting feature vectors.

This approach integrates all enrollment samples into a single sequence, which can encode all information available in the enrollment samples. Therefore, the variance is also embedded into the model and not only the threshold as by the choosing the cheapest sample approach.

4.3.3.3 Multiple Templates

A very different approach to use multiple samples is proposed by Ogawara et al. by extending DTW (Ogawara, et al., 2001). This approach extends DTW to enable temporal alignment between more than two sequences. This is achieved by increasing the dimensionality of the cost matrix to number of sequences to align and using an adapted cost function. No additional computation to create a multidimensional model is required, but the computational complexity increases exponentially with the number of used samples. Therefore, this approach is discarded in this thesis, because it is by far too complex for mobile devices and it is further not guaranteed that it would lead to better results.

4.3.4 Variants

In the following, the variants of DTW are presented that were assumed to be useful for the gesture-based authentication mechanism of this thesis. They were evaluated in the user study, which is presented in chapter 5 and the results are presented in chapter 6.

In Table 2 all parameters that were varied are shown. The parameters for DTW are the slope constraint, the parameter k of the Minkowski Norm, the Training Mode and the non-diagonal alignment penalty. This makes overall 42 variants of DTW. For all variants the Model Acceptance Function and the function to calculate the threshold as presented in section 4.3.1 was used.

Slope constraint	No ($R = 30$) / 1 / 2	
Minkowski parameter k	1 / 2	
Training Mode	Choose cheapest enrollment sample as model	0
	Calculate model using averaging and duplication	1
	Calculate model using summation and distribution	2
Non-diagonal alignment penalty H	0 / 5	

Table 2: Parameters of the variants of DTW.

If no slope constraint, i.e. $P = 0$, is used the possible paths are restricted with an adjustment window condition R of 30 and the non-diagonal alignment H is set to 5. Therefore, all applied DTW variants restrict the difference in length of unknown samples and the used model. The adjustment window condition is fair with regard to variation in length, i.e. neither favoring longer nor shorter samples. Applying a slope constraint introduces unfairness regarding variation in length, because samples are favored that are longer than the model M over shorter samples. For example, a sample must be in range between the half and double length of M for a slope constraint of 1. This problem is addressed by adding the Length Constraint into the Acceptance Function, which is presented in section 4.5.

The only expected result of the presented DTW variants was that a slope constraint should perform better, because the local path is restricted and therefore does require that an unknown sample needs be more similar in timing.

4.4 Hidden Markov Models

For the gesture-based authentication mechanism HMMs were chosen as alternative to DTW, which were not applied in prior work. In fact, HMMs were successfully applied for gesture recognition based upon accelerometers (Pylvänäinen, 2005; Niezen, 2008) and for written signature verification using online information (Fierrez, et al., 2007), which are both closely related fields. Thus, HMMs may be promising for the gesture-based authentication mechanism. In the following HMMs and related algorithms are presented.

HMMs are generative statistical models based upon Markov processes (Rabiner, 1989). A Markov process is a stochastic process, where the next state depends on the current state only, i.e. the process is memoryless. This is called Markov property (Bilmes, 2006). A Markov process can be stationary, i.e. allow a state transition from the current to the current state (Dietterich, 2002 p. 9). In HMMs the underlying Markov process is hidden and can be observed indirectly through another stochastic process only (Rabiner, 1989).

A HMM λ is depicted by the number of states N, the initial state distribution π, the state transition probabilities A and emission probabilities B, i.e. $\lambda = \{\pi, A, B\}$ (Rabiner, 1989). N is implicitly encoded in π, A and B. π is the probability for each state that underlying hidden process starts in this state. Therefore, π sums to 1. The probability of a state transition from the current state to any state is encoded by A, which sums to 1 for each state. B describes the chance of a possible emission depending on the current state only. The emission of a HMM can either be discrete, i.e. only a set of finite symbols can be emitted, or continuous (Rabiner, 1989). To model continuous observations continuous probability functions are used (Rabiner, 1989 p. 258).

Given a HMM λ and an observed emission sequence O, three computations are possible (Rabiner, 1989 p. 261). First, compute the probability that the model λ generated the sequence O, i.e. compute the likelihood: $P(O|\lambda)$. Calculating the likelihood requires $N^2 * length\,(S)$ calculations (Rabiner, 1989). Second, adapt the parameters of the model λ to λ', so the likelihood $P(O|\lambda')$ is maximized, i.e. train the model. Third, calculate the sequence of hidden states, which most likely explains O.

Training a HMM is done with a variant of the Expectation-Maximization-Algorithm (EM-Algorithm) namely the Baum-Welch-Algorithm. A complete introduction to the EM-Algorithm is given in (Bishop, 2007 p. 439 ff.) and for HMM also in (Bishop, 2007 p. 615 ff.). The EM-Algorithm is prone to find local maxima (Rabiner, 1989). Therefore, the initial guess of parameters needs to be done carefully. The initial parameters can either be chosen randomly using prior knowledge or by apply clustering algorithms on the data to find the parameters of likely states. The complexity of training depends on the number of parameters. This property depends on the number of states and on the allowed state transitions. N parameters need to be estimated for π and N^2 for A for ergodic HMMs. The number of parameters with regard to the emission probability also depends on the number of states. In fact, the more parameters need to be estimated, the more training examples are necessary to fit the model well (Rabiner, 1989).

Depending on the sequences to model, restrictions to HMMs can be useful. For sequential data usually Left-to-right HMMs[9] are applied, if it can be assumed that the sequence is not composed of repetitions of subsequences (Bishop, 2007 p. 614f.). Left-to-right refers to the underlying Markov process by limiting state transitions to successor states or to the same state.

[9] Left-to-right HMMs are also called Bakis-Model after R. Bakis (Rabiner, 1989).

Therefore, it is assumed that the underlying process starts in the first state and traverses successive through the states until the last one is reached. For the gesture-based authentication mechanism it is assumed that a First-order Markov process is suited as visualized in Figure 3. First-order means that the underlying process can only stay in the current state or transition to the direct successor state (Bilmes, 2004). As a consequence, only N parameters for A need to be estimated and π is fixed.

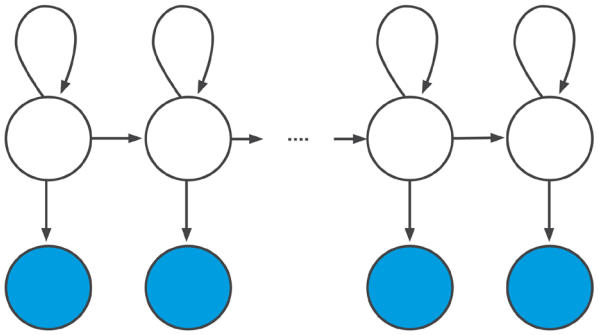


Figure 3: First-order Markov process (white) with state transitions (grey) and resulting observations (blue).
(Bishop, 2007 p. 609)

Characteristic for First-order HMMs is the number of states, because steps backward to prior states are not allowed. Therefore, the number of state transitions is limited to $N - 1$. The optimal number of states depends on the expected number of state transitions in the hidden Markov process that is to be modeled. Too few states are problematic, because not all states of the underlying stochastic process can be learned specifically and therefore some characteristics may not be learnable. Too many states require more parameters to estimate and thus require more training samples. The optimal number of states is usually unknown and needs to be chosen or estimated carefully.

Pylvänäinen showed that gesture recognition using a 3-dimensional accelerometer and HMMs with continuous emission is feasible (Pylvänäinen, 2005). As alternative the raw measurements could be quantized. Quantization means that the continuous data is mapped to a discrete alphabet, which are used for a HMM with discrete emissions. A quantized approach for accelerometer-based gesture recognition is presented in (Schlömer, et al., 2008). However, quantization removes information and is therefore discarded in this thesis.

4.4.1 Model Acceptance Function

The Model Acceptance Function of the HMM is analogical to the one used for DTW. As similarity metric the normalized log-likelihood is used. The likelihood is the probability that a sample S is generated by λ. To avoid numerical underflow the likelihood is used logarithmically: $\log P(S|\lambda)$. The likelihood depends on the length of a given sample and penalizes longer sequences. To achieve fairness the log-likelihood is normalized by dividing it through the length of S. The Model Acceptance Function for HMM is defined as:

$$MA_{\lambda,Tr}(S) \coloneqq -\frac{\log P(S|\lambda)}{length(S)} < Tr. \tag{12}$$

Tr denotes the threshold. In contrast to DTW for HMM a static threshold, i.e. independent of the enrollment samples, was used. This was not intended, but required due to reasons related to HMMs. In some cases the computed log-likelihood was greater than zero. This should be impossible as the likelihood is a probability. However, this was not a bug in the used implementation of HMM, but dedicated to the Gaussian distribution.[10]

HMMs were evaluated using the open source software *Probabilistic Modeling Toolkit 3* (PMTK3), which is an add-on for Mathworks Matlab and Octave. A detailed description of PMTK3 and the interfaces can be found in the Appendix ii.ii.

4.4.2 Variants

For the gesture-based authentication mechanism First-order HMMs with a multivariate Gaussian distribution were used with different number of states. HMMs were evaluated using 4, 8, 12 and 14 states. In fact, more states make training and also evaluation of the Model Acceptance Function more complex.

[10] Thanks to Kevin Murphy for pointing out that the likelihood can be greater than 1 for a Gaussian distribution and that it is not an issue of PMTK3. A Gaussian distribution can lead to a probability greater than 1: [MATLAB command]: *normpdf(0, 0, 0.9/sqrt(2*pi))*.

4.5 Acceptance Function

In this section the algorithm independent Acceptance Function G is presented, which uses the Model Acceptance Function and the Length Constraint to decide, if an unknown sample might be genuine. Before presenting the Acceptance Function in detail, the Length Constraint is introduced. The Length Constraint is required, because the variants of DTW and HMM handle variations in length very differently. DTW with a slope constraint restricts the length unfair by favoring longer sequences. The normalized log-likelihood used as similarity metric for HMM does not impose any restriction at all. However, the length is regarded as an important characteristic of a sample for the gesture-based authentication, because it is assumed that the genuine user will repeat the gesture in a similar manner. Therefore, the differences in length of genuine samples should be relatively small. It is further assumed that all genuine samples of gesture are scattered symmetrically around an unknown mean length L. L is approximated by the mean length of the enrollment samples T:

$$\hat{L} = \emptyset\big(length(T)\big). \tag{13}$$

The Length Constraint is defined as:

$$inRange_{\hat{L},E}(S) \coloneqq \hat{L} * (1 - E) < length(S) < \hat{L} * (1 + E). \tag{14}$$

The parameter E defines the maximal proportional difference between the length of the sample S and $\hat{L}$. An unknown sample is accepted as genuine, if it satisfies the Length Constraint and the Model Acceptance Function. This leads to the Acceptance Function G, which is defined as:

$$G_{MA,T,E}(S) \coloneqq inRange_{\hat{L},E}(S) \wedge MA(S). \tag{15}$$

To avoid unnecessary computations the Length Constraint is evaluated first as it has constant complexity. The Model Acceptance function is evaluated only, if the Length Constraint is satisfied.

4.6 Performance

The implemented gesture-based authentication mechanism uses a machine learning algorithm to decide if an unknown is genuine and authentication may succeed. The performance of the used algorithm is important for the achievable overall security performance of the mechanism. The security performance of an authentication mechanism can be denoted by the FAR and FRR. The FAR and FRR should be as small as possible, i.e. in the optimal case 0. However, there is usually trade-off between the FAR and FRR and decreasing the FRR usually increases the FAR.

A variant of an algorithm outperforms another variant or algorithm, if it achieves a lower FRR and FAR. In fact, this comparison is only meaningful, if all variants were evaluated with the same set of enrollment, validation and forgery samples, which should be likely for the desired implementation scenario.

5 User Study

A user study was conducted to evaluate the gesture-based user authentication mechanism. The user study was divided in two stages to study different dimensions thoroughly. In stage 1 the feasibility, usability and the user perception was studied. This part is called *Proof-of-Concept-Study*. All participants of this stage were in the role of a genuine user. The most important goal of this stage is to study the feasibility and usability. In the second stage the participants tried to forge selected gestures of the Proof-of-Concept-Study. This stage is called *Forgery-Study*.

Before both stages are explained in detail, the tools used in the user study are described. The tools consist of the Gesture Recorder, a set of designed gestures and three questionnaires. At the end of this chapter the Proof-of-Concept-Study (section 5.4) and Forgery-Study (section 5.5) are presented in detail.

5.1 Gesture Recorder

The gesture-based user authentication mechanism is designed for usage on mobile devices. For the user study it was required to measure the gestures with a tool - called Gesture Recorder -, which is in size and form similar to a mobile device that is designed for one-handed interaction and has built-in a 3-dimensional accelerometer and a 3-dimensional gyroscope. The iPhone 4 was chosen to implement the Gesture Recorder, because fulfills the requirements and it was available on the market.[11] The graphical user interface of the Gesture Recorder is very simple as it consists of the push-to-gesture button only that covers almost the complete screen. In Figure 4 the orientation of the sensor axis of the iPhone 4 is shown. The accelerometer measures the acceleration along each axis and the sign marks the direction of the force along one axis. The gyroscope measures the rotation around each axis. Rotating the device around one axis into the direction of the corresponding arrow - in Figure 4 (right) - yields positive readings. Using both sensors simultaneously a maximal frequency of 80 Hertz was achieved. The accelerometer readings are in *g* and the gyroscope readings in *radians per second*.

[11] At the beginning of this thesis the iPhone 4 was the only mobile device available on the market including accelerometer and gyroscope. The accelerometer sensor in the iPhone 4 is a STMicro STM33DH and the gyroscope a STMicro L3G4200D (edepot.com, 2010).

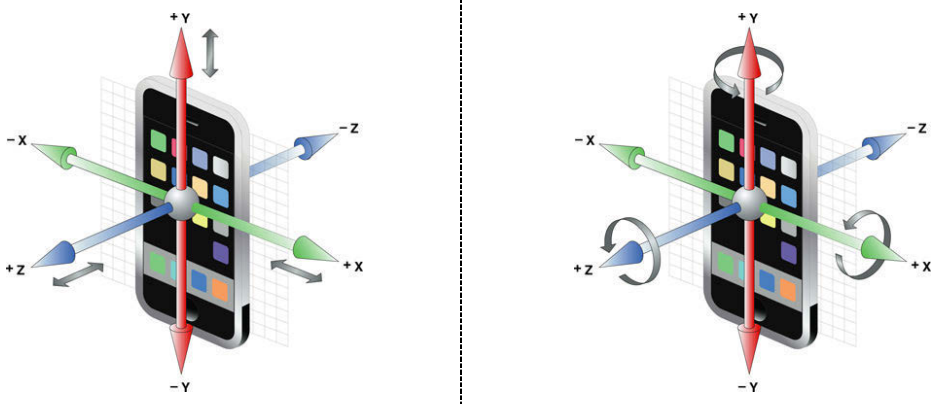


Figure 4: The direction of the x-, y- and z-axis of the embedded sensors in an iPhone 4. The accelerometer measures the force in direction of each axis (left). The gyroscope measures the rotation around each axis (right). Source: (Apple Inc., 2011).

An exemplary record of the triangle gesture (presented in the following section) is shown in Figure 5. The accelerometer readings include the overall measured acceleration, which also includes gravity. The effect of gravity is mainly encoded in the z-axis, because the gesture is designed, so the back of the device points mainly to the floor. A detailed technical description of the Gesture Recorder is to be found in the Appendix ii.i.

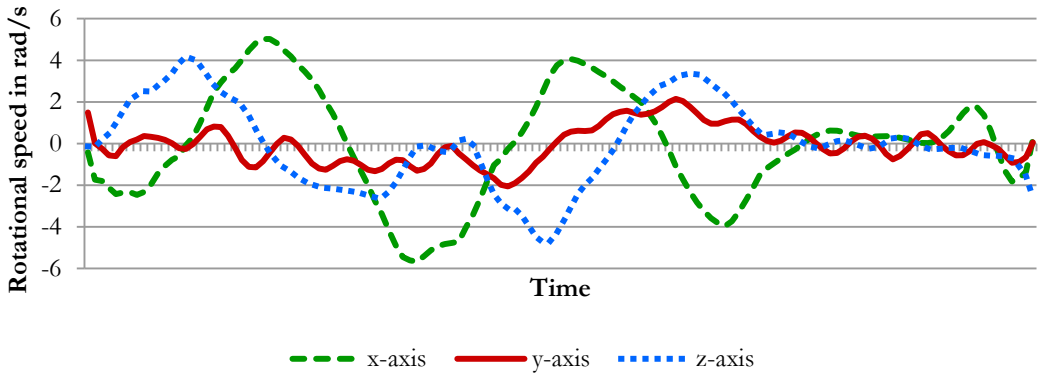


Figure 5: Exemplary recording of a triangle gestures. For accelerometer (above) and gyroscope (below) for all three-axis 136 measurements at 80 Hertz.

5.2 Designed Gestures

In the following, the gestures used in the Proof-of-Concept-Study are described. These gestures were designed a priori, because choosing an individual gesture that is appropriate for daily life might be a complex task. Prior designed gestures allow the participants to concentrate on the gesture-based authentication mechanism. Nevertheless, the designed gestures are only roughly sketched and the participants therefore required interpreting each gesture on their own.

5.2.1 Requirements

In the following, the requirements for useful gestures are discussed, because not all physiological possible hand gestures are usable. First of all, it is required that the user holds the device tightly during the complete gestural input, because otherwise the device may be thrown away. This could damage the device, harm nearby people and be a general threat to the surrounding environment. Second, it is required that mobile devices should be usable almost everywhere, at any time and in any context. It cannot be assumed that the devices are always used in private places or settings. In some places or situations some gestures may be inappropriate. A gesture is only usable, if its execution is acceptable in most settings and social contexts. It should not be offending, threatening, aggressive, displeasing or ridiculous to other persons. This excludes inappropriate, hectic and also overly larges gestures. Third, the performance of a gesture should be satisfying to the user. This also limits maximal duration of useful gestures. It is assumed that a longer gesture is more exhausting and possible fatiguing to the user than a shorter one. In addition, the time required to decide if the current sample is accepted as genuine depends on the length and waiting might be annoying to the user. Therefore, it is assumed that a gesture should last between 0.5 seconds and 1.5 seconds at most.

5.2.2 Design

The designed gestures take into account that the iPhone 4 is the basis for the Gesture Recorder. The main user interface of the device is a touch screen and user interaction with the device is often done one-handed. For interaction the device lies usually in the palm of one hand with the screen upward and the user use his fingers to hold the device.

In this posture the user can look at the touch screen interact with the touch screen using his thumb. The general posture is shown in . It is likely that the user will hold the device in this way, when authentication is required. Afterwards, it is also likely that the user likes to continue interaction in the same posture. Therefore, this posture was chosen as initial and final position of all designed gestures, so the authentication does not require any additional movements before interaction is possible, which are not part of the used gesture.

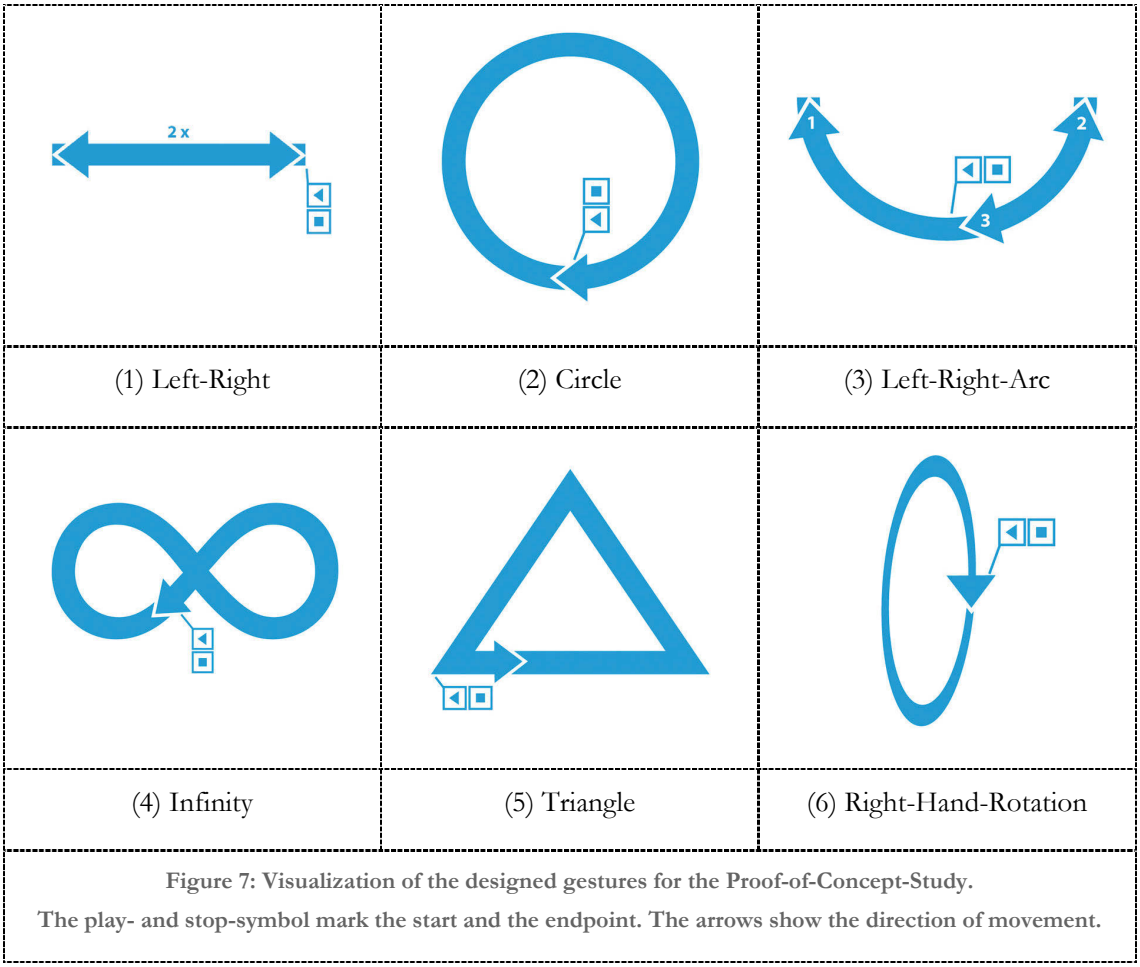
For gestural input it is required that the user grasps the device tightly, which he can do by closing his hand around the device and using his thumb on the touch screen to fixate it. Then, the device is secured between his thumb and palm. Therefore, the push-to-gesture-button is implemented for thumb interaction using a large area on the touch screen and the button needs to be pressed continuously during gestural input.

Gestures for the gesture-based authentication mechanism should take place in the room around the user, which he is aware of and can sense actively. All designed gestures take place in front of the upper body, because this space is perceived well. Furthermore, the gestures should not be too large, fast or hectic. It is assumed that a gesture without edges is perceived as more pleasant by spectators and the user, because edges require sharp changes in the direction of the movement. With regard to these basic restrictions possible movements of the hand and arm were evaluated, which are the basis for the designed gestures.

Figure 6: Posture for interaction with the iPhone 4.

The designed gestures were presented as a static image to avoid that the participants of the Proof-of-Concept-Study simply mimic the movement performed by another person. The images visualize the general idea and path of the gestures. The start and the end position are marked with a play- and stop-symbol and arrows are used to visualize the direction. To provide additional information about a gesture a short textual description (see Appendix ii.v - German only) was provided, which describes the general shape and the direction of the gesture. Two gestures required additional information in the visualization as some parts of the path were used multiple times in different directions.

For the Proof-of-Concept-Study six gestures were designed. The gestures (1) to (5) are based upon geometric figures. This makes them easily understandable as they can be described in words and visualized well. Gesture (6) is different as it is a natural movement, i.e. designed by natural possible and satisfying movements. This gesture uses the ability to rotate the hand using the wrist and the underarm, which leads to a small and short movement. The visualization used in the study is shown in Figure 7.



(1) Left-Right	(2) Circle	(3) Left-Right-Arc
(4) Infinity	(5) Triangle	(6) Right-Hand-Rotation

Figure 7: Visualization of the designed gestures for the Proof-of-Concept-Study. The play- and stop-symbol mark the start and the endpoint. The arrows show the direction of movement.

The geometrical gestures (1)-(4) are edgeless and start into the direction of the hand that does not hold the device. Those gestures are designed to be performed in front of the user and parallel to the upper body. Gesture (5) is also geometrically, but different. It symbolizes a triangle that per definition has edges and it is performed parallel to the floor in front of the user. Thus, the gesture should require more space. This gesture starts in the opposite direction of the other gestures, i.e. away from the center of the body. It might be that this gesture is perceived as uncomfortable and impractical, but this is an unproven assumption. In fact, this makes this gesture interesting for the user study.

Each participant of the Proof-of-Concept-Study was required to interpret the designed gestures on their own. Therefore, the resulting movements should be very different with regard to the path, space and also timing.

5.3 Questionnaires

An authentication mechanism is not only required to be feasible and provide security, but it also needs to be usable.[12] To gather information about the user perception of the implemented gesture-based authentication mechanism three questionnaires were used. They are described in the following and can be found in the Appendix ii.vi (German only).

5.3.1 First Questionnaire

The first questionnaire aims at acquiring demographic information about the participants and their usage of mobile devices. This questionnaire also contains questions and statements to estimate the security perception.

In the first section the participant provides demographic information like gender, age and profession. In the following, two sections questions about the mobile device that the participant uses were asked. These questions include the capabilities, e.g. GPS included, and the characteristics of the user interface. Section 4 contains questions about the interaction with the device like how often he interacts with it per day and how long the average interaction lasts. The ranges of the possible answers are based upon Falaki et al. (Falaki, et al., 2010). This section also includes questions about the usage of the key lock and authentication mechanisms. Section 5 and 6 are about the usage of the mobile device with regard to applications and tasks. Section 7 is about the experience of losing a mobile device. Section 8 is about the usage of passwords, which includes questions like how often the participant changes his passwords, reuse it for multiple accounts and if he passes it to other persons. In section 9 the participant was asked to name the authentication mechanisms available on their mobile device. Section 10 consists of eleven statements about the usage of mobile devices, the perception of stored data and the assessment of security etc. In this section all items use a Likert-scale with 7 categories. In all other sections most of the questions can be answered with *yes/no/unknown* or *available/unavailable*. For the items about the interaction with the mobile device (section 5) an ordinal scale was used.

This questionnaire was used in the Proof-of-Concept-Study and the Forgery-Study, which are presented at the end of this chapter.

[12] Referring to question 3 from chapter 3: How usable is a gesture-based authentication mechanism from the user perspective?

5.3.2 Second Questionnaire

The second questionnaire was created for the Proof-of-Concept-Study. It aims at acquiring information about the participant's perception of the implemented gesture-based authentication mechanism. For each of the 6 designed gesture the participant is asked to rate his agreement or disagreement to 13 statements on a Likert-scale with 7 categories. These statements include the perceived complexity, usability in daily etc.

In addition, the user is asked to estimate the security for each gesture and if he would prefer the current gesture over a password. At the end of the questionnaires the participant judges general statements about the implemented gesture-based authentication mechanism. The questionnaire closes with the task to order the gestures according to the participant's preference.

5.3.3 Third Questionnaire

The third questionnaire was designed for the Forgery-Study exclusively. This questionnaire aims at the perception of the gesture-based authentication mechanism from the perspective of spectators and the perspective of forgers. For each gesture to forge 6 Likert-scaled statements need to be judged. These statements include the perceived usability in daily life and if the presented movement ridiculous is perceived as ridiculous. In addition, the participant is asked if the gesture is perceived as complex and/or difficult to learn.

5.4 Proof-of-Concept-Study

For the Proof-of-Concept-Study 6 female and 9 male persons were recruited. All were right-handed students and aged from 20 to 32 ($\emptyset = 24.3$, $\sigma = 2.8$). 9 participants were experienced with gestural interfaces like the Nintendo Wii or Microsoft Kinect.

This stage of the user study was conducted with one participant at a time. First of all, the participant completed the first questionnaire. Afterwards, the prepared hand-out was given to the participant that contains general information about test procedure and the principle of the gesture-based authentication mechanism (see Appendix ii.iii - German only). All participants were instructed how the mechanism works in general. Afterwards, the designed gestures were then shown to the participant using the visualization and textual as descripted in section 5.2.2. For interpretation and training of the designed gestures the participant had 10 minutes time.

Afterwards, the recording of the gestures started. For each gesture the participant provided 25 samples of his interpretation. The samples of a gesture were not recorded at once, but rather in groups of 5 to verify that the interpretation can be memorized. This also avoids that the movement is over tuned during recording. Such an effect may take place due to training, exhaustion or boredom of the participant. The groups of 5 samples were recorded in the numerical order of the gestures, i.e. (1) to (6). This is in the following denoted as iteration and was overall repeated 5 times. Iterations 1-4 were recorded while the participant was standing.

The 5th iteration was recorded while the participant was sitting on a chair. After the 2nd iteration the participant filled out the corresponding part on the second questionnaire. In complete, the Proof-of-Concept-Study lasted 90 minutes. 30 minutes were required to complete the first questionnaire and introduce the participant into the gesture-based authentication mechanism and the designed gestures. The recording of the gestures including the second questionnaire required 60 minutes. Overall, the 15 participants provided 2247 valid samples.

The gesture recording part was recorded on video using 3 cameras. The setup is sketched in Figure 8. The cameras were placed at 0°, 45° and 90° around the position of the participant with focus on the participant. The cameras on 45° and 90° recorded a single synchronic video stream with lower resolution and frame rate than the camera at 0°. This setup should allow capturing the characteristics of the genuine movements precisely enough for qualitative forgery.

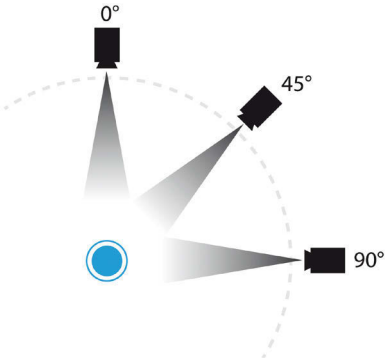


Figure 8: Camera setup during gesture recording in the Proof-of-Concept-Study.

5.5 Forgery-Study

For the Forgery-Study 6 men and 4 women were recruited aging from 24 to 35 ($\emptyset = 27.5, \sigma = 3.9$) that were all right-handed. 8 of them had experience with gestural interfaces.

The goal of the Forgery-Study was to evaluate the chance of creating successful forgeries. The participants of this stage are in the following denoted as forgers. The Forgery-Study also started with the first questionnaire and was conducted with one forger at a time. For the Forgery-Study also a hand-out was given to the forgers, which describes the test procedure and the general functional principle of the implemented gesture-based authentication mechanism (see Appendix ii.iv - German only). This includes that the Gesture Recorder needs to be moved and that only the movement of the Gesture Recorder needs to be mimicked. It was made clear that this also includes the rotation of the device and not only the path. Furthermore, the forgers were informed that the timing of the movement is important and that the interaction of the push-to-gesture-button is an essential part of the gesture.

12 individual interpretations of the Proof-of-Concept-Study were chosen for forgery, i.e. for each designed gesture two using the video recordings of the 3rd iteration. The videos of the gesture to forge were shown to the forger once. Afterwards, the forger had approximately 1 minute to understand and train the gesture. Then 5 forgeries were recorded. This procedure was repeated afterwards once, so the forger had a chance to see and train the movement more precisely. The first 5 samples are referenced as 1st iteration and the second 5 as 2nd iteration of forgeries. None of the forgers were given the visualization, description or even the name of the designed gesture.

Overall, 100 forgeries per attacked interpretation were created. No external feedback was provided to the forgers about their forgery performance. Therefore, the forgers could only rely on their own perception of the similarity. After each forgery attempt the attacker completed the corresponding part of the third questionnaire. The Forgery-Study lasted overall 90 minutes: 15 minutes for the questionnaires and 75 minutes for recording the forgeries.

6 Results

In this chapter the results of the user study are presented starting with the first questionnaire as described in section 5.3.1. Afterwards, the Forgery Classes and the chosen Length Constraint are presented. The detailed results of the Proof-of-Concept-Study are presented in section 5.4 and the results of the Forgery-Study in section 5.5.

6.1 Survey: Security on Mobile Devices

The results of the first questionnaires are presented in the following. This questionnaire was completed by all participants of the Proof-of-Concept-Study and the Forgery-Study. It should make no difference for this questionnaire in which stage of the user study a participant take part in. Thus, the evaluation is done without considering the stage.

First of all, the results of the capabilities of the mobile device and the usage are presented. Then the perceived security results and afterwards the security perception are presented.

6.1.1 Mobile Device Capabilities and Usage

All participants of the user study own at least one mobile device and use it on a daily basis. 13 use a smartphone whereas 12 use a regular mobile phone. All devices were equipped with a microphone and a camera. 12 of the devices were equipped with GPS. 10 participants said that they can install Apps and already did it to extend the functionality of their devices. All of the devices were able to access the internet, but only 7 participants used mobile internet. 9 devices were equipped with motion sensors. 8 devices were equipped with a 3-dimensional accelerometer and one iPhone 4, provided also a 3-dimensional gyroscope. This is an interesting finding with regard to the implementation of a gesture-based authentication as the sensors are already broadly available.

All participants answered that they use their mobile device for communication like telephony and short messages, but also as telephone book. The participants used their mobile device also as calendar (18), as note book (17) and for to-do lists (9). The 9 participants that use the mobile internet capability also use their device to read and write e-mails, participate on social networks and read online news. All participants agreed to the statement that mobile devices should support and assist them in their daily life. To the statement that the mobile devices offers access to important services agree 13 persons whereas 6 answered neutral and 6 disagree. 18 people answered that they use their device for work-related tasks.

13 persons strongly disagree with the statement that the device is more important than the stored data whereas 12 persons softly agree.

6.1.2 Mobile Device Security

21 of 25 participants answered that they are aware of the fact that the device stores personal data. 22 said that they know which data is stored and 16 participants value this data as private. Interestingly, only 10 participants believe that the personal data stored on their mobile device is interesting to others, but 24 participants said that privacy is important to them. 20 participants regard security as more important than usability.

All participants use a key lock to avoid unintentional interaction with the device whereas only 11 use an authentication mechanism. Therefore, all participants are aware that unintentional interaction might happen and the effects are well understood. In fact, most of the participants (22) are aware of risks due to unauthorized usage, because they pay special attention to their device in public places. So, they are aware of potential risks due to lost and theft. However, some participants seem to believe that they can protect their device by carefully paying attention to it. The results of the items regarding security and privacy contradict with the small numbers of person, who use an authentication mechanism. Only 4 of 7 participants, who either lost a mobile device or lost it dedicated to theft, use an authentication mechanism on their current device. This leads to the assumption that risks and potential costs of exposing personal data are underestimated even if a potential attack has been experienced. In the survey no evidence was found that the usage of authentication mechanisms depends on either if a mobile phone or a smart phone is used, i.e. the capabilities and also the costs of the device seem to make no difference.

The first questionnaire allows very limited conclusions about the security perception and also security on mobile devices. However, it was found that mobile devices may reveal a lot of information about the owner. On the one hand, it may be that the participants are not aware what data and services are available through the device and therefore the need for user authentication is underestimated. On the other hand, it may be that the participants are aware of the risks but simply ignoring them for unknown reason.

6.2 Enrollment and Validation Samples

In the Proof-of-Concept-Study 15 participants provided 25 samples in 5 iterations à 5 samples of their interpretation for each of the 6 designed gestures as described in detail in section 5.4. The participants were regarded as genuine users of the implemented mechanism. It was decided that the samples of the 3rd iteration of the genuine interpretation were best as enrollment samples. Those should be most representative of all iterations, because in this iteration the user had average training. For validation the samples of iteration 2, 4 and 5 were used only. The 1st iteration was left out completely, because it was assumed that these samples were prone to be outliers as the movement may have evolved due to the influence of training. Therefore, 15 samples were available for validation.

6.3 Forgery Classes

In this section the classes of forgery are presented. The Forgeries are categorized depending on the information available to an attacker about the genuine proof: *Naïve Forgery*, *Semi-naïve Forgery* and *Visual Forgery*. The names were chosen after the presented classes of attacks to behavioral biometric proofs as presented in section 2.2.5.

Naïve and Semi-naïve Forgeries were created by the participants of the Proof-of-Concept-Study. To the first belong all samples that are not based-upon the same gesture. These samples are regarded as naïve attacks, because an attacker has no information about the genuine proof. He can only do random guessing. Nevertheless, some properties of a genuine gesture may be guessable like the start and end position including the orientation of the device. This information was indeed available to all participants of the Proof-of-Concept-Study. Semi-naïve Forgeries are all samples that are based-upon the same gesture, but are samples of another participant. Therefore, more information is available to an attacker, i.e. the visualization and description. An attacker knows at least the general path in addition to the start and end position of the gesture. However, this information does not specify an interpretation completely as the genuine user interpreted it individually. Information about the timing, the real path and also the rotation of the device during the performance are not available.

All samples gathered in the Forgery-Study are Visual Forgeries of the genuine interpretation of a gesture, which was visual disclosed to the forgers. Therefore, qualitative information about the genuine movement is available potentially including the interaction with the push-to-gesture-button. The visualization and description of the gesture were not made available to the forgers. So, they were not supported in any way to attach meaning to the gesture for easier memorization. Visual Forgeries are skilled forgeries, because qualitative information was revealed to an attacker and time for training was available.

The probability to create a successful forgery depends on the knowledge about the proof. Therefore, Visual Forgeries should be more successful than Semi-naïve Forgeries and naïve Forgeries should be the least successful. The probability of a successful forgery does not exclusively depend on information available to a forger and his skills, but also on the precision of the genuine user to repeat the gesture, i.e. the allowed variance.

6.4 Length Constraint

The Length Constraint was created, because the variants of DTW and also HMM handle variations in the length of samples very differently. It limits the maximal allowed difference around the mean length of the enrollment samples as described in section 4.5.

The influence of the Length Constraint parameter E from 30% down to 5% on the exclusion of enrollment, validation and forgeries of the user study is shown in Figure 9. The classes of Naïve Forgery and Semi-naïve Forgery are combined for this analysis, because no information about the timing was available to the attackers.

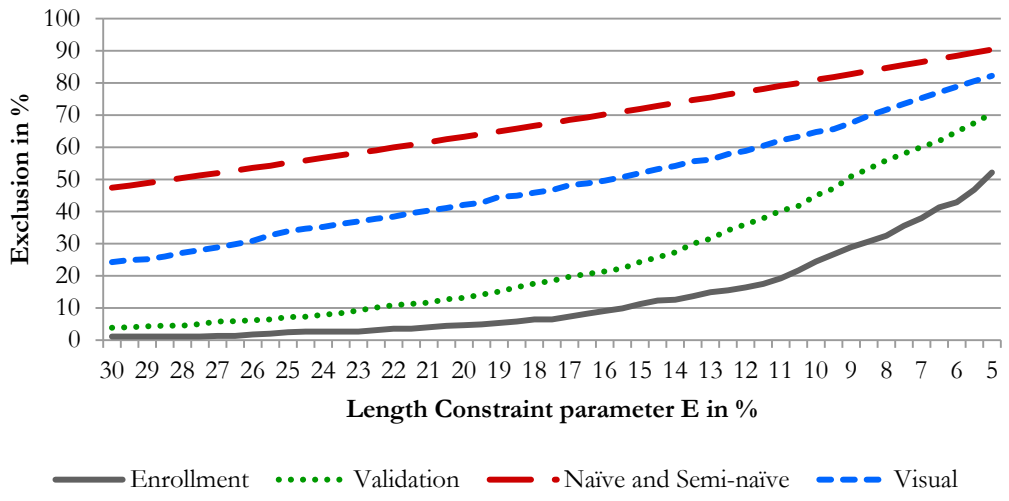


Figure 9: Length Constraint parameter E.
Exclusion of enrollment, validation and forgery samples.

As expected the enrollment samples perform very well with regard to the Length Constraint. The rate of exclusion increases very fast for E smaller than 12%. The validation samples perform slightly worse. As expected the class of Visual Forgery performs better than the combination of Naïve and Semi-naïve Forgery, because timing information was revealed to the forgers. This also includes the interaction with the push-to-gesture-button, which is very important to the Length Constraint. However, a great gap between validation samples and Visual Forgery is apparent, which leads to the assumption that the forgers were not able to learn the timing very well. Overall, the Length Constraint is very successful, because a much higher rate of forgeries is excluded than genuine samples.

The Length Constraint parameter E was set to 23%, because this value includes 97.3% of the enrollment and 90.7% of the validation samples, but excludes 58.3% of the Naïve and Semi-naïve Forgeries, and 36.9% of the Visual Forgeries. The chosen Length Constraint excludes 12.9% of the corresponding validation samples regarding the 12 attacked models of the Forgery-Study. Interestingly, 3 of the 12 models are responsible for 79.1% of the excluded genuine examples.

6.5 Proof-of-Concept-Study

In the following, the results of the Proof-of-Concept-Study using the chosen Length Constraint are presented. Overall, 90 individual interpretations of the designed gestures were recorded in the Proof-of-Concept-Study. First of all, the results of the second questionnaire and then the results of HMM and DTW are presented.

6.5.1 User Survey: User Acceptance

The second questionnaire was completed by the 15 participants of the Proof-of-Concept-Study. It contains questions about their general impression of the gesture-based authentication mechanism, the perceived usability and if it is suited for daily-use as described in section 5.3.2.

None of the participants perceived the implemented gesture-based authentication mechanism as exhausting or annoying with regard to the designed gestures. Only 1 participant perceived the time required to input the gestures as too long. 12 said that they could memorize the gestures easily. However, the participants were asked to interpret the designed gestures and thus qualitative visualization was available to support memorization. In a real use case, the user should pick their individual gestures, which potentially cannot be visualized in a good manner.

All gestures except the Triangle (5) were perceived as appropriate for daily use by the majority of participants. The Triangle is perceived as impractical, not intuitive and ridiculous. This supports the assumption that gestures with hard edges that are directed away from the user are not perceived as usable. The Right-Hand-Rotation gesture (6) is at most preferred with a small advantage over the Infinity gesture (4) by all participants. The Left-Right gesture (1) and the Circle gesture (2) share the 3rd rank. The Left-Right-Arc gesture (3) is on rank 4, but by far more preferred than the Triangle (5). Overall, the Left-right-Arc (3), Infinity (4), Triangle (5) and Right-Hand-Rotation (6) gesture would be used by more than 8 persons as replacement for a password. The reason for the popularity of the Triangle gesture (5) as password replacement remained unknown.

The majority of the participants believed that gesture-based authentication is less secure than password-based mechanism. All designed gestures were not perceived as complex and 7 participants assumed that forgery is easy. It is assumed that this is dedicated to the fact that the Gesture Recorder did not provide feedback during this stage of the user study. Therefore, the participants may not be able to learn and fully understand the inner working of the mechanism. This might prevented the participants to build up trust in the mechanism.

6.5.2 Hidden Markov Model

In the following, the results of HMM are presented. First-order HMMs were evaluated with 4, 8, 12 and 14 states with a model independent threshold for the gesture-based authentication mechanism as described in section 4.4. During the evaluation of the Proof-of-Concept-Study data it was found that the performance of HMMs is increased by smoothing the samples as described in section 4.2 using a span of 5.

The interval of useful thresholds for the Model Acceptance Function of HMM is between 9 and 17. As expected it was found that the performance of First-order HMMs depends on the number of states. Increasing the number of states from 4 to 8 and also from 8 to 12 states improves the overall performance strongly, whereas models with 14 states perform only slightly better than 12 states. Figure 10 shows the ROC diagram of all genuine interpretations against Naïve Forgery and Semi-Naïve Forgery for First-order HMMs with 12 and 14 states. Interestingly, the difference for Semi-naïve Forgery is more pronounced than for Naïve Forgery.

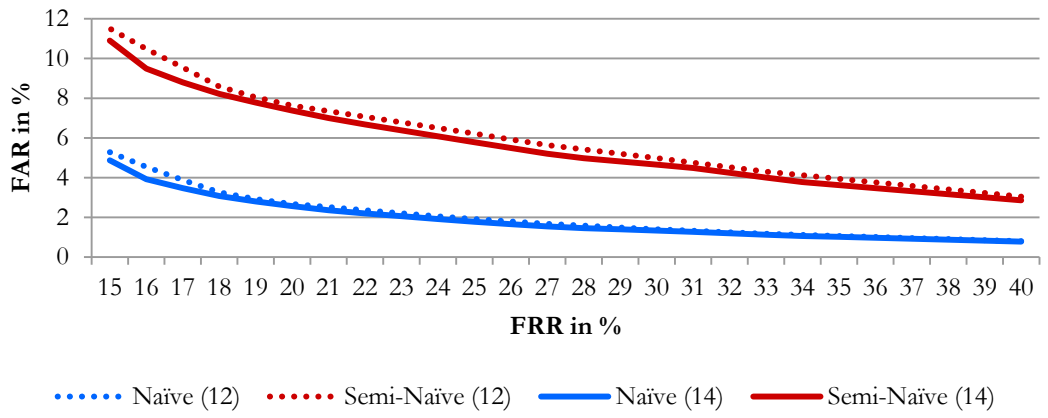


Figure 10: Plot of the FAR against the FRR for 12 and 14 state First-order HMMs for Naïve and Semi-naïve Forgery.

HMMs are promising for authentication with regard to Naïve Forgery. A FAR of 2% can be achieved by a FRR of 23%. However, regarding Semi-naïve Forgery HMMs did not perform well, because even a FRR of 40% leads to a FAR of 3%. A FAR of 3% may be regarded as secure enough for certain non-critical authentication situations, but the high FRR makes HMMs unusable. It is important to note that Semi-naïve Forgeries are based upon the visualization and description only. An attacker had no knowledge about the individual interpretation.

The First-order HMMs performed very badly due to the usage of the likelihood as metric for the similarity. The likelihood that a given sequence O is generated by a HMM is calculated by computing the summed up probability of all paths Q through the HMM with regard to the emission probabilities (Rabiner, 1989):

$$P(O|\lambda) := \sum_{q}^{Q} (P(O|\lambda, q) * P(\lambda, q)). \tag{16}$$

A First-order HMM imposes two requirements on possible paths: a path needs to start in the first state and only state transitions to the direct successor are allowed. However, these restrictions allow a very broad range of possible paths. Assume that an attacker is able to create a sample with constant values for each stream and that the constant values fit very well to the emission distribution of a certain state. The likeliest path would consist of a sequence of state transitions to reach the desired state and afterwards a sequence of state transition to the same state. In this case the likelihood will be fairly high even if the sample is very different. Genuine sample should produce similar paths, which starts in the first state and then traverses through the model and reaches the last state. Therefore, the likelihood alone is not a good similarity metric.

As similarity metric the likelihood of the likeliest path q_{max} through the HMM of a given sample might be more suited. This might be further enhanced by limiting the set of possible paths for q_{max}. However, this issue cannot be studied further in this thesis and is left for future work.

6.5.3 Dynamic Time Warping

In the following, the results of DTW are presented. The cost functions used for in the studied variants of DTW for the implemented gesture-based authentication mechanism are based upon the Minkowski Norm with a weighting function to achieve a certain level of fairness for different scaled dimensions. The minimal and maximal readings of the 3-dimensional accelerometer and the 3-dimensional gyroscope in the user study are shown in Table 3.

	Acceleration in *g*		**Rotation rate in *radians per second***	
Axis	min	max	min	max
x	-2,66	3,05	-16,5	16,64
y	-1,95	3,73	-20,37	20,97
z	-2,08	3,77	-18,14	19,92

Table 3: Minimal and maximal measured values with the Gesture Recorder of acceleration and rotational measurements for the x-, y- and z-axis.

The range of the rotational readings exceeds the range of the acceleration by far in each dimension. However, no a priori knowledge about the importance of the different sensors and also the axis to distinguish genuine and non-genuine samples was available. Therefore, all dimensions were assumed to be equally important. To achieve fairness the acceleration dimensions are rescaled by factor 6. The rotational dimensions were rescaled by using the factor 1. These factors are used in the Weighted Minkowski Norm.

For the gesture-based authentication mechanism the Manhattan Norm $(k = 1)$ and the Euclidian Norm $(k = 2)$ were studied. The Manhattan Norm calculates the absolute difference for each dimension separately and sums up the results whereas the Euclidian Norm calculates the length of the vector, which would connect the compared feature vectors to a triangle. In the evaluation of the results it was found that the Manhattan Norm outperforms the Manhattan Norm in every case. Therefore, the Euclidian Norm is not suited and is omitted in the following. This might be due to the fact that the distance between two vectors calculated with the Manhattan Norm is always greater or equal to the distance calculated by the Euclidian Norm.

The threshold for DTW is calculated with the enrollment samples using the mean and the standard deviation, but using only the mean lead to better performance. It was found that the interval from 1.3 to 3.0 is sufficient. However, the variants of DTW perform very different for a same Q, so this parameter required separate evaluation for each variant. Furthermore, it was found that the authentication performance of the DTW variants is very different. The parameters of the three best performing variants are shown in Table 4. It was found that variants with no slope constraint, i.e. $(P = 0)$, outperformed almost all other variants and only one variant performed comparable.

Variant	1	2	3
Slope constraint	0 (R=30)	0 (R=30)	1
Training mode	1 (mean/dup)	2 (distribution/sum)	1 (mean/dup)
Non-diagonal alignment penalty H	5	5	0
Minkowski parameter k	1	1	1

Table 4: Parameters of the best performing variants of DTW.

Two very different approaches were used to create a model: choosing the cheapest enrollment sample (Training Mode 0) or compute the model based upon integrating all enrollment samples (Training Mode 1 and 2) as described detailed in section 4.3.3.2. Training Mode 1 and 2 calculate the model based upon the enrollment samples using DTW to locally align the samples. Training Mode 1 uses duplication for expansion and averaging for compression, whereas Training Mode 2 does expansion by distribution and compression by summation. Both procedures performed very similar, if no slope constraint was applied. With a slope constraint the Training Mode 1 performed slightly better. Choosing the cheapest sample performed poorly. Another interesting result is that all variants with an applied slope constraint and non-diagonal alignment penalty of 5 were outperformed by the similar variants without penalty.

The ROC diagram of the three best performing variants for Naïve Forgery and Semi-naïve Forgery is shown in Figure 11. For a FRR greater than 20% all variants perform very similar with regard to Naïve Forgery. Variant 3, i.e. slope constraint 1, performed by far best for a FAR under 20%. For Semi-naïve Forgery the performance of variant 1 and 2 is very similar for a FRR greater than 20%. The performance of variant 3 is slightly worse in this range.

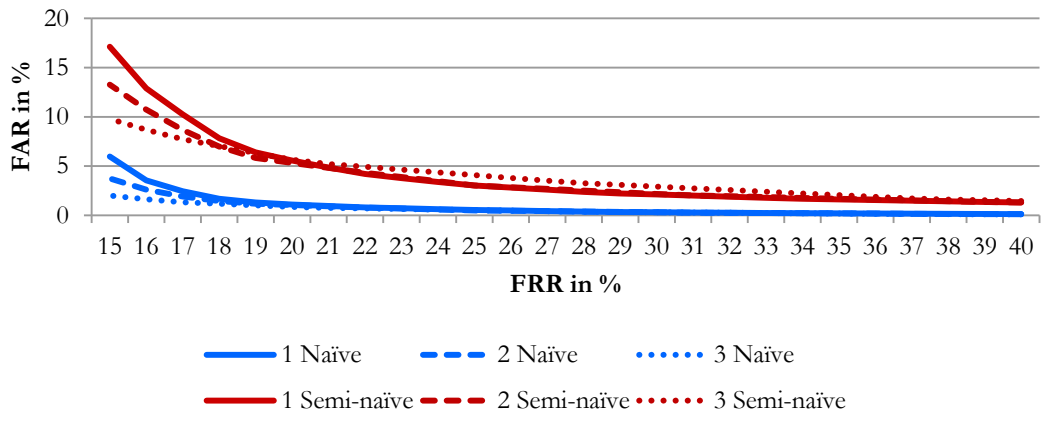


Figure 11: ROC diagram for the best DTW variants for Naïve Forgery and Semi-naïve Forgery of all Proof-of-Concept-Study models.

The results of the variant 1 and 2 regarding Naïve Forgery and Semi-naïve Forger are very similar with a slight advance of variant 2. The only difference of these variants is the Training Mode. Therefore, Training Mode 2 seems to be more suited in this constellation for DTW with no slope constraint, because variant 2 seems to be superior for a low FRR.

6.6 Forgery-Study

In the prior sections, the three best performing variants of HMM and DTW regarding Proof-of-Concept-Study were presented. However, Naïve Forgery and Semi-naïve Forgery samples are mere accidental than motivated attacks. In the Forgery-Study a realistic attack scenario was studied based upon the fact that a gesture is visually disclosed in the authentication process. Therefore, an attacker may be able acquire information about the genuine gesture and can use this knowledge to train his ability to create qualitative forgeries. For the Forgery-Study 12 interpretations of the designed gestures were chosen and the videos of the enrollment samples shown to the forgers. The evaluation of this attack allows reasoning about security of gesture-based authentication mechanisms more strongly than the naïve attacks. In the following, the results of the 14-state HMM and the three best performing DTW variants with regard to the Visual Forgeries are presented starting with HMM. At the end of this section the influence of training and the results of the third questionnaire are discussed. Visual Forgeries should perform better than Semi-naïve Forgeries, because more qualitative information is available to an attacker.

The 14-state HMM with regard to Naïve Forgery and Semi-naïve Forgery did not perform well. This is also true with regard to the 12 attacked interpretations of the designed gestures. The ROC diagram of the 14-state HMM for the studied classes of forgery is shown in Figure 12.

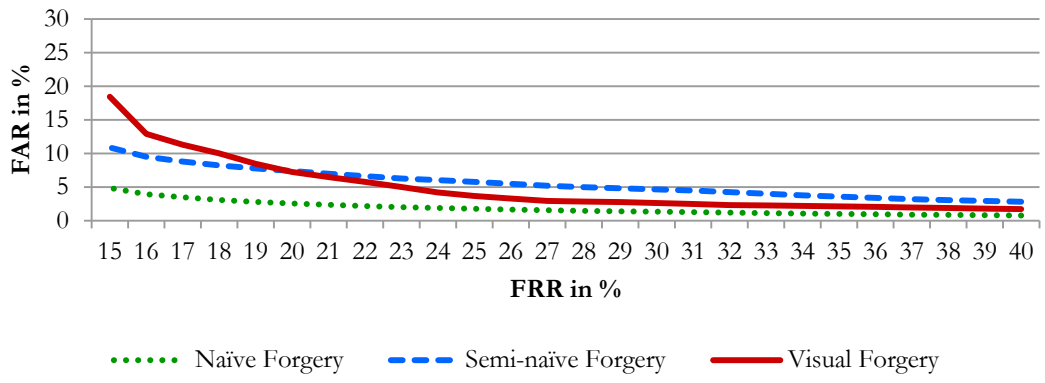


Figure 12: ROC diagram of 14-state First-order HMMs for Naïve Forgery, Semi-naïve Forgery and Visual Forgery.

An unexpected result is encountered: Semi-naïve Forgery achieves a higher FAR than Visual Forgery for a FRR greater than 20%. This means, an attacker knowing less about the genuine gesture is able to create forgeries that are more likely to be accepted as genuine. Visual Forgery outperforms Semi-naïve Forgery for FRR less than 20% only. This supports the assumption that the likelihood is unsuited as similarity metric. However, it is very important to notice that the achieved FAR with regard to the FAR of Visual Forgery is fairly low. This very interesting and astonishing effect is not noticed for the variants of DTW as shown in Figure 13. The different forgery classes perform as expected. The more information is available to an attacker the higher the probability to create successful forgeries. It is interesting that the three variants perform very similar regarding the attacked gestures only.

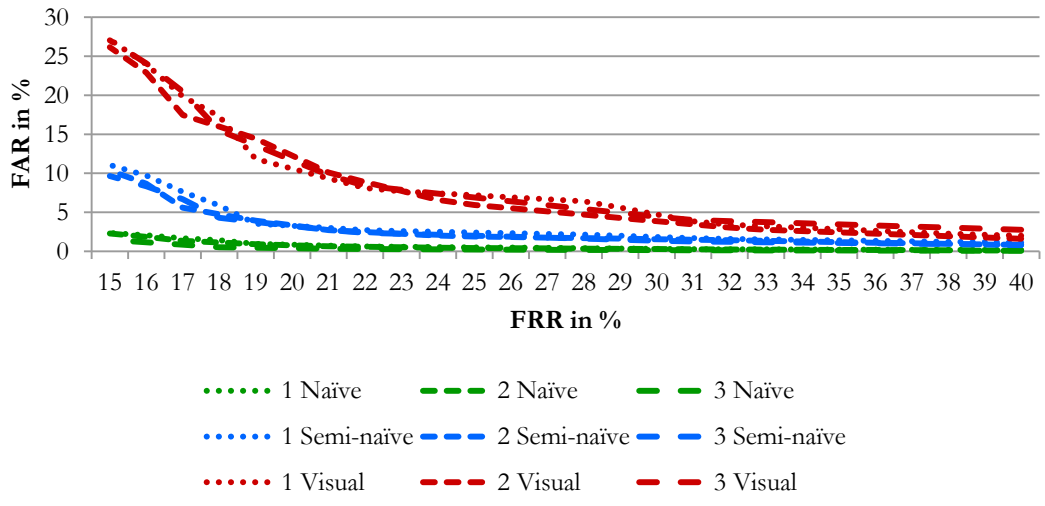


Figure 13: ROC diagram of best DTW variants (1-3) for Naïve Forgery, Semi-naïve Forgery and Visual Forgery.

The effect that variant 3 outperforms the other two variants regarding Naïve and Semi-naïve Forgery for a FRR smaller than 20% is noticeable. Regarding Visual Forgery only the three variants perform very similar with a slight advance of variant 2. In fact, only 100 Visual Forgery samples by 10 forgers per attacked gesture were created and the difference therefore might not be significant.

6.6.1 Comparison DTW and HMM

In the following, the performance of the 14-states HMM and DTW variant 2 are compared. DTW variant 2 was chosen as it seems to be performing slightly better. The ROC diagram for these variants of DTW and HMM is shown in Figure 14.

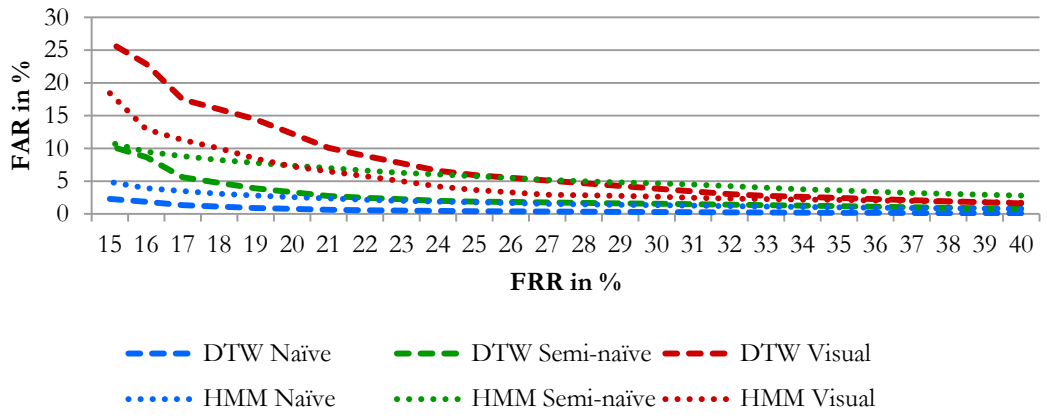


Figure 14: ROC diagram for DTW variant 2 and 14-state HMM for Naïve Forgery, Semi-naïve Forgery and Visual Forgery.

For Naïve Forgery and Semi-naïve Forgery DTW variant 2 performs by far better than HMM. Even for a FRR of 20% DTW accepts less than 0.8% Naïve Forgery samples and less than 3.3% Semi-naïve Forgery samples. However, this is very different for Visual Forgery. For this realistic class of attacks HMM performs by far better than DTW. The difference decreases but does not vanish for a greater FRR. Even at a FRR of 40% HMM achieves a FAR of 1.7%, whereas DTW only achieves 1.9%. Therefore, the 14-state First-order HMM in combination with the likelihood as similarity metric is more suited to reject Visual Forgery samples than DTW. The reason for this effect is unknown and requires further work by studying the characteristics of HMMs and the used similarity metric.

So far all attacked gestures were evaluated together. This does not allow any conclusion about the quality, i.e. the similarity of enrollment and validation samples, of the individual interpretations. To gain a more detailed view the attacked gestures were evaluated with DTW variant 2 using the Visual Forgery samples. The quality of each attacked gesture is estimated by calculating how many of the 15 validation samples are accepted before the first Visual Forgery sample would be accepted. The result of this analysis is shown in Figure 15.

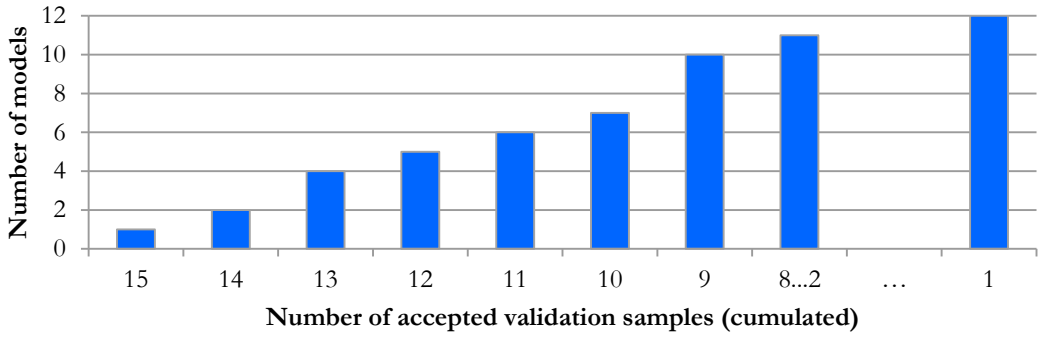


Figure 15: Detailed analysis of attacked gestures
Comulated sum of attacked gestures that accept number of validation samples without accepting a Visual Forgery sample (DTW variant 2).

As it can be seen one of the attacked interpretations is perfect, i.e. no Visual Forgery is accepted by accepting all validation samples. Half of the attacked gestures achieved a FRR smaller or equal to 26.6% (11 of 15). 11 out of 12 gestures achieved a FRR greater than 50% without accepting a Visual Forgery. One of the gestures is not usable at all, because one validation sample is accepted only (FRR 93.3%) before the first Visual Forgery sample is accepted.

The results of the separate analysis are promising as it was found that most of the attacked gestures achieved a certain quality. This supports the assumption that people can memorize and repeat a gesture similar enough, so gesture-based authentication is feasible on handheld mobile devices.

6.6.2 Influence of Training

In the Forgery-Study each forger created 10 samples of each attacked interpretation. Before creating the first 5 forgeries, the video recordings of the enrollment samples were shown once to a forger. After limited time the forger created the first set of forgeries. This procedure was repeated to record the 2nd iteration. Thus, the 2nd iteration should perform better as more time for learning and training was available.

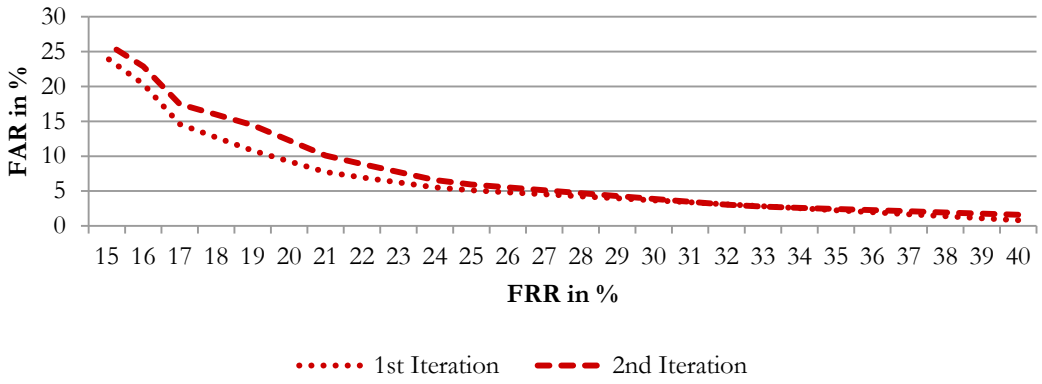


Figure 16: ROC diagram of DTW variant 2 for Visual Forgery depending on the iteration.

Figure 16 shows the ROC diagram of DTW variant 2 depending on the iteration of recording. The 2nd iteration performs slightly better than samples of iteration 1. This supports the assumption that additional training increases the ability of creating successful forgeries. However, one gesture could not be forged successfully at all and therefore training did not improve the performance of the forgeries with regard to this gesture. In fact, amount of time for training was very limited and therefore does not allow drawing significant conclusions.

6.6.3 Survey: Forger's Perspective

The participants of the Forgery-Study completed the third questionnaire. For each of the 12 attacked gestures 6 statements needed to be judged. 3 statements are of special interest with regard to forgery: the perceived complexity, judged learnability and the estimated ability to create successful forgeries.

A gesture that is perceived as highly complex should be perceived as less learnable than a less complex one. Thus, more training should be required to learn a more complex perceived gesture. The assessed perceived complexity of the attacked gestures is very diverse. Only one interpretation based upon the Triangle (5) shows a clear trend that it is perceived as complex. 3 interpretations - based upon Left-Right (1), Circle (2) and Triangle (5) - are perceived by the majority as not so complex. The results of the judged learnability are more clearly, but inconsistent to the perceived complexity. It was found that all gestures except one are perceived as easy to learn. All forgers more or less agree to this even for the one, which was perceived as complex. The estimated ability to create exact and successful forgeries leads to similar results. The majority of forgers agreed with this statement for 9 of the 12 attacked gestures.

The results achieved by the used machine learning algorithms with regard to the Forgery-Study are very different. In combination with the results of the results of this questionnaire it seems as if the forgers overestimated their skills to create successful forgeries.

7 Conclusion

In this thesis a gesture-based authentication mechanism for mobile devices using a 3-dimensional accelerometer in combination with a 3-dimensional gyroscope was proven as being theoretically as well as practically feasible. In addition, it was shown that a gesture-based mechanism is also perceived as usable.

In this chapter the results of this thesis are summarized and future work is pointed out.

7.1 Summary

The user survey affirmed that authentication mechanisms on mobile devices are often. Multiple reasons may apply for that including unawareness of exposable data, accessible services, ignorance, and underestimation of risks. In addition, authentication mechanisms are often regarded as avoidable obstacles. The implemented gesture-based authentication mechanism does not necessarily make users aware of threats and risks, but seems to be more suited for mobile devices due to limitations of the user interface and the interaction style. The general feasibility of the implemented gesture-based authentication mechanism was proven in the Proof-of-Concept-Study. It was shown that participants were able to learn and repeat a gesture similar enough to be sufficient for authentication. Furthermore, they were able to repeat a gesture similar enough in different posture, i.e. standing and sitting. Interestingly, even the interpretation based upon the same gesture, i.e. Semi-naïve Forgery, differed enough, to be useful for authentication. Furthermore, it was shown Naïve Forgery samples were rejected very successfully.

A realistic attack to the gesture-based authentication mechanism was studied in the Forgery-Study. The forgers were shown video recordings of the enrollment samples to enable them to create Visual Forgery samples. As expected the Visual Forgeries were more successful than the other two classes, because the forger was able to acquire knowledge about the interpretation. However, even for Visual Forgery the performance of the mechanism was promising. All in all, the studied machine learning algorithms had their advantages and drawbacks. The variants of DTW performed well for Naive Forgery and Semi-naïve Forgery. Regarding these forgery classes DTW outperforms HMM by far. An unexpected finding was encountered regarding Visual Forgery. 14-state First-order HMMs performed better for Visual Forgery than for Semi-naïve Forgery and also outperforms DTW for Visual Forgery. However, DTW is considered more useful as it rejects simpler forgeries more precisely. The studied video recording attack is realistic, but the videos available to the forgers would be very hard to acquire in daily-life as they were recorded under laboratory conditions.

The gesture-based authentication mechanism is also promising from the usability perspective. 12 of 15 participants of the Proof-of-Concept said that the gestures could be memorized and recalled easily. 10 of them would use the gesture-based authentication mechanism in public places. Furthermore, 4 of the designed gestures were considered useful as alternatives for passwords.

7.2 Future Work

Future work is necessary to refine the gesture-based authentication mechanism in multiple dimensions before it may be implemented successfully in daily life. The studied machine learning algorithms performed very different depending on the forgery classes. HMMs achieved a lower FAR for Visual Forgery and DTW performed better with the other forgery classes. Nevertheless, HMMs showed that even Visual Forgery samples can be rejected successfully. To achieve a better overall performance it should be possible to integrate both algorithms or refine one algorithm alone. Furthermore, the push-to-gesture-button approach might be refined to allow small variations as the implemented approach used only the measurements, if the button was pressed. This might be unsuited, because the user might not hit the button perfectly every time and therefore introduce additional penalty.

Not only the algorithms need to be refined, but also further user studies are required. The conducted user study was done under artificial conditions in a laboratory and the Proof-of-Concept-Study only lasted 90 minutes. Therefore, it can be assumed that physical and mental condition of the participants was almost constant, which is not true in daily life. It needs to be studied how the mental and physical state of a user influences the performance of his gesture. The physical state is not restricted to the body, but also includes wearing different clothes or carrying a bag. Furthermore, the user study was only long enough to show that a user is able to train and repeat a gesture in a short period of time. This does not allow inferring the user is able to memorize and repeat the movement similar enough later on. This is indeed very important for the authentication mechanism, because otherwise the user would be required to repeat the enrollment process. An approach to adapt the learned characteristics slowly after a successful authentication like presented in Matsuo et al. can might be promising (Matsuo, et al., 2007).

Furthermore, the Proof-of-Concept-Study was conducted with the Gesture Recorder, which does not provide feedback about the similarity. This is a problem in two ways. First of all, the participants of the Proof-of-Concept-Study using the mechanism for the first time needed to rely on their self-perception of similarity to estimate their ability to repeat the gesture. Automatic feedback would support the user to understand the characteristics of the mechanism. The same problem applies to the Forgery-Study. The forgers had the same problem as their forgery performance could only be assessed by themselves. Therefore, it is required to implement a fully functional prototype. In further user studies such a prototype can be used to determine usability and security perception and the resilience against attacks with feedback in more detail. Further studies are also required with regard to resistance against video recording attacks with extensive training and feedback.

A gesture-based authentication can only guarantee a certain level of security if the used gesture is complex. The automatic assessment of gesture complexity requires future work as the usage of a not complex gesture would break security. This may be done by developing a metric similar to password quality metrics. Using such a metric during the enrollment process allows the rejection of gestures, which are too simple and therefore regarded as insecure.

Future work is also required as the user study was done under artificial conditions related regarding the motion sensors. During the recording the participants were standing or sitting. The accelerometer and gyroscope only measured the user movement and the gravity. However, in many situations other sources of acceleration and rotation might influence the measurements. For example if the user is walking or riding a bus. To be useful on mobile devices the mechanism needs to be also reliable in such situations. Therefore, the gesture-based authentication mechanism needs to be studied in realistic situations.

Nevertheless, gesture-based user authentication using a 3-dimensional accelerometer and a 3-dimensional gyroscope implemented and studied in this thesis is promising for usage on mobile devices. The gesture-based mechanism is well suited for mobile devices, because it overcomes the limitations of currently available authentication mechanisms imposed by the form factor of mobile devices. Furthermore, authentication can be done in natural manner and eyes-free. In addition, the mechanism allows the input of a complex secret in a fast manner and the genuine user rather trains than memorizes.

Future work will have to prove the potentials of gesture-based user authentication especially for mobile devices.

Thanks to all persons, who made this work possible
and supported it.

Thanks to Brandenburg.

Thanks to Regine.

i Bibliography

Abdulla, Waleed H., Chow, David and Sin, Gary. 2003. Cross-words Reference Template for DTW-based Speech Recognition Systems. *TENCON 2003. Conference on Convergent Technologies for Asia-Pacific Region.* October 15-17, 2003, Vol. 4, pp. 1576-1579.

Adams, Anne and Sasse, Martina Angela. 1999. Users are not the Enemy. *Communications of the ACM.* December 1999, Vol. 12, 42, pp. 41-46.

Alpcan, Tansu, et al. 2008. A Lightweight Biometric Signature Scheme for User Authentication over Networks. *SecureComm '08 Proceedings of the 4th international conference on Security and privacy in communication networks.* September 2008.

Apple Inc. 2011. iOS Developer Library: Event Handling Guide for iOS. [Online] 2011. [Cited: January 7, 2011.] http://developer.apple.com/library/ios/#documentation/EventHandling/Conceptual/EventHandlingiPhoneOS/MotionEvents/MotionEvents.html.

Authorless. 2004. *The new international Webster's Student Dictionary International Encyclopedic Edition.* Köln : Bellavista, 2004. ISBN 3898939804.

Aviv, Adam J., et al. 2010. Smudge Attacks on Smartphone Touch Screens. [ed.] USENIX Association. *WOOT '10 Proceedings of the 4th USENIX conference on Offensive technologies.* August 11-13, 2010.

Ballard, Lucas, Lopresti, Daniel and Monrose, Fabian. 2007. Forgery Quality and ist Implications for Behavioral Biometric Security. *IEEE Transactions on Systems, Man, and Cybernetics.* October 2007, Vol. 37, 4, pp. 1107-1118.

Baumgarten, Uwe and Eckert, Claudia. 2001. Mobile, but Nevertheless secure? *it+ti - Informationstechnik und Technische Informatik.* 2001, Vol. 43, 5, pp. 254-263.

Bichler, David, Stromberg, Guido and Huemer, Mario. 2005. Security and Privacy for Pervasive Computing. [ed.] Falko Dressler and Jürgen Kleinöder. Erlangen, Germany : Institut für Informatik, Friedrich-Alexander-Universität Erlangen-Nürnberg, 2005, Vol. 38, pp. 48-55.

Bilmes, Jeff A. 2006. What HMMs can do. [ed.] Oxford University Press. *IEICE - Transactions on Information and Systems.* March 2006, Vols. E89-D, 3.

—. 2004. What HMMs can't do. *Beyond HMM: Workshop on Statiatical Modeling Approach for Speech Recognition.* December 2004.

Bishop, M. Christopher. 2007. *Pattern Recognition and Machine Learning (Information Science and Statistics).* Secaucus, NJ, USA : Springer, 2007. Vol. 1st ed. 2006. Corr. 2nd printing. ISBN 9780387310732.

Bishop, Matt. 2002. *Computer Security: Art and Science.* s.l. : Addison-Wesley Professional, 2002. ISBN 9780201440997.

Castelluccia, Claude and Mutaf, Pars. 2005. Shake Them Up! A movement-based pairing protocol for CPU-constrained devices. *MobiSys '05 Proceedings of the 3rd international conference on Mobile systems, applications, and services.* 2005, pp. 51–64.

Chong, Ming Ki and Marsden, Gary. 2009. Exploring the Use of Discrete Gestures for Authentication. [ed.] Tom Gross, et al. *Lecture Notes in Computer Science: Human-Computer Interaction – INTERACT '09.* Heidelberg, Germany : Springer, 2009, Vol. 5727, pp. 205-213.

Chong, Ming Ki. 2009. *Usable Authentication For Mobile Banking.* Cape Town, South Africa : University of Cape Town, 2009.

Chong, Ming Ki, Marsden, Gary and Gellersen, Hans. 2010. GesturePIN: Using Discrete Gestures for Associating Mobile Devices. *MobileHCI '10 Proceedings of the 12th international conference on Human computer interaction with mobile devices and services.* September 7-10, 2010.

Clarke, Nathan L., Furnell, Steven M. and Reynolds, Paul L. 2002. Biometric Authentication for Mobile Devices. *Proceedings of the 3rd Australian Information Warfare and Security Conference.* November 28-29, 2002, pp. 61-69.

Dietterich, Thomas G. 2002. Machine Learning for Sequential Data: A Review. *Proceedings of the Joint IAPR International Workshop on Structural, Syntactic, and Statistical Pattern Recognition.* 2002.

Dworschak, Manfred. 2011. Im Netz der Späher. *Der Spiegel.* January 10, 2011, 2/2011.

edepot.com. 2010. *iPhone, iPad, and iPod Touch Secrets.* [Online] 2010. [Cited: January 14, 2011.] http://www.edepot.com/iphone.html.

Eren, Evren and Detken, Kai-Oliver. 2006. *Mobile Security, Risiken mobiler Kommunikation und Lösungen zur mobilen Sicherheit.* Munich, Germany : Hanser Fachbuchverlag, 2006. ISBN 9783446404588.

Falaki, Hossein, Mahajan, Ratul and Kandula, Srikanth. 2010. Diversity in Smartphone Usage. *MobiSys '10 Proceedings of the 8th international conference on Mobile systems, applications, and services.* June 15-18, 2010.

Fang, Ping, et al. 2005. Improved DTW Algorithm for Online Signature Verification Based on Writing Forces. [ed.] Guang-Bin Huang, Xiao-Ping Zhang and De-Shuang Huang. *Lecture Notes in Computer Science: Advances in Intelligent Computing.* Heidelberg, Germany : Springer, 2005, Vol. 3644, pp. 631-640.

Farella, Elisabetta, et al. 2006. Gesture Signature for Ambient Intelligence: A Feasibility Study. [ed.] Kenneth Fishkin, et al. *Lecture Notes in Computer Science: Pervasive Computing.* Heidelberg, Germany : Springer, 2006, Vol. 3968, pp. 288-304.

Fierrez, Julian, et al. 2007. HMM-Based On-Line Signature Verification: Feature Extraction and Signature Modeling. *Journal Pattern Recognition Letters.* December 2007, Vol. 28, 16.

Guerra Casanova, J., et al. 2010. A Real-Time In-Air Signature Biometric Technique Using a Mobile Device Embedding an Accelerometer. [ed.] Filip Zavoral, et al. *Communications in Computer and Information Science: Networked Digital Technologies.* Heidelberg, Germany : Springer, 2010, Vol. 87.

Hayashi, Eiji, et al. 2008. Use Your Illusion: Secure Authentication Usable Anywhere. *SOUPS '08 Proceedings of the 4th symposium on Usable privacy and security.* July 23-25, 2008.

Haze, Timothy J., et al. 2007. Multi-model Face and Speaker Identification on a Handheld Device. [ed.] Riad Hammoud, Besma Abidi and Mongi Abidi. *Signals and Communication Technology: Face Biometrics for Personal Identification.* Heidelberg, Germany : Springer, 2007, pp. 123-138.

Hinckley, Ken, et al. 2005. Foreground and Background Interaction with Sensor-Enhanced Mobile Devices. *ACM Transactions on Computer-Human Interface (TOCHI).* March 2005, Vol. 12, 1, pp. 1-22.

Hinckley, Ken, et al. 2000. Sensing Techniques for Mobile Interaction. *UIST '00 Proceedings of the 13th annual ACM symposium on User interface software and technology.* November 5-8, 2000.

Holmquist, Lars Erik, et al. 2001. Smart-Its Friend: A Technique for Users to Easily Establish Connections between Smart Artefacts. [ed.] Gregory Abowd, Barry Brumitt and Steven Shafer. *Lecture Notes in Computer Science: Ubicomp 2001: Ubiquitous Computing.* Heidelberg, Germany : Springer, 2001, Vol. 2201, pp. 116-122.

Hulsebosch, Robert, et al. 2003. Pioneering advanced mobile privacy and security. [ed.] Chris J. Mitchell. *IET Telecommunications Series: Security for Mobility.* s.l. : The Institution of Engineering and Technology, 2003, pp. 383-432.

Igarashi, Kei, et al. 2004. Biometric Identification Using Driving Behavioral Signals. *ICME '04 Multimedia and Expo.* June 27-30, 2004, pp. 65-68.

Itakura, Fumitada. 1975. Minimum Prediction Residual Principle Applied to Speech Recognition. *IEEE Transactions on Acoustics, Speech, and Signal Processing.* February 1975, 23, pp. 67-72.

Jansen, Wayne. 2004. Authenticating Mobile Device Users Through Image Selection. *Data Security 2004.* 2004.

—. 2003. Authenticating Users on Handheld Devices. *Proceedings of the Canadian Information Technology Security Symposium.* 2003.

Jones, Neil C. and Pevzner, Pavel. 2004. *An Introduction to Bioinformatic Algorithms.* s.l. : MIT Press, 2004. ISBN 9780262101066.

Kainda, Ronald, Flechais, Ivan and Roscoe, A.W. 2010. Security and Usability: Analysis and Evaluation. *ARES '10 International Conference on Availability, Reliability, and Security.* February 15-18, 2010, pp. 275-282.

Kallio, Sanna, et al. 2009. Turn-Based Gesture Interaction in Mobile Devices. [ed.] Stephen Hailes, Sabrina Sicari and George Roussos. *Lecture Notes of the Institute for Computer Sciences, Social Informatics and Telecommunications Engineering: Sensor Systems and Software.* Heidelberg, Germany : Springer, 2009, Vol. 24, pp. 11-19.

Karlson, Amy K., Bernheim Brush, A.J. and Schechter, Stuart. 2009. Can I Borrow Your Phone? Understanding Concerns When Sharing Mobile Phones. *CHI '09 Proceedings of the 27th international conference on Human factors in computing systems.* April 4-9, 2009.

Ketabdar, Hamed, et al. 2010. MagiSign: User Identification/Authentication Based on 3D Around Device Magnetic Signatures. *UBICOMM '10 The Fourth International Conference on Mobile Ubiquitous Computing, Systems, Services and Technologies.* October 25, 2010, p. 30.

Klemmer, Scott R., Hartmann, Björn and Takayama, Leila. 2006. How Bodies Matter: Five Themes for Interaction Design. *DIS '06 Proceedings of the 6th conference on Designing Interactive systems.* June 26-28, 2006.

Liu, Jiayang, et al. 2009. User Evaluation of Lightweight User Authentication with a Single Tri-Axis Accelerometer. *MobileHCI '09 Proceedings of the 11th International Conference on Human-Computer Interaction with Mobile Devices and Services.* September 15-18, 2009.

Lopresti, Daniel P. and Raim, Jarret D. 2005. The Effectiveness of Generative Attacks on an Online Handwriting Biometrics. [book auth.] Takeo Kanada, Anil Jain and Nalini Ratha. [ed.] Springer. *Lecture Notes in Computer Science: The Effectiveness of Generative Attacks on an Online Handwriting Biometric.* Heidelberg, Germany : Springer, 2005, Vol. 3546, pp. 267-310.

Mäntyjärvi, Jani, et al. 2004. Identifying Users of portable Devices from Gait Pattern with Accelerometers. *ICASSP ' 05 IEEE International Conference on Acoustics, Speech, and Signal Processing.* March 18-23, 2004, pp. 973-976.

Marsland, Stephan. 2009. *Machine Learning: An Algorithmic Perspective.* s.l. : Chapman and Hall/CRC, 2009. ISBN 9781420067187.

Matsuo, Kenji, et al. 2007. Arm Swing Identification Method with Template Update for Long Term Stability. [ed.] Seong-Whan Lee and Stan Li. *Lecture Notes in Computer Science: Advances in Biometrics.* Heidelberg, Germany : Springer, 2007, Vol. 4642, pp. 211-221.

Mayrhofer, Rene and Gellersen, Hans. 2009. Shake Well Before Use: Intuitive and Secure Pairing of Mobile Devices. *IEEE Transactions on Mobile Computing.* 8, 2009, Vol. 8, 6, pp. 792-806.

Mayrhofer, Rene. 2007. Pervasive Computing Security. [Online] August 3, 2007. [Cited: January 7, 2011.] http://www.mayrhofer.eu.org/downloads/presentations/2007_08_03-BaCaTec-SummerSchool-PervasiveHealth.pdf.

Mitra, Sushmita and Acharya, Tinku . 2007. Gesture Recognition: A Survey. *IEEE Transactions on Systems, Man, and Cybernetics, Part C: Applications and Reviews.* May 2007, Vol. 37, 3, pp. 311-324.

Ni, Xudong, et al. 2009. DiffUser: Differentiated User Access Control on Smartphones. *MASS '09 6th International Conference on Mobile Adhoc and Sensor Systems.* November 7-10, 2009.

Niezen, Gerrit. 2008. *Parametric Hidden Markov Models Gesture Recognition.* Pretoria, South Africa : University of Pretoria, 2008.

Ogawara, Koichi, et al. 2001. Extraction of fine motion through multiple observation of human demonstration by DP matching and combined template matching. *ROMAN '01 10th IEEE International Workshop on Robot and Human Interactive Communication.* September 18-21, 2001.

Okumura, Fuminori, et al. 2006. A Study on Biometric Authentication based on Arm Sweep Action with Acceleration. *ISPACS '06. International Symposium on Intelligent Signal Processing and Communications.* December 12-15, 2006, pp. 219-222.

Patel, Shwetak N., Pierce, Jeffrey S. and Abowd, Gregory D. 2004. A Gesture-based Authentication Scheme for Untrusted Public Terminals. *UIST '04 Proceedings of the 17th annual ACM symposium on User interface software and technology.* October 24-27, 2004.

Prekopcsák, Zoltán. 2008. Accelerometer Based Real-Time Gesture Recognition. *POSTER 2008: Proceedings of the 12th International Student Conference on Electrical Engineering.* 2008.

Pylvänäinen, Timo. 2005. Accelerometer Based Gesture Recognition using continuous HMMs. [ed.] Jorge Marques, Nicolás Pérez de la Blanca and Pedro Pina. *Lecture Notes in Computer Science: Pattern Recognition and Image Analysis.* Heidelberg, Germany : Springer, 2005, Vol. 3522, pp. 413-430.

Rabiner, Lawrence R. 1989. A Tutorial on Hidden Markov Models and Selected Applications in Speech Recognition. *Proceedings of the IEEE.* 1989, Vol. 77, 2, pp. 257-286.

Redžić, Milan, et al. 2010. Multimodel Identification of Journeys. *IMPRESS 2010 - 1st International Workshop on Interactive Multimodal, Pattern Recognition in Embedded Systems in conjunction with DEXA.* September 1, 2010.

Renaud, Karen . 2005. Evaluating User Authentication Mechanisms. [ed.] Lorrie, Faith Cranor and Simson Garfinkel. *Security and Usability. Designing Secure Systems that People Can Use Editors.* s.l. : O'Reilly, 2005, pp. 103-128.

Rico, Julie and Brewster, Stephen. 2010. Usable Gestures for Mobile Interfaces: Evaluating Social Acceptability. *CHI '10 Proceedings of the 28th international conference on Human factors in computing systems.* April 10-15, 2010.

Ross, Arun and Jain, Anil K. 2007. Human Recognition Using Biometrics: An Overview. *Annals of Telecommunications.* 2007, Vol. 62, 1/2, pp. 11-35.

Saevanee, Hataichanok and Bhatarakosol, Pattarasinee. 2008. User Authentication using Combination of Behavioral Biometrics over the Touchpad acting like Touch screen of Mobile Devices. *ICEE '08 International Conference on Computer and Electrical Engineering.* July 27-30, 2008, pp. 82-86.

Sakoe, Hiroaki and Chiba, Seibi. 1978. Dynamic Programming Algorithm Optimization for Spoken Word Recognition. *IEEE Transactions on Acoustics, Speech and Signal Processing.* 1978, Vol. 26, 1, pp. 43-49.

Salvador, Stan and Chan, Philip. 2007. Toward Accurate Dynamic Time Warping in Linear Time and Space. *Intelligent Data Analysis.* 2007, Vol. 11, 5.

Schlömer, Thomas, et al. 2008. Gesture Recognition with a WII Controller. *TEI '08 Proceedings of the 2nd international conference on Tangible and embedded interaction.* February 18-20, 2008.

Schmidt, Albrecht, Beigl, Michael and Gellersen, Hans-W. 1999. There is more to Context than Location. *Computers & Graphics Journal.* December 1999, Vol. 23, 6, pp. 893-902.

Schwiderski-Grosche, S. and Knospe, H. 2002. Secure mobile commerce. *Electronics & Communication Engineering Journal.* 2002, pp. 228-238.

Stallings, William and Brown, Lawrie. 2007. *Computer Security: Principles and Practice.* s.l. : Prentice Hall, 2007. ISBN 9780136004240.

Weiss, Roman and De Luca, Alexander. 2008. PassShapes - Utilizing Stroke Based Authentication to Increase Password Memorability. *NordiCHI '08 Proceedings of the 5th Nordic conference on Human-computer interaction: building bridges.* October 20-22, 2008.

Westermann, Tilo. 2010. *I'm Home - Smartphone-enabled Gestural Interaction with Multi-Model Smart-Home Systems.* Berlin : Berlin Institute of Technology, 2010.

Williamson, John, Murray-Smith, Rod and Hughes, Stephen. 2007. Shoogle: Excitatory multimodal interaction on mobile devices. *CHI '07 Proceedings of the SIGCHI conference on Human factors in computing systems.* April/May 28/3, 2007.

Wilson, Andrew D. and Bobick, Aaron F. 1999. Parametric Hidden Markov Models Gesture Recognition. *IEEE Transactions on Pattern Analysis and Machine Intelligence.* 1999, pp. 884-900.

Wobbrock, Jacob O. 2009. TapSongs: Tapping Rhythm-Based Passwords on a Single Binary Sensor. *UIST '09 Proceedings of the 22nd annual ACM symposium on User interface software and technology.* October 4-7, 2009.

Yampolskiy, Roman V. and Govindaraju, Venu. 2008. Behavioural biometrics: a survey and classification. *International Journal of Biometrics.* 2008.

Yampolskiy, Roman V. 2008. Mimicry Attack on Strategy-Based Behavioral Biometric. *ITNG '08 Proceedings of the Fifth International Conference on Information Technology: New Generations.* April 7-9, 2008.

Zoebisch, Frank and Vielhauer, Claus. 2003. A test tool to support brut-force online and offline signature forgery tests on mobile device. *ICME '03. International Conference on Multimedia and Expo.* July 6-9, 2003, Vol. 3.

ii Appendix

ii.i Implementation Gesture Recorder

In this section the implementation of the Gesture Recorder is documented. The Gesture Recorder is implemented in Objective-C for Apple iOS 4.2. It is designed to run on an iPhone 4. The device provides a 3-dimensional accelerometer and a 3-dimensional gyroscope. Access to the sensors is done via the Core Motion Framework. The developer documentation is available in (Apple Inc., 2011). The base class is the *CMMotionManager*. It allows pulling or requesting periodic push of sensor readings. The latter is used for the Gesture Recorder. As update interval a maximum of 100 Hertz is proposed in the developer documentation. During development of the Gesture Recorder 80 Hertz were achieved only.

The interaction with the CMMotionManager is shown in Listing 1. At first CMMotionManager is instantiated and the update interval selected, which need to be provided in seconds. Then a block is passed to the CMMotionManager, which functions as handler to incoming sensor readings. The readings are packed into an instance of *CMDeviceMotion*. The handler passes the object asynchronously to the application thread via the *doDeviceMotion* function. The reference to the object is stored for later usage. Instances of CMDeviceMotion contain the raw gyroscope data, the user acceleration, the gravity, the estimated attitude and a timestamp. The acceleration measurements are automatically filtered to distinguish between user and gravity force.

```
1  #import <CoreMotion/CoreMotion.h>
2  CMMotionManager motionManager = [[CMMotionManager alloc] init];
3  motionManager.deviceMotionUpdateInterval = 1.0 / 100;
4  [motionManager startDeviceMotionUpdatesToQueue withHandler:^
   (CMDeviceMotion *motion, NSError *error) {
5   [self performSelectorOnMainThread:@selector(doDeviceMotion:)
    withObject:motion waitUntilDone:NO];
7  }];
9  -(void) doDeviceMotion:(CMDeviceMotion *) motion {
10  //Store motion data.
11 }
```

Listing 1: Interaction with CMMotionManager

The user interface of the Gesture Recorder consists of the push-to-gesture-button and further a switch to manage the push of motion data. If the switch brought into "On" position, the motion updates are subscripted and un-subscripted, when turned "Off". The push-to-gesture-button is only enabled, when the switch is "On" position. In this way, small delays in data recording due management overhead are prevented. The user interface of the Gesture Recorder is shown in Figure 17. The recorded motion data stored in the CMDeviceMotion objects can be exported as comma-separated-value file via e-mail. This includes all recorded, i.e. switch in "On" position. The information if the push-to-gesture-button is pressed is added as an extra column to the file.

iOS 4.2 is capable of multitasking and thus resources can be used by applications in the background. To avoid influence during recording all applications were closed and only the Gesture Recorder is allowed to run. Furthermore, the flight-mode should be enabled to prevent interruptions.

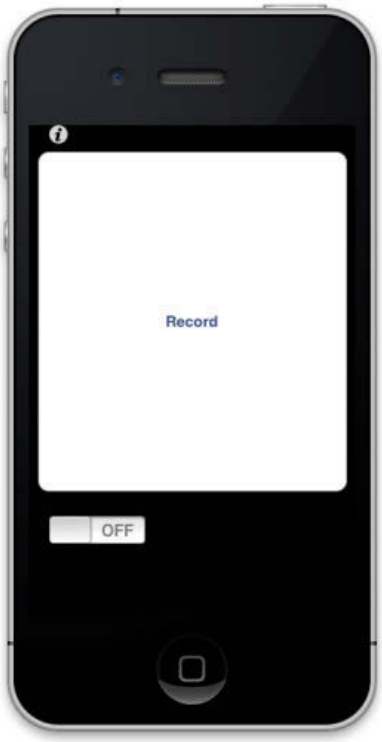


Figure 17: User interface of the Gesture Recorder.

ii.ii Probabilistic Modeling Toolkit 3

The Probabilistic Modeling Toolkit 3 (PMTK3)[13] is an open-source toolkit for Mathworks Matlab or the open-source pendant Octave. It adds support for many machine learning algorithms not included in Matlab. This includes static vector machines, Gaussian processes and also latent variables models. From PMTK3 only the functionality for continuous HMMs was used. The results presented in section 4.4 were derived with the PMTK3 version from 28. February 2011.

The used functions and parameters are presented in the following. The function *hmmFit* allows to fit a HMM to training samples. By default this function uses Maximum-a-posteriori estimation of the parameters. At least the number of states and the emission type must be specified. The emission distribution was chosen to be a multivariate Gaussian. First-Order HMMs can be used by explicitly specifying the initial guess for the initial state distribution π_0 and initial transition matrix A_0. In addition, the prior for the state distribution π_{Prior}, transition matrix A_{Prior} and needs to be specified. π_0 and π_{Prior} are similar vectors, where only the first entry is one and the rest are zero. For A_0 a stationary state transition probability of 90% was used and therefore the probability of a transition to the successor state is 10%. The initial guess for the multivariate Gaussian emission distribution was set to a mean of 0 and a covariance of 1. The function *hmmFit* returns the trained model, which can be used as parameter for the function *hmmLogprob*. This function calculates the likelihood that a given sequence is generated by the HMM. The likelihood is encoded logarithmically.

Thanks to Kevin Murphy and Matt Dunham for developing PMTK3 and their support to get me used to the framework.

[13] PMTK3 is hosted at Google Code under *http://code.google.com/p/pmtk3/*.

ii.iii Hand-out Proof-of-Concept-Study

Sehr geehrte(r) Versuchsteilnehmer(in),
vielen Dank für deine Hilfe bei dieser wissenschaftlichen Studie.

Untersuchungsgegenstand

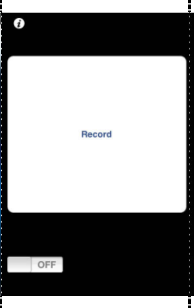
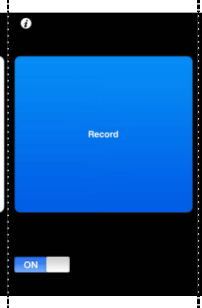
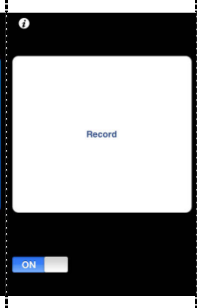
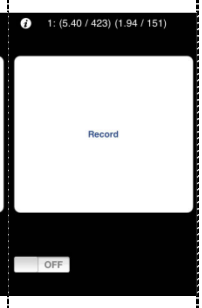
Die Nutzung von mobilen Geräten wie u. a. Smartphone, Handy und Musik-Player unterwegs und in der Öffentlichkeit ist mittlerweile Normalität. Bisher authentifizieren sich Nutzer an mobilen Geräten mit Passwörtern und PINs. Wir untersuchen in dieser Studie die Authentifizierung eines Nutzers an mobilen Geräten mittels Gesten. Im Gegensatz zu passwortbasierten Methoden bietet dies den Vorteil, dass eine Geste einfacher zu merken ist. Zudem beeinflussen die biometrischen Merkmale eines Nutzers die Art und Weise, wie eine Geste ausgeführt wird. Damit kann die Geste nicht einfach von anderen Personen erlernt werden, wie dies bei Passwörtern der Fall ist.

Ablauf des Versuches

Der Versuch besteht aus drei Teilen und dauert ungefähr 90 Minuten. Nach der Begrüßung, Einleitung und deinem Einverständnis, an dem Versuch teilzunehmen, beginnt der Versuch mit dem Ausfüllen eines kurzen Fragebogens zu deiner persönlichen Nutzung mobiler Endgeräte. Im Anschluss daran findet der eigentliche Versuch statt. Bei diesem wird ein mobiles Endgerät von dir auf sechs unterschiedliche Arten bewegt. Deine Bewegungen werden mit den Bewegungs- und Lagesensoren des Gerätes aufgezeichnet. Dabei wirst du per Kamera aufgenommen. In einem zweiten kurzen Fragebogen bewertest Du zum Schluss die von dir durchgeführten Gesten.
Bitte sei bei den Fragebögen ehrlich und scheue dich nicht, offen Kritik zu üben. Es gibt bei diesem Versuch keine falschen Antworten – wir wollen deine Meinung.

Datenaufnahme

Gesten-Rekorder

Starten des Rekorders	Sensoren deaktiviert	Sensoren aktiviert	Aufzeichnung läuft	Aufzeichnung beendet	Sensoren deaktiviert
					

Für das iPhone wurde eine Anwendung entwickelt, mit welcher Bewegungsdaten aufgezeichnet werden. Die Anwendung wird mit dem weißen Icon gestartet. Auf dem Hauptbildschirm der Anwendung befinden sich drei Steuerelemente. Bevor Du eine Geste beginnst, müssen die Senso-

ren aktiviert werden. Bringe hierfür den Schieberegler unten links in die „On"-Position. Ab diesem Zeitpunkt werden Bewegungsdaten aufgezeichnet. Kurz bevor Du die Bewegung beginnst, drücke den „Record"-Button und halte ihn während der Eingabe gedrückt. Höre auf, den Button zu drücken, wenn Du deine Bewegung beendet hast. Nach dem Aufzeichnen einer Bewegung muss der Schieberegler unten links wieder in die „Off"-Stellung gebracht werden. Damit werden die Bewegungssensoren deaktiviert. Im oberen Bereich werden dann die Kenndaten der letzten Aufzeichnung angezeigt.

Beachte:

- **Der „Record"-Button muss bei den verschiedenen Aufzeichnungen ungefähr zum gleichen Zeitpunkt gedrückt werden.**
- **Warte nach dem Aktivieren und vor dem Deaktivieren der Sensoren ungefähr 1 Sekunde.**
- **Halte das Gerät auf eine für dich natürliche Art.**

Szenario

Die Situation während des Versuches wird natürlich künstlich sein. Bewege dich bitte dennoch so, dass du die Geste auch in der Öffentlichkeit ausführen würdest. Du musst die Geste unter Umständen in Situation mit wenig Freiraum ausführen. Stelle dir beispielsweise, vor Du sitzt in einer vollen S-Bahn.

Gesten

Du hältst das Gerät vor dem Körper in der rechten Hand. Stelle dir vor, Du willst mit der Geste die Tastensperre des Gerätes deaktivieren und wirst es gleich im Anschluss benutzen. Bevor Du die Geste beginnst, halte das Gerät so, dass es bequem und fest in deiner Hand liegt. Wähle die Ausgangsposition so, dass sie bequem und realistisch ist. Alle Gesten enden in der Position in der Du sie begonnen hast. Die sechs Gesten werden dir nur grafisch und schriftlich erläutert, aber nicht durch den Versuchsleiter vorgeführt. Probiere die Gesten aus und finde eine Interpretation. Du kannst die Beschreibungen frei interpretieren. Du solltest in der Lage sein, diese Bewegung mehrfach zumindest ähnlich zu reproduzieren. Natürlich kannst Du dich an den Versuchsleiter wenden, falls dir die Beschreibung einer Geste unklar ist.

Aufzeichnung der Gesten

Nachdem Du dir eine Bewegung für die Geste überlegt hast und mit der Bewegung zufrieden bist, beginnt die Datenaufzeichnung mit dem Gesten-Recorder. Um ausreichend Daten zu sammeln, muss deine Bewegung für jede Geste mehrfach aufgezeichnet werden. Die Aufzeichnung einer Geste besteht aus fünf Wiederholungen. Mache zwischen den Bewegungen eine kurze Pause und lege das Gerät aus der Hand. Dies entspricht der realistischen Nutzung eines mobilen Endgerätes. Nach der Aufzeichnung werden die aufgenommen Daten vom Versuchsleiter exportiert und im Anschluss mit der Aufzeichnung der nächsten Geste begonnen. Für jede Geste werden mehrere Aufzeichnungen durchgeführt.

ii.iv Hand-out Forgery-Study

Sehr geehrte(r) Versuchsteilnehmer(in),

vielen Dank für deine Hilfe bei dieser wissenschaftlichen Studie.

Untersuchungsgegenstand

Die Nutzung von mobilen Geräten wie u. a. Smartphone, Handy und Musik-Player unterwegs und in der Öffentlichkeit ist mittlerweile Normalität. Bisher authentifizieren sich Nutzer an mobilen Geräten mit Passwörtern und PINs. Wir untersuchen in dieser Studie die Authentifizierung eines Nutzers an mobilen Geräten mittels Gesten. Im Gegensatz zu passwortbasierten Methoden bietet dies den Vorteil, dass eine Geste einfacher zu merken ist. Zudem beeinflussen die biometrischen Merkmale eines Nutzers die Art und Weise, wie eine Geste ausgeführt wird. Damit kann die Geste nicht einfach von anderen Personen erlernt werden, wie dies bei Passwörtern der Fall ist.

Ablauf des Versuches

Der Versuch besteht aus drei Teilen und dauert ungefähr 90 Minuten. Nach der Begrüßung, Einleitung und deinem Einverständnis an dem Versuch teilzunehmen, beginnt der Versuch mit einem kurzen Fragebogen zu deiner persönlichen Nutzung mobiler Endgeräte. Im Anschluss daran findet der eigentliche Versuch statt.

Bei diesem werden Dir Videos von Versuchspersonen gezeigt. Auf diesen bewegen die Versuchspersonen ein mobiles Gerät auf individuelle Art und Weise. Deine Aufgabe ist es, diese Bewegung möglichst exakt nachzumachen. Mittels der Bewegungs- und Lagesensoren werden die Bewegungen aufgezeichnet und später mit den originalen Daten verglichen.

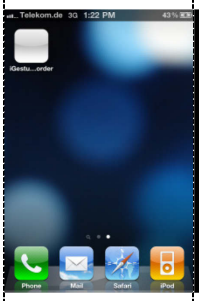
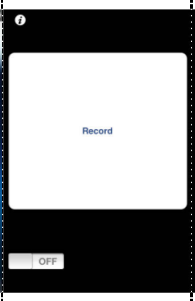
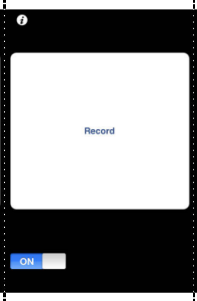
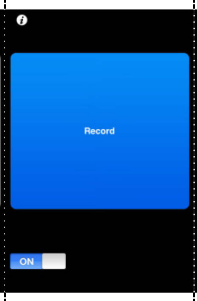
In einem zweiten kurzen Fragebogen bewertest Du zum Schluss die von dir durchgeführten Gesten. Bitte sei bei den Fragebögen ehrlich und scheue dich nicht, offen Kritik zu üben. Es gibt bei diesem Versuch keine falschen Antworten – wir wollen deine Meinung.

Szenario

Du beobachtest jemanden, wie er sein mobiles Gerät mittels einer Geste entsperrt. Aus irgendwelchen Gründen bekommst Du das Gerät in die Hände und willst es benutzen, obwohl Du dazu nicht berechtigt bist. Du musst dafür die Geste des Besitzers möglichst genau nachmachen. Hierzu zählt die Bewegung und Rotation im Raum, das Tempo des Gerätes und das Drücken bzw. Loslassen des Buttons.

Datenaufnahme

Gesten-Rekorder

Starten des Rekorders	Sensoren deaktiviert	Sensoren aktiviert	Aufzeichnung läuft	Aufzeichnung beendet	Sensoren deaktiviert
					

Für das iPhone wurde eine Anwendung entwickelt, mit welcher Bewegungsdaten aufgezeichnet werden. Die Anwendung wird mit dem weißen Icon gestartet. Auf dem Hauptbildschirm der Anwendung befinden sich drei Steuerelemente. Bevor Du eine Geste beginnst, müssen die Sensoren aktiviert werden. Bringe hierfür den Schieberegler unten links in die „On"-Position. Ab diesem Zeitpunkt werden Bewegungsdaten aufgezeichnet. Kurz bevor Du die Bewegung beginnst, drücke den „Record"-Button und halte ihn während der Eingabe gedrückt. Höre auf, den Button zu drücken, wenn Du deine Bewegung beendet hast. Nach dem Aufzeichnen einer Bewegung muss der Schieberegler unten links wieder in die „Off"-Stellung gebracht werden. Damit werden die Bewegungssensoren deaktiviert. Im oberen Bereich werden dann die Kenndaten der letzten Aufzeichnung angezeigt.

Beachte:

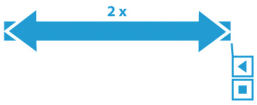
- **Mache die Bewegung möglichst exakt nach. Hierzu zählen Dauer, Geschwindigkeit, Drehung des Geräts und auch die Bewegung im Raum.**
- **Warte nach dem Aktivieren und vor dem Deaktivieren der Sensoren ungefähr 1 Sekunde.**

Aufzeichnung deiner nachgemachten Gesten

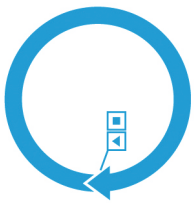
Das Nachmachen besteht aus zwei Phasen. Zuerst wird Dir nur die nachzumachende Bewegung aus drei Perspektiven gezeigt. Du kannst dir jede Perspektive einmal anschauen, wobei fünf Wiederholungen gezeigt werden. Im Anschluss kannst Du währenddessen die Bewegung kurz üben. Dann wird die Bewegung fünfmal aufgezeichnet. Mache zwischen den Bewegungen eine kurze Pause. In der zweiten Phase kannst du dir die Videos noch einmal anschauen, um sie weiter zu üben.

Insgesamt müssen zwölf Bewegungen nachgemacht werden.

ii.v Designed Gestures

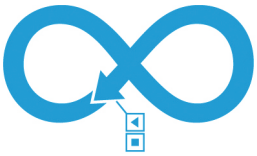


Bewegen Sie das Telefon entsprechend der Pfeilform **parallel zu Ihrem Oberkörper**. Starten und beenden Sie die Bewegung auf der Höhe des Play-/Stop-Symbols.

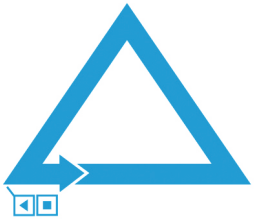


Bewegen Sie das Telefon entsprechend der Pfeilform **parallel zu Ihrem Oberkörper**. Starten und beenden Sie die Bewegung auf der Höhe des Play-/Stop-Symbols.

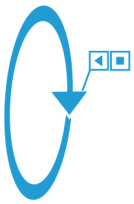
Bewegen Sie das Telefon entsprechend der Pfeilform **parallel zu Ihrem Oberkörper**. Starten und beenden Sie die Bewegung auf der Höhe des Play-/Stop-Symbols.



Bewegen Sie das Telefon entsprechend der Pfeilform **parallel zu Ihrem Oberkörper**. Starten und beenden Sie die Bewegung auf der Höhe des Play-/Stop-Symbols.



Bewegen Sie das Telefon entsprechend der Pfeilform **parallel zum Fußboden**. Starten und beenden Sie die Bewegung auf der Höhe des Play-/Stop-Symbols.



Lassen Sie das Telefon **um Ihr rechtes Handgelenk kreisen**.
Beginnen Sie mit der Drehung nach unten zu Ihrem Körper hin. Öffnen Sie anschließend die eingedrehte Hand mit einer Bewegung nach oben. Das Handy liegt nun auf Ihrer flachen Hand. Diese ist von Ihrem Körper weg gerichtet.

ii.vi Questionnaires

Am Fachbereich *Quality & Usability* an der TU Berlin untersuchen wir die Verwendung von Gesten als Alternative zu Passwörtern zur Anmeldung an mobilen Geräten. Im Fragebogen werden Sie zu Ihrer Nutzung von mobilen Geräten wie Handys, Smartphones, iPods usw. befragt.

1 BEFRAGUNG ZUR NUTZUNG MOBILER GERÄTE

1.1 PERSÖNLICHE ANGABEN

Geschlecht	☐ männlich ☐ weiblich
Alter	_____
Beruf	_____
Bevorzugte Hand	☐ links ☐ rechts
Haben Sie Erfahrungen mit Gestensteuerung (z. B. Nintendo Wii, Kinnect)?	☐ ja ☐ nein ☐ unbekannt

1.2 MOBILE GERÄTE

Nennen Sie das Gerät, welches Sie am häufigsten nutzen. (Modell bzw. Art) Alle folgenden Fragen beziehen sich auf dieses.	_____
Nutzen Sie mobiles Internet?	☐ ja ☐ nein ☐ unbekannt
Nutzen Sie Apps?	☐ ja ☐ nein ☐ unbekannt
Hat Ihr Gerät GPS?	☐ ja ☐ nein ☐ unbekannt

1.3 EINGABEMÖGLICHKEITEN IHRES MOBILEN GERÄTES

Physische Tastatur	☐ Nicht vorhanden ☐ Nummernpad ☐ Vollständige Tastatur
Touchscreen	☐ ja ☐ nein ☐ unbekannt
Bewegungs- und Lagesensoren	☐ ja ☐ nein ☐ unbekannt
Spracherkennung	☐ ja ☐ nein ☐ unbekannt

1.4 ANGABEN ZUR VERWENDUNG MOBILER GERÄTE

Wie häufig nutzen Sie das Gerät im Schnitt pro Tag? Dazu zählt jede Interaktion mit dem Gerät.	☐ weniger als 20 mal ☐ 20 bis 50 mal ☐ 50 bis 100 mal ☐ 100 bis 150 mal ☐ mehr als 150 mal
Wie lange dauert eine solche Nutzung im Durchschnitt?	☐ weniger als 30 s ☐ 30 s bis 60 s ☐ mehr als 60 s
Nutzen Sie die Tastensperre?	☐ ja ☐ nein
Müssen Sie ein Passwort eingeben oder ist Ihr Gerät auf eine andere Art und Weise geschützt?	☐ ja ☐ nein ☐ andere: _____

1.5 NUTZUNG MOBILER GERÄTE

Bitte geben Sie an, wie Sie Ihr mobiles Gerät benutzen und welche Aufgaben Sie damit erledigen.

☐	E-Mail	☐	e-Commerce: z. B. Amazon, eBay
☐	SMS/MMS	☐	Online-News lesen
☐	Telefonie	☐	Microblogging: z. B. Twitter
☐	Uhr	☐	Soziale Netzwerke: z. B. Facebook, StudiVZ, Xing
☐	Online-Banking	☐	Im Internet surfen
☐	MP3/Video-Player	☐	Kamera (Video und Foto)
☐	Spiele	☐	Bücher lesen
☐	Navigation	☐	Chat
☐	Telefonbuch	☐	Kalender
☐	Notizbuch	☐	TODO-Liste
☐	VOIP: z. B. Skype	☐	Sonstiges: _____

1.6 DATEN AUF MOBILEN GERÄTEN

Befinden sich persönliche Daten auf dem Gerät?	☐ ja ☐ nein
Sind diese Daten privat bzw. vertraulich?	☐ ja ☐ nein
Wissen Sie, welche Daten auf dem Gerät gespeichert sind?	☐ ja ☐ nein
Nutzen Sie das Gerät für berufliche Kommunikation?	☐ ja ☐ nein
Sind die Daten auf dem Gerät verschlüsselt?	☐ ja ☐ nein ☐ unbekannt

1.7 VERLUST VON MOBILEN GERÄTEN

Haben Sie bereits ein mobiles Gerät verloren?	☐ ja ☐ nein
Wurde Ihnen bereits ein mobiles Gerät entwendet?	☐ ja ☐ nein ☐ unbekannt

1.8 PASSWÖRTER ALLGEMEIN

Haben Sie Passwörter, welche Sie für mehrere Dienste bzw. Accounts benutzt?	☐ ja ☐ nein
Ist die Eingabe aufwändig?	☐ ja ☐ nein
Dauert die Eingabe zu lang?	☐ ja ☐ nein
Wie häufig ändern Sie Ihr Passwort?	☐ nie (> 1 Jahr) ☐ selten (> 6 Monate) ☐ häufig (> 1 Monat) ☐ sehr häufig (< 1 Monat)
Schreiben Sie Passwörter auf?	☐ ja ☐ nein
Haben Sie schon einmal ein Passwort vergessen?	☐ ja ☐ nein
Geben Sie Passwörter weiter?	☐ ja ☐ nein
Geben Sie Passwörter an andere Personen bewusst weiter?	☐ ja ☐ nein ☐ unbekannt
Achten Sie bei der Eingabe auf fremde Personen in Ihrer Umgebung?	☐ ja ☐ nein

1.9 SICHERHEITSMECHANISMEN

Welche Möglichkeiten bietet Ihr mobiles Gerät zum Schutz vor unautorisiertem Zugriff durch andere Personen?

☐	PIN	☐	Passwort
☐	Fingerabdruck	☐	„Wisch-Geste“
☐	Unbekannt	☐	Sonstiges:_____

1.10 EINSCHÄTZUNG

Bitte bewerten Sie die folgenden Aussagen.
Wählen Sie zwischen **++** für volle Zustimmung und **- -** für vollständige Ablehnung.

	Vollständige Zustimmung								**Vollständige Ablehnung**
Einfachheit ist wichtiger als Sicherheit.	++	☐	☐	☐	☐	☐	☐	☐	- -
Mobile Geräte sollen mich unterstützen.	++	☐	☐	☐	☐	☐	☐	☐	- -
Meine Daten sind für andere uninteressant.	++	☐	☐	☐	☐	☐	☐	☐	- -
Privatsphäre ist mir wichtig.	++	☐	☐	☐	☐	☐	☐	☐	- -
Meine Daten sind nur für mich wichtig.	++	☐	☐	☐	☐	☐	☐	☐	- -
Ich achte in der Öffentlichkeit immer auf mein Gerät.	++	☐	☐	☐	☐	☐	☐	☐	- -
Bei Verlust ist das Gerät wichtiger als die Daten.	++	☐	☐	☐	☐	☐	☐	☐	- -
Mein Gerät bietet mir Zugang zu wichtigen Diensten.	++	☐	☐	☐	☐	☐	☐	☐	- -
Meine Daten sind privat bzw. vertraulich.	++	☐	☐	☐	☐	☐	☐	☐	- -
Die Eingabe eines Passwortes stört mich.	++	☐	☐	☐	☐	☐	☐	☐	- -
Die Daten auf meinem Gerät sind wichtig.	++	☐	☐	☐	☐	☐	☐	☐	- -

1 BEFRAGUNG ZUM VERSUCH

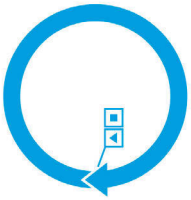
1.1 BEWERTUNG DER EINZELNEN GESTEN

1.1.1 LINKS-RECHTS

	Vollständige Zustimmung							Vollständige Ablehnung
Die Geste ist intuitiv.	++	☐	☐	☐	☐	☐	☐	--
Die Geste ist kompliziert.	++	☐	☐	☐	☐	☐	☐	--
Die Geste ist mir zu aufwändig.	++	☐	☐	☐	☐	☐	☐	--
Die Geste benötigt viel Platz.	++	☐	☐	☐	☐	☐	☐	--
Die Geste ist lächerlich.	++	☐	☐	☐	☐	☐	☐	--
Die Geste ist zu lang.	++	☐	☐	☐	☐	☐	☐	--
Die Geste ist angenehm.	++	☐	☐	☐	☐	☐	☐	--
Die Ausführung der Geste vor fremden Personen wäre mir peinlich.	++	☐	☐	☐	☐	☐	☐	--
Die Geste kann ich mir gut merken.	++	☐	☐	☐	☐	☐	☐	--
Die Ausführung der Geste bereitet mir Probleme.	++	☐	☐	☐	☐	☐	☐	--
Die Geste ist alltagstauglich.	++	☐	☐	☐	☐	☐	☐	--
Ich würde die Geste einem Passwort gegenüber bevorzugen.	++	☐	☐	☐	☐	☐	☐	--
Ich halte die Geste für sicher genug.	++	☐	☐	☐	☐	☐	☐	--

Anmerkungen zur Geste:

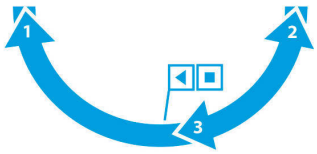
1.1.2 KREIS



	Vollständige Zustimmung							Vollständige Ablehnung
Die Geste ist intuitiv.	++	☐	☐	☐	☐	☐	☐	--
Die Geste ist kompliziert.	++	☐	☐	☐	☐	☐	☐	--
Die Geste ist mir zu aufwändig.	++	☐	☐	☐	☐	☐	☐	--
Die Geste benötigt viel Platz.	++	☐	☐	☐	☐	☐	☐	--
Die Geste ist lächerlich.	++	☐	☐	☐	☐	☐	☐	--
Die Geste ist zu lang.	++	☐	☐	☐	☐	☐	☐	--
Die Geste ist angenehm.	++	☐	☐	☐	☐	☐	☐	--
Die Ausführung der Geste vor fremden Personen wäre mir peinlich.	++	☐	☐	☐	☐	☐	☐	--
Die Geste kann ich mir gut merken.	++	☐	☐	☐	☐	☐	☐	--
Die Ausführung der Geste bereitet mir Probleme.	++	☐	☐	☐	☐	☐	☐	--
Die Geste ist alltagstauglich.	++	☐	☐	☐	☐	☐	☐	--
Ich würde die Geste einem Passwort gegenüber bevorzugen.	++	☐	☐	☐	☐	☐	☐	--
Ich halte die Geste für sicher genug.	++	☐	☐	☐	☐	☐	☐	--

Anmerkungen zur Geste:

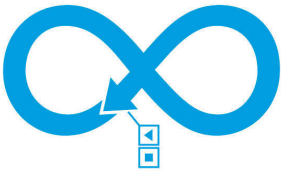
1.1.3 HALBKREIS



	Vollständige Zustimmung							Vollständige Ablehnung
Die Geste ist intuitiv.	++	☐	☐	☐	☐	☐	☐	☐ --
Die Geste ist kompliziert.	++	☐	☐	☐	☐	☐	☐	☐ --
Die Geste ist mir zu aufwändig.	++	☐	☐	☐	☐	☐	☐	☐ --
Die Geste benötigt viel Platz.	++	☐	☐	☐	☐	☐	☐	☐ --
Die Geste ist lächerlich.	++	☐	☐	☐	☐	☐	☐	☐ --
Die Geste ist zu lang.	++	☐	☐	☐	☐	☐	☐	☐ --
Die Geste ist angenehm.	++	☐	☐	☐	☐	☐	☐	☐ --
Die Ausführung der Geste vor fremden Personen wäre mir peinlich.	++	☐	☐	☐	☐	☐	☐	☐ --
Die Geste kann ich mir gut merken.	++	☐	☐	☐	☐	☐	☐	☐ --
Die Ausführung der Geste bereitet mir Probleme.	++	☐	☐	☐	☐	☐	☐	☐ --
Die Geste ist alltagstauglich.	++	☐	☐	☐	☐	☐	☐	☐ --
Ich würde die Geste einem Passwort gegenüber bevorzugen.	++	☐	☐	☐	☐	☐	☐	☐ --
Ich halte die Geste für sicher genug.	++	☐	☐	☐	☐	☐	☐	☐ --

Anmerkungen zur Geste:

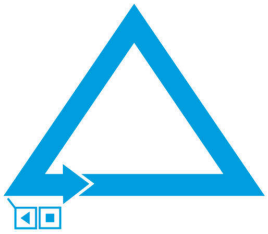
1.1.4 UNENDLICH



	Vollständige Zustimmung							Vollständige Ablehnung
Die Geste ist intuitiv.	++	☐	☐	☐	☐	☐	☐	--
Die Geste ist kompliziert.	++	☐	☐	☐	☐	☐	☐	--
Die Geste ist mir zu aufwändig.	++	☐	☐	☐	☐	☐	☐	--
Die Geste benötigt viel Platz.	++	☐	☐	☐	☐	☐	☐	--
Die Geste ist lächerlich.	++	☐	☐	☐	☐	☐	☐	--
Die Geste ist zu lang.	++	☐	☐	☐	☐	☐	☐	--
Die Geste ist angenehm.	++	☐	☐	☐	☐	☐	☐	--
Die Ausführung der Geste vor fremden Personen wäre mir peinlich.	++	☐	☐	☐	☐	☐	☐	--
Die Geste kann ich mir gut merken.	++	☐	☐	☐	☐	☐	☐	--
Die Ausführung der Geste bereitet mir Probleme.	++	☐	☐	☐	☐	☐	☐	--
Die Geste ist alltagstauglich.	++	☐	☐	☐	☐	☐	☐	--
Ich würde die Geste einem Passwort gegenüber bevorzugen.	++	☐	☐	☐	☐	☐	☐	--
Ich halte die Geste für sicher genug.	++	☐	☐	☐	☐	☐	☐	--

Anmerkungen zur Geste:

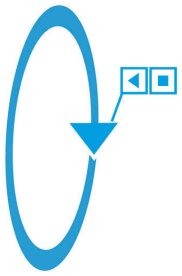
1.1.5 DREIECK



	Vollständige Zustimmung							Vollständige Ablehnung
Die Geste ist intuitiv.	++	☐	☐	☐	☐	☐	☐	--
Die Geste ist kompliziert.	++	☐	☐	☐	☐	☐	☐	--
Die Geste ist mir zu aufwändig.	++	☐	☐	☐	☐	☐	☐	--
Die Geste benötigt viel Platz.	++	☐	☐	☐	☐	☐	☐	--
Die Geste ist lächerlich.	++	☐	☐	☐	☐	☐	☐	--
Die Geste ist zu lang.	++	☐	☐	☐	☐	☐	☐	--
Die Geste ist angenehm.	++	☐	☐	☐	☐	☐	☐	--
Die Ausführung der Geste vor fremden Personen wäre mir peinlich.	++	☐	☐	☐	☐	☐	☐	--
Die Geste kann ich mir gut merken.	++	☐	☐	☐	☐	☐	☐	--
Die Ausführung der Geste bereitet mir Probleme.	++	☐	☐	☐	☐	☐	☐	--
Die Geste ist alltagstauglich.	++	☐	☐	☐	☐	☐	☐	--
Ich würde die Geste einem Passwort gegenüber bevorzugen.	++	☐	☐	☐	☐	☐	☐	--
Ich halte die Geste für sicher genug.	++	☐	☐	☐	☐	☐	☐	--

Anmerkungen zur Geste:

1.1.6 RECHTE-HAND-ROTATION



	Vollständige Zustimmung							Vollständige Ablehnung
Die Geste ist intuitiv.	++	☐	☐	☐	☐	☐	☐	--
Die Geste ist kompliziert.	++	☐	☐	☐	☐	☐	☐	--
Die Geste ist mir zu aufwändig.	++	☐	☐	☐	☐	☐	☐	--
Die Geste benötigt viel Platz.	++	☐	☐	☐	☐	☐	☐	--
Die Geste ist lächerlich.	++	☐	☐	☐	☐	☐	☐	--
Die Geste ist zu lang.	++	☐	☐	☐	☐	☐	☐	--
Die Geste ist angenehm.	++	☐	☐	☐	☐	☐	☐	--
Die Ausführung der Geste vor fremden Personen wäre mir peinlich.	++	☐	☐	☐	☐	☐	☐	--
Die Geste kann ich mir gut merken.	++	☐	☐	☐	☐	☐	☐	--
Die Ausführung der Geste bereitet mir Probleme.	++	☐	☐	☐	☐	☐	☐	--
Die Geste ist alltagstauglich.	++	☐	☐	☐	☐	☐	☐	--
Ich würde die Geste einem Passwort gegenüber bevorzugen.	++	☐	☐	☐	☐	☐	☐	--
Ich halte die Geste für sicher genug.	++	☐	☐	☐	☐	☐	☐	--

Anmerkungen zur Geste:

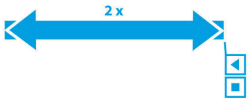
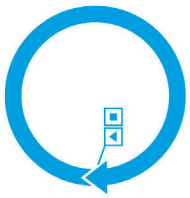
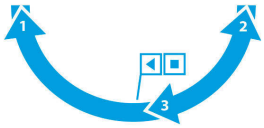
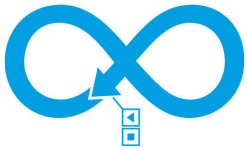
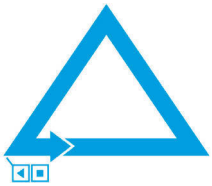
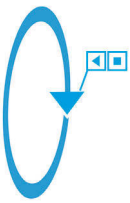
1.2 EINSCHÄTZUNG DER GESTEN

Bitte bewerten Sie die folgenden Aussagen.
Wählen Sie zwischen **++** für volle Zustimmung und **- -** für vollständige Ablehnung.

	Vollständige Zustimmung								**Vollständige Ablehnung**
Die Eingabe der Gesten ist intuitiv.	++	☐	☐	☐	☐	☐	☐	☐	- -
Ich würde gerne eine eigene Geste nutzen.	++	☐	☐	☐	☐	☐	☐	☐	- -
Gesten sind genauso sicher wie Passwörter.	++	☐	☐	☐	☐	☐	☐	☐	- -
Die Ausführung der Gesten an öffentlichen Orten würde mir keine Probleme bereiten.	++	☐	☐	☐	☐	☐	☐	☐	- -
Ich kann mir Gesten gut merken.	++	☐	☐	☐	☐	☐	☐	☐	- -
Gesten können leicht nachgemacht werden.	++	☐	☐	☐	☐	☐	☐	☐	- -
Die Ausführung der Gesten dauert zu lang.	++	☐	☐	☐	☐	☐	☐	☐	- -
Gesten reichen für die Sicherung meines mobilen Gerätes.	++	☐	☐	☐	☐	☐	☐	☐	- -
Die Eingabe eines Passwortes ist einfacher als die einer Geste.	++	☐	☐	☐	☐	☐	☐	☐	- -
Die Eingabe der Gesten ist anstrengend.	++	☐	☐	☐	☐	☐	☐	☐	- -
Die Ausführung der Gesten vor fremden Personen ist peinlich.	++	☐	☐	☐	☐	☐	☐	☐	- -

1.3 BEVORZUGTE GESTE

Welche Geste bevorzugen Sie?
Bringen Sie die Gesten in Ihre persönliche Reihenfolge.
Platz 1 steht für **„Finde ich am besten“** und Platz 6 für **„Gefällt mir am wenigsten“**.

		Platz
Links-Rechts		_____
Kreis		_____
Halbkreis		_____
Unendlich		_____
Dreieck		_____
Rechte-Hand-Rotation		_____

1.4 EIGENE GESTE

Welche weitere(n) Geste(n) würden Sie gerne nutzen?
Bitte skizzieren und beschreiben Sie diese entsprechend.

1.5 ALLGEMEINES FEEDBACK

Bitte schreiben Sie Ihre Anregungen und Anmerkungen zu dem Versuch auf.

Am Fachbereich *Quality & Usability* an der TU Berlin untersuchen wir die Verwendung von Gesten als Alternative zu Passwörtern zur Anmeldung an mobilen Geräten. Im Fragebogen werden Sie zu Ihrer Nutzung von mobilen Geräten wie Handys, Smartphones, iPods usw. befragt.

1 BEFRAGUNG ZUM VERSUCH

1.1 BEWERTUNG DER EINZELNEN GESTEN

1.1.1 GESTE 1

	Vollständige Zustimmung							Vollständige Ablehnung
Die Geste sieht alltagstauglich aus.	++	☐	☐	☐	☐	☐	☐	☐ --
Die Geste sieht unnatürlich aus.	++	☐	☐	☐	☐	☐	☐	☐ --
Die Geste ist lächerlich.	++	☐	☐	☐	☐	☐	☐	☐ --
Die Geste ist komplex.	++	☐	☐	☐	☐	☐	☐	☐ --
Die Geste ist schnell erlernbar.	++	☐	☐	☐	☐	☐	☐	☐ --
Ich kann die Geste genau nachmachen.	++	☐	☐	☐	☐	☐	☐	☐ --

Anmerkungen:

1.1.2 GESTE 2

	Vollständige Zustimmung							Vollständige Ablehnung
Die Geste sieht alltagstauglich aus.	++ ☐	☐	☐	☐	☐	☐	☐	--
Die Geste sieht unnatürlich aus.	++ ☐	☐	☐	☐	☐	☐	☐	--
Die Geste ist lächerlich.	++ ☐	☐	☐	☐	☐	☐	☐	--
Die Geste ist komplex.	++ ☐	☐	☐	☐	☐	☐	☐	--
Die Geste ist schnell erlernbar.	++ ☐	☐	☐	☐	☐	☐	☐	--
Ich kann die Geste genau nachmachen.	++ ☐	☐	☐	☐	☐	☐	☐	--

Anmerkungen:

1.1.3 GESTE 3

	Vollständige Zustimmung							Vollständige Ablehnung
Die Geste sieht alltagstauglich aus.	++ ☐	☐	☐	☐	☐	☐	☐	--
Die Geste sieht unnatürlich aus.	++ ☐	☐	☐	☐	☐	☐	☐	--
Die Geste ist lächerlich.	++ ☐	☐	☐	☐	☐	☐	☐	--
Die Geste ist komplex.	++ ☐	☐	☐	☐	☐	☐	☐	--
Die Geste ist schnell erlernbar.	++ ☐	☐	☐	☐	☐	☐	☐	--
Ich kann die Geste genau nachmachen.	++ ☐	☐	☐	☐	☐	☐	☐	--

1.1.4 GESTE 4

	Vollständige Zustimmung							Vollständige Ablehnung
Die Geste sieht alltagstauglich aus.	++ ☐	☐	☐	☐	☐	☐	☐	--
Die Geste sieht unnatürlich aus.	++ ☐	☐	☐	☐	☐	☐	☐	--
Die Geste ist lächerlich.	++ ☐	☐	☐	☐	☐	☐	☐	--
Die Geste ist komplex.	++ ☐	☐	☐	☐	☐	☐	☐	--
Die Geste ist schnell erlernbar.	++ ☐	☐	☐	☐	☐	☐	☐	--
Ich kann die Geste genau nachmachen.	++ ☐	☐	☐	☐	☐	☐	☐	--

Anmerkungen:

1.1.5 GESTE 5

	Vollständige Zustimmung							Vollständige Ablehnung
Die Geste sieht alltagstauglich aus.	++ ☐	☐	☐	☐	☐	☐	☐	--
Die Geste sieht unnatürlich aus.	++ ☐	☐	☐	☐	☐	☐	☐	--
Die Geste ist lächerlich.	++ ☐	☐	☐	☐	☐	☐	☐	--
Die Geste ist komplex.	++ ☐	☐	☐	☐	☐	☐	☐	--
Die Geste ist schnell erlernbar.	++ ☐	☐	☐	☐	☐	☐	☐	--
Ich kann die Geste genau nachmachen.	++ ☐	☐	☐	☐	☐	☐	☐	--

1.1.6 GESTE 6

	Vollständige Zustimmung							Vollständige Ablehnung
Die Geste sieht alltagstauglich aus.	++ ☐	☐	☐	☐	☐	☐	☐	--
Die Geste sieht unnatürlich aus.	++ ☐	☐	☐	☐	☐	☐	☐	--
Die Geste ist lächerlich.	++ ☐	☐	☐	☐	☐	☐	☐	--
Die Geste ist komplex.	++ ☐	☐	☐	☐	☐	☐	☐	--
Die Geste ist schnell erlernbar.	++ ☐	☐	☐	☐	☐	☐	☐	--
Ich kann die Geste genau nachmachen.	++ ☐	☐	☐	☐	☐	☐	☐	--

Anmerkungen:

1.1.7 GESTE 7

	Vollständige Zustimmung							Vollständige Ablehnung
Die Geste sieht alltagstauglich aus.	++ ☐	☐	☐	☐	☐	☐	☐	--
Die Geste sieht unnatürlich aus.	++ ☐	☐	☐	☐	☐	☐	☐	--
Die Geste ist lächerlich.	++ ☐	☐	☐	☐	☐	☐	☐	--
Die Geste ist komplex.	++ ☐	☐	☐	☐	☐	☐	☐	--
Die Geste ist schnell erlernbar.	++ ☐	☐	☐	☐	☐	☐	☐	--
Ich kann die Geste genau nachmachen.	++ ☐	☐	☐	☐	☐	☐	☐	--

1.1.8 GESTE 8

	Vollständige Zustimmung							Vollständige Ablehnung
Die Geste sieht alltagstauglich aus.	++ ☐	☐	☐	☐	☐	☐	☐	--
Die Geste sieht unnatürlich aus.	++ ☐	☐	☐	☐	☐	☐	☐	--
Die Geste ist lächerlich.	++ ☐	☐	☐	☐	☐	☐	☐	--
Die Geste ist komplex.	++ ☐	☐	☐	☐	☐	☐	☐	--
Die Geste ist schnell erlernbar.	++ ☐	☐	☐	☐	☐	☐	☐	--
Ich kann die Geste genau nachmachen.	++ ☐	☐	☐	☐	☐	☐	☐	--

Anmerkungen:

1.1.9 GESTE 9

	Vollständige Zustimmung							Vollständige Ablehnung
Die Geste sieht alltagstauglich aus.	++ ☐	☐	☐	☐	☐	☐	☐	--
Die Geste sieht unnatürlich aus.	++ ☐	☐	☐	☐	☐	☐	☐	--
Die Geste ist lächerlich.	++ ☐	☐	☐	☐	☐	☐	☐	--
Die Geste ist komplex.	++ ☐	☐	☐	☐	☐	☐	☐	--
Die Geste ist schnell erlernbar.	++ ☐	☐	☐	☐	☐	☐	☐	--
Ich kann die Geste genau nachmachen.	++ ☐	☐	☐	☐	☐	☐	☐	--

1.1.10 GESTE 10

	Vollständige Zustimmung							Vollständige Ablehnung
Die Geste sieht alltagstauglich aus.	++ ☐	☐	☐	☐	☐	☐	☐	--
Die Geste sieht unnatürlich aus.	++ ☐	☐	☐	☐	☐	☐	☐	--
Die Geste ist lächerlich.	++ ☐	☐	☐	☐	☐	☐	☐	--
Die Geste ist komplex.	++ ☐	☐	☐	☐	☐	☐	☐	--
Die Geste ist schnell erlernbar.	++ ☐	☐	☐	☐	☐	☐	☐	--
Ich kann die Geste genau nachmachen.	++ ☐	☐	☐	☐	☐	☐	☐	--

Anmerkungen:

1.1.11 GESTE 11

	Vollständige Zustimmung							Vollständige Ablehnung
Die Geste sieht alltagstauglich aus.	++ ☐	☐	☐	☐	☐	☐	☐	--
Die Geste sieht unnatürlich aus.	++ ☐	☐	☐	☐	☐	☐	☐	--
Die Geste ist lächerlich.	++ ☐	☐	☐	☐	☐	☐	☐	--
Die Geste ist komplex.	++ ☐	☐	☐	☐	☐	☐	☐	--
Die Geste ist schnell erlernbar.	++ ☐	☐	☐	☐	☐	☐	☐	--
Ich kann die Geste genau nachmachen.	++ ☐	☐	☐	☐	☐	☐	☐	--

1.1.12 GESTE 12

	Vollständige Zustimmung							Vollständige Ablehnung
Die Geste sieht alltagstauglich aus.	++	☐	☐	☐	☐	☐	☐	☐ --
Die Geste sieht unnatürlich aus.	++	☐	☐	☐	☐	☐	☐	☐ --
Die Geste ist lächerlich.	++	☐	☐	☐	☐	☐	☐	☐ --
Die Geste ist komplex.	++	☐	☐	☐	☐	☐	☐	☐ --
Die Geste ist schnell erlernbar.	++	☐	☐	☐	☐	☐	☐	☐ --
Ich kann die Geste genau nachmachen.	++	☐	☐	☐	☐	☐	☐	☐ --

Anmerkungen:

1.2 ALLGEMEINES FEEDBACK

Bitte schreiben Sie Ihre Anregungen und Anmerkungen zu dem Versuch auf.