Gesellschaft für Informatik (GI)

publishes this series in order to make available to a broad public recent findings in informatics (i.e. computer science and information systems), to document conferences that are organized in cooperation with GI and to publish the annual GI Award dissertation.

Broken down into the fields of
• Seminars
• Proceedings
• Dissertations
• Thematics
current topics are dealt with from the vantage point of research and development, teaching and further training in theory and practice. The Editorial Committee uses an intensive review process in order to ensure high quality contributions.

The volumes are published in German or English.

Information: http://www.gi-ev.de/service/publikationen/lni/

ISSN 1617-5468
ISBN 978-3-88579-258-1

The proceedings of the BIOSIG 2010 include scientific contributions of the annual conference of the Special Interest Group on Biometrics and Electronic Signatures (BIOSIG) of the Gesellschaft für Informatik (GI).The conference took place in Darmstadt, 09.-10. September 2010. Within two days mainly the advances of biometrics research and new developments in the biometric application fields beyond security applications have been presented and discussed by biometrics and security professionals.

Arslan Brömme, Christoph Busch (eds.): BIOSIG 2010: Biometrics and Electronic Signatures

164

GI-Edition

Lecture Notes in Informatics

Arslan Brömme, Christoph Busch (Eds.)

BIOSIG 2010: Biometrics and Electronic Signatures

Proceedings of the Special Interest Group on Biometrics and Electronic Signatures

09.–10. September 2010, Darmstadt, Germany

Proceedings

Arslan Brömme, Christoph Busch (Eds.)

BIOSIG 2010

Proceedings of the Special Interest Group on Biometrics and Electronic Signatures

09.-10. September 2010 in Darmstadt, Germany

Gesellschaft für Informatik e.V. (GI)

Lecture Notes in Informatics (LNI) - Proceedings
Series of the Gesellschaft für Informatik (GI)

Volume P-164

ISBN 978-3-88579-258-1
ISSN 1617-5468

Volume Editors
Arslan Brömme
GI BIOSIG, Gesellschaft für Informatik e.V.
Ahrstraße 45, 53175 Bonn
Email: arslan.broemme@aviomatik.de
Christoph Busch
Hochschule Darmstadt
CASED
Haardtring 100, D-64295 Darmstadt

printed by Köllen Druck+Verlag GmbH, Bonn

Chairs' Message

Welcome to the annual international conference of the Special Interest Group on Biometrics & Electronic Signatures (BIOSIG) of the Gesellschaft für Informatik (GI) e.V.

GI BIOSIG was founded as an experts' group for the topics of biometric person identification/authentication and electronic signatures and its applications. Over the last nine years the annual conference in strong partnership with the Competence Center for Applied Security Technology (CAST) established a well known forum for biometrics and security professionals from industry, science, and representatives of the national governmental bodies working in these areas.

The BIOSIG 2010 international conference is jointly organized by GI BIOSIG, the Competence Center for Applied Security Technology (CAST) e.V., the German Federal Office for Information Security (BSI), the European Commission Joint Research Centre (JRC), the Biometric European Stakeholders Network (BEST Network), the Center for Advanced Security Research Darmstadt (CASED), and the TeleTrusT Deutschland e.V.

The international program committee accepted full scientific papers strongly along the LNI guidelines within a scientific review process of about five reviews per paper in average.

Furthermore, the program committee has created a very interesting program including selected contributions of strong interest (invited and further conference contributions) for the outlined scope of this conference. We would like to thank all authors for the contributions and the numerous reviewers for their work in the program committee.

Darmstadt, 09th September 2010

Arslan Brömme
GI BIOSIG, GI e.V.

Christoph Busch
Hochschule Darmstadt

Chairs

Arslan Brömme
GI BIOSIG, GI e.V., Bonn, Germany

Christoph Busch
Hochschule Darmstadt, Germany

Program Committee

Harald Baier, Björn Brecht, Arslan Brömme, Patrick Bours, Julien Bringer, Christoph Busch, Victor-Philipp Busch, Henning Daum, Nicolas Delvaux, Farzin Deravi, Bernadette Dorizzi, Martin Drahansky, Julian Fierrez, Simone Fischer-Hübner, Davrondzhon Gafurov, Rüdiger Grimm, Patrick Grother, Olaf Henniger, Detlef Hühnlein, Heinrich Ihmor, Klaus Keus, Ulrike Korte, Bernd Kowalski, Michael Kreutzer, Herbert Leitold, Stan Li, Luigi Lo Iacono, Jan Löscher, Udo Mahlmeister, Davide Maltoni, Tony Mansfield, Gisela Meister, Johannes Merkle, Emilio Mordini, Alexander Nouak, Hisao Ogata, Michael Peirce, Ioannis Pitas, Fernando Podio, Reinhard Posch, Kai Rannenberg, Marek Rejman-Greene, Raul Sanchez-Reillo, Heiko Roßnagel, Günter Schumacher, Takashi Shinzaki, Max Snijder, Elham Tabassi, Till Teichmann, Cathy Tilton, Massimo Tistarelli, Dimitrios Tzovaras, Markus Ullmann, Michiel van-der-Veen, Raymond Veldhuis, Jim Wayman, Andreas Wolf, Bian Yang, Xuebing Zhou

Hosts

Special Interest Group on Biometrics and Electronic Signatures (**BIOSIG**)
of the Gesellschaft für Informatik (GI) e.V.
http://www.biosig.org

Competence Center for Applied Security Technology (**CAST**) e.V.
http://www.cast-forum.de

Bundesamt für Sicherheit in der Informationstechnik (**BSI**)
http://www.bsi.bund.de

Biometric European Stakeholders Network (**BEST Network**)
http://www.best-nw.eu/

European Commission Joint Research Centre (**JRC**)
http://ec.europa.eu/dgs/jrc/index.cfm

Center for Advanced Security Research Darmstadt (**CASED**)
http://www.cased.de/

TeleTrusT Deutschland e.V.
http://www.teletrust.de/

BIOSIG 2010 – Biometrics and Electronic Signatures

"**Biometrics and Electronic Signatures – Research, Applications, and Beyond**"
09th -10th September 2010

Biometrics research in high quality imaging of fingerprint minutiae, improvement for fingerprint image quality evaluation, and the usage of biometric fingerprint information in cryptographic keys are still hot topics for intensive research in this area. Three research contributions are addressing this year those aspects.

But what if the basic research on biometric fingerprint information is sufficient for governmental usage one day?

If one assumes that biometric data in sufficient quality is available for various applications, the interest in those high quality, unprotected biometric data for third party usage is very significant. So, the protection of biometric templates due to privacy reasons and to avoid unauthorized third party usage is of high interest for current research. Biometric template protection is addressed this year by two research papers.

What have the users of biometric systems experienced?

Availability of sufficient reliable biometric technology leads to core project management challenges by introducing such systems to a broader public. These challenges are addressed this year in a report from a user's perspective for national ID schemes.

What is beyond security applications?"

Legal aspects of new application scenarios regarding the prevention from criminal or in other ways dangerous behaviour are considered in another research contribution.

Finally, an information system will know who is using it. And this enables a broad spectrum of applications beyond security. One should think and discuss about it.

Several further contributions are focussing on approaches for fingerprint recognition, biometric authentication services, best practices for enrolment, activity-related biometrics, biometric verification based on 3D face and vein patterns, aspects of phone security for mobile biometrics, and an insight into the safe and reliable use of biometrics.

BIOSIG 2010 offers you again a platform for experts' discussions on biometric research, security applications, and beyond.

Table of Contents

BIOSIG 2010

Regular Research Papers

Facing the Challenge of Enrolment in National ID Schemes

Ali M. Al-Khouri

Emirates Identity Authority,
Abu Dhabi, UAE.
ali.alkhouri@emiratesid.ae

Abstract: This article presents the approach followed in the United Arab Emirates (UAE) national ID scheme to register its population for the new smart ID card it launched in 2005. It presents how the organisation reengineered its operations to achieve its strategic objectives. It also presents some of the experienced challenges, and how they were dealt with. Some key management consideration areas were also listed for the purpose of sharing knowledge and experience in the field.

Keywords: *National ID, ID Card, population enrolment, process reengineering.*

1 Introduction

Governments around the world have been very much attracted to National ID programs. These programs are globally justified on the basis of building an identity management system to achieve primarily two objectives: support national security and improve access to services [Ak07]. More than 30 countries have initiated smart ID card programs in the last decade with a total value of those projects exceeding $24 billion. Besides, more than 15 countries are in the process of upgrading their current ID cards to biometric based systems.

GCC countries have been among the first countries to launch biometric based smart ID card initiatives. Due to nature and complexity of such schemes, these initiatives have been challenged to meet its specified projects scope, timelines, and budgets. Table 1 below shows the progress of smart ID card schemes in GCC countries and the percentage of population registered so far.

Our observations of national ID card projects show that many countries are struggling with the enrolment of population in their ID schemes. Apart from the technical complexity of such projects, the most significant challenge lies in the fact that these programs include biometric acquisition which entails the presence of individuals. Some countries capture only two fingerprints, others capture a full set of fingerprints including palm prints and writers, while others use a variety of biometric identification systems such facial, iris, and fingerprints.

Table 1: ID card projects in GCC countries [GCC09]

Country	Program Start Year	Total Population	Registered population	% to total population	Biometrics
Saudi Arabia	2004	28,686,633	1.2 million	4.2%	2x Flat prints
UAE	2005	8,200,000	1.8 million	22.0%	Rolled 10 prints, palm & writer prints
Kuwait	2009	2,691,158	200,000	7.4%	Rolled ten prints
Bahrain	2005	1,039,297	800,000	77.0%	2x Flat prints
Qatar	2007	833,285	100,000	12.0%	2x Flat prints and Iris
Oman	2004	3,418,085	3 million *	90%**	2x Flat prints

* biometric capture is not mandatory for females

** not all registered have biometrics in the database.

The practice of biometric acquisition was previously limited to forensic and traditional law enforcement applications. For obvious reasons, developments of systems like fingerprints compared to other biometric systems and hence the maturity of the overall technology, did not take into consideration higher levels of customer or service satisfaction since the intended users were in forensic and police jurisdictions. Therefore, and based on the biometrics and verification procedures, the registration process can be time consuming and inconvenient. A well thought through enrolment plan that captures an understanding of population demographics and cultural elements, and follows a modular approach of gradual registration based on geographical distribution and other segmentation factors, is likely to yield more successful results.

This article presents a case study of the process followed to develop an enrolment plan to register the population of the United Arab Emirates. It touches upon a broader organisational scope, and presents essential lessons learned and important building blocks for government officials working in this field. Though the project size and targeted population is considered relatively small in comparison to other countries, the presented processes and overall thoughts are believed to contribute and advance existing knowledge.

2 Emirates Identity Authority

Emirates Identity Authority was established in 2004, as a federal government authority tasked to build an identity management system, by enrolling and issuing ID cards to more than 5 million people at the time. The organisation relied primarily on a social marketing strategy to enroll the population and its copious developed strategies only succeeded to enroll less than 20% of the total population over a 5 year period. This represented a challenge to overcome and a difficulty to justify the heavy budget expenses and no clear return on investment (ROI) upshots. Altogether, this forced the organisation to go through muscular change process to address this problem area.

A four-staged change process was developed to guide the change implementation, as depicted in Figure 1 below. The change process was instigated to enact an organisational mindset change with the aim of developing a service driven and result oriented organisation. It also aimed to increase accountability, improve efficiency, overall performance and high quality services.

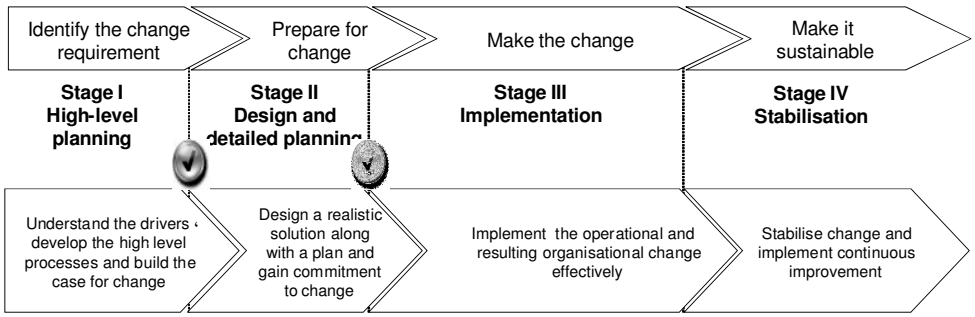

Figure 1: Change management program components

The initial phase of the change process dealt with the identification of the change requirements and building the overall case for change. The second phase was more of a planning phase, and included detailed assessment of the impact of change to the overall organisation. The third phase was about implementing the change according to the plan, and the fourth was more of an improvement and sustainability stage.

The outcome of the first phase was the development of an operating model that captured the fundamental and evolving functions of the organisation. It provided the foundation and flexibility required to execute the organisation's initiatives. As depicted in Figure 2, the primary function that needed to be addressed at first was population enrolment. As the organisation progresses, the function of enrolment will shrink down to become less than 20% of the overall operation. The organisation's role will turn gradually into a service delivery function related to authentication and identification. This model is considered to be a valuable knowledge to existing literature in the field, as it is generic and applicable to all ID card programs.

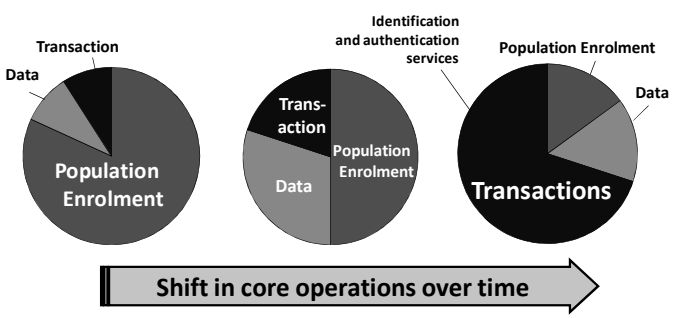

Figure 2: The operating model

Another outcome of the first phase was the development of the core pillars of strategic directions that would determine the success of the overall program. They were later used in the development process of the corporate strategy and the design of the consequent initiatives. These pillars included:

- **Effective population enrolment strategy:** develop strategies to increase population enrolment, that incorporates marketing, outreach, program, and staff development efforts to increase enrolment in an effective manner.
- **Integration/Interface with key government organisations:** keeping the population register database timely updated, is essential to the overall success of the program. Connecting to the databases of "data owners" is therefore inevitable. Six government entities were identified: (1) Ministry of Interior: immigration; (2) Health Ministry: birth and death; (3) Labour Ministry (4) Justice: marriage and divorce; (5) Education, and (6) Higher Education.
- **Supporting e-Government:** to develop secure and robust infrastructure to support Governmental electronic services, in relation to the validation and authentication of online identities in electronic transactions.
- **Customer Focus:** to become a customer focused organisation, and complement enrolment strategy through renewed attention to the customers' interface with the organisation.

3 Registration: The Status Quo!

The existing process implied that the applicants needed to fill an application form at a typing centre or on the internet. They then may choose to take an appointment by the available online system, or go directly to the registration centres. The actual registration time varied from 15 minutes to 20 minutes, but waiting queues lasted from at least 4 hours to 8 hours before they get registered. Reasons for such deficiency included factors related to lack of flow management procedures at registration centres, unstudied media campaigns that attracted higher population to registration centres than their actual capacities, untrained staff, etc. The overall process caused public frustration and media criticism.

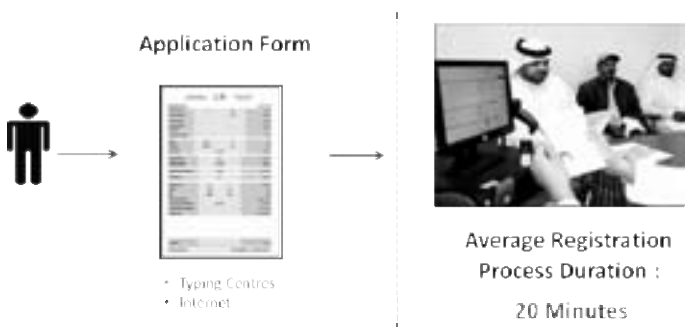

Figure 3: Registration process prior to re-engineering

Some of the *quick fixes* adopted by the organisation were to cancel the presence of the children to the registration centres, and rely on the supporting documents presented by the parents. The information was verified with the Ministry of Interior's database for validation. The second process change was related to how registration equipment were organised. The registration process at first, required applicants to go through three enrolment stations:

(1) verification of documents and fee collection,

(2) portrait and signature capture and scan in documents, and

(3) fingerprinting. This process provided a smooth management of applicants flow.

For reasons related to lack of resources, management at the time decided previously to merge some functions together, i.e., second and third functions and as depicted in Table 2. This poorly studied change resulted in longer and process "*locked in"* applicants. As the first process took normally 3 minutes to complete, the new combined process of taking portrait, electronic signature, scan in document, and fingerprinting, took almost 15 to 20 minutes, that created long waiting queues inside these offices. This also led to more data entry errors by operators.

The introduced change here included changing the process to keep fingerprinting as a separate function, and merge all others in a separate workstation. This allowed a better flow management as illustrated in Table 2 below.

Table 2: Example of tactical process changes

+portrait /verification fingerprinting		portrait +verification \fingerprinting		Criteria
2		2		Steps
fingerprinting	Verification and portrait	Portrait and fingerprinting	verification	Time
7	8	12	3	
2		2		Space
v		v		Privacy
equal		equal		Accuracy
equal		equal		Machinery Cost
equal		equal		Human Resources
organised		Not organised		Customer Flow
fingerprinting	Verification and portrait	Portrait and fingerprint	verification	Work Load
Equal	Equal		incomplete	
Best Choice		Secondary choice		Decision

These two change tactics provided temporary fixes, and supported better management of flows at registration centres. The next section will shed light on the developed enrolment strategy.

4 The Need for an Enrolment Strategy

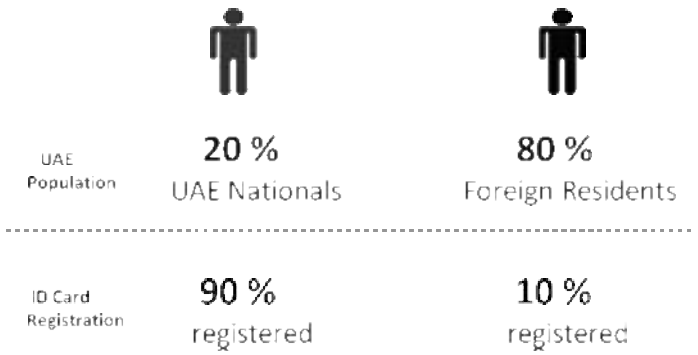

Figure 4: Registration vs. total population

According to the original enrolment strategy, it was envisaged that a total of 5 million people will be registered by the end of 2010. However, and towards the end of 2009, only 20% of this number were registered.

A study conducted to evaluate and forecast enrolment, showed that it would take Emirates ID more than 10 years to register the population with existing enrolment rates. As depicted in Table 3, the organisation needed to have a capacity of 20,000 enrolment per day (new and renewal) in order to achieve its objectives in the shortest and practical timeframe.

Table 3: Challenge of enrolment

Equipment and outsourcing	outsourcing	More equipment	available	Existing	
20,000	12,000	8,000	4,500	3,200	Daily registration capacity
4,800,000	2,880,000	1,920,000	1,080,000	768,000	Cards per year
1.6	**2.7**	**4.1**	**7.2**	**10.2**	**Time needed to register 7.8 million people**
12.1	11.5	9.6	8.7	0	Time savings (years)

Another factor that forced the development of an enrolment strategy was the increasing financial cost to the organisation and reprehensible revenues. The cost of the card have gone up more than 30% higher than the fees paid by the applicants, as the cost of the card is dependent on specific annual registration.

This meant that for each card, the organisation issued, it lost around 250 dirhams. In fact, the organisation needed to produce 1.6 million cards a year to make the breakeven point, and as depicted in Figure 5. All together, these factors forced the organisation to rethink its value proposition, and rework the overall enrolment strategy, which is discussed next.

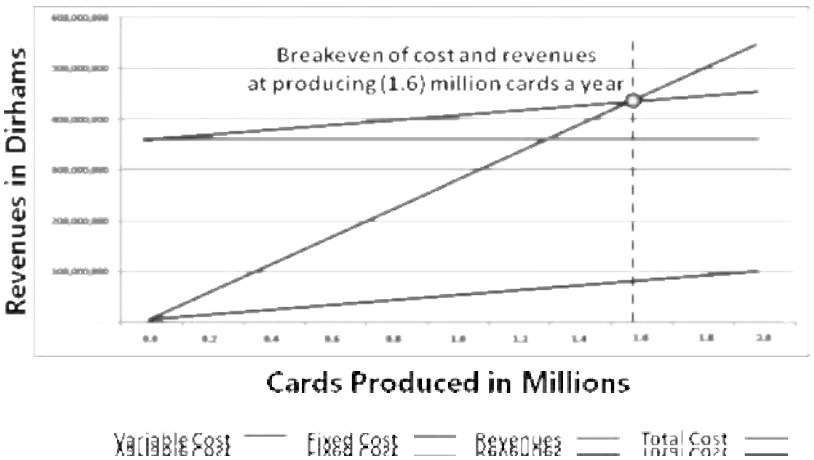

Figure 5: Cost and revenues

5 The New Enrolment Strategy

As indicated earlier, the previous enrolment strategy adopted in the organisation was a marketing based. The fundamental thinking that guided the development of the new enrolment strategy was to follow a process driven approach. The principles of this approach were based on the relationships between business processes that would promote public participation. The new strategy consisted of three main focus themes:

5.1 The New Process: Reengineering of the Enrolment Process

The new process divides the registration into three segregated functions. More than 3000 typing centres in the government were equipped with a new application form allowed them to key in personal information, scan in documents, scan in photos for those below 15's, and accept payments, and automatically generate appointments to applicants. All these functions apart from the application form were previously done at registration centres. Registration centres new role was limited to do portrait and biometric acquisition only.

This implied an 8-10 minute process compared to 20 to 25 minute previously. Applicants' data is then transferred electronically to the internal audit office (back office) which verifies the complete dossier against the Ministry of Interior's database, and authorises or rejects applications. The new process made more than half of the previous procedures invisible to the applicants, as they were shifted to either back end or typing centres.

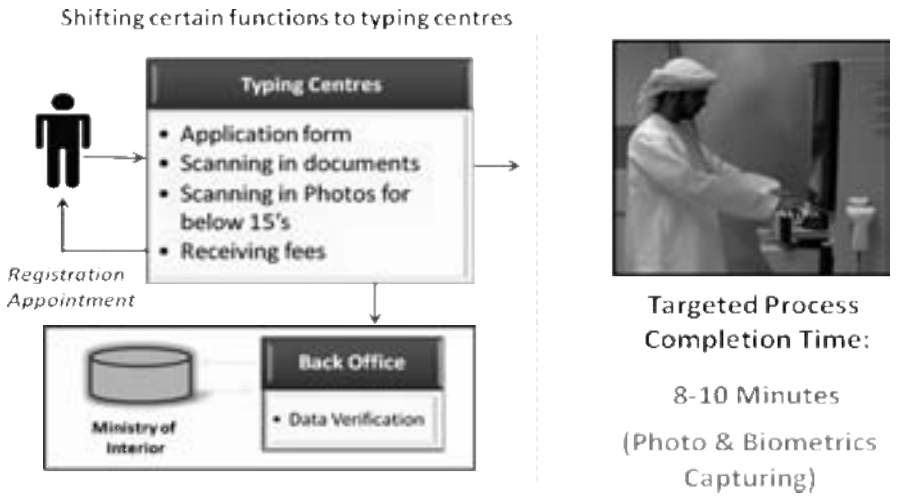

Figure 6: New registration process

The new process also had a great impact on the existing registration centres layout. As depicted in Figure 7, the new process was considered as a *one stop shopping office,* and allowed higher capacity in terms of enrolment rates, and space utilisation.

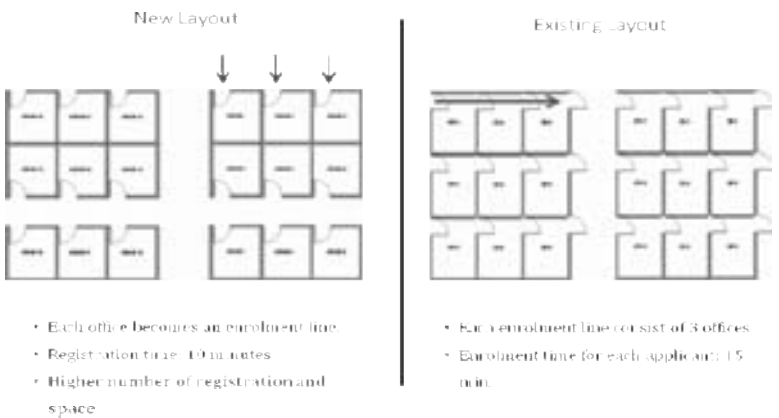

Figure 7: New process impact on registration centres layout

5.2 Linking Registration with Immigration Procedures

The second focus theme of the enrolment strategy was to link the ID registration with the issuance and renewal of residency permits. Taking into consideration that the maximum validity of residency permits is 3 years, then it was assumed that all residents will be enrolled in this timeframe given that all registration sites are operational.

In order to make the process more convenient to the applicants, new registration centres were envisaged to be built near existing preventative medicine centres; responsible to issue medical fitness certificates to complete the residency procedures. According to statistics, there were around 9,000 to 15,000 daily transactions of new and renewal of residencies in the UAE. This process merge between ID card and residency permit, was envisaged to enforce and increase the daily registration rate remarkably.

It was also noted in this focus area, that the residents during their application for issuance and renewal of residency permits fill different application forms for different entities, e.g., immigration form, labour form, and ID card form. Comparing the three forms, it was fond that they were almost identical. It was then decided to merge the three forms to be a unified form for the three entities, besides the preventative medicine which also issues separate forms. This step would contribute to prevent double implementation of such procedures and promote data accuracy. The new 3+1 form will also include the feature of central fees collection for all four entities, payable at typing centres. The fees will be automatically transferred to the beneficiary authorities through an electronic clearance system.

5.2.1 Registration Process

The registration process starts with the applicant or a representative visiting the typing centre to fill the unified application form. The form will also include the new functionalities described in section (5.1). Applicants aged 15 and above will go to the preventative medicine centre for medical check up and go through the ID card registration office for portrait and biometric acquisition. Upon the acceptance of the issuance/renewal of residency permit, the immigration database at the Ministry of Interior, electronically notifies the ID card database, which will trigger card printing request, and dispatch it to the applicant through a registered courier. For the purpose of unification, ID card validity is linked with the residency permit. It is envisaged that once the process is streamlined, and reached to a satisfactory level, the residency sticker and labour card will be replaced with the ID card, as a single identity document for residents.

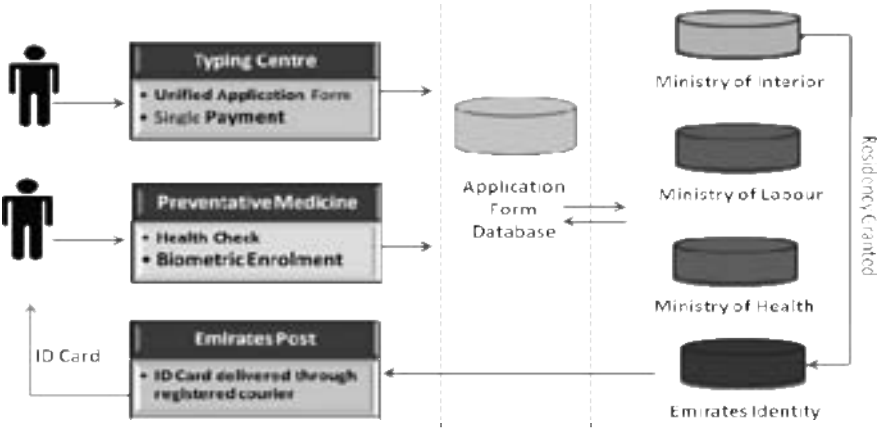

Figure 8: Process of registration merged with immigration processes

5.3 Labour Registration

The third focus theme is the registration of labour population through mobile registration devices at labour campus or their workplaces. This will relax the traffic at existing registration centres. Existing statistics refer that the UAE has around two million unskilled labour population. The registration of this category was planned with the Ministry of Labour to ensure prompt registration and enforcement through their employing companies. Statistics also show that large number of labour camps have been developed in the past five years, with average residents in those camps ranging from 5,000 to 50,000 people.

Having presented the components of the enrolment strategy, the next two sections will briefly discuss the three remaining pillars of the strategic directions presented in section 2.

6. Integration with Key Organisations

One of the most strategic objectives of building an identity management system in the UAE was to make a central identity reference repository for the UAE government about population demographics, timely available census and statistical surveys. This database was also foreseen to provide decision makers with key data to enable informed planning decisions. Maintaining an up to date and accurate population database is considered an impossible objective without a centralised e-information infrastructure to bring different databases together into one centralised repository. An initiative was developed called citizen data-hub that aimed to connect six key government databases together that were considered the *"primary data owners*".

The secondary objective of this initiative was to establish dynamic and real-time links between administrative government departments across the country, thus enabling information sharing that ultimately contributes to the better administration of the country and provision of service delivery. See also Figure 9.

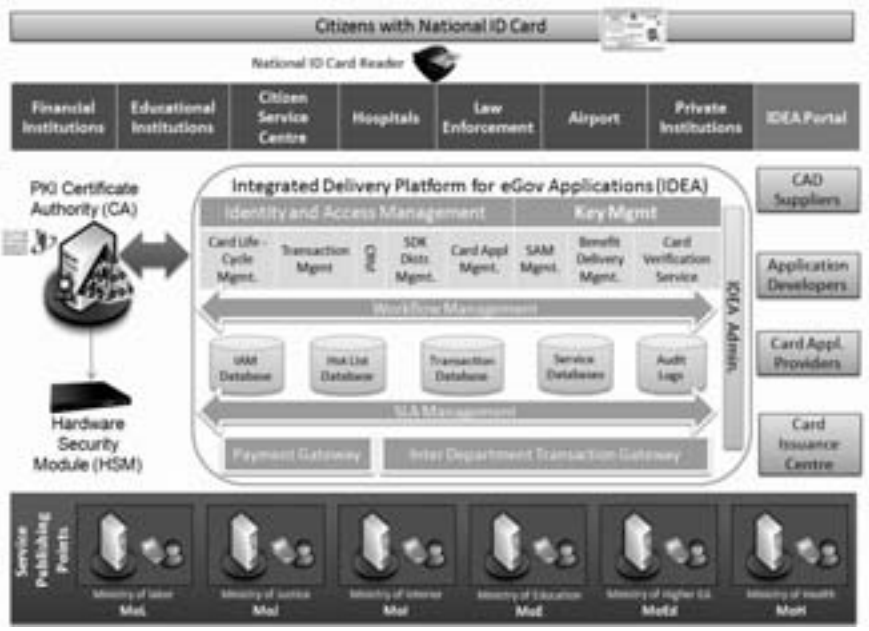

Figure 9: ID Card solutions architecture

7. Supporting e-Government

Development of a national population infrastructure should consist of enabling the basis for online authentication of users. It should address the overall requirements of trust, identity management and privacy and in the context of electronic governance. The federated identity management initiative was designed to facilitate implementation of e-Government services within the United Arab Emirates. This is envisaged to support advanced development of e-government specifically in areas related to e-inclusion and e-participation, as well as the end-to-end integrated government work processes.

8. Customer Service Orientation

Given the challenges the UAE ID card program is facing, it is confronted with key building blocks represented accelerating enrolment rates, meeting stakeholders expectations, improving quality of service, etc. The new organisation thinking as explained above shifted more towards a customer driven business organisation.

The aim was to positively embrace a customer focused culture, where core competencies are identified and developed to deliver value for customers. A customer service standard was developed based on guidelines specified in the International customer service institute [CSI10]. This focus area described a management culture that emphasised centrality of the citizen or customer in the process, as well as accountability for results.

This section concluded the change management program and the enrolment strategy overview developed at Emirates ID. The next section presents some key management consideration areas that require management attention.

9. Management Considerations

9.1 Change Management and Communication Plan

Change management as a discipline has grown tremendously over the last few years in the Gulf region. Our close interactions with government organisations in the region show us that a large number of public sector organisations used consultancy firms to develop and implement structured approach to managing change programs.

Indeed, a carefully planned change management program is imperative to the overall success of any strategic endeavour. Figure-10 shows that the success of a change program is determined by the awareness of the involved individuals or groups of the need and objectives of the change program. A change program is likely to be associated with vagueness, rumors, distrust even among those involved in the change process. Strong and consist leadership is needed to draw a clear path and set out performance and expectations of outcomes.

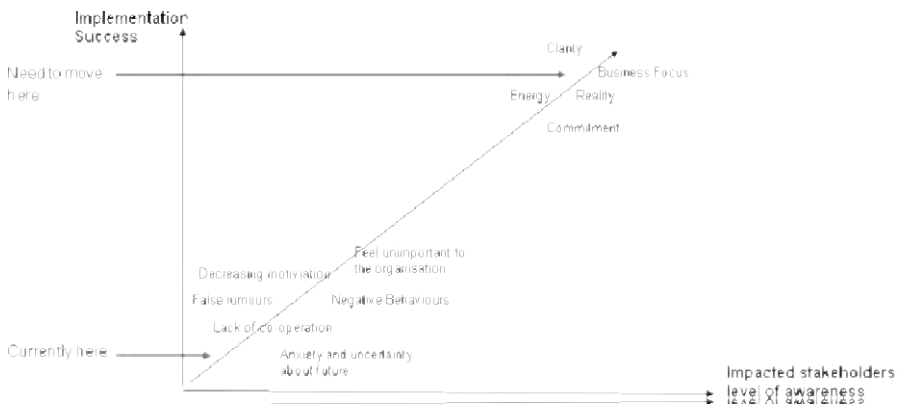

Figure 10: The need for clear communication plan

Management would prefer to implement change and expect least resistance and with the most buy-in as possible. For this to occur, change must be applied with a structured approach so that transition from one type of behaviour to another organisation wide will be smooth. Management need to carefully assess employees' reaction to an implemented change and attempt to understand the reaction to it. Although change programs are implemented to achieve organisational goals and objectives, certain changes do sometimes produce tremendous amount of resistance at several operational and management levels. Management is expected to provide support throughout the process of these changes, which are at times very difficult. Managing changes especially in public sector organisations requires a broad set of skills like political, analytical, communication, people, system, and business skills.

9.2 Organisational Development Principles

Due to the enormous pressure on management to create value and bring out tangible results, it is easily found that we get distracted with day to day operations. A commendable framework management need to always keep in mind is the EFQM model (see also Figure 11). The model was found to sustain a management focus on key governance perspectives. It is a good management assessment tool to measure the strengths and improvement areas of an organisation across all business operations, and to define the organisation's capability and performance.

Figure 11: EFQM excellence framework

The three main elements that were considered crucial to the success of the overall organisation strategy were: (1) to become a result driven organisation, and focus on (2) employing and developing highly qualified and trained staff, who should enable and promote (3) creativity, innovation and learning organisation culture. The framework supported management to rethink values, policies, and controls and a restructuring that reflected a renewed sense of mission.

9.3 Management Dash Board

It is important for management to develop a dash board that gives an overview of the strategy and projects status. The use of simple graphical charts and maps, make it easier for management to understand and interpret business information, rather than wading through masses of numbers and spreadsheets. The management dash board need to be real-time reporting, to support executives and managers take actions at the first sign of a problem, instead of waiting for monthly or quarterly meetings or reports.

The management dashboard need to some degree to include drill-down capabilities, to reveal more associated graphs and breakdowns. Developing an electronic KPI dashboard as an active organisational messaging platform, should increase the visibility of key performance indicators for informed decisions that should in turn improve overall performance.

9.4 Users training

User training is a critical success factor for. The routine nature of work at registration centres caused a shortage of workers with the necessary skills to cope with the rapid growth and expansion of centres. This shortage forced the organisation to continuously hire and train new employees who lack adequate technology skills, and to accept the chore of constantly retraining present employees.

In ID card schemes, fingerprint quality has huge impact on the identification/verification system. Therefore, and to meet these challenges, organisations need to develop a system to manage end-user training, and focus to enhance fingerprint capture quality.

9.5 Media and Marketing Strategy

National ID schemes have been a very much subject to controversial debate on international levels [Ag01][Ld06]. It is seen by privacy advocates to be a *'massive invasion*' of their liberty and freedom rights, and promotes the concept of setting up *'big brother*' or '*big government'*. It was therefore important for the organisation to develop a social media marketing strategy to better understand community interests by running customer and market surveys within the social communities, and promote engagement and social participation into the project value proposition.

The second component of the media strategy was related to building visibility about the program through information sharing and interactions. The communication strategy included specific aspects that considered the cultural diversity of the target society *(eg. multiple language communication, information leaflets etc).*

10. Conclusion

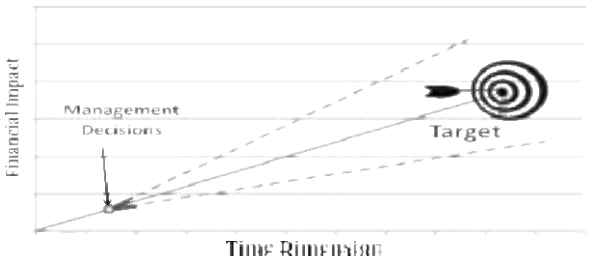

Figure 12: Impact of unfocused management decisions

Without a clear blueprint and plan, organisations are more likely to drift and run in different directions. Management critical decisions that are not based on solid understanding of impact and well-deliberated calculations will most probably yield to an unknown outcome - see also [FL93][PBJ93]. Public sector projects are to a great degree involve risk and uncertainty.

This article was written in an attempt to reveal some of the challenges experienced in the implementation of a strategic and large scale government program. National ID schemes and due to their size and complexity need scrupulous planning to achieve their audacious goals. Population enrolment in such schemes is considered a challenging chore. The presented case study expounded how the UAE government reacted to this challenge.

Though it could be argued that population size in the UAE is lower than many other larger initiatives in other countries, the presented approach in this article is believed to provide a virtuous thinking path to address similar issues. Besides, the presented management consideration areas are assumed to be important knowledge building blocks for those in the field to address fundamental organisational and project management rudiments.

References

[Ag01] **Agar,** J., 2001, "Modern Horrors: British Identity and Identity Cards," appears in Documenting Individual Identity: The Development of State Practices in Modern World, J. Caplan & J. Torpey (eds.), Princeton & Oxford, pp. 101-20.

[Ak07] Al-Khouri, A.M., 2007, "UAE National ID Programme Case Study," *International Journal of Social Sciences*, vol. 1, no.2, pp.62-69.

[BCZ07] Barton, B., Carlton, D. and Ziehm, O., 2007, "Identity management in the 21st century: Balancing safety, security and liberty in a global environment." [Online]. IBM Global Services. USA. Available: http://www-

935.ibm.com/services/us/gbs/bus/pdf/g510-7859-00-id-mgt.pdf [Accessed 14 April 2010].

[CSI10] The International Customer Service Institute [Online] http://www.ticsi.org/. Accessed 05 January 2010.

[FL93] French, S. and Liang, Y., 1993, "Decision support systems: a decision analytic perspective," appears in Developments in Operational Research, Norman, J. (ed.), Operational Research Society, Birmingham.

[GCC09] 6th GCC ID Cards Interoperability Committee Meeting, Riyadh, Saudi Arabia, November, 2009

[Gr08] Greenleaf, G., 2008, "Hong Kongs Smart ID Card: Designed to be Out of Control," appears in *Playing the Identity Card*, Bennett, C. and Lyon D. (eds.), Routledge.

[Ld06] Lyon, D., 2006, "The border is everywhere: ID cards, surveillance and the other," appears *in Global Surveillance and Policing: Borders, Security, Identity,* Cullompton: Willan Publishing, pp. 66-82.

[NW02] Neumann, P. G. & Weinstein, L., 2001, "Risks of National Identity Cards," *Communications of the ACM* 44, p. 176.

[PBJ93] Payne, J.W., Bettman, J.R. and Johnson, E.J., 1993, The adaptive decision maker." Cambridge University Press.

About the Author

Dr. Al-Khouri is a senior government official working in the United Arab Emirates. He is currently working with Emirates Identity Authority as a Director General. He has been involved in the UAE ID Card project since its early conception phases. He obtained his doctorate from Warwick University in the UK in field of project management of large scale initiatives in the public sector. He has been an active researcher in the field of strategic management in public sector, and has published many articles in the last few years.

Towards Improving the NIST Fingerprint Image Quality (NFIQ) Algorithm

Johannes Merkle, Michael Schwaiger, Marco Breitenstein

secunet Security Networks AG,
D-45128 Essen,
Germany

Oliver Bausinger, Kristina Elwart, Markus Nuppeney

Bundesamt für Sicherheit in der Informationstechnik,
D-53175 Bonn,
Germany

Abstract: The NIST Fingerprint Image Quality (NFIQ) algorithm has become a standard method to assess fingerprint image quality. However, in many applications a more accurate and reliable assessment is desirable. In this publication, we report on our efforts to optimize the NFIQ algorithm by a re-training of the underlying neural network based on a large fingerprint image database. Although we only achieved a marginal improvement, our work has revealed several areas for potential optimization.

1 Introduction

Assessing the quality of fingerprint images is crucial for many biometric applications. In particular, in the enrollment process it must be ensured that the provided fingerprint images are of sufficiently high quality. In the context of biometric applications, the quality is defined by the expected recognition accuracy (in terms of low error rates) in fingerprint verification with respect to this image. Consequently, the quality of a fingerprint image depends on the feature extraction and template comparison algorithms used. A lot of fingerprint verification software comes with a specific quality assessment algorithm for fingerprint images, which aims at predicting the recognition accuracy of the images when using the feature extractor and template comparison algorithm included in the software package. However, in applications where the fingerprint verification software of various manufacturers can be deployed, such a vendor-specific quality assessment is of limited use and a vendor-independent quality measure is needed.

In August 2004, NIST issued the NIST Fingerprint Image Quality (NFIQ) algorithm as part of the NIST Biometric Image Software (NBIS) [WGT+06]. The NFIQ algorithm is an open source tool for measuring the quality of fingerprint images independent of the fingerprint verification software used. The NFIQ measures quality by 5 classes, where class one refers to "excellent" and class five to "poor", and the "NFIQ value" output by NFIQ algorithm refers to the class number of the input fingerprint. As explained in [TWW04], the NFIQ algorithm is based on an artificial neural network that tries to predict the quality class from 11 features of the image. These features include the numbers of minutiae and image blocks with quality index exceeding several thresholds. The neural network has been trained with a large number of fingerprint images and the corresponding comparison statistics obtained with different fingerprint verification software.

While the present NFIQ algorithm is extremely useful, there still is demand for optimization. First, the division in five quality classes is quite rough: empirical tests reveal that for typical fingerprints more than 45% have NFIQ class 1. Certain applications, e.g. template protection techniques, may require particularly high quality images that can not be recognized by the NFIQ. Second, practical experience shows that the accuracy of the assessment is limited. On the other hand, an evaluation of the methodology followed for the development of the NFIQ algorithm [TWW04] indicates that there is still room for improvements: in particular, 40% of the fingerprint images used for the training of the neural network were not obtained from live scans but from inked impressions; clearly this type of fingerprint image is not relevant for authentication systems based on electronic capture devices. Furthermore, the training set consisted only of two imprints per finger, which rendered the estimation of the genuine matching performance difficult and inherently imprecise.

In a first attempt to improve the NFIQ algorithm, we conducted a complete re-training of the neural network based on a large database of live scan fingerprint images. Although the main focus was to improve the significance of the quality assessment for recognition performance by using a better data basis, we detected and implemented other optimizations.

This document is structured as follows. In Section 2 we outline the approach followed by NIST to implement the original NFIQ algorithm. In Section 3 we describe our efforts to re-train the neural network to obtain an improved NFIQ+ algorithm. In Section 4 we report the evaluation results of the re-trained NFIQ+ algorithm and its comparison to the original NFIQ. Eventually, in Section 5 we summarize the lessons learned and point out the potential for future optimizations.

2 Training of the Original NFIQ Algorithm

According to [TWW04], the training of the original NFIQ algorithm was based on the similarity score statistics of 3900 fingerprints with 3 fingerprint software development kits (SDK) from Cogent, Sagem and NIST. Of these 3900 fingerprints in the training set, approximately 40% were scanned inked imprints, comprising both rolled and plain impressions. Furthermore, the training set contained two imprints per finger.

Subsequently, we use the following notations: we denote the fingerprints in the training set by $x_{f,i}$, where f enumerates the fingers (we do not distinguish persons but only the fingers) and i enumerates the imprints of this finger. For fixed SDK, we denote the similarity scores output by the comparison operation of $x_{f,i}$ and $x_{f',i'}$ by $S(x_{f,i}, x_{f',i'})$.

The approach of NIST was to determine a quality measure (the NFIQ) of the fingerprints by statistical evaluation of their similarity scores, and then to use these (statistically determined) NFIQ values for a training of a neural network that aims at predicting the NFIQ from certain image features representing relevant characteristics of the fingerprint. The overall process can be summarized as follows:

- For each pair of fingerprints in the training set, a comparison operation was conducted with each SDK; for each comparison operation the SDK outputs a similarity score indicating the level of similarity between the compared fingerprints. The matrix of the similarity scores resulting from all comparison operations with a specific SDK is called *similarity matrix*.

- For fixed SDK, the quality of each fingerprint was measured by (bins of) a function called *normalized match score* that depends on all its similarity scores. The normalized match score was defined as

$$o(x_{f,i}) := \frac{s_m(x_{f,i}) - \mu_a(x_{f,i})}{\sigma_a(x_{f,i})}, \qquad (1)$$

  where $s_m(x_{f,i}) = S(x_{f,i}, x_{f,i'})$ was the (only) genuine (match) score of $x_{f,i}$ in the test set, and $\mu_a(x_{f,i})$ and $\sigma_a(x_{f,i})$ denote the mean value and standard deviation of its impostor (non-match) scores, i.e. the $S(x_{f,i}, x_{f',i'})$ with $f \neq f'$. This definition is based on the assumption that a good fingerprint yields high genuine scores, low impostor scores, and a small deviation in impostor scores. The normalized match scores were computed for each SDK individually.

- The fingerprints were binned into five classes (the NFIQ) according to their normalized match scores, where class one covers the fingerprints with highest score and class five those with lowest score. The bins are defined by unevenly selected quantiles of the frequency distribution of the normalized match score. By selection, the training set contained only fingerprints where the class is unambiguous over all SDKs, i.e. fingerprints for which the class varied between the considered SDKs were excluded from the training set.

- For all fingerprints, 11 *features* were computed using the NBIS (NIST Biometric Image Software) package [WGT+06]. These features were supposed to represent the characteristics of the image relevant for its recognition performance, and were based on the quality values computed by the MINDTCT feature extraction algorithm (which is part of the NBIS package) for minutiae and image regions (blocks). In particular, the features contained

- the number of pixels covered by the fingerprint in the image,
- the total number of minutiae detected in the fingerprint,
- for several thresholds, the number of minutiae with quality value exceeding this threshold, and
- for several thresholds, the number of blocks with quality value exceeding this threshold.

The quality values of minutiae and blocks are computed from three maps created during feature extraction marking areas with low contrast, without dominant ridge flow, or with high curvature of the ridge flow. These properties are assumed to reduce the reliability of detection of minutiae and other local fingerprint features (e.g. ridge flow orientation) utilized by comparison algorithms.

- A neural network was trained with the features and NFIQ classes of the fingerprints in the training set. In particular, a 3-layer feed forward perception network was used that takes as input the 11 features of the fingerprint and outputs an approximation of the NFIQ class. Using the real NFIQ (determined from the normalized match scores) the deviation from the correct answer was computed and the internal weights of the neural network were corrected accordingly.[1] A second set of fingerprints along with their (statistically determined) NFIQ and image features was used to continuously check the prediction accuracy and to stop training as soon as the accuracy was settled; well-timed ceasing of the training is crucial to prevent over-specialization of the neural network with respect to the training set.

Particular attention was given to the treatment of the genuine scores, as they may not equally depend on the quality of both fingerprints. In particular, in [TWW04] the plausible hypothesis was established that the genuine score depends on the fingerprint having lower quality; for instance, the comparison of a fingerprint of high quality and a poor quality imprint of the same finger will result in a low score. As a consequence, a genuine score is hardly significant for assessing the quality of the better quality fingerprint, at least, if the quality of both fingerprints differ considerably. For this reason, NIST used the trained neural network to assess for each finger which of the two fingerprints has higher quality and to perform a second training of the neural network, where to each fingerprint a *pattern weight* was assigned, which determined its influence on the network training:

- for fingerprints that were assessed to be the higher quality imprint of this finger a pattern weight of zero was assigned, implying that it had no influence to the training;

1 The neural network and the training algorithm deployed were originally developed and deployed for fingerprint pattern classification by the PCASYS algorithm in the NBIS package [WGT+06].

- for fingerprints that were assessed to be the lower quality imprint of this finger a pattern weight of one was assigned, implying that it had full influence to the training;

- if both imprints of a finger were assessed to be of equal quality, both were assigned with pattern weight 0.5.

Many more details about this process can be found in [TWW04]; the most important ones for optimization will be discussed in Section 3.

3 Re-Training the Algorithm (NFIQ+)

3.1 Fingerprint Images and SDKs Used

We used a database consisting of 9 imprints of 8784 fingers from 1098 individuals. The fingerprints were taken with 3 different optical sensors, each of which used to capture 3 images per finger. We used the imprints of 4392 fingers for training and the other 4392 fingers for testing. Furthermore, we used 5 SDKs from Dermalog, L1, Neurotechnologija, NIST, and NEC (in alphabetical order). For the sake of anonymity, the algorithms will subsequently be referred to as SDK A-E with random order.

For each SDKs, we computed 450 similarity scores for each fingerprint, 442 of which were impostor scores with imprints of 442 different fingers and 8 were the genuine scores with the imprints of the same finger. Based on these scores we computed the match score statistics for each SDK separately.

3.2 Treatment of Genuine Scores

The computation of 8 genuine scores per fingerprint (and SDK) allowed a more accurate consideration of the genuine scores as compared to the development of the original NFIQ.

Obviously, multiple samples allow a more accurate estimation of the expectation value. However, following the hypothesis of [TWW04] a genuine score is mainly significant for the lower quality fingerprint. With several genuine scores per fingerprint available, we were able to á priori (i.e. prior to the neural network training) rank the imprints of a finger according to their quality. We could then use this (preliminary) ranking to identify which genuine scores are significant for which fingerprints and finally could compute the normalized match score considering only the relevant (i.e. significant for the fingerprint) genuine scores. In particular, we implemented the following process:

1. For each fingerprint we computed the average over all its genuine scores, and used this value to compute a preliminary normalized match score and its corresponding preliminary rank among all fingerprints (over all fingers).

2. We used this ranking to identify for each fingerprint the set of *significant* genuine scores, which were all genuine scores with imprints of this finger having at least approximately the same quality. Precisely, we regarded only the scores with those

fingerprints, whose preliminary rank was at most 1/6 below the rank of the fingerprint in question.

3. We re-computed the normalized match scores, this time (for each fingerprint) considering only the significant genuine scores.

In the third step, the number of significant genuine scores could be quite small, in extreme cases even zero. This happened in particular for fingers where the quality of the 9 fingerprints varied significantly. In such a case, special treatment of this fingerprint was necessary, as described in Section 3.3.

As a second optimization potential, we were able to measure the dispersion of the genuine scores. It is clearly desirable that a fingerprint does not only produce high average genuine scores but also yields these high scores consistently over all imprints of comparable or better quality. Therefore, we modified the definition of the normalized match score as described in Section 3.3.

3.3 Defining the Normalized Match Score

The definition of NIST for the normalized match score (1) was based on the limitation that only one genuine score is available per fingerprint. Therefore, we had to generalize it to the case of several genuine scores by replacing the (only) genuine score $s_m(x_{f,i})$ in (1) by the mean of the genuine scores that are significant for this fingerprint (see previous section for a discussion). As a result, we obtained the following function

$$o_1(x_{f,i}) := \frac{\tilde{\mu}_m(x_{f,i}) - \mu_n(x_{f,i})}{\sigma_n(x_{f,i})} \tag{2}$$

where $\tilde{\mu}_m(x_{f,i})$ denotes the mean of the genuine scores that are significant for fingerprint $x_{f,i}$.[2]

On the other hand, as argued in Section 3.2, it is reasonable to regard the dispersion of the genuine scores as well. An obvious approach would be to use the standard deviation of the significant genuine scores of a finger in the modified function $o_2(x_{f,i})$ for the normalized match score, e.g. by replacing the denominator in (1) with $\sigma_n(x_{f,i}) + \tilde{\sigma}_m(x_{f,i})$ (as suggested in the annex of [ISO2]) or $\sqrt{\sigma_n(x_{f,i})^2 + \tilde{\sigma}_m(x_{f,i})^2}$, where $\tilde{\sigma}_m(x_{f,i})$ denotes the standard deviation of the significant genuine scores of fingerprint $x_{f,i}$. However, although there are robust estimators for the standard deviation from very small samples [RV02], we have to expect considerable inaccuracy in an estimation from at most 8 values. As the genuine scores are generally much greater than the impostor scores, this will result in significant noise in the denominator. Furthermore, the deviation of the genuine scores would predominate the deviation of the impostor scores in the denominator.

Therefore, we decided to choose a definition of the normalized match score, where the dispersion of the significant genuine scores (in particular, the deviation below the mean

2 We use the tilde on top of the symbol μ to distinguish from the mean value over all genuine scores.

value) is represented in the nominator. We found such a definition, by replacing the mean of the significant genuine scores $\hat{\mu}_m(x_{f,i})$ in (1) with an appropriate α-quantile of the fingerprint's genuine score distribution with $\alpha < 0.5$. The selection of the quantile size α is a compromise between accuracy and statistical significance: A large quantile, e.g. $\alpha = 0.3$, will not sufficiently represent the dispersion, while a small quantile, say $\alpha = 0.05$, can hardly be estimated from at most 8 samples. As a (heuristic) compromise we set $\alpha = 0.15$ and arrived at the definition

$$o_3(x_{f,i}) := \frac{Q_{0.15}(s_m(x_{f,i})) - \mu_n(x_{f,i})}{\sigma_n(x_{f,i})}, \tag{3}$$

where $Q_{0.15}(s_m(x_{f,i}))$ denotes the threshold that is (expected to be) exceeded by 85% of all significant genuine scores of fingerprint $x_{f,i}$. The estimation of this quantile from at most 8 samples is still error-prone, but its value is usually relatively close to $\hat{\mu}_m(x_{f,i})$. Hence, if the number of samples is not too small, say 4, the impact of the noise from the estimation is relatively mild.

For all fingerprints, where at least 4 significant genuine scores were available, we used its rank with respect to the normalized match score according to definition (3) as a basis for the NFIQ+ classes. In all other cases, reasonable accurate estimation of the quantile is hardly possible, and thus, we used its rank according to definition (2) as a fall-back solution. The rank-based fusion accommodates the fact that the two functions $o_1(x_{f,i})$ and $o_3(x_{f,i})$ produce incomparable values.

3.4 Robust estimation of statistical measures

The estimation of the mean and standard deviation of a fingerprint's impostor score distribution (over random comparisons within the world's population) from 442 samples is no problem because this sample size is sufficiently large to ensure a quite accurate estimation by means of the arithmetic mean and the sample standard deviation.

For a fingerprint's genuine scores, the situation is different, as we have at most 8 samples and in many cases, in particular for high quality fingerprints, the number of significant genuine scores is much lower. In order to estimate the mean and 15%-quantile of the significant genuine scores of a fingerprint from such small samples, we need robust estimators for these statistical measures. As shown in [WW05], the distribution of genuine and impostor scores is not Gaussian and even varies from vendor to vendor, and thus, we cannot assume any specific distribution of the genuine scores. We decided to use the median as the mean value, and to estimate the median (which is the 50% quantile) and the 15% quantile of the significant genuine scores by the quantile estimator of Harrell and Davis [HD82], which has been shown to be particularly robust for small samples (see [DLP94]). The estimator computes an approximation of a quantile of an arbitrary distribution by a weighted sum $Q_n = \sum_{i=1}^{n} w_{n,i} x_i$ of the samples, where the samples $x_1 \leq \ldots \leq x_n$ are ordered by size and the constants $w_{n,i}$ are given by

$$w_{n,i} = \frac{B(i/n, (n+1)\alpha, (n+1)(1-\alpha)) - B((i-1)/n, (n+1)\alpha, (n+1)(1-\alpha))}{B(1, (n+1)\alpha, (n+1)(1-\alpha))},$$

with $B(z, a, b) = \int_0^z y^{a-1}(1-y)^{b-1}\,\mathrm{d}\,y$ being the incomplete beta function.

3.5 Multi-Algorithmic Fusion and Definition of NFIQ+ Classes

For the training of the original NFIQ algorithm, NIST restricted the training set to those fingerprints that had the same NFIQ class for all 3 considered SDKs. However, the neglected fingerprints might have specific properties, which are the reason for their algorithmic-dependent performance and, hence, are not equally present in those fingerprints that exhibit similar performance for all SDKs. Consequently, the NIST approach for fusion may have resulted in an NFIQ that shows poorer performance on such fingerprints. Therefore, we decided to follow a different approach and to implement a fusion method for the quality measures obtained for the individual SDKs.

Since the similarity scores from different vendors have different ranges and distributions, the resulting values of the normalizes match score are incomparable. On the other hand, fusion on the basis of vendor-specific NFIQ+ classes seemed to rough. Therefore, we deployed a rank-based fusion: for each fingerprint, we computed the arithmetic mean of its ranks obtained with the individual SDKs, and then computed a final rank of all fingerprints over these mean values.

In order to evaluate the impact of the number and distribution of the NFIQ+ classes to the quality assessment accuracy by the neural network, we tried different classifications, in particular, 5 and 10 classes with uniform and non-uniform distributions of fingerprints among the classes. For each of these class definitions, we trained the neural network as described in Section 3.6 and evaluated the results.

3.6 Image Feature Selection and Neural Network Training

For our current investigation, we decided to keep the feature vectors as they are. Selection of appropriate feature vectors is quite complex and should be conducted independently of other optimizations to allow a step-by-step evaluation of the impact of individual improvements. The feature vectors were computed using a dedicated tool that is included in the NBIS package (see [WGT+06]).

The neural network training was conducted using the training software in the NBIS package. This software executes the neural network on a set of feature vectors of fingerprints from the training set, and uses its outputs and the true class numbers to determine the error according to a configured error function. It then uses an optimization method to determine the changes to the internal network weights that would best reduce the error for the processed fingerprints. After each training run, an equally large number of data of fingerprints of an independent test set is processed to measure the prediction accuracy. The number of training runs can be configured, but the training can also stop as soon as the rate, at which the classes of fingerprints from the test set are correctly predicted, does not improve significantly.

The neural network and the training software have many parameters and alternative options, some of which are discussed below.

- ***Number of nodes.*** While the number of input nodes and output nodes are predetermined by the number of feature vector components and the number of NFIQ+ classes, respectively, the number of nodes in the hidden (i.e. middle) layer can be freely configured. In the original NFIQ training, NIST used 22 nodes in this layer.

- ***Optimization method.*** For optimization of the network weights with respect to the error, the NBIS training software supports the Scaled Conjugate Gradient (SCG, see [M93]), as well as the Broyden Fletcher Goldfarb Shanno (BFGS, see [NW06]) method.

- ***Error function.*** The error function measures the deviation of the output of the neural network, which is given be the activations (each output node has an activation value which is normalized to range [0,1]) of the output nodes, from the perfect output, which would be an activation of one at the correct node (corresponding to the true NFIQ+ class of the input) and zero activation at all other output nodes. The NBIS package implements three different error functions, of which the Mean Squared Error (MSE) function is recommended by [WGT+06]. It measures the error of the activations $(x_1, \ldots, x_5)$ by

  $$(1-x_i)^2+\sum_{j\neq i} x_j^2,$$

  where i is the true NFIQ+ class. Since the NBIS neural network software was developed for classification problems and not for approximation of a function, the error functions neglect the distance of the individual output nodes from the correct node. However, the NFIQ+ is a gradual measure, and for a true NFIQ+ class 1 the activation of output node 2 is "less wrong" than the activation of node 5. Therefore, we modified the MSE error function so that it takes into account the distance between the individual output node and the correct node. In particular, we defined a modified error function MSE that computes the error as

  $$\left((1-x_i)^2+\sum_{j\neq i} |i-j|x_j^2\right)/C, \qquad (4)$$

  where the constant C keeps the average error at the same value as for the original MSE.[3]

- ***Regularization factor.*** The regularization is used to limit the size of the network weights and to avoid over-fitting of the network to the training set.

- ***Boltzmann pruning.*** In order to improve the computational performance of the neural network, edges with very low weights can be removed. This pruning is

3 Larger average error values can result in global decrease of the network weights which can affect the prediction performance, in particular in the context with Boltzmann pruning.

performed in a random pruning process resembling thermodynamic processes using the *Boltzmann temperature* as a parameter.

4 Evaluation of Results

4.1 Inter-algorithmic Consistency of Score Statistics

In order to analyze the consistency of the statistical measures used for quality assessment between the deployed SDKs, we evaluated correlations. As shown in Table 1, the rank-based correlation of normalized match scores between all used comparison algorithms is predominantly consistent. The ranks calculated with the individual SDKs exhibit high correlation.

The 15% quantile of the genuine scores are also very consistent between the different SDKs with correlation coefficients between 0.82 and 0.9. This high correlation is presumably the main reason for the consistency of the normalized match score. In contrast, the average impostor scores for the individual software packages are much less consistent, with correlations being predominately in the range between 0 and 0.36. However, as the impostor scores are typically much smaller than the genuine scores, in (3) the nominator is dominated by the 15% quantile of the genuine scores, and consequently, the inconsistency of the average impostor scores should only have very limited impact to the normalized match scores.

	SDK A	SDK B	SDK C	SDK D	SDK E
SDK A	1				
SDK B	0.86	1			
SDK C	0.75	0.75	1		
SDK D	0.75	0.77	0.62	1	
SDK E	0.61	0.59	0.70	0.48	1

Table 1: Rank-based correlation of normalized match scores.

Much worse for the inter-algorithmic consistency of the quality assessment is the standard deviation of impostor scores, which is the only term in the denominator of (3) and turned out to be quite inconsistent among the SDKs. As the values from Table 2 show, the rank-based correlation of impostor scores is very low in most cases. In the case of SDK D, the inconsistency is presumably due to fact that an internal threshold is applied and very frequently the same minimal impostor score 0 is output, which results in an unnatural low standard deviation of impostor scores. (Most matching algorithms perform impostor normalization internally but this process annihilates information needed for the training.)

	SDK A	SDK B	SDK C	SDK D	SDK E
SDK A	1				
SDK B	0.07	1			
SDK C	0.43	-0.10	1		
SDK D	-0.03	-0.01	-0.22	1	
SDK E	0.38	0.14	0.69	-0.19	1

Table 2: Rank-based correlation of the standard deviation of impostor scores.

4.2 Results of Re-Training

After training the neural network on 10 uniformly distributed classes, it predicted the NFIQ+ class correctly for 18.3% of all fingerprints in our test set. This prediction rate was reached and remained stable after just a few training runs.

Subsequently, we tested several optimizations of the neural network training to achieve better results in terms of predicting the correct class of the fingerprint.

- ***Variation of the regularization factor:*** Mainly performance increases in re-training time were observed by changing the default regularization factor of 10^{-4}. The recognition performance did not change significantly.

- ***Changing numbers of nodes in the hidden layer:*** No increases of recognition performance was observed by increasing and decreasing the default number of hidden layers (22).

- ***Usage of BFGS optimization:*** Using the BFGS method for optimization instead of the SCG method (see Section 3.6) resulted in a significant worse recognition performance. We also tried combining both methods (in succession), but this approach did also not yield any improvement over using SCG only.

- ***Adaption of class distribution:*** Changing the class distribution to the distribution of the original NFIQ algorithm (non-uniform distribution) leads to a prediction rate of about 48 %, which is slightly better than the 46% achieved by the original NFIQ algorithm. Note that the increase of the prediction rate from 18.3% to 48% is not necessarily an improvement as 5 non-uniform classes are much easier to predict than 10 uniform classes.

- ***Modification of the error function:*** Out of three available error functions for training the neural network Mean Squared Error (MSE) returns the best results in terms of class prediction rate. Our optimization of the error function according to (4) results in a marginal reduction of the rate at which the class is correctly predicted but also significantly reduces the frequency of larger deviations from the correct output. Table 3 summarizes the statistic of deviation of the predicted

class from the correct value for the original MSE function and our optimized error function[4].

Deviation of predicted classes from correct value	Frequency *MSE*	Frequency *Optimized MSE*
0 (correctly predicted class)	35.1 %	34.7 %
1	36.3 %	41.1 %
2	21.0 %	19.9 %
3-4	7.6 %	4.3 %

Table 3: Deviation of predicted classes for original and optimized error function MSE.

4.3 Biometric Performance Results

In order to determine the biometric performance of our re-trained NFIQ+ algorithm and to quantify the improvements compared to the original NFIQ algorithm, the following simulation was conducted using the test set of the available data basis (which is disjoint from the set used for training). For each finger, the "best" imprint in terms of fingerprint quality was selected using the quality assessment methods to be compared (e.g. the NFIQ and NFIQ+ algorithms). (If quality assessment resulted in multiple imprints having the same quality, of these the imprint with lowest index in the database was selected.) The 450 similarity scores (442 impostor and 8 genuine scores) calculated for the statistical determination of the NFIQ+ classes (see Section 3.1) were then used to determine the false match rate (FMR) and false non-match rates (FNMR) of this imprint[5]. The calculated error rates were then used to plot a Detection Error Trade-Off (DET) curve, which displays the biometric recognition performance that can be expected if during enrollment the best fingerprint was selected by the corresponding quality assessment method.

We also calculated the error rates for the case that the best imprint is chosen from each triple according to the statistically determined classes. The DET curve of this imaginary *perfect class predictor* shows the optimal achievable error rates for the given statistical NFIQ+ definition. As a result, by comparing the DET curve of the NFIQ+ algorithm and the curve of this perfect class predictor, the prediction accuracy of the neural network can be assessed according to a very application-relevant metric.

Figure 1 shows DET curves for the re-trained NFIQ+ algorithm using 5 non-uniformly distributed classes and for the original NFIQ algorithm, as well as for the perfect class predictor (labeled "statistical classification") for 5 non-uniform classes. Unfortunately, the improvement by the re-training is quite small: when using the re-trained NFIQ+ algorithms, the equal error rate (EER) decreased to 6.38% from 6.50% for the NIST NFIQ algorithm. However, due to time constraints, we could not evaluate an NFIQ+ algorithm trained with the modified error function (see Section 3.6).

4 5 uniformly distributed classes were used to obtain these values.

5 Bozorth3 from the NBIS package [WGT+06] was used as comparison algorithm.

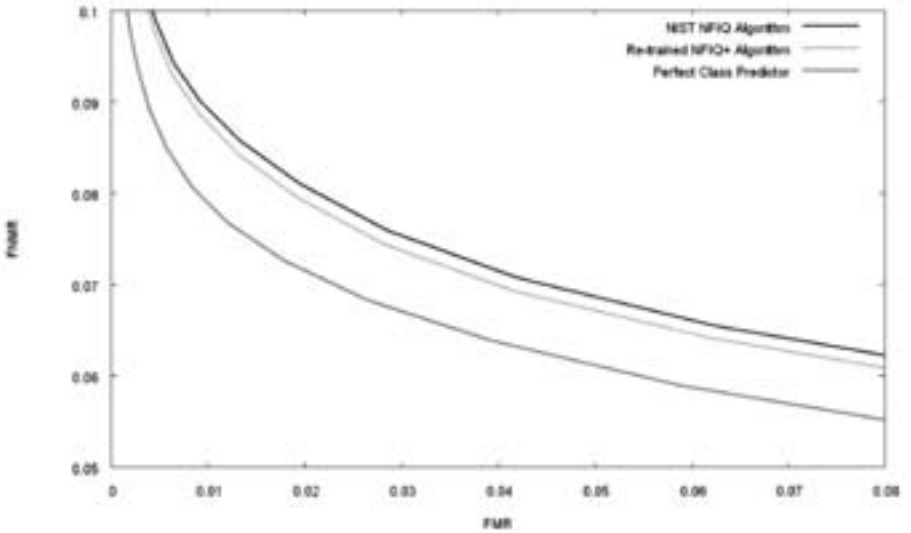

Figure 1: DET curves of for 5 non-uniform classes.

By comparing the DET curves of the perfect class predictor for different class definitions (varying number and distribution of classes), we can determine which class determination bears the most potential with respect to error rates assumed that the class can be predicted with comparable accuracy by a trained neural network. We evaluated the DET curves for the following three statistical classifications:10 classes with uniform distribution, 5 classes with uniform distribution, and 5 classes resembling the (non-uniform) distribution of the original NFIQ. The result depicted in Figure 2 shows that the biometric performance is better for 10 classes than for 5 classes. On the contrary, the distribution of the classes seems to have relatively little impact on the biometric performance.

The approach of evaluating the DET curves of the imaginary perfect class predictor for varying class definitions could also allow assessment of the optimization potentials that influence the class definitions, e.g. the training set used, the statistical functions used, the fusion method, etc. We will discuss this approach in more detail in Section 5.

5 Conclusions and Future Work

In our re-training, the main focus was to investigate the optimization potential by a better data basis. Rolled and inked impressions, which were (among others) used in the original training of the NFIQ algorithm, are not relevant for authentication systems based on electronic capture devices. By using only plain live scans captured by optical fingerprint sensors the data basis was adopted for this purpose. Further optimization potentials arose from this new data basis. As more than 2 imprints per finger were available, a more accurate and comprehensive treatment of genuine match scores was possible, which required a modification of the statistical functions used for quality assessment.

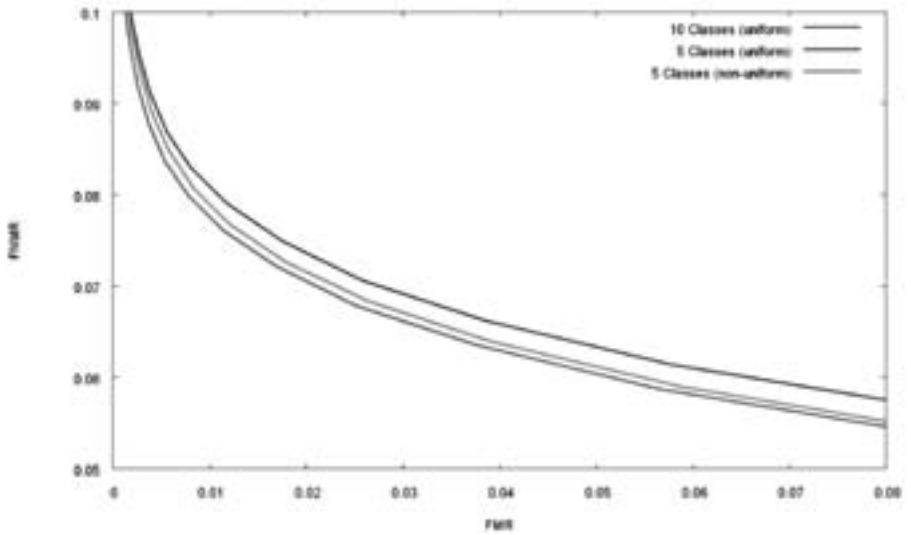

Figure 2: DET curve for perfect class predictors of 3 different class definitions.

Furthermore, since 2004, new fingerprint software has been developed and are deployed. Thus, using up-to-date feature extractors and comparison algorithms makes the re-trained NFIQ algorithm fit for current applications.

Finally, we increased the number of classes as it seemed that the current categorization by the NFIQ is not exact enough and does not sufficiently differentiate good fingerprints. Conducting a rank-based fusion of all comparison algorithms was also a new approach considered within this re-training process.

Unfortunately, the evaluation of our re-trained NFIQ+ algorithm only shows a slight improvement over the original NFIQ algorithm of NIST. As potential reasons, we suspect the internal thresholding of one SDK and a non-optimal neural network training. Nevertheless, a lot of experience was gained and lessons were learned regarding the statistical measurement of fingerprint image quality and neural network training. Moreover, several potentials for optimization were detected, which can be structured according to the 5 main steps in the general re-training process:

- ***Fingerprint data basis:*** The selection of the fingerprint data basis used for training is essential for the results of the re-training. Special data sets may lead to specialized behavior of the re-trained algorithm, which may or may not be desired (see end of this section for a discussion). Key factors of this area are (beside the number of fingerprints) the type of fingerprints (plain or rolled, inked or live-captured), the sensors used for capturing, and the number of imprints per finger.

- ***Calculation of similarity score statistics:*** The selection of the fingerprint SDKs (feature extraction and comparison algorithms) used for the re-training will

impact the applicability of the resulting NFIQ algorithm. As with the fingerprint data basis (see above) specialization for specific SDKs may or may not be desired. Furthermore, the definition of the normalized match scores determines the relevance of the NFIQ class definition to the actual application scenario.

- ***Definition of NFIQ classes:*** The number and (according to our evaluation, to a lower extent) the distribution of the NFIQ classes can have a significant influence on the results of the re-training. Furthermore, it is important to define how to merge normalized match scores obtained from different SDKs.

- ***Image feature selection:*** We believe that the current definition of the image features used to predict the NFIQ leaves great optimization potential. However, the features must be chosen with utmost care to ensure that they are highly significant for the NFIQ. The number of features is also crucial as too few features may not reflect sufficient information about the image quality, and too many features may render training of a neural network difficult.

- ***Neural network training:*** Training the neural network may be strongly influenced by adapting and changing various parameters of the neural network. This is particularly true for a larger number of features. Key factors for successful neural network training are the size of the network, the error function but also other optimization methods and parameters described in Section 3.6.

Figure 3 illustrates the areas and key factors for optimizations.

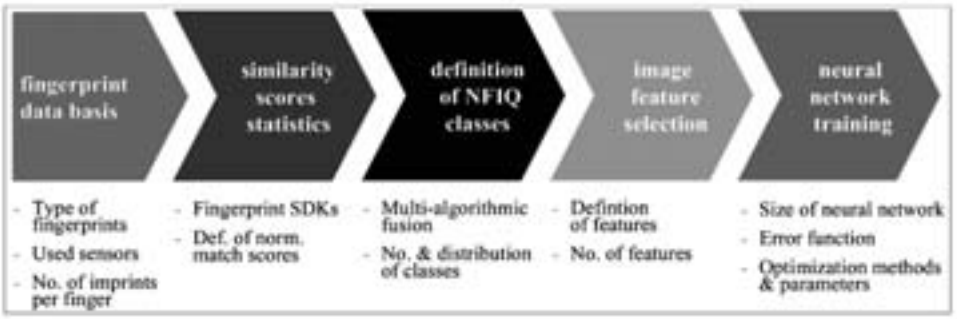

Figure 3: Areas and key factors for optimization of the NFIQ.

For future attempts to optimize the NFIQ algorithm, we propose the following general two-step approach:

1. ***Optimization of classification:*** In a first step, the statistical classification, i.e. the NFIQ classes definition, should be optimized, as it constitutes the basis for the prediction by the neural network. This comprises the first three areas depicted in Figure 3, in particular, selection of the data basis and appropriate SDKs, calculation of the similarity score statistic, and its translation into NFIQ classes.

The potential of a candidate classification can be evaluated by the DET curve of the perfect class predictor as shown in Figure 2.

2. ***Optimization of prediction:*** In the second step, the most promising NFIQ classification identified in step 1 is used as a basis to develop a prediction algorithm. Within this step an eligible set of image features is selected, and a neural network is trained using various parameters and methods to achieve an optimal prediction accuracy.

Developing several specialized NFIQ algorithms is also an option. Specialized versions with respect to certain biometric sensors or to certain fingerprint comparison algorithms may be useful in some application scenarios. Nevertheless, a generalized version of the NFIQ algorithm is still required. For instance, in case of electronic passports and identity cards, where fingerprint image data is stored for verification of the document holder at border control posts such a generalized version of NFIQ is crucial.

References

[WGT+06] C. Watson, M. Garris, E. Tabassi, C. Wilson, R. McCabe, S. Janet, K. Ko, "User's Guide to NIST Biometric Image Software (NBIS)", National Institute of Standards and Technology, 2006

[TWW04] E. Tabassi, C. Wilson, C. Watson, "Fingerprint Image Quality", NIST Internal Report 7151, National Institute for Standards and Technology, 2004

[ISO2] International Organization for Standardization, "ISO/IEC 29794-1 - Information technology -- Biometric sample quality - Part 1: Framework", 2009

[RV02] P. J. Rousseeuw, S. Verboven, "Robust estimation in very small samples", Computational Statistics & Data Analysis, 40, 2002

[WW05] J. C. Wu, C. Wilson, "Nonparametric Analysis of Fingerprint Data", 7226, National Institute for Standard and Technology, 2005

[HD82] F. E. Harrell, C. E. Davis, "A new distribution-free quantile estimator", Biometrika, 69, 1982

[DLP94] T. E. Dielman, C. Lowry, R. Pfaffenberger, "A comparison of quantile estimators", Communications in Statistics - Simulation and Computation, 23, 1994

[M93] M. F. Møller, "A scaled conjugate gradient algorithm for fast supervised learning", Neural Networks, 6, 1993

[NW06] J. Nocedal, S.J. Wright, "Numerical Optimization (2nd ed.)", Springer-Verlag, 2006

A Reference Framework for the Privacy Assessment of Keyless Biometric Template Protection Systems

Tom Kevenaar[1)], Ulrike Korte[2)], Johannes Merkle[3)], Matthias Niesing[3)], Heinrich Ihmor[2)], Christoph Busch[4)], Xuebing Zhou[5)]

[1)] priv-ID B.V.
High Tech Campus 9
5656 AE Eindhoven,
the Netherlands
tom.kevenaar@priv-id.com

[2)] BSI
Postfach 20 03 63
53133 Bonn, Germany
ulrike.korte@bsi.bund.de
heinrich.ihmor@bsi.bund.de

[3)] secunet
Kronprinzenstraße 30
45128 Essen, Germany
johannes.merkle@secunet.com
matthias.niesing@secunet.com

[4)] Hochschule Darmstadt-CASED
Mornewegstraße 2
64293 Darmstadt, Germany
christoph.busch@h-da.de

[5)] Fraunhofer IGD
Fraunhoferstraße 5
64283 Darmstadt, Germany
xuebing.zhou@igd.fraunhofer.de

Abstract[1]: Over the past decades, a number of methods have been reported in the scientific literature to protect the privacy of biometric information stored in biometric systems. Keyless Biometric Template Protection (KBTP) methods aim to protect biometric information without the use of long-term secrets by deploying one-way functions. These KBTP methods are currently developed to an extent that commercial products have become available. When assessing and comparing different KBTP methods it is important to have a common and generic approach. Therefore, in this paper we present a reference framework that can be used in assessing and comparing the privacy properties of KBTP systems.

1 Introduction

Biometric systems are becoming increasingly popular because they may offer more secure solutions than traditional means for authentication such as PIN codes, passwords and security badges because a biometric characteristic is tightly linked to an individual. For the same reason, biometrics can prevent the use of multiple identities by a single individual. Finally, in many applications biometric authentication is also considered to be more convenient.

Biometric technologies are, however, not without their challenges [JRP06]. Although accuracy, speed and interoperability remain important, this paper focuses on the security of biometric systems as well as privacy issues related to the biometric information stored

[1] This work is part of the BioKeyS project which is supported by the Bundesamt für Sicherheit in der Informationstechnik (BSI), Germany.

in these systems. Many of these challenges are related to the special properties of biometrics as compared to traditional means for authentication:

- Biometric characteristics are tightly coupled to an individual which makes revocation and re-issuing of authentication information unfeasible. In contrast, PIN codes, passwords, tokens, etc. can easily be revoked and re-issued;
- Biometric data is personal and in many cases contains sensitive information. For example, it might contain information on the health status of an individual [Pe65], gender, ethnicity, age, etc. Therefore, in contrast to PIN codes and passwords, in many countries biometric data are considered to be Personally Identifiable Information and use of biometrics is governed by privacy legislation (e.g. [Eu08]);
- Each individual has a limited number of instances for each biometric characteristic (e.g., one face, two irises, ten fingers) while the number of possible passwords or token identifiers is several orders of magnitude higher. As a consequence, an individual will have to re-use the same characteristic in different applications which can lead to cross-matching of applications;
- Biometric measurements are affected by noise and other forms of variability while authentication protocols based on passwords and the like rely on 'bit-exactness' of the authentication information. This variability limits the distinctiveness of biometric features. Although this limitation also applies to, say, 4-digit PIN codes, passwords and token identifiers allow for a higher level of distinctiveness than single biometric modalities.

These special properties of biometric characteristics and measurements have an impact on security and privacy considerations of biometric authentication systems. Keyless Biometric Template Protection (KBTP) technologies can make an important contribution in solving some of these vulnerabilities [CS07]. In this paper we define a framework to assess the privacy of KBTP methods. In Section 2 we will define security and privacy for biometric systems and define the objective of KBTP methods. In Section 3 an overview of practical KBTP methods will be given and an abstraction will be made to allow for a generic framework. Finally, in Section 4, the reference framework will be given illustrating how the privacy assessment of KBTP methods could be done.

2 Security and privacy

Figure 1 gives a high-level overview of a biometric system where, without loss of generality, we consider a fingerprint verification system. During enrolment, a fingerprint sensor SENS generates the image sample of a fingerprint. After processing the image and extracting relevant features in the feature extraction block FE, a template representing the fingerprint is stored as reference in the biometric system (STOR). During verification, an individual claims an identity, and a so-called probe image from this individual is obtained. This image is transformed into a template and compared (COMP) with the template stored in the biometric system corresponding to the claimed identity. The comparison subsystem produces a similarity score and applying a threshold T to this score leads to an Accept or Reject message.

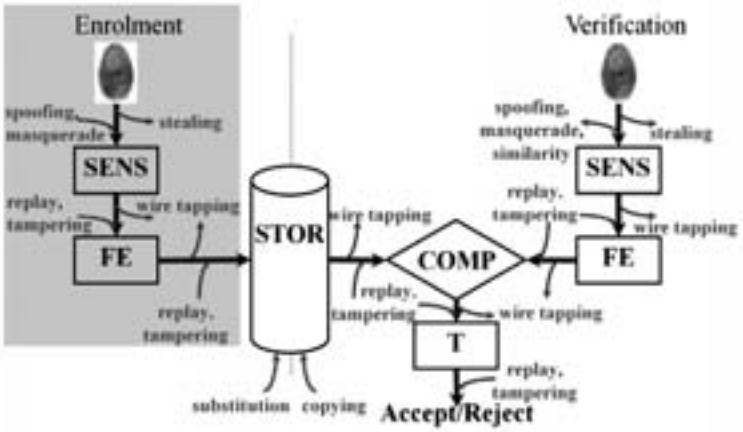

Figure 1 Security and privacy vulnerabilities in a biometric system

Figure 1 also depicts some important vulnerabilities of a biometric system [Bu08] [CS07] [RCB01]. Although in the literature a large number of potential attacks are mentioned, we propose to group them into the following two categories:

- **Insertion**, depicted with ingoing arrows in Figure 1,
- **Eavesdropping**, depicted with outgoing arrows in Figure 1.

As illustrated by the gray rectangle in Figure 1, in many cases the enrolment functionality can be assumed to operate in a secure environment such that the most important vulnerabilities are concerned with the storage and verification functionality.

In the context of biometric systems we associate the *security* of a biometric system with *insertion* vulnerabilities while the *privacy* of a biometric system is associated with *eavesdropping* vulnerabilities. Thus, the *security* of a biometric system defines how difficult it is to be illegitimately accepted by the system. In contrast, the *privacy* of a biometric system is related to the difficulty to obtain any relevant information from a provided biometric characteristic other than a verification decision.

2.1 A Perfectly Private Biometric system

If we assume that enrolment takes place in a trusted environment, a Perfectly Private Biometric (PPB) system can be defined as the gray rectangle in Figure 2 (see also [Br09] [ISO10]). A PPB system contains a sensor SENS, a feature extractor FE, a state-of-the-art comparator COMP, a threshold module T and storage STOR. When offering a fingerprint, possibly in combination with an identity claim, the system outputs an Accept or Reject decision. This PPB system is sealed and perfectly private in the sense that it outputs the minimal required amount of information in the form of a binary Accept/Reject decision. Furthermore, assuming a sensor leaving no latent prints (e.g. a touchless sensor or swipe sensor), the system has no eavesdropping vulnerabilities.

On the other hand, if the comparator COMP is not perfect, the system will occasionally incorrectly generate an Accept message due to a wrong decision of the comparator and in that sense the system is not perfectly secure.

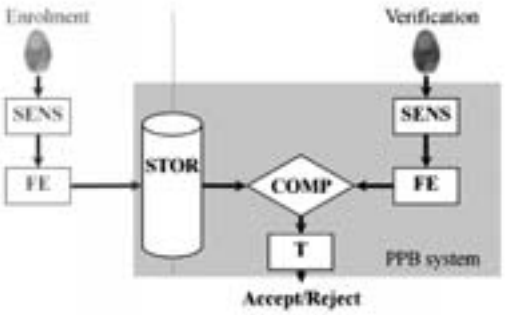

Figure 2 A Perfectly Private Biometric system

In practical systems, apart from the non-perfect comparator COMP, many security and privacy vulnerabilities can be prevented using standard cryptographic protocols. For example, replay and tampering attacks can be prevented by combining message uniqueness (by, for example, time stamps or sequence numbers) and verifying message authenticity using Message Authentication Codes (MACs) or digital signatures. A substitution attack can be foiled by adding a digital signature over a template and the wiretapping attack can be thwarted using secure channels which can be generated by the Diffie-Hellman key exchange protocol [DH76]. Although these methods require at least mid-term secret keys, management of these authentication keys is much easier than management of keys used to protect, for example, stored information.

The copying attack is very similar in nature to a wiretapping attack by interpreting storage as a communication channel and it can be prevented by a secure channel (encryption). In a (storage) communication channel the key must be retained for at least as long as the message is in storage. This means that in order to protect biometric information from a copying attack using standard cryptographic protocols, a long-term secret is required.

There are two approaches for storing and handling these long-term secret keys. The first is to keep the keys inside the biometric system under control of the biometric system owner. In a practical situation, this will lead to protocols and access rights that are limited to trusted operators only. An important drawback is that if these biometric systems scale up, the protocols become vulnerable to "incidents" by sloppy execution, change in regulations or legislation, human mistakes or intentional misuse.

The second method to handle long-term secrets is to store the keys outside the system. In many proposed systems the user will have control over the key. The key can be stored on a smartcard or token, or it can be derived from a password, PIN code, pass phrase, etc. An advantage of this method is that the user is in control of his biometric information, and the system owner does not have access to the decrypted biometric data. Drawbacks are that it reduces the convenience advantage of biometrics and it does not allow biometric identification (1-to-many), thus limiting the scope of biometric applications.

The purpose of Keyless Biometric Template Protection (KBTP) technology is to eliminate the drawbacks of both approaches by obviating long-term keys. This prevents

key abuse by the biometric system owner while simultaneously allowing higher convenience and biometric identification. Thus, the goal of (KBTP) technology can be formulated as to *prevent relevant biometric information to be obtained from storage facilities in biometric systems without the need for long-term secrets.*

Ongoing ISO standardization activities [ISO10] more specifically state that safeguarding the privacy of the data subject comprises i) preventing anyone to have access to biometric data or derived attributes thereof that are not required and agreed upon by the data subject and ii) ensuring that third parties and external observers have no access to the biometric references.

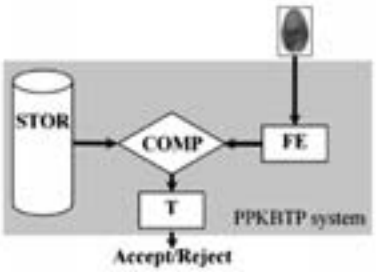

Figure 3 A Perfectly Private Keyless Biometric Template Protection system

2.2 Perfectly Private Keyless Biometric Template Protection

Based on the discussion above, a Perfectly Private Keyless Biometric Template Protection (PPKBTP) system can be depicted as in Figure 3 where, without loss of generality, it is assumed that feature extraction is part of the KBTP technology. It contains a feature extractor FE, a state-of-the-art comparator COMP, a threshold module T and storage STOR. It accepts the image of a fingerprint, possibly in combination with an identity claim, and outputs an Accept or Reject decision. This (conceptual) PPKBTP system does not need any secrets while at the same time, it does not leak any biometric information (other than the binary Accept/Reject decision) and in that sense is the highest achievable goal of any practical KBTP system. It is important to notice that vulnerabilities that are common between a PPKBTP system and a practical implementation thereof cannot be attributed to limitations of the implementation but must be attributed to the limitations of the use of biometrics. These intrinsic biometric vulnerabilities need to be addressed by proper system design.

3. Practical KBTP methods

Practical KBTP systems aim at implementing the PPKBTP system. In the scientific literature a number of methods are proposed to protect biometric information stored in a biometric system. In this section, a brief overview will be given and, following the reasoning in the previous chapter, we will only consider methods that do not require secret information to achieve this protection. Therefore methods as presented in, for example, [TNG04] [Br02][Ti02][SKK04] will not be considered.

The essence of all KBTP systems is that a biometric template, before it is stored in the biometric system, is first transformed into a representation from which it is impossible to retrieve any biometric information. On a high level of abstraction, all practical KBTP methods use the same format to represent the protected biometric information consisting of a Pseudonymous-Identifier (*PI*) and Auxiliary Data (*AD*) [ISO10]. The *PI* is generated using a (keyless) one-way function (e.g. a hash function $h(.)$) which forms the basis of the protection mechanism. The *AD* essentially contains variability information and/or randomization data. During verification *AD* is combined with a probe biometric measurement to generate a candidate Pseudonymous-Identifier *PI**. During verification, *PI** is compared with *PI* leading to an Accept or Reject message. Thus, practical KBTP protected templates consists of the pair (*AD*, *PI*) and KBTP methods differ in the way that the *PI* and *AD* are generated. Next, a brief overview of KBTP methods is given.

- Mytec [So98] was the first practical KBTP system. The method works directly on (fingerprint) images and protects the image by multiplying the phase part of the Fourier transform of a (fingerprint) image $F(\omega)$ with a random phase function $\varphi(\omega)$ and $H(\omega)=F(\omega)\varphi(\omega)$ is stored as auxiliary. A secret S is embedded by pointing to certain bits in $c(x)$, which is the inverse Fourier transform of the random phase function $\varphi(\omega)$ multiplied by the magnitude part of the (fingerprint) image "optimal" filter. *PI* is defined as $h(S)$ where h is a cryptographic hash function (e.g. SHA256).
- In [RCB01] the authors introduce an approach known as cancelable biometrics[2]. During enrolment, the image of a biometric is distorted using a parameterized one-way geometric distortion function before storing it in a biometric system. The parameter determining the distortion function is stored as *AD*. The function is made such that from the distorted image it is difficult to retrieve the original image. The distorted image is stored as *PI*.
- The fuzzy vault [JS02] is a general cryptographic construction that allows storing a secret S in a "vault" that can be locked using an unordered set X. The secret S can only be retrieved from the vault using a set Y if the sets X and Y have sufficient overlap. The "vault' is stored as *AD* while *PI* is set to $h(S)$. The use of unordered sets makes the method well suited to be used with minutiae fingerprints (see e.g. [UPJ05]).
- In the recently proposed Biotope™ scheme [Bo06], each component x of a feature vector is translated by t and scaled by s to obtain $v = (x-t)/s$. The resulting value v is separated into the integer $g = \lfloor v/E \rfloor$ and the residual $r = v \bmod E$ such that $v = g+r$ where E is a parameter. The entities t, s, r are stored as *AD*, while the integer part g is first passed through a one-way function to obtain *PI* which is then stored.
- The Norwegian company Genkey has developed an approach referred to as Biocryptics™ [LLO06]. The approach works directly on continuous features. In order to cope with noise and other variabilities, a correction vector, stored as *AD*, is used to shift a measured feature to the middle of a quantization interval that defines one bit of a binary string S to be embedded in the biometric template. *PI* is interpreted as a public key derived from S. The method resembles the so-called

[2] The term 'cancelable biometrics' is somewhat misleading because clearly the biometric itself cannot be cancelled. In the context of KBTP methods, the terms 'cancelable' and 'revocable' refer to the property that authentication information can be changed and revoked.

shielding functions as proposed in [LT03].

- The Fuzzy Commitment scheme [JW99] is considered most suitable for biometrics that have a template in the form of an ordered string or fixed length feature vector. A biometric X represented as a binary string is XORed with a codeword C of an (arbitrary) error correcting code. $C \oplus X$ is stored as AD while PI is set to $h(C)$.

In this brief overview it was shown that all practical KBTP systems generate a private representation of a biometric in the form of the KBTP template (AD, PI) which is stored in the biometric system. In the following section a framework for the assessment of the privacy of such systems will be given.

4 Privacy of KBTP systems

4.1 Privacy requirements

In Section 2, a high level notion of privacy was introduced in terms of a (conceptual) PPB and PPKBTP system that leaks no information about biometric templates. The concept of a PPB system is also described in [Br09] which serves as a basis for a new ISO standard that is currently being developed [ISO10]. This ISO standard provides guidance for the protection of biometric information under various requirements for confidentiality, integrity, availability and renewability/revocability. More specifically the standard proposes the following privacy goals for biometric information:

- irreversibility: To prevent the use of biometric data for any other purpose than originally intended, biometric data shall be transformed in such a way that the biometric sample or a deductible attribute that does not serve the agreed purpose of the identity management system cannot be retrieved from the transformed representation;
- unlinkability: The stored biometric references shall not be linkable across applications or databases;
- confidentiality: To protect biometric references against access by external observers resulting in a privacy risk, biometric references shall be kept confidential;
- data minimization: minimizing irrelevant and/or undesired processing of personal data, including during the verification of a person's identity.

The standard does not prescribe the mechanisms of how to achieve these requirements but as a framework standard it is applicable to a much wider range of techniques than KBTP techniques including traditional encryption of the template. In the assessment of a KBTP method, it must be verified to what extent it obtains these privacy goals or how much effort an adversary should invest in order to thwart one of these goals. Clearly, whether or not an adversary can thwart one of the privacy goals depends on his capabilities. The adversary capabilities are formalised in the following section.

4.2 Adversary capabilities

In the assessment of security systems and KBTP systems, it is essential to define the capabilities of an adversary.

A first high level notion in adversary capabilities is to assume a black-box model or a white-box model [Wy09]. In the black-box model it is assumed that an adversary knows all the details of the algorithm. During operation, the adversary has access to the inputs and/or the outputs of the algorithm but not to the internal intermediate computation results. In contrast, the white-box model assumes that an adversary, besides all the black-box capabilities, also has access to the implementation of the algorithm and is able to observe and modify intermediate computation results.

The white-box assumption is a very strong. For example, most implementations of traditional ciphers (such as RSA, AES, etc.) and security systems are not secure under the white-box model and it is customary in the assessment of such systems to adopt the black-box model. Therefore it seems reasonable to also assess KBTP methods under the black-box model.

A second notion that is important in the assessment of security systems is the efficiency of an attack. If the (minimum) required effort to thwart a certain security goal of a system (e.g. secrecy, privacy, authenticity) is higher, then the system is considered to have a better security concerning this specific goal. The security is commonly expressed as a number of bits which is the logarithm (to the base 2) of the required effort. This notion of the efficiency of an attack for a certain security goal should also be used in the assessing the privacy properties of KBTP systems.

A third important notion in the assessment of security algorithms is that the best-known-attack against a certain security goal defines the security of the algorithm for this goal. For example, if the General Number Field Sieve (GNFS) is the best (i.e., in terms of required effort/resources) known algorithm to break RSA, then the security of the RSA algorithm is directly related to the required effort of the GNFS to factor the public RSA modulus in its two primes. The notion of best-known-attack should also be adopted in assessing KBTP solutions.

4.3 Possible attacks

In defining adversary capabilities one can distinguish between high level and low level attacks. High level attacks are independent of the algorithmic details of the underlying KBPT method while low level attacks target specific properties of the KBTP method.

4.3.1 High Level Attacks

FAR attack Being a high level attack, the FAR attack does not exploit algorithm-specific knowledge. Instead it uses the fact that practical biometric systems have a non-zero False Accept Rate (FAR). The FAR is the probability that the biometric system will incorrectly accept an unauthorized user in a verification setting. Thus, given a KBTP private template, the attack consists of trying sufficient biometric images until an Accept

message is obtained. If the comparator is operating at FAR=α and the required effort for a single comparison is N_{FAR} then the expected required effort for a successful FAR attack is N_{FAR}/α.

It is important to note that the FAR attack is applicable also to the PPKBTP system introduced in Section 2.2 and therefore, it does not exploit a vulnerability of the KBTP implementation per se. Still, it allows the adversary to obtain information about the protected biometric information in the sense that a successful trial image is in some sense similar to the image that was used to generate the KBTP template. Thus, the FAR attack has an impact on the privacy goal of 'irreversibility'. If different applications use a PPKBTP system, the FAR attack can also be exploited to link templates across applications. Thus, the FAR attack also puts a limit on the 'unlinkability' goal of [ISO10]. Therefore it is essential that the system design incorporates a strategy to prevent the FAR attack.

Hill climbing In traditional biometric systems, hill climbing exploits the continuity of a similarity score as a function of changes in the input image [Ma06]. Regarding the ISO privacy goals, this threat is similar to the FAR attack: the adversary obtains an image that is in some sense close to the image that was used to generate the KBTP template.

Referring to Section 3 it can be seen that KBTP systems are traditionally implemented such that they do not output a similarity score but just a one-bit Accept/Reject decision (or a hashed version of some stable value *S*) which thwarts the high level hill climbing threat. For KBTP systems, hill climbing allows obtaining a working image using the FAR attack but it does not allow increasing the image quality.

4.3.2 Low Level Attacks

Hash inversion In most KBTP systems, the *PI* is computed from a secret bit string *S* using a strong one-way hash function (see Section 3). In this case, a first low level threat is inverting the hash in the *PI* of a private template. For good hash functions, the best-known-attack for hash inversion is to perform an exhaustive search (dictionary attack) which means that the required effort for inversion is proportional to $2^{|S|}$. For example, in case of the FCS, the adversary would have to try all possible codewords *C* of the applied error correcting code. Since the one-wayness of the hash function in KBTP systems is an essential part of the privacy mechanism, a successful inversion of the hash function will at least leak some information on the biometric that was used to generate the KBTP template (*AD*, *PI*) and will affect the 'irreversibility' and 'unlinkability' goals of [ISO10]. In view of the notion of the best-known-attack, hash inversion should be significantly more difficult than the FAR attack where is should be noted that hash inversion does not necessarily brings the adversary the same information (a ‘working’ biometric characteristic) as a FAR attack.

Using AD In a KBTP protected template (*AD*, *PI*), the auxiliary data *AD* contains user-specific information. Therefore, in information theoretical sense it is expected that *AD* will leak information on the biometric that was used to generate the KBTP template. On the other hand, it can be shown that if robustness against variability is required, some

privacy leakage cannot be avoided [DRS04]. If and how this privacy leakage can be exploited depends on the specific KBTP system. For example, in case of the FCS, if the code word C is chosen from an (n,k) error correcting code, then k information bits of the biometric are protected and, in an information theoretic sense, AD leaks n-k bits of information about the biometric. However, this information theoretic representation does not indicate how this leakage can be exploited by an adversary to learn dedicated information about the biometric or to thwart the 'irreversibility' and 'unlinkability' goals.

4.4 Information theoretic treatment

As opposed to the assessment of a KBTP method by known attacks, many scientific publications use information theoretic measures of privacy. Although these measures do not always point towards a practical attack, they are useful in assessing the required effort for certain attacks.

In terms of the unified KBTP template format (AD, PI), while assuming that PI leaks no information because it is protected by a hash function, it is interesting to consider $H(X|AD)$ which is the remaining entropy in the biometric information X after observation of AD. Two definitions of entropy have been considered for measuring the information leakage of a KBTP system.

- The Shannon entropy $\mathbf{H}$. Due to its rich mathematical theory, this measure allows a very comprehensive analysis of information leakage in a KBTP system [Ig09]. The conditional Shannon entropy $\mathbf{H}(X|AD)$ can be used to measure the average information content of X after observation of AD.
- The min-entropy $\mathbf{H}_\infty$. This defines an upper bound for the success probability of an attacker who tries to guess X from AD. The average min-entropy $\hat{\mathbf{H}}_\infty(X|AD)$ provides an upper bound for an attacker's success probability for average AD [DRS04].

If the entropy is a measure for the required effort for certain attacks, one could be interested in the relation between $\mathbf{H}(X|AD)$ or $\mathbf{H}(S|AD)$ and the FAR of the system (where S is the embedded key). Some publications state that $\mathbf{H}(S|AD) \leq -\log_2(\mathrm{FAR})$ which bound is derived assuming the Fuzzy Commitment scheme [JW99] where X is a perfectly random independent and identically distributed (i.i.d.) binary string [Pl07]. For some special choices of the entropy function (e.g. the average min-entropy $\hat{\mathbf{H}}_\infty$ [Ko08]), it has been shown that $\hat{\mathbf{H}}_\infty(S|AD)=\hat{\mathbf{H}}_\infty(X|AD)$ and $\hat{\mathbf{H}}_\infty(X|AD) \leq -\log_2(\mathrm{FAR})$ which holds for arbitrary distributions of the biometric strings X. Moreover, it is expected that similar bounds will hold for any KBTP system. The details of using entropy measures to estimate the required effort of a practical attack are a point of further research.

5 Summary

In this paper we presented a reference framework that can be used in assessing and comparing the privacy properties of KBTP systems. KBTP methods are a building block

in larger biometric systems and in the privacy assessment of KBTP systems it is important to differentiate between, on one hand, threats against the system and, on the other hand, specific threats against the KBTP method. This has lead to the concept of a Perfectly Private KBTP system and to the goal of KBTP systems to protect biometric information without the use of long-term secrets. In order to set up a generic framework, an abstraction of KBTP templates was taken from [ISO10] in the form (*AD*, *PI*). Based on this uniform format, after defining the adversary capabilities, attacks can be defined that affect the privacy goals as defined in [ISO10]. High level attacks are independent of the algorithmic details of the underlying KBPT method while low level attacks must be targeted at a specific KBTP method. The presented reference framework can be used as a first step to set up practical methods assess and compare the privacy properties of commercial KBTP systems.

Bibliography

[Bo06] T. Boult: Robust distance measures for face recognition supporting revocable biometric tokens, Proc. 7th Int. Conf. on Automatic Face and Gesture Recognition (FGR), IEEE Computer Society, Washington, DC, pp560-566, 2006.

[Br02] M. Braithwaite, U.C. von Seelen, J. Cambier, J. Daugman , R. Glass, R. Moore, and I. Scott: Application-specific biometric templates, IEEE Workshop on Automatic Identification Advanced Technologies, Tarrytown, NY, March 14-15, pp167-171, 2002.

[Br09] Jeroen Breebaart, Bian Yang, Ileana Buhan-Dulman, Christoph Busch: Biometric template protection, the need for open standards. Datenschutz und Datensicherheit - DuD, Vol. 33, No 5, May 2009, Vieweg Verlag, pp299-304.

[Bu08] Ileana Buhan: Cryptographic keys from noisy data, theory and applications, PhD Thesis, University of Twente, the Netherlands, 2008.

[CS07] Ann Cavoukian, Alex Stoianov: Biometric Encryption, A Positive-Sum Technology that Achieves Strong Authentication, Security AND Privacy, 2007, see http://www.ipc.on.ca/images/Resources/bio-encryp.pdf.

[DH76] W.Diffie, M.E.Hellman: New directions in cryptography. IEEE Trans. Inform. Theory, IT-22, 6, 1976, pp644-654.

[DRS04] Y. Dodis, M. Reyzin, and A. Smith: Fuzzy extractors: How to generate strong keys from biometrics and other noisy data, Proc. Eurocrypt 2004, Lecture Notes in Computer Science, Vol. 3027, pp523-540, Springer-Verlag, New York, 2004.

[Eu08] European Parliament and European Council: Directive 1995/46/EC of the European Parliament and of the Council of 24 October 1995 on the protection of individuals with regard to the processing of personal data and on the free movement of such data.

[Ig09] T. Ignatenko: Secret-Key Rates and Privacy Leakage in Biometric Systems, PhD thesis, Technical University of Eindhoven, 2009.

[ISO10] ISO/IEC JTC1 SC27 2nd CD 24745 – Information technology- Security techniques- Biometric template protection.

[JRP06] A.K. Jain, A.Ross, S. Pankanti: Biometrics: A Tool for Information Security. IEEE Trans. Information Forensics And Security, v.1, No.2, pp125-143, 2006.

[JW99] A. Juels and M. Wattenberg: A fuzzy commitment scheme, Sixth ACM Conf. on Computer and Communications Security, p28-36. ACM Press, N, 1999.

[JS02] A. Juels, M. Sudan: A fuzzy vault scheme. Proc. IEEE Int. Symp. on Information Theory, p408. IEEE Press, Lausanne, Switzerland, 2002.

[Ko08] Ulrike Korte, Michael Krawczak, Ullrich Martini, Johannes Merkle, Rainer Plaga, Matthias Niesing, Carsten Tiemann, Han Vinck: A cryptographic biometric authentication system based on genetic fingerprints, Proc. Sicherheit 2008, in Lecture Notes of Informatics, pp263-276, LNI P-128, Springer-Verlag, 2008.

[LT03] J.-P. Linnartz and P. Tuyls: New shielding functions to enhance privacy and prevent misuse of biometric templates. In Proc. of the 4th Int. Conf. on Audio and Video Based Biometric Person Authentication, pp393-402, Guildford, UK, 2003.

[LLO06] J. Lyseggen, R. A. Lauritzsen, and K. G. S. Oyhus: System, Portable device and method for digital authenticating, crypting and signing by generating short-lived cryptokeys, US Patent Application 2006/0198514 A1, Sep. 7, 2006.

[Ma06] M. Martinez-Diaz, J. Fierrez-Aguilar, F. Alonso-Fernandez, J. Ortega-Garcia and J. A. Siguenza, Hill-climbing and brute-force attacks on biometric systems: A case study in Match-on-Card fingerprint verification", Proc. IEEE Intl. Carnahan Conf. on Security Technology, ICCST, pp151-159, Lexington, USA, October 2006

[Pe65] L. Penrose: Dermatoglyphic topology. Nature, 205:545–546, 1965.

[Pl07] R. Plaga: Biometrics and cryptography - On biometric keys, their information content and proper use, Conference on Biometric Feature Identification and Analysis, Göttingen,7 September 2007.

[RCB01] N. K. Ratha, J. H. Connell, R. M. Bolle: Enhancing security and privacy in biometrics-based authentication systems. IBM Systems Journal, 40(3):614–634, 2001.

[SKK04] M. Savvides, B.V.K.Vijaya Kumar and P.K.Khosla: Cancelable biometric filters for face recognition, Proceedings of the 17th International Conference on Pattern Recognition (ICPR'04), Cambridge, England. v.3,pp922-925, 2004.

[So98] C. Soutar, D. Roberge, A.V. Stoianov, R. Gilroy, and B. V. K. Vijaya Kumar: Biometric Encryption using image processing, in Proc. SPIE, Optical Security and Counterfeit Deterrence Techniques II, v. 3314, pp178-188, 1998.

[TNG04] A. Teoh , D. Ngo, and A. Goh: Biohashing: two factor authentication featuring fingerprint data and tokenised random number, Pattern Recognition, v. 37, pp2245-2255, 2004.

[Ti02] M. Tiberg:A Method and a System for Biometric Identification or Verification, Swedish patent 0202147-5, PCT patent appl. WO 2004/006495, PCT/SE2003/001181. US Patent Application US2005/0210269 A1, Sep. 22, 2005.

[UPJ05] U. Uludag, S. Pankanti, A. K. Jain: Fuzzy vault for fingerprints, in Lecture Notes in Computer Science. Vol. 3546, pp270-319, Springer-Verlag, 2005.

[Wy09] B. Wyseur: White-Box Cryptography, PhD thesis, Catholic University of Leuven.

Performance of the Fuzzy Vault for Multiple Fingerprints (Short Version)*

Johannes Merkle, Matthias Niesing, and Michael Schwaiger

secunet Security Networks AG,
D-45128 Essen,
Germany

Heinrich Ihmor and Ulrike Korte

Bundesamt für Sicherheit in der Informationstechnik (BSI),
D-53175 Bonn,
Germany

Abstract: The fuzzy vault is an error tolerant authentication method that ensures the privacy of the stored reference data. Several publications have proposed the application of the fuzzy vault to fingerprints, but the results of subsequent analyses indicate that a single finger does not contain sufficient information for a secure implementation.

In this contribution, we present an implementation of a fuzzy vault based on minutiae information in several fingerprints aiming at a security level comparable to current cryptographic applications. We analyze and empirically evaluate the security, efficiency, and robustness of the construction and several optimizations. The results allow an assessment of the capacity of the scheme and an appropriate selection of parameters.

1 Introduction

Biometric authentication requires the storage of reference data for identity verification, either centrally (e.g., in a database) or locally (e.g., on the users token). However, the storage of biometric reference data poses considerable information security risks to the biometric application and concerns regarding data protection. As a potential solution to this dilemma, *biometric template protection* systems [BBGK08] use reference data which reveal only very limited information on the biometric trait. Another term frequently used for these schemes is *biometric encryption*. One of the most prominent approaches is the fuzzy vault scheme [JS02],

*This contribution represents a shortened version. Additional investigations, pseude codes and more detailed results are presented in the full paper [MIK+10b].

where the sensitive information (the biometric reference data) is hidden among random *chaff points*.

In [CKL03, UPJ05, UJ06, NJP07], the application of the fuzzy vault scheme to minutiae (*fuzzy fingerprint vault*) has been proposed. However, a subsequent analysis in [MMT09] has revealed that the parameters suggested do not provide security beyond 50 bit cryptographic keys. One of the suggestions in [MMT09] was to enhance security by using multiple fingers. This idea is supported by the observation in [Pla09] that a single fingerprint cannot provide enough entropy to implement a secure biometric template protection.

In this paper we present an implementation of a fuzzy vault based on the minutiae data of several fingerprints. We investigate the security and robustness of the scheme and of several optimizations applied, some of which have already been proposed in the previous constructions [CKL03, UPJ05, UJ06, NJP07]. In particular, we evaluate the impact of the basic parameters and optimizations to error rates, efficiency and security, and we derive suggestions for parameter selection.

This article is structured as follows: In Section 2, we specify the fuzzy multi-fingerprint vault and its optimizations and justify our design decisions. Section 3 assesses the security of the fuzzy fingerprint vault. In Section 4, we report the results obtained in evaluation with real fingerprints. Finally, in Section 5, we draw conclusions and identify open issues for future investigations.

2 Design of the scheme

2.1 Underlying biometric template protection scheme

In the fuzzy vault scheme [JS02], a polynomial is used to redundantly encode a set of (pairwise distinct) private attributes $m_1, \ldots, m_t$ (e.g., biometric feature data) using a variant of Reed-Solomon decoding. First, a random (secret) polynomial $P(z)$ over a finite field $\mathbf{F}_q$ with degree smaller than k is chosen. Then, each attribute m_i is encoded as element x_i of the finite field, i.e. $x_i = E(m_i)$, where E is an arbitrary injective map from the space of attributes to $\mathbf{F}_q$. Each of these elements x_i is evaluated over the polynomial, resulting in a list of (pairwise distinct) pairs $(x_i, y_i) \in \mathbf{F}_q^2$ with $y_i = P(x_i)$. In order to hide the private attributes, $r - t$ *chaff points* $x_{t+1}, \ldots, x_r \in \mathbf{F}_q$ are randomly selected so that $x_i \neq x_j$ for all $1 \leqslant i < j \leqslant r$. For each chaff point x_i, a random $y_i \in \mathbf{F}_q$ with $y_i \neq P(x_i)$ is chosen. The list of all pairs $(x_1, y_1), \ldots, (x_r, y_r)$, sorted in a predetermined order to conceal which points are genuine and which are the chaff points, is stored as the *vault*.

For authentication and recovery of the secret polynomial, another set of attributes (the query set) has to be presented. This set is compared with the stored fuzzy vault $(x_1, y_1), \ldots, (x_r, y_r)$, and those pairs (x_i, y_i) are selected for which x_i corresponds to an attribute in the query set. The selected points are then used to try to recover the secret polynomial using Reed-Solomon decoding.

If the number of genuine points among the identified correspondences (*correct matches*) is at least k, the secret polynomial can be recovered. However, if the set of correspondences also comprises chaff points (*false matches*), the number of correct matches must be greater than k, or the decoding must operate on subsets of the matches resulting in many trials. Details are given in Section 2.6.

In the original fuzzy vault scheme, correspondence between points in the query set and the fuzzy vault means equality (of the encodings in the finite field). However, for the application of the fuzzy vault to fingerprints the definition of this correspondence is usually adjusted to allow a compensation of noise in the measurement of the minutiae. Following the approach of [NJP07] and [UJ06], we define correspondence as mappings determined by a minutiae matching algorithm (see Section 2.3).

The fuzzy vault scheme is error tolerant with respect to the set difference metric, which covers exactly the errors introduced to (naively encoded) minutiae information by insertions, omissions, and permutation of minutiae. The deployment of a minutiae matching algorithm for identifying correspondences between the query set and the fuzzy vault adds robustness with respect to global rotations and translations or non-linear deformations of the fingerprint. Since the matching algorithms included in standard fingerprint software only output a match score and not the list of corresponding minutiae, we use our own matching algorithm (see Section 2.3).

2.2 Multi-biometric fusion

In order to obtain sufficient information for a secure scheme, we use the imprints of $f \geqslant 2$ fingers of each person. We implemented feature level fusion by encoding the minutiae of all fingers in one feature vector. In this vector each minutiae is encoded as a triplet (ℓ, a, b), where $\ell \in \{1, \ldots, f\}$ is an index of the finger on which the minutiae was detected, while a and b denote the Cartesian coordinates of the minutiae location in the fingerprint image. Chaff points are encoded analogously.

A justification for using feature level fusion is given in the full paper [MIK+10b].

2.3 Minutiae matching algorithm

We need to identify matching minutiae between fingerprints for enrollment and for verification: during enrollment minutiae matching is used to identify the most reliable minutiae from several measurements. During verification we have to identify a sufficiently large set of genuine minutiae within the vault to recover the secret polynomial.

The matching is performed for each finger separately by a simple matching algo-

rithm that identifies minutiae correspondences between two sets A and B of points (a, b) (minutiae or chaff points) in the fingerprint image. (We do not use minutiae orientation due to the intrinsic correspondences that render security assessment much more difficult.) The algorithm tries to maximize the number of correspondent points between the sets by finding a suitable global rotation and translation transformation T and tolerates (Euclidean) distances $||\cdot||_2$ between two points smaller than δ, where δ is a parameter of the algorithm.

A pseudo code description and more details are given in the full paper [MIK+10b].

The tolerance parameter δ varies: we use a greater value $\delta = \delta_e$ for enrollment than the value $\delta = \delta_V$ for verification to increase the number of reliable minutiae.

2.4 Optimizations

In this section we introduce several optimizations to the scheme. The impact of these optimizations are evaluated in Section 4.

2.4.1 Restriction of fingerprint area

If the sensor area is sufficiently large, minutiae rarely occur in the corners of the image. In order to ensure that the distribution of the randomly selected chaff points resembles that of genuine minutiae, we restrict both the chaff points and the minutiae considered for the vault to an area $\mathcal{M}$ with sufficiently high minutiae occurrence. A statistical evaluation of the minutiae positions of 82800 fingerprints having 500 DPI revealed that 7/8 of all minutiae occurred in an area defined by a centered ellipse that covers approximately 87000 pixels, which roughly corresponds to 2.25 cm^2. A figure visualizing the determined distribution of minutiae locations and the ellipse can be found in [MIK+10a]. Consequently, for the vault we choose minutiae and chaff points only from the set $\mathcal{M}$ composed of the union of these ellipses on the considered fingers.

2.4.2 Reliability filtering during enrollment

In order to minimize minutiae insertion and omission errors, we use only the most reliable minutiae for the feature vector. For this reason, we use multiple measurements during enrollment and consider only those minutiae in the feature vector that have been detected in all measurements. Details are given in Section 2.5.

2.4.3 Enforcing minimum distance

Due to the deviations in the measured minutiae locations, it can happen during authentication that a minutia in the query fingerprint is closer to a chaff point than to the corresponding minutiae in the vault. The frequency of such assignment errors

can drastically increase if the chaff points are selected too close to genuine minutiae. Therefore, we select the chaff points with a minimum distance d to the genuine minutiae (from the same finger) with respect to the Euclidean distance. Furthermore, in order to prevent that an adversary can exploit this minimum distance to distinguish chaff points from genuine minutiae, we also enforce the minimum distance among the minutiae and chaff points. In particular, if the Euclidean distance between two minutiae of the same finger is smaller than d, one of them is randomly disregarded, and chaff points are selected with minimum Euclidean distance d to all other chaff points and minutiae from this finger.

2.4.4 Quality filtering during authentication

On average, the number of *false matches* of minutiae with chaff points increases with the average number of surplus minutiae (i.e., minutiae not matching with real minutiae in the reference template) per query fingerprint. However, an increase of false matches requires stronger error correction by lowering the degree k of the secret polynomial, which decreases the security of the scheme.

In order to limit the average number s of surplus minutiae per query fingerprint, we filter the minutiae from the query fingerprint using the quality index value output by the minutiae extraction algorithm. Precisely, we define a minimum quality value Q and provide to the matching algorithm only those minutiae of the query fingerprint that have a quality value of at least Q. In our concrete implementation we used the MINDTCT algorithm of NIST [WGT+07] for minutiae extraction which outputs minutiae quality values in the range between 0 and 1.

2.4.5 Enforcement of minimum number of minutiae per finger

One of the main sources for failures during authentication is the difficulty to correctly align the query fingerprints with respect to the stored minutiae. This task is performed by the minutiae matching algorithm (described in Section 2.3) for each finger by identifying the isometry (rotation and translation) that maximizes the number of matches between the minutiae extracted from the query fingerprint and the points (representing minutiae or chaff point) stored in the reference data. However, this approach can only be successful if the reference template contains a sufficient number of minutiae of each finger; otherwise, i.e., if for one of the fingers the reference template contains only few minutiae, the number of wrong matches (with chaff points) resulting by chance from an incorrect isometry may be higher than the number of matches for the correct isometry. In practice, such cases can easily occur if one of the fingerprints captured during enrollment is of relatively poor quality.

For this reason, we require that the reference template computed during authentication contains at least χ minutiae from each finger, where χ is an additional parameter. Since this constraint reduces the number of possible reference templates its impact on security must be analyzed. We provide an estimation of this

reduction in Section 3.2.

2.5 Enrollment

Let $f \geqslant 2$ be the number of fingers used per person, q a prime power, $k < t < r \leqslant q$, and $\chi \leqslant t/f$.

For each user, a random polynomial P of degree less than k over the finite field $\mathbf{F}_q$ is selected. The coefficients of this polynomial represent the secret of the scheme. Then, for each finger u imprints are taken and the minutiae correspondences between these instances are identified using the matching algorithm with tolerance parameter $\delta = \delta_e$. Minutiae outside the considered set $\mathcal{M}$, i.e., with position outside the ellipse $\mathcal{E}$ on the respective finger, are neglected (see Section 2.4.1). Then, t of those minutiae that have been detected in all u imprints of the respective finger are selected at random so that at least χ minutiae are taken from each finger and each pair of chosen minutiae from the same finger has a minimum distance of d. This set T of t reliable minutiae can be considered the biometric template to be protected by the fuzzy vault scheme. The template T is amended with random chaff points, resulting in a set R of r points containing t genuine minutiae and $r-t$ chaff points, so that each point in R has a minimum distance of d to all other points on that finger. Furthermore, in order to ensure that minutiae and chaff points within the vault are not distinguishable by their order, they are lexicographically ordered.

In contrast to the original fuzzy vault scheme [JS02], the secret polynomial is redundantly encoded not by evaluating it on the biometric data itself but only on the minutiae's indexes in the ordered list. Precisely, for all $1 \leqslant i \leqslant r$ we (re-)define $x_i = E(i)$, where E is an injective embedding from the set $\{1, \ldots, r\} \subset \mathbf{Z}$ to $\mathbf{F}_q$. Further, we set $y_i = P(x_j)$, if $\mathbf{m}_i$ is a genuine minutia, and choose a random value $y_i \neq P(x_j)$, if $\mathbf{m}_i$ is a chaff point, where j is the index of $\mathbf{m}_i$ after applying the lexicographic order. This optimization allows a reduction of the field size to the range of r. The *vault* Y is given by the ordered list of minutiae and chaff points, paired with the corresponding y_j values. The vault and a hash value H of the polynomial's coefficients are stored in the database.

A pseudo code description of the enrollment is given in the full paper [MIK+10b].

2.6 Recovery of the polynomial

The unlocking of the vault (during authentication) requires the recovery of the secret polynomial P from a set of points (x_{j_i}, y_{j_i}), some of which (those resulting from *correct matches* with minutiae) lie on the polynomial, while others (resulting from *false matches* with chaff points) do not. For this task, a Reed-Solomon decoder RSDecode is used that receives as input a set of w points $(x_{j_1}, y_{j_1}), \ldots, (x_{j_w}, y_{j_w}) \in \mathbf{F}_q^2$ with $w \geqslant k$ and outputs $e_0, \ldots, e_{k-1} \in \{0, \ldots, q-1\}$,

so that $y_{j_i} = P(x_{j_i})$ holds for at least k of the (x_{j_i}, y_{j_i}) with $P(z) = \sum_{i=0}^{k-1} e_i z^i$, if such a polynomial exists. We assume that the Peterson-Berlekamp-Massey-decoder is used as suggested in [JS02]. This technique is successful, if at least $(w + k)/2$ of the w points handed over to the decoder are correct.

Evaluation reported in [MIK+10a] revealed that setting $w = 2m_c - k$ and $k \approx m_c - m_f$ can provide a good balance between efficient decoding and security, where m_c and m_f are the expected numbers of correct and false matches, respectively. However, if the match rate disperses considerably, it may by necessary to slightly deviate from this value, in order to reduce the False Rejection Rate. As we will see in Section 4.2, this is the case.

2.7 Authentication

We only implement an authentication in the verification scenario, where the identity of the (alleged) user is known a priori.

In order to verify the identity of a user, a query fingerprint is taken for each considered finger. The minutiae are extracted and matched with the minutiae and chaff points contained in the vault stored for the alleged user. (Thereby, the tolerance parameter δ_v used for the minutiae matching algorithm can differ from that used during enrollment.) The indices of those minutiae and chaff points in the vault matching with minutiae in the query fingerprint are identified; the encoded indices $x_i = E(i)$ along with the corresponding y_i values are given to RSDecode (see Section 2.6) to recover the secret polynomial P. If the number of genuine minutiae among these points is sufficiently high (see Section 2.6 for a discussion), the polynomial can be recovered. Finally, the correctness of the recovered polynomial is verified using the hash value stored in the database. Optionally, the coefficients of the recovered polynomial can be used for further cryptographic applications, e.g., as seed in a key derivation function.

A pseudo code description of the verification is given in the full paper [MIK+10b].

3 Security analysis

In this contribution we consider the security of the fuzzy vault for multiple fingerprints with respect to attacks that try to recover the minutiae or, equivalently, the secret polynomial from the vault. It is understood that there are other types of attacks against biometric template protection schemes to which the fuzzy vault is susceptible [SB07]. In particular, the cross matching of the vaults from several independent enrollments of a user represents a serious threat to the fuzzy vault. However, a comprehensive analysis of all potential attacks against the fuzzy vault would go beyond the scope of this paper.

3.1 Polynomial reconstruction attack

The most efficient method to recover the minutiae or the secret polynomial from the vault was published by Mihailescu [MMT09]. This brute force attack is designed to break the implementations of [CKL03] and [UJ06]; in the context of our scheme it is even slightly more efficient as the correctness of the recovered polynomial can be verified using the hash value of the secret coefficients and does not require additional evaluations of the polynomial. With this adaptation the attack systematically searches through all subsets $\{i_1, \ldots, i_k\}$ of $\{1, \ldots, r\}$, computes the unique polynomial P satisfying $P(E(i_j)) = y_{i_j}$ by Lagrange interpolation, and checks the correctness of this polynomial with the stored hash value. According to [MMT09], the number of trials needed is $1.1(r/t)^k$ and each trial requires $6.5k \log^2(k)$ arithmetic operations over $\mathbf{F}_q$. However, in the latter estimation an explicit constant of 18 for multiplication of the polynomials (see Corollary 8.19 in [GG03]) has been overlooked, and thus, we end up with a total number of approximately $129k \log^2(k)(r/t)^k$ arithmetic operations.

If the number of chaff points is close to the maximum possible, the attack described in [CST06] can be more efficient than brute force. The basic idea is that the free area around chaff points is smaller than around genuine minutia. Assuming a density 0.45 for random sphere packings [CKL03], the maximum number of chaff points per finger would be $0.45 \cdot 87000/V_d$, where V_d is the number of integer point in the sphere of radius d.

3.2 Entropy loss by the minimum number of minutiae per finger

Whereas the enforcement of a minimum number χ of minutiae per finger (see Section 2.4.5) aims at reducing the false rejection rate it also decreases the security of the scheme by narrowing the set of possible templates. This applies to the lower bound on attacks according to [MIK+10a] as well as to the practical attack of [MMT09]. The subsequent analysis quantifies this reduction of security.

We will assume that the minutiae chosen are independently and uniformly distributed among the F fingers. This assumption can be fulfilled by a suitable probabilistic selection method of the template T from the set of reliable minutiae during enrollment.

Using this assumption and the inclusion-exclusion-principle, we can estimate the probability $\zeta(t, \chi)$ that a template with t minutiae includes for each finger f at least χ minutiae by

$$\zeta(t,\chi) = 1 - f^{-t} \sum_{\ell=1}^{f} (-1)^{\ell} \binom{f}{\ell} \cdot \sum_{i_1,\ldots,i_\ell=0}^{\chi} \binom{t}{i_1, \ldots i_\ell, t - \sum_j i_j} (f-\ell)^{t-\sum_j i_j},$$

where $\binom{a}{b_1,\ldots,b_m}$ with $b_1 + \cdots + b_m = a$ denotes the multinomial coefficient.

On the other hand, the conditional probability p that a particular instance of a template T is chosen, if a minimum number of χ minutiae per finger is enforced, can be calculated from the probability p' that this instance is chosen, if no minimum number of minutiae per finger is enforced, by the equation $p = p'/\zeta(t, \chi)$. Therefore, the search space for an attack is narrowed by the factor $\zeta(t, \chi)$, and consequently, the best known attack could be adapted to require at most $129\zeta(t, \chi)k \log^2(k)(r/t)^k$ operations.

4 Results

In this section, we summarize the results of empirical parameter evaluations, the impact of our optimizations and the general performance of the scheme.

We used a test set of 864 fingerprints taken from 18 persons in the course of this research using an optical sensor, each person providing 6 imprints of 8 fingers (little fingers were excluded). In our experiments, we used 6 or all 8 fingers per person (without or with thumbs), but results referring to single fingers were averaged over all finger types.

For minutiae extraction, we used the MINDTCT algorithm of NIST [WGT+07]. We stress that other feature extraction algorithms may exhibit a different performance, and therefore, the resulting statistics may deviate from ours.

4.1 Size of feature vector

First, we determined how large the feature vector can be in dependence of the number u of measurements and the tolerance parameter δ_e used during enrollment. We did this by evaluating the number of minutiae per finger that are reliably (i.e., u times) detected in u measurements. Since this number varies considerably among individuals and measurements, acceptable Failure To Enroll (FTE) rates can only be achieved, if the required number of reliable minutiae is considerably lower than its average value. Therefore, we evaluated the maximum number M_r of reliable minutiae that is achieved in at least 80% of all measurements. The results of this evaluation are listed in Table 1.

4.2 Minutiae matching rates

In order to configure the error correction capabilities of our scheme appropriately, it is necessary to determine the rate at which the genuine minutiae in the vault are identified during authentication. For various tolerance parameters $\delta_v = \delta_e$, we computed the biometric template set T, containing t minutiae reliably detected in

Table 1: Number M_r of reliable minutiae per finger that is found in 80% of all measurements.

u	$\delta_e = 5$	$\delta_e = 7$	$\delta_e = 10$	$\delta_e = 15$
1	63	63	63	63
2	23	32	39	43
3	18	24	31	35
4	9	16	22	27
5	6	9	15	18

u measurements and matched them with the minutiae of an (independent) query fingerprint using our matching algorithm. We did not add chaff points to the template T. The average match rate, i.e. the average ratio between the number of matches found and t, are given in Table 2.

Similarly to the number of reliable minutiae, the match rate varies considerably between different measurements. Moreover, in the presence of chaff points, the match rates slightly decrease depending on the expected number of false matches (with chaff points), as the chaff points render the correct mapping of the minutiae more difficult for the matching algorithm. (This aspect is further discussed in Section 4.5.) Therefore, a reasonably small FRR can only be achieved if k is selected slightly smaller than the expected value of $m_c - m_f$ (see Section 2.6). Our empirical evaluation suggests to set k 10%-20% smaller than this value.

For $2 \leqslant u \leqslant 4$, we obtain good match rates at a reasonable number of minutiae. Therefore, we will subsequently focus on these cases.

4.3 Effect of quality filtering during verification

As argued in Section 2.4.4, quality filtering of the minutiae in the query fingerprints aims to reduce the number of surplus minutiae, i.e., minutiae in the query fingerprints that do not match with genuine minutiae in T. We evaluated the effectiveness and eligible configuration of the filtering based on the minutiae quality

Table 2: Average match rate (in percentage) in the absence of chaff points for $\delta_e = \delta_v$.

u	$\delta_v = 5$	$\delta_v = 7$	$\delta_v = 10$	$\delta_v = 15$
1	40	50	58	64
2	66	72	78	82
3	75	81	84	87
4	81	85	88	90
5	85	89	91	92

Table 3: Expected number of minutiae in a query fingerprint after filtering with minimum quality value Q.

Minimum quality Q	Av. no. τ of minutiae
0	70
0.1	67
0.2	52
0.3	48
0.4	41
0.5	33
0.6	32

values output by the MINDTCT algorithm of NIST [WGT+07].

The average number of minutiae detected in a single fingerprint depends on the sensor used, the feature extractor algorithms, the quality of the images, and even the finger type (e.g. thumbs contain more minutiae than other fingers). In our tests, we detected an average number of 84 minutiae per fingerprint (excluding thumbs) inside ellipse $\mathcal{E}$. Based on this number and the distribution of quality values, we can estimate the expected number τ of minutiae in a query fingerprint after filtering with minimum quality value Q. The results are listed in Table 3.

The reduced average number τ of minutiae per query finger given to the matching algorithm results in a decreased average number s of surplus minutiae per finger and in less false matches (i.e., matches with chaff points). On the other hand, it may also reduce the number of correct matches (and likewise the match rate) because the minutiae filtered out could have matched with genuine minutiae in the vault. For different sets of parameters we empirically determined the decrease of the number of correct and false matches resulting from the quality filtering. An example plot is presented in Figure 1.

For larger r and smaller δ_v, quality filtering with higher values of Q results in a more drastic reduction of the correct matches. Nevertheless, for various parameters we consistently found a value Q between 0.2 and 0.3 to be optimal, reducing the false matches by approximately 30% while decreasing the number of correct matches by less than 3%.

4.4 Effect of minimum number of minutiae per finger

If the tolerance parameter δ_v is set appropriately as described in Section 4.5, the number of correct matches typically exceeds the number of false matches. On the other hand, if the matching algorithm fails to identify the correct isometry, the number of correct matches is typically significantly lower than the number of false matches. As explained in Section 2.4.5, the enforcement of a minimum number χ of

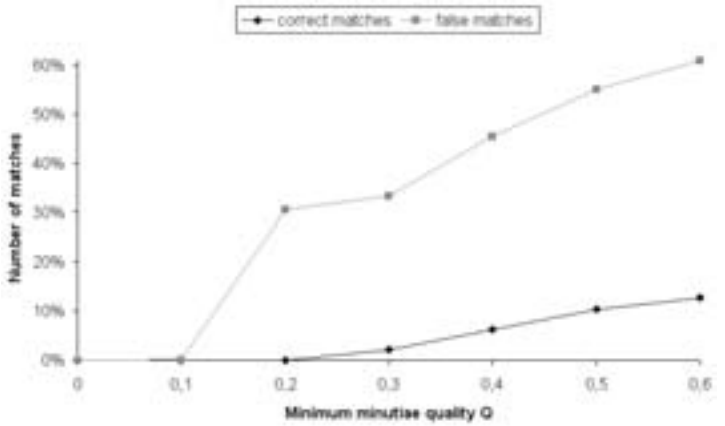

Figure 1: Reduction of the correct and false matches by quality filtering for $f = 6$, $t = 120$, $r = 400$ and $\delta_v = 7$.

minutiae per finger in the template T aims at reducing the frequency of such cases. We evaluated the effectiveness and reasonable configuration of this optimization by determining the ratio of fingers for which the number of false matches exceeded the number of correct matches for various values of the parameter χ. Furthermore, we analyzed the influence of this optimization to the FTE by determining the rate at which a finger contained at least χ minutiae and, hence, would succeeded to enroll. The results of this evaluation are displayed in Figure 2 by the curves of the match rate and the rate of successful enrollment. (Other failures of enrollment, particularly cases, where the fingers of a person contained less than t minutiae in total, were neglected.) Obviously, $\chi = 9$ already yields a considerable improvement with only moderately increased FTE rates.

The impact of the value of χ becomes particularly strong as the average number of false matches approaches the number of correct matches. As shown in Figure 3 for $f = 6$, $u = 4$, $t = 100$, $r = 600$, $\delta_e = 10$, $\delta_v = 7$ and $Q = 0.3$, where even for $\chi = 15$ the fraction between the average numbers of correct and false matches was 2.1 (as opposed to a fraction of 2.9 for the parameters of Figure 2), the average match rate steadily and considerably increases until $\chi = 15$. The decrease of the successful enrollment rate is similar to the case of Figure 2. This finding indicates that in these cases it may be worth to choose χ larger than 9 at the cost of higher FTE rates.

4.5 Balancing correct and false matches

In order to enable the minutiae matching algorithm to determine the correct isometry by which the query fingerprint is correctly aligned to the minutiae in the

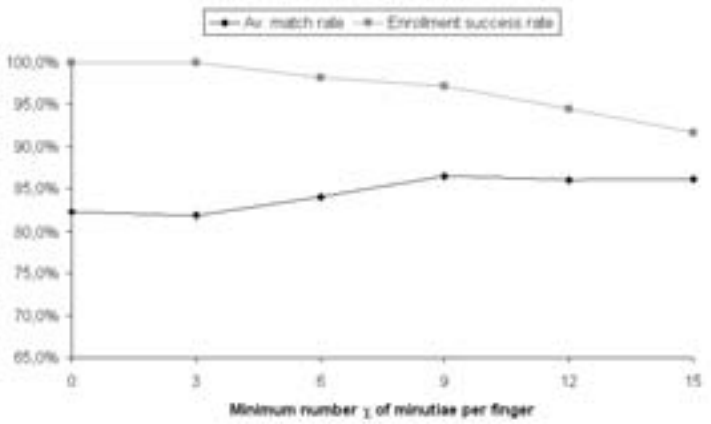

Figure 2: Impact of enforcing a minimum number χ of minutiae per finger in T to match rate and rate of successful enrollment for $f = 6$, $u = 4$, $t = 120$, $r = 400$, $\delta_e = 10$, $\delta_v = 7$ and $Q = 0$.

vault, we must ensure that, on average, the number of correct matches considerably exceeds the number of false matches. The results of Section 4.4 indicate that a fraction of 2 between the average numbers of correct and false matches already requires large values for χ which considerably increases the FTE.

In [MIK+10a], the expected number m_f of false matches is estimated by $(r - t)sV_{\delta_v}/|\mathcal{E}|$, where $V_\delta = 1 + 4\sum_{i=1}^{\lceil\delta-1\rceil}\left\lceil\sqrt{\delta^2 - i^2}\right\rceil$ is the number of integer points in the 2-dimensional plane with Euclidean norm smaller than δ and s is the average number of surplus minutiae (i.e., minutiae not matching with genuine minutiae) per query fingerprint. On the other hand, we can estimate $s \approx \tau - \mu t/f$, where τ is the number of minutiae of the query fingerprint after quality filtering.

Our experiments show that for typical parameters the average number of false matches is 20%-60% larger than these estimations imply, depending on the specific parameters. The deviation is presumably due to those outliers resulting from an incorrect determination of the isometry: if the matching algorithm is unable to detect the correct alignment, its optimization strategy with respect to the number of matches will yield extraordinary many false matches. Based on this observation, we adjust our above estimation to

$$m_f \approx 1.4(r - t)(\tau - \mu t/f)V_{\delta_v}/|\mathcal{E}|. \tag{1}$$

Yet, we expect the number of false matches to grow linearly with V_{δ_v}, which is a quadratic function in δ_v.

On the other hand, the average number m_c of correct matches is given by μt, where μ is the match rate, and therefore, grows slowly with increasing δ_v as shown in Table 2. Therefore, the selection of δ_v should carefully balance the expected numbers of correct and false matches. For $f = 3$, $u = 4$, $t = 66$, $r = 320$ and

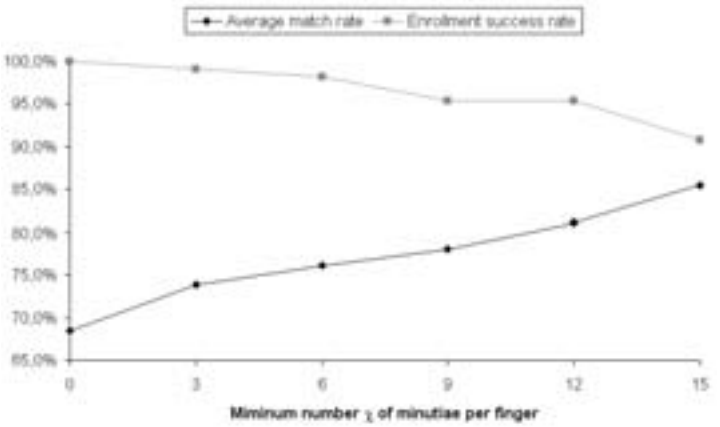

Figure 3: Impact of enforcing a minimum number χ of minutiae per finger in T to match rate and rate of successful enrollment for $f = 6$, $u = 4$, $t = 100$, $r = 600$, $\delta_e = 10$, $\delta_v = 7$ and $Q = 0.3$.

$Q = 0.3$, and for $5 \leqslant \delta_v \leqslant 15$ we estimated the number of false matches by (1) and the number of correct matches as μt using the match rates empirically determined. The results show that, for these parameters, $\delta_v \leqslant 8$ should be selected to ensure that the average number of correct matches is at least twice the number of correct matches.

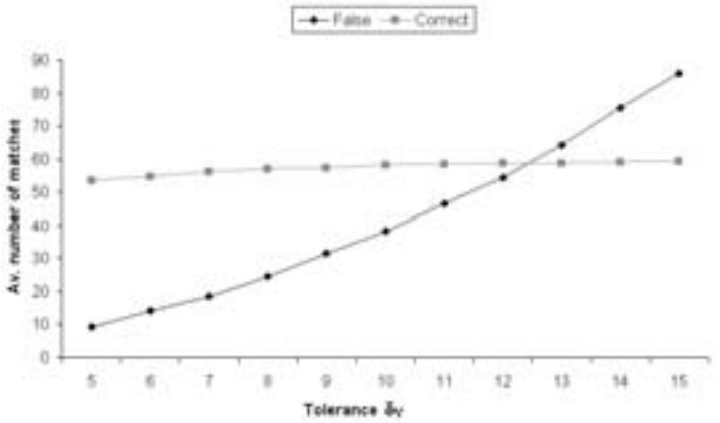

Figure 4: Impact of the tolerance parameter δ_v on the estimated average number of correct and false matches for $f = 3$, $u = 4$, $t = 66$, $r = 320$ and $Q = 0,3$.

For a smaller ratio r/t, the curves meet at higher values of δ_v, but still the accelerating growth of the number of false matches implies that $\delta_v \leqslant 8$ is a good choice.

Table 4: Parameters for a security level of 2^{Sec}.

f	u	δ_e	δ_v	t	r	k	Sec
2	2	7	5	26	240	27	70
3	2	7	5	90	351	41	97
3	3	7	5	70	360	34	97

4.6 Achievable Security

Based on our experiences gained, the example parameters listed in Table 4 have been determined to provide the indicated security level against existing attacks. We did not experimentally determine real error rates during enrollment and verification; therefore, these parameters are mere suggestions which require practical validation. We set $d = \lfloor 3/2 \cdot \delta_v \rfloor$, $Q = 0.3$, and $\chi = 9$. Furthermore, we choose $r < \lfloor 0.2 \cdot 87000/V_d \rfloor$ to avoid the attack described in [CST06] (see Section 3.1) that could significantly reduce security.

5 Conclusions

Our analysis shows that a fuzzy vault for multiple fingerprints can be very secure against template recovery from the helper data, if appropriate optimizations are applied. Filtering minutiae for reliability during enrollment and for quality during verification turn out to be particularly effective. Furthermore, enforcing a minimum number of minutiae per finger in the template significantly increases matching performance. Both optimizations are very sensitive to the respective thresholds, which must be carefully set on the basis of empirical data.

Finally, we would like to stress that our security analysis only covered template recovery attacks. Other types of attacks have been published [SB07] and need to be addressed before the scheme can be considered ready for use. We encourage research on methods to harden the fuzzy fingerprint vault against these attacks.

Acknowledgments

This work was conducted as part of the projects “BioKeyS-Multi” and “BioKeyS Pilot-DB” of the Bundesamt für Sicherheit in der Informationstechnik.

The matching algorithm was designed and implemented by Stefan Schürmans.

References

[BBGK08] Jeroen Breebaart, Christoph Busch, Justine Grave, and Els Kindt. A Reference Architecture for Biometric Template Protection based on Pseudo Identities. In *Proc. BIOSIG 2008*, volume 137 of *LNI*, pages 25–38. Gesellschaft für Informatik, 2008.

[CKL03] Charles Clancy, Negar Kiyavash, and Dennis Lin. Secure smartcardbased fingerprint authentication. In *WBMA '03: ACM SIGMM workshop on Biometrics methods and applications*, pages 45–52, 2003.

[CST06] Ee-Chien Chang, Ren Shen, and Francis Weijian Teo. Finding the original point set hidden among chaff. In *Proc. ACM Symp. on Information, Computer & Communication Security (ASIACCS)*, pages 182–188, 2006.

[GG03] Joachim von zur Gathen and J$^{'}$urgen Gerhard. *Modern Computer Algebra*. Cambridge University Press, 2nd edition edition, 2003.

[JS02] Ari Juels and Madhu Sudan. A Fuzzy Vault Scheme. In *Proc. IEEE Int. Symp. Inf. Theory*, page 408, 2002.

[MIK+10a] J. Merkle, H. Ihmor, U. Korte, M. Niesing, and M. Schwaiger. Provable Security for the Fuzzy Fingerprint Vault. In *Proc. 5th Int. Conf. Internet Monitoring and Protection (ICIMP 2010)*, pages 65–73, 2010.

[MIK+10b] Johannes Merkle, Heinrich Ihmor, Ulrike Korte, Matthias Niesing, and Michael Schwaiger. Performance of the Fuzzy Vault for Multiple Fingerprints (Extended Version). eprint arXiv:1008.0807, 2010.

[MMT09] Preda Mihailescu, Axel Munk, and Benjamin Tams. The Fuzzy Vault for fingerprints is Vulnerable to Brute Force Attack. In *Proc. BIOSIG 2009*, volume 155 of *LNI*, pages 43–54. Gesellschaft für Informatik, 2009.

[NJP07] Karthik Nandakumar, Anil Jain, and Sharath Pankanti. Fingerprint-Based Fuzzy Vault: Implementation and Performance. *IEEE Trans.Inf. Forensics Security*, 2(4):744–757, 2007.

[Pla09] Rainer Plaga. Biometric keys: Suitable Uses and Achievable Information Content. *Int. J. Inf. Secur.*, 8(6):447–454, 2009.

[SB07] Walter J. Scheirer and Terrance E. Boult. CRACKING Fuzzy Vaults and BIOMETRIC ENCRYPTION. In *Proc. IEEE Biometrics Symposium*, pages 1–6, 2007.

[UJ06] Umut Uludag and Anil Jain. Securing Fingerprint Template: Fuzzy Vault with Helper Data. In *IEEE Workshop on Privacy Research In Vision*, pages 163–169, 2006.

[UPJ05] Umut Uludag, Sharath Pankanti, and Anil Jain. Fuzzy Vault for Fingerprints. In *Proc. Int. Conf. Audio- and Video-Based Biometric Person Authentication (AVBPA)*, volume 3546 of *LNCS*, pages 310–319. Springer, 2005.

[WGT+07] Craig Watson, Michael Garris, Elham Tabassi, Charles Wilson, Michael McCabe, Stanley Janet, and Kenneth Ko. *User's Guide to NIST Biometric Image Software (NBIS)*. National Institute of Standards and Technology, 2007.

Template Protection for Biometric Gait Data

Claudia Nickel[1], Xuebing Zhou[2], Christoph Busch[1]

[1] Hochschule Darmstadt - CASED, [2] Fraunhofer IGD
c.nickel@fbi.h-da.de, xuebing.zhou@igd.fraunhofer.de, christoph.busch@h-da.de

Abstract: Biometric gait recognition is a well suited method for authentication on mobile devices as it is unobtrusive and concurrent. Hence, in contrast to PIN authentication it is no extra-effort for the user. The characteristic gait of a subject can be recorded using accelerometers which are nowadays already contained in many mobile devices. From this data biometric feature vectors can be extracted and stored as reference data on the device. Only if the user is not recognized by his walk an active authentication via PIN is necessary.
As the number of attacks on mobile devices increases it cannot be assumed that the data stored on the device is under constant control of the subject. Therefore, template protection techniques should be applied to secure biometric data. As biometric gait recognition is a new field of research no specific template protection methods have been developed so far. This paper describes a new method for securing biometric gait features based on histograms and using the earth mover's distance for comparison. The method is tested with gait data of 48 subjects recorded using a mobile phone and the results are compared to the ones obtained without template protection.

1 Introduction

A survey by Furnell and Clarke [CF05] shows that data in mobile devices is often insufficiently protected. When turning on the phone, entering a PIN is only necessary in 66% of the cases and after a stand-by phase this is only required at 18% of the devices (either because the phone does not offer this setting or because the owner did not select it). This implies that in most cases everybody who has physical access to the device can directly access all stored information. As the proportion of sensitive information (contacts, emails, ...) saved in mobile devices grows, this is becoming critical. 30% of the respondents of the survey consider PIN authentication to be inconvenient. But most mobile devices do not offer a suitable alternative. Accelerometer-based gait recognition is such an alternative. In contrast to PIN authentication no active input of the user is necessary. Most smartphones do contain accelerometers for games or changing the orientation of the display. These accelerometers can directly be used to record the specific gait of a subject. This means that no special hardware is needed to collect the gait data which is a great advantage to other biometric modalities like fingerprint. When a subject is walking with his phone he is directly authenticated based on his gait. Recently, Gafurov et al. [GS09, GHS06] and Ailisto et al. [ALM$^+$05] have suggested methods for extracting feature vectors from accelerometer data. Using data collected with dedicated accelerometers (i.e. not accelerometers

contained in mobile devices) they report equal error rates up to 6.4%.
While biometric identification and authentication provides considerable convenience and also some security benefits over token- or password-based methods, other security and privacy concerns unique to biometrics must be taken into account. These include identity theft, cross-matching, and the exposure, often irrevocable, of sensitive private information, as well as traceability of individuals.
This has stimulated research on the protection of stored biometric data in recent years, primarily focusing on preventing information leakage. Template protection techniques, also referred to as biometric encryption, untraceable biometrics, cancelable or revocable biometrics, have been developed. These convert biometric data elements into multiple (ideally) uncorrelated references, from which it is infeasible to retrieve the original information and in some cases have already been integrated into existing systems [gen, pri]. [ZWBK09] gives an overview and security analysis of existing template protection techniques, which have been already developed for different modalities like finger [UJ04, RCCB07], face [VKjS+06, Zho07], iris [WHNB08] and vision based gait recognition [ATIS09]. Template protection is a generalized and efficient method to preserve privacy and to enhance security of biometric data by limiting the exposure of template data which cannot be revoked. They exhibit the following key properties:

One-Way and Robustness A secure reference can be computed efficiently from a biometric datum (template) while it is either computationally hard or impossible to deduce the template from such a reference. The derivative references can be compared to a biometric datum under similarity metrics for the underlying biometric template. This allows the successful comparison of measurements exhibiting small variations or measurement errors to a derivative reference.

Diversity and Randomness Template protection can create numerous secure references from one biometric feature with the references independent on each other, i.e. knowledge of one reference does not yield information on other references derived from the same template. This eliminates the problem of cross-matching and tracebility.

The resulting various references are also called pseudo identifiers [BBGK08]. Different methods to protect the biometric data exist, which can be classified into four categories: cancelable biometrics, biometric salting, fuzzy encryption and biometric hardening passwords.
Although the research on accelerometer based biometric gait recognition shows that it offers a promising way to provide a more convenient method for authentication on mobile devices, no research has been done so far in the area of template protection for biometric gait data collected using accelerometers. One reason for this might be, that biometric data stored on the mobile device seems to be under control of the subject (similar to systems using on-card biometric comparison, see [iso]). Nevertheless several attacks on mobile devices have been reported [HJO08, Win] which make clear that data stored on the mobile devices should be protected.
The paper is structured as follows. The next section gives an overview over the collected gait data and the extracted feature vectors. Section 3 describes the developed template protection method and section 4 explains the test and states the obtained results. A summary

and conclusions are given in section 5.

2 Biometric Gait Data

As our focus on the application of the proposed technique is the protection of biometric gait data collected with and stored on mobile devices, a database containing suitable test data had to be created. Our database consists of data of 48 subjects collected using a standard mobile device (T-Mobile G1) which contains accelerometers for measuring acceleration in three directions as indicated in figure 1. The phone was carried in a bag which was attached to the belt of the walking subject at the right hip (see figure 1). All subjects had

Figure 1: Phone attached to subject and the three axes in which acceleration is measured.

to walk about 37 meters down the hall, wait for a short time, turn around, wait again and walk back. Subjects were told to walk at a normal speed, which results in different walking speeds for different subjects depending on what is assumed to be normal. An example of the accelerometer data recorded during one of these sessions is given in figure 2.

From this data, the parts where the subject is walking (called *walk* and indicated by the dashed lines in figure 2) have been extracted. As two sets of data have been collected from each subject at two different days, four walks of each subject are available for tests. The first walk is used to compute the reference data, the further three walks provide the probes. The movements of a subject approximately repeat every two steps (see figure 3), which suggests the extraction of these *cycles* and determination of a cycle which is typical for a specific subject and hence can be used as feature vector. As most of the gait recognition methods proposed in literature are based on a cycle extraction method (e.g. [GS09, GHS06, ALM+05]), our proposed template protection method should be applicable to these features (cycles). Therefore, to get appropriate test data, cycles have been manually extracted from the collected data. The most suitable cycles of each subject have been selected by determining the reference cycle (from the first walk) and one probe cycle of each of the other walks in such a way that sum of the distance from the reference to the probes is minimal. The data is interpolated to a fixed sampling rate (100 Hz) and

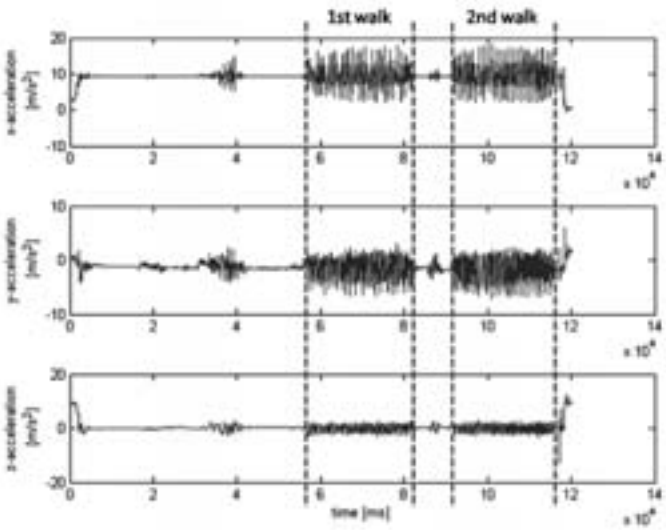

Figure 2: Recorded accelerometer data.

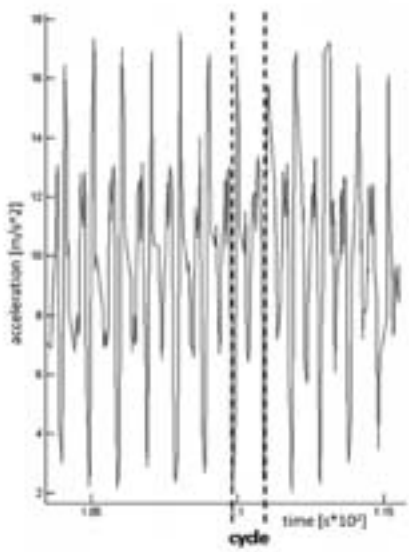

Figure 3: Example of gait data. The cyclic repetition can be clearly seen.

centered around zero. After these preprocessing steps our data base consisted of 192 cycles (feature vectors) of length between 62 and 138 samples.

3 Horizontal Projection

Biometric template protection techniques are used to create pseudo identifiers from the biometric template. This pseudo identifier should only reveal limited information about the original biometric characteristic and in addition the generation of different pseudo identifiers from the same biometric template should be possible. To transform the template in a way such that no private information can be obtained, one-way functions can be used. We propose using a horizontal projection of the values in the feature vector to the y-axis as one-way function which is the same as calculating the histogram of the accelerometer data. Figures 4 and 5 show examplary four cycles of two different subjects and the obtained histograms. Crosses on the cycles show the accelerometer values contained in the feature vector. Only these values contribute to the histogram. The direction of projection is indicated at the upper left cycle.

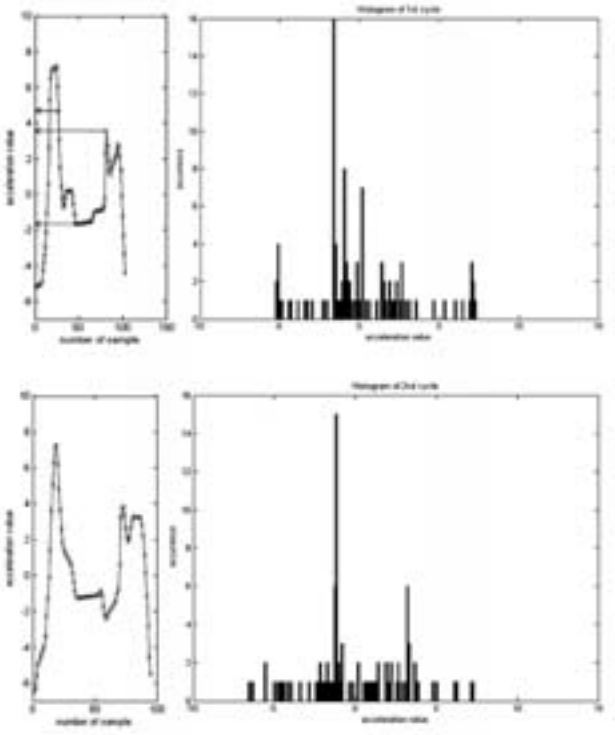

Figure 4: Two different cycles and corresponding histograms of subject 1.

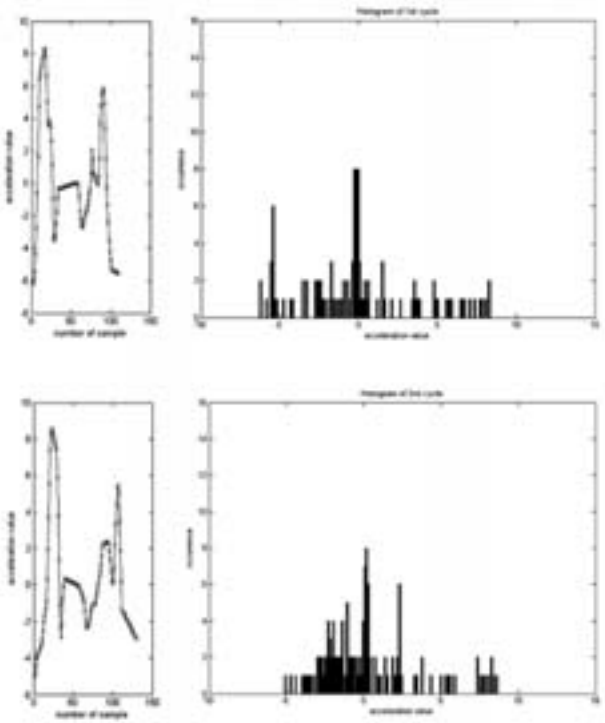

Figure 5: Two different cycles and corresponding histograms of subject 2.

The figures already indicate that the histograms obtained from cycles of the same subject are more similar than histograms of different subjects. This assumption has been tested using the data described in section 2. Test and results are given in the following section. The histogram generation removes private information (e.g. about health conditions) from the template and makes it impossible to reconstruct the original gait cycle. The histogram reveals information about the minimal and maximal value of the gait cycle (corresponding to the first and last non-zero bin), but furthermore no relevant information about the shape of the cycle can be obtained. To give the possibility of creating different protected templates from one biometric feature, it is possible to permutate the bins of the histogram. This permutation can be chosen differently for different applications to prevent cross matching but it has to be considered when calculating the distance between the histograms.

4 Tests and Results

Cycles obtained as described in section 2 are used to test the proposed method. From each cycle a histogram was created. Positions of the used bins have been the same for all cycles.

200 bins were used as this resulted in best recognition rates, but varying the number of bins did not have significant influence. This resulted in protected templates of length 200.
For comparision of the templates the earth mover's distance (EMD) [RTG98] returned the best results, which is a well known distance for comparing histograms. To illustrate this metric, bins of one of the histograms are assumed to be piles of soil and the bins of the second histogram are assumed to correspond to holes in which the soil should be filled. EMD measures the minimum cost needed to fill the holes with the soil, where the cost is the amount of soil transported times the distance by which it is moved. In our basic case the distance between bins at the same position will be zero, neighboured bins have distance one and so on. When the bins have been permuted, the distance matrix has to reflect this permutation.
The obtained results are given as Detection Error Trade-off curve (DET curve) by the dotted line in figure 6. The result is compared with the one obtained without template

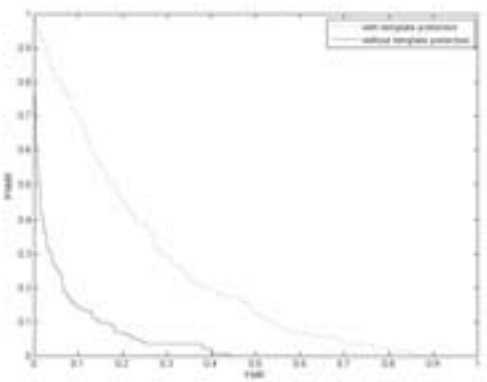

Figure 6: DET-curves obtained with and without template protection.

protection indicated by the continuous line in figure 6. The distance used for comparison in that case is dynamic time warping (DTW) [MÖ07]. Without using template protection the equal error rate (EER) is 12.85%. By using the proposed template protection method this increases to 29.47%. The reason for this will be the loss of information introduced by the histogram calculation, due to which no temporal information remains.

5 Summary and Conclusion

Having in mind the increasing amount of sensitive information stored on mobile devices and the increasing number of attacks on those devices, the need for secure, user-friendly authentication methods becomes clear. Accelerometer based gait recognition is such a method as it is able to authenticate a subject unobtrusively whithout his intervention. Up to now, no publications about template protection for accelerometer based gait recognition exist. This paper proposes a template protection method for cycle-based gait recognition

techniques, as this is the mainly chosen approach applied in existing feature extraction methods.
The feature vectors are converted into protected templates via histogram generation. Diversibility is obtained by applying different permutations to the template for different applications. The resulting templates are compared using the earth mover's distance. This technique does increase the EER significantly from 12.85% to 29.47% which indicates that the temporal distribution of acceleration values, which gets lost by histogram computation, does contain major information. Future work will focus on developing template protection methods which keep this information to guarantee a lower EER.

6 Acknowledgement

This work was supported by CASED (www.cased.de).

References

[ALM+05] Heikki J. Ailisto, Mikko Lindholm, Jani Mäntyjärvi, Elena Vildjiounaite, and Satu-Marja Mäkelä. Identifying people from gait pattern with accelerometers. *Biometric Technology for Human Identification II*, 5779(1):7–14, 2005. VTT Electronics, Finland.

[ATIS09] Savvas Argyropoulos, Dimitrios Tzovaras, Dimosthenis Ioannidis, and Michael G. Strintzis. A channel coding approach for human authentication from gait sequences. *Transactions on Information Forensics and Security*, 4(3):428–440, 2009.

[BBGK08] Jeroen Breebaart, Christoph Busch, Justine Grave, and Els Kindt. A reference architecture for biometric template protection based on pseudo identities. In *BIOSIG 2008: Biometrics and Electronic Signatures*, 2008.

[CF05] N.L. Clarke and S.M. Furnell. Authentication of users on mobile telephones - A survey of attitudes and practices. *Computers & Security*, 24(7):519 – 527, 2005.

[gen] http://genkeycorp.com/.

[GHS06] Davrondzhon Gafurov, Kirsi Helkala, and Torkjel Søndrol. Biometric Gait Authentication Using Accelerometer Sensor. *Journal of Computers*, 1(7), 2006.

[GS09] Davrondzhon Gafurov and Einar Snekkenes. Gait recognition using wearable motion recording sensors. *EURASIP J. Adv. Signal Process*, 2009:1–16, 2009.

[HJO08] S.M. Habib, C. Jacob, and T. Olovsson. A practical analysis of the robustness and stability of the network stack in smartphones. In *Computer and Information Technology, 2008. ICCIT 2008. 11th International Conference on*, pages 393 –398, 24-27 2008.

[iso] ISO/IEC FCD 24787: Information technology – Identification cards – On-card biometric comparison.

[MÖ7] Meinard Müller. *Information Retrieval for Music and Motion*, chapter 4 - Dynamic Time Warping. Springer-Verlag New York, Inc., Secaucus, NJ, USA, 2007.

[pri] http://www.priv-id.com/.

[RCCB07] Nalini K. Ratha, Sharat Chikkerur, Jonathan H. Connell, and Ruud M. Bolle. Generating Cancelable Fingerprint Templates. In *IEEE Transactions on Pattern Analysis and Machine Intelligence*, volume 29, April 2007.

[RTG98] Y. Rubner, C. Tomasi, and L.J. Guibas. A metric for distributions with applications to image databases. In *Computer Vision, 1998. Sixth International Conference on*, pages 59 –66, 4-7 1998.

[UJ04] U. Uludag and A. Jain. Fuzzy fingerprint vault. In *Workshop: Biometrics: Challenges Arising from Theory to Practice*, August 2004.

[VKjS+06] Michiel Van Der Veen, Tom Kevenaar, Geert jan Schrijen, Ton H. Akkermans, Fei Zuo, and Prof Holstlaan. Face Biometrics with Renewable Templates. In Ping Wah Wong Edward J. Delp III, editor, *Proceedings of SPIE: Security, Steganography, and Watermarking of Multimedia Contents VIII*, volume 6072, 2006.

[WHNB08] Zhifang Wang, Qi Han, Xiamu Niu, and Christoph Busch. A Novel Template Protection Algorithm for Iris Recognition. *Intelligent Systems Design and Applications, International Conference on*, 2:340–345, 2008.

[Win] WinCE/Infojack. Windows Mobile trojan sends unauthorized information and leaves device vulnerable. online, last visited May 2010. http://www.avertlabs.com/research/blog/index.php/2008/02/26/windows-mobile-trojan-sends-unauthorized-information-and-leaves-devicevulnerable/.

[Zho07] Xuebing Zhou. Template Protection and its Implementation in 3D Face Recognition Systems. In *in Proceedings of SPIE Conference on Biometric Technology for Human Identification*, pages 214–225, 2007.

[ZWBK09] Xuebing Zhou, Stephen Wolthusen, Christoph Busch, and Arjan Kuijper. A Security Analysis of Biometric Template Protection Schemes. In *International Conference on Image Analysis and Recognition (ICIAR09)*, pages 429–438, 2009.

Biometric systems in future preventive Scenarios – Legal Issues and Challenges

Dr. Gerrit Hornung, LL.M., Monika Desoi, Matthias Pocs, LL.M.

Projektgruppe verfassungsverträgliche Technikgestaltung (provet)
Universität Kassel
Wilhelmshöher Allee 64-66
34109 Kassel
gerrit.hornung@uni-kassel.de
m.desoi@uni-kassel.de
matthias.pocs@uni-kassel.de

Abstract: The privacy and data protection challenges posed by biometric systems have been discussed in detail in the last years. Both security opportunities and privacy risks however may develop and change with the technical enhancement of the respective systems, which also induces the emergence of new application scenarios. One group of such new scenarios appears to be the prevention of criminal or in other ways dangerous behaviour. From a legal point of view, this brings about new challenges which go well beyond the problems of authentication as such. While some of the features of the scenarios discussed below may not be feasible in the short term, it is apparent that the associated fundamental rights and data protection law problems will have to be addressed in the future. This applies to the international plane as well as to national legal orders, for which Germany will serve as an example in the following.[1]

1 Biometrics, behavioural Pattern Analysis and the Law

From the very beginning of the technical development of biometric systems, this technology has been put into question from the privacy and data protection[2] point of view. This is however in no way a sole characteristic. Rather, it appears that virtually

[1] Acknowledgement: The work in this paper has been funded in part by the German Federal Ministry of Education and Science (Bundesministerium für Bildung und Forschung, BMBF) through the Research Programmes under Contract No. 13N10820 – "Digitale Fingerspuren" (Digi-Dak), and Contract No. 13N10814 – "Verteilte, vernetzte Kamerasysteme zur in situ-Erkennung Personen-induzierter Gefahrensituationen (CamInSens)".

[2] For the purpose of this article, it is not necessary to discuss the different notions of these terms. It has been argued that the two concepts differ to a considerable extend. According to [HeGu06] [GuHe08], privacy should be understood as an opacity tool, guaranteeing non-interference in individual matters by the state and private actors. On the contrary, data protection is construed as an transparency tool, meant to compel government and private actors to "good practices" by focusing on the transparency and accountability of governmental or private decision-making and action. However, this useful distinction may be misleading as regards specific legal provisions, which may serve both or other purposes. In addition, "privacy" may have considerable different meanings in different legal orders, and there are further conceptions such as the German right to informational self-determination ([HoSc09]), which may not fall in just one of the two categories.

every new technology which processes personal data brings about issues of personal privacy and governmental control, of autonomous decision-making and heteronomy of the individual, of societal transparency and clandestine collection of information.

Given this fact however, there are some privacy and data protection risks particularly associated with biometric data, which are due to their inherent characteristics. In short, the most relevant of these risks appear to be the following [Al03, 152 ff.] [Ho04] [Ho05, 85 ff., 179 ff.] [Me07, 1089 ff.]:

- "Identity Theft", i.e. the unlawful capture of biometric characteristics in public or from a database, followed by the use of other persons than the data subject,
- The processing of "additional information" (e.g. on illnesses or likeliness for developing an illness, personal origins and current psycho-social constitution) which may be included particularly in biometric samples (raw data),
- The tracking and continuous surveillance of people's behaviour through frequent biometric identification,
- The collection of biometric data and surveillance without notice of the data subject,
- The linking of several databases using biometric data as a common single identifier, and
- Decision errors (false acceptance and false rejections) which lead to subsequent measures or expectations addressing the "wrong" person.

2 New technological Developments

Biometric systems are subject to permanent development. Biometric systems get better as such (i.e. better failure rates), new biometric characteristics may be included, new application scenarios may become feasible or existing scenarios may shift from the authentication of single data subjects to the authentication of larger social groups. Not all of these developments pose new legal questions. However, even the plain enhancement of the comparison algorithm of a given system or the improvement of its spoof prevention mechanisms may require a new legal assessment, because this leads to an ever stronger link between the data subject and the respective biometric samples. Interestingly, while this significantly reduces some of the aforementioned privacy and data protection risks, other risks may increase at the same time. A strong link may most notably lower the risk of identity theft, but add to the possibilities of tracking and surveillance.

In the future, these new technological developments could enable police authorities to introduce biometric systems for the prevention of crime. Subsequently, two examples of possible scenarios will be given. Some of the features of these scenarios may not be

feasible for the short term, but it is apparent that the associated fundamental rights and data protection law problems will have to be addressed in the future.

2.1 Recognition of Fingerprints on Baggage and Freight

To date, fingerprint recognition appears to take place in two scenarios. On the one hand, there is the "old-fashioned", forensic way of manually collecting fingerprints at crime scenes in order to compare the captured data with existing databases such as the AFIS or with the fingerprints of a known suspect. On the other hand, there are digital biometric authentication systems, where the biometric data of a present person is collected and compared with the reference data on a one-to-one or one-to-many basis.

Technological development appears to allow for a combination of the two in order to digitally collect fingerprint data at everyday objects, i.e. without knowing where exactly the fingerprints are or even whether there are any fingerprints at all.[3] This could significantly enhance police work at crime scenes. At the same time, biometric fingerprint systems may even play a role in new preventive scenarios. The systems could, among other things, even be able to find and scan fingerprints on baggage and freight in the airport in order to singling out dangerous materials in the baggage and freight. To this end, it could automatically detect and collect fingerprints and even further proceed by comparing them with a list of dangerous persons. This procedure is only viable because in respect of fingerprints, one could make use of the already existing automation of fingerprint comparison conducted by national police offices such as the German Federal Criminal Police Office (*Bundeskriminalamt*).

2.2 CCTV, behavioural Pattern Analysis, and Identification

Quite similarly, it appears to be an "old-fashioned", first generation of CCTV systems whereby one or several cameras observe a public space and transmit the data to a control room. Clearly, these systems have ever improved, allowing for higher image quality and the analogous or now digital storage of the data for later analysis.

The next technological step however could bring about major changes as regards both the security opportunities and the privacy risks of CCTV. "Smart" cameras, based on video content analysis may use methods of behavioural pattern recognition to monitor large public areas, e.g. airports, football stadiums and railway stations, and to automatically identify acute threat situations at the moment of their development.[4]

Smart cameras combine image sensors and microcomputers to analyse video content in a single device, so that they are the basis of new video systems. Smart cameras renounce image transmission in favour of essential abstracted environmental information.

[3] The technical and legal issues of such systems are currently being scrutinised within the project "Digitale Fingerspuren (Digi-Dak)" (see above n. 1), http://omen.cs.uni-magdeburg.de/digi-dak/.
[4] The technical and legal issues of such systems are currently being scrutinised within the project "Verteilte vernetzte Kamerasysteme zur in situ-Erkennung Personen-induzierter Gefahrensituationen in öffentlichen Räumen (CamInSens)", (see above n. 1), http://www.caminsens.org/.

Therefore disadvantages of central video system architectures (disaster tolerance, scalability) can be overcome. These smart cameras shall be able to identify moving objects, track them and simultaneously compare their motion to common patterns. If it differs from these common patterns and is subsequently identified as a security threat, private or state security services could be alarmed automatically.

Likewise, smart cameras can be interconnected. Long track logs can be created by linking shorter track logs of several camera places. These long tracks can be interpreted to detect conspicuous movement patterns.

In future, smart camera surveillance systems could easily be combined with biometric techniques of facial recognition or other biometric or non-biometric means of personal identification. There are large numbers of possible applications such as the automatic comparison with watch lists of people under restraining orders or missing person´s reports. Meanwhile technical problems have to be solved. Recognition and tracing of temporary or partly hidden persons and high dynamic scenarios (such as different perspectives and lightings) are issues that must be resolved.

3. New and enhanced Privacy and Data Protection Risks

Both scenarios show that biometric authentication may well go beyond specified situations in which the individual is able to control or at least become aware of the use of his/her biometric data. This brings about problems with the data protection principle of transparency. In particular, this principle protects the individual by requiring the controller to inform the data subject about the collection of personal data. This requirement is enshrined in Articles 10 and 11 of the European Data Protection Directive 95/46/EC (DPD) and the respective national data protection acts. For Germany in particular, police legislation of the German *Länder* likewise provides for a precedence of direct over indirect collection of personal data [Pi07, 226 f.]. As regards biometric systems, this precedence is particularly important since data subjects unintentionally leave their fingerprints on objects and facial data may be captured by cameras in the public domain. Further, indirect and covert collection of biometric data also disables the individual to seek legal remedy against unjustified processing of personal data.

Additionally, large-scale applications may significantly influence the legal assessment as regards the principle of proportionality. While this is already a problem in 1:1 verification scenarios ([WP03, 6 ff.]), it becomes critical in preventive scenarios. By way of example, the German Constitutional Court (*Bundesverfassungsgericht*) has frequently ruled that the question of how many citizens are subject to a technical surveillance measure is of utmost importance for the assessment of its constitutionality (scatter or "*Streubreite*", see e.g. [BV08, para. 78]). In respect of prevention, police authorities may exploit their investigative powers in dangerous and endangered places [Pi07, 244 ff.] and automatically collect data related to numerous persons. In relation to fingerprints, national AFIS throughout the world allow for automated recognition of fingerprints of criminals and immigrants. As to both fingerprints and the human face, most states built up databases for passport and ID card registers, or plan to do so in the future. Germany

appears to be a special case, as fingerprint data has to be deleted after the issuance of the documents, and databases with biometric facial data are being built not centrally, but only at local passport and ID card registers. For the time being, the automatic transfer of this data to German police authorities is restricted to single cases of urgency in which the passport or ID card authorities are not reachable [Ho07, 185]. This follows from both Section 22a of the German Passport Act (*Passgesetz* [PG09]) and Section 25 of the German ID Card Act (*Personalausweisgesetz* [PA10] entering into force in November 2010). It remains to be seen whether this will change in the future.

Associated with the new type of preventive large-scale applications, there may also be a tendency towards 1:N identification. Biometric systems for preventive law enforcement necessitate this identification functionality in order to determine the suspect, that is, they need to include a significant group of the population in order to have successful searches. In general however, identification leads to greater privacy problems than 1:1 verification, because it could allow for the non-transparent surveillance of a large group of people ([Al03, 162 ff.] [Bi02, 44] [WP03, 6 f.] [GoPr03, 69 f., 72] [WHO03, 40] [Ho05, 191 ff.]). Biometric encryption, a means of biometric template protection, is suitable to reduce privacy threats. This approach avoids the storage of biometric data and template data by encrypting a random number on the basis of the collected biometric data (since [TSS96]; lately [Br09]). It is however crucial that the random numbers are not stored in the same database, in order to avoid 1:N identification. Otherwise any biometric characteristic could be combined with every random number and the result of that combination could be compared with reference data. This comparison establishes the association of a biometric characteristic with a data set which may identify a person.

Further, the principle of purpose specification is at stake. Biometric data do not as such tie the processing to a certain purpose. For instance, ID card registries process facial data for the purpose of issuing ID cards and certain CCTV surveillance cameras process data for crime prevention purposes. If biometric data can be extrapolated from the video images, ID card images could be used to identify persons located by means of the camera system. If a fingerprint scanning system would be introduced in airports, the data contained in the national AFIS for the purpose of preventing crimes and illegal claims of asylum and residence could be utilised to identify persons of this group that are located by the fingerprint scanner. In both cases, interoperability appears to lead to additional privacy and data protection concerns.

Data subjects may also become subject to further security measures. As regards the identification after an incident, the severity of this risk depends on the reliability of the biometric system. In prevention scenarios which are based on behavioural analysis however, the decision on further security measures may be influenced or even decided by the technical analysis of people's individual behaviour and the comparison with generalised, "dangerous" types of behaviour. Smart cameras may be able to observe people and their motion and to compare this motion to common patterns in order to alarm private or state security services in the event that the motion differs and is identified as a security threat. In result, concrete measures against persons may take place just because of an automatic process.

The pressure of permanent identification and behavioural analytics by smart cameras may lead to a risk of incursions into the freedom of action and the freedom of decision. The reason is that data subjects who feel like being watched, may abstain from deviant behaviour patterns and accommodate themselves to behavioural adaptations. The new intelligent and self-organising smart cameras could become able to track human routes, so that complete targeted monitoring and tracing in public places becomes feasible.

Clearly, this brings about major legal and ethical problems regarding the general possibilities to describe deviant behaviour, the reliability of the system, its decision structure, and the possibilities of ultimate human decision-making. Therefore, among others, the data protection authorities of the Member States recognise that "the challenges for data protection are immense [and a] future legal framework should in any event address [the tendency] towards a more or less permanent surveillance of all citizens [for example] the combined use of intelligent CCTV-cameras and other tools" [WP09, para. 107].

4. Legal Requirements and Challenges

The risks of the use of biometric systems in future preventive scenarios may lead to violations of several human rights protected by the EU Charter of Fundamental Rights, the European Convention of Human Rights and national constitutions such as the German *Grundgesetz*. These laws do not only include the rights to privacy and data protection, as well as national particularities such as the German right to informational self-determination. Additionally, human dignity may be concerned if biometric characteristics are used as single identifiers by state authorities for treatment of data subjects as mere "objects." Moreover, the special protection of sensitive data (Article 8 DPD) such as health and ethnic information may be applicable at least to some forms of biometric data. Finally, the right to travel and the freedom of movement could be at risk in case that data subjects are tracked and continuously monitored in different places. In addition, property as a fundamental right (in case of confiscation), the right to innocence until proven guilty (if the system or its design suffer from errors), the right to judicial review (in non-transparent systems), and the prohibition of arbitration (in case of unspecified purpose of use) may be violated.

This plurality of possible infringements on basic rights causes difficulties to discuss the use in conformity with privacy and data protection requirements. On the other hand, the general legal data protection requirements have been well discussed and may be applied to new biometric systems as well. As those systems in principle process personal data within the meaning of Article 2 (a) DPD, they are subject to the respective national Data Protection Acts which implement this Directive. Albeit differing in detail, the national acts follow common principles due to the harmonising effects of the European legislation. These principles are frequently (but not in all countries) fostered by fundamental rights of national constitutions such as the German right to informational self-determination, recognised by the *Bundesverfassungsgericht* since [BV83] (on the concept see e.g. [HoSc09]). Thus the following principles are in general also vested with the power of constitutional rights in Germany and other national legal orders.

Accordingly, each citizen has the right to, in principle, decide for him/herself which personal information is to be disclosed in his/her social environment. In short, the processing needs to be based on legislation or effective consent by the data subject, personal data may only be collected and used for specified purposes, the data must be anonymised or deleted once this purpose is accomplished, data must not be processed beyond the absolute minimum required (data minimisation), the interference with personal privacy must be proportional to the purpose, the data processing must be transparent for the data subject, proper organisational and technical security measures must be in place to protect the data, and the data subject enjoys certain rights, e.g. to get their data rectified, locked or erased under certain circumstances. Additionally, there are significant restrictions for the use of "sensitive" data. According to Article 8 DPD, this includes "personal data revealing racial or ethnic origin, political opinions, religious or philosophical beliefs, trade-union membership, and the processing of data concerning health or sex life".

As regards the new scenarios above, there are some common legal challenges which apply to all EU Member States. First of all, these security applications cannot be based on the consent of the data subjects because they are used to take preventive police measures against the will, and possibly even without notice, of the data subject. Thus there is the need of legislation specifying in detail the circumstances and requirements of the processing of biometric data. In Germany in particular, it is in both cases very doubtful whether the existing police and data protection laws allow for the use of biometrics in the described manner. In contrast, this may be possible in the case of criminal investigation using digital capturing of fingerprints, as this appears to merely replace the analogue measures used hitherto.

In preventive scenarios such as the (even routinely) scanning of baggage and freight in airports and the behavioural analysis and personal tracking of visitors through the use of "intelligent" CCTV systems however, no concrete suspicion of a crime can be established before processing personal data related to a myriad of persons. Police intervention in endangered places requires the establishment of such a suspicion beforehand [Pi07, 244.]. However, since it cannot automatically be determined whether or not a person is dangerous without collecting personal data, a decision of the *Bundesverfassungsgericht* could solve this conflict. In the case of number plate recognition [BV08], the Court ruled that collecting data from car number plates does not interfere with the right to informational self-determination if the data stay anonymous and are instantly and untraceably deleted in case that the comparison with the police search database is negative [BV08, para. 68]. For fingerprint recognition, this could mean that the German legislator is allowed to provide for the scanning of baggage and freight in airports for preventive purposes. It remains to be seen whether other national courts and data protection authorities of other countries will take up this approach.

There has so far been no occasion for the *Bundesverfassungsgericht* to deliberate on the preventive use of a biometric system, and the same appears for other national constitutional courts. Nonetheless, existing data protection legislation might be applicable to preventive biometric scenarios. While the DPD does not include provisions on biometrics or CCTV, some national data protection acts do. By way of example,

following Section 6b of the German Federal Data Protection Act (*Bundesdatenschutzgesetz*, BDSG) the monitoring of publicly accessible areas using optoelectronic devices under statutory requirement is allowed.[5] Monitoring publicly accessible areas using optoelectronic devices shall be lawful only as far as necessary (1) for public bodies to perform their duties, or (2) to exercise the right to determine who shall be allowed or denied access, or (3) to pursue legitimate interests for specifically defined purposes. Additionally, there must be no indications of overriding legitimate interests of the data subject. Furthermore, suitable measures shall be taken to indicate that the area is being monitored and to identify the data controller.

Section 6b (3) BDSG governs the processing and use of the collected data. Whether the authorisation of this section provides for biometric comparison of the collected data or not has not been investigated so far. Currently under investigation too is the consequence of the new characteristics of the data collected by "smart" CCTV systems. The new technology aims at detecting behavioural patterns, but may also detect sensitive data such as disabilities or ethnic groups on the basis of behavioural patterns or appearance. As Section 6b BDSG was not drafted for these cases, the question will have to be answered whether the law provides for this new level of data processing.

Another common feature of public, large-scale scenarios is the unclear situation as regards the competence for the collection of biometric data, its processing and the possible subsequent danger prevention measures. Even in today's airports, railway stations, sport stadiums and other open venues, the borders between "public" and "private" spheres have been blurred to a considerable extend, leading to complicated models of public-private-partnerships in the domain of public security. From a legal point of view, this raises severe questions of competence and accountability [Gu01] [St97]. It could also come into conflict with the concept of informational separation of powers recognised by the *Bundesverfassungsgericht* in the population census decision (*Volkszählungsurteil*) [BV83, 69]. As regards CCTV systems and fingerprint recognition systems, there need to be clear legal provisions about the data controller, the data collection and transmission between public authorities and possible private actors.

For instance, several bodies may be in charge to control baggage and freight in airports. In Germany, these are the German Federal Police (*Bundespolizei*) as regards controlling baggage, the aviation company as to the freight control, the police authority of the respective *Land* as to the control of airport staff, and both the captain and the *Bundespolizei* as to the control in the airplane [Gi07, 49/54/55/75]. Further, private security firms may be obliged to carry out these control measures (see Sections 8 and 9 of the German Aviation Security Act, *Luftsicherheitsgesetz*). This diversity of the parties involved in the implementation of security measures at national level is recognised by the EU legislator (see Recital 9 of the EC Regulation 2320/2002).

[5] For the United Kingdom, protection against CCTV surveillance is guaranteed by the Data Protection Act 1998 and the CCTV Code of Practice 2008 by the Information Commissioner´s Office. As to France, the loi n°95-73 du 21 janvier 1995, the décret n°96-926, the arrêté du 26 septembre 2006 and especially the décret du 3 aout 2007 apply.

In future, another major legal challenge within the context of biometric systems will relate to court evidence. One example for this could be the conviction of a subject on the basis of the outcome of a biometric comparison: On which threshold of a biometric system could this be based in different settings? In Germany, the Federal Criminal Court (*Bundesgerichtshof*) ruled that even the highly secure DNA analysis must not be the only evidence for the conviction of the accused [BG98]. For prevention scenarios in particular, the requirement for subsequent measures is usually a threat to public safety. So far, it remains completely unclear under which circumstances the existence of such a threat may be solely based on the outcome of a biometric process or technical behavioural analysis.

In the end, the new systems must comply with legal requirements concerning the privacy-friendly technical design. According to the principle of data minimisation (see Article 6 (1) (c) DPD and Section 3a BDSG respectively), the processing of data must not be excessive in relation to the purposes for which they are collected. Thus anonymous or at least pseudonymous data must be used wherever possible. For example, this may be possible in the case of the behavioural analysis in the CCTV setting, where the technical analysis itself can be conducted without personal data and the pictures in control rooms could blur people's faces as long as there is no incident [St05].

5. Conclusion

It has become clear that new technical possibilities of biometric systems lead to new challenges for personal privacy and data protection. At least in Germany, data protection laws do not specifically cover biometric data, which for example represent fingerprints or behavioural patterns of data subjects. In addition, data subjects are so far often not aware of the information quality that is revealed from latent fingerprints and bodily movements.

The amount of the new challenges depends on the respective technical design. Automated collection of biometric data enables law enforcement authorities not only to prosecute the accused but also proactively collect information about an unspecified group of persons. Hence, entire societies may be posed under suspicion if there are no technical and legal safeguards in place. In consequence, citizens may feel of being watched and adapt their behaviour so that they hide individual characteristics.

Technology may on the other hand, depending on the design, also preserve personal privacy and data protection. However, this may be limited in certain scenarios if the sole aim of a biometric system is the identification of an unknown person in a large group of other persons or the mapping of a large amount of newly captured biometric data on a biometric database.

For certain purposes such as prevention of serious crime, a biometric system might be a pressing need for the society. Further, secret collection of data may be necessary for police work. Prevention however, by the very nature of the concept, cannot be restricted to a group that only consists of dangerous persons. Thus, one has to establish procedural

rules that enable data subjects to seek judicial review. In addition, places where data are collected secretly may require a notice to the data subject in order to enable him/her to decide where he/she can behave freely without worrying about the interpretation of that behaviour at the other end of the surveillance system.

The specific aspects of biometric characteristics necessitate a very cautious approach because of the unique and durable relation to the data subject. This is caused by the fact that biometric characteristics can be used to single someone out and do, in principle, not change during the course of the data subject's life. Thus, the data subject may be deprived of choosing his/her role according to the respective situation because others may interconnect different sets of data about him/her. Hence, biometric template protection regimes that utilise renewable and irreversible representations of biometric data may be an option to ensure that data can only be used by a certain authority for a certain purpose in a certain biometric system. This sort of purpose limitation by design could also prevent or reduce the risk of identity theft. Moreover, long-term storage of biometric data may require regular data security measures, for example, fresh re-encryption since state-of-the-art cryptosystems may become ineffective after a certain period of time.

Finally, users of biometric systems for law enforcement purposes need to take error rates into account. From this it follows that the data subjects singled out by the biometric system must not be subject to particularly burdensome consequences of police measures on the sole basis of the biometric recognition or rejection. Rather, the police legislator and the executing officer always have to be aware of the actual efficiency of the recognition mechanisms and the contents of the data pools in use to avoid sanctioning the "wrong" person by putting him/her under suspicion or hindering that person to travel, move freely, and keep his/her property.

The function of the law – particularly the fundamental rights to privacy and informational self-determination, as well as the respective data protection acts – needs to be the protection of the personal rights of the data subjects. To this end, specific provisions for certain application scenarios may be necessary in the future. Furthermore, a legally compliant technology design is able to significantly reduce privacy and data protection risks. The earlier in the process of research and development this takes place, the better the potential outcome for privacy enhancing technologies.

Bibliography

[Al03] Albrecht, A.: Biometrische Verfahren im Spannungsfeld von Authentizität im elektronischen Rechtsverkehr und Persönlichkeitsschutz, Nomos, Baden-Baden 2003.

[Bi02] Bizer, J.: Selbstauthentifizierende Ausweiskarte, Datenschutz und Datensicherheit 2002, p. 44.

[BG98] Bundesgerichtshof, Decision of 2 September 1997, Neue Zeitschrift für Strafrecht 1998, p. 97.

[Br09] Breebaart, J.; Yang, B.; Buhan-Dulman, I. ; Busch, C.: Biometric Template Protection. The need for open standards, (Datenschutz und Datensicherheit) 2009, pp. 299-304.

[BV83] Bundesverfassungsgericht, BVerfGE (Collection of decicions), volume 65, pp. 1-71 („Volkszählungsurteil“).

[BV08] Bundesverfassungsgericht, BVerfGE (Collection of decisions), volume 120, pp. 378-433 (number plate recognition), press release in English at http://www.bundesverfassungsgericht.de/pressemitteilungen/bvg08-027en.html.

[Gi07] Giemulla, E.; Rothe, B.: Recht der Luftsicherheit, Springer Verlag, Berlin 2008.

[GoPr03] Golembiewski, C.; Probst, T.: Datenschutzrechtliche Anforderungen an den Einsatz biometrischer Verfahren in Ausweispapieren und bei ausländerrechtlichen Identitätsfeststellungen (Gutachten des Unabhängigen Landeszentrums für Datenschutz Schleswig-Holstein für das Büro für Technikfolgenabschätzung beim Deutschen Bundestag), available at http://www.datenschutzzentrum.de/download/Biometrie_Gutachten_Print.pdf, Kiel 2003.

[Gu01] Gusy, C.: Polizei und private Sicherheitsdienste im öffentlichen Raum – Trennlinien und Berührungspunkte, Verwaltungsarchiv 2001, pp. 344-367.

[GuHe08] Gutwirth, S.; De Hert, P.: Regulating Profiling in a Democratic Constitutional State. In (Hildebrandt, M.; Gutwirth, S. Eds.): Profiling the European Citizen. Cross-Disciplinary Perspectives, Springer Verlag, 2008, 271-293.

[HeGu06] De Hert, P.; Gutwirth, S.: Privacy, data protection and law enforcement. Opacity of the individual and transparency of power. In (Claes, E.; Duff, A.; Gutwirth, S. Eds.): Privacy and the criminal law, Intersentia, Antwerp/Oxford 2006, pp. 61-104.

[Ho04] Hornung, G.: Biometrische Systeme – Rechtsfragen eines Identifikationsmittels der Zukunft, Kritische Justiz 2004, pp. 344-360.

[Ho05] Hornung, G.: Die digitale Identität. Rechtsprobleme von Chipkartenausweisen: digitaler Personalausweis, elektronische Gesundheitskarte, JobCard-Verfahren, Nomos, Baden-Baden 2005.

[Ho07] Hornung, G.: Fingerabdrücke statt Doktortitel: Paradigmenwechsel im Passrecht. Der Gesetzesentwurf der Bundesregierung zur Änderung des Passgesetzes und weiterer Vorschriften, Datenschutz und Datensicherheit 2007, pp. 181-185.

[HoSc09] Hornung, G.; Schnabel, C., Data protection in Germany I: The population census decision and the right to informational self-determination, Computer Law & Security Review 2009, pp. 84-88, also available at http://cms.uni-kassel.de/unicms/fileadmin/groups/w_030405/Gerrit_Hornung/Hornung___Schnabel__Data_protection_in_Germany_I__CLSR_2009__84.pdf.

[Me07] Meints, M.; Biermann, H.; Bromba, M.; Busch, C.; Hornung, G.; Quiring-Kock, G.: Biometric Systems and Data Protection Legislation in Germany. In (Pan, J.-S.; Niu, X.-M.; Huang, H.-C.; Jain, L. C. Eds.): 2008 Fourth International Conference on Intelligent Information Hiding and Multimedia Signal Processing, IEEE 2008, pp. 1088-1093.

[PA10] German ID Card Act (*Personalausweisgesetz*), of 18.6.2009, Federal Law Gazette I, p. 1346.

[PG09] German Passport Act (*Passgesetz*) of 20.7.2007, Federal Law Gazette I, p. 1566.

[Pi07] Pieroth, B.; Schlink, B.; Kniesel, M.: Polizei- und Ordnungsrecht, 3rd ed., Verlag C.H. Beck, Munich 2005.

[St97] Stober, R.: Staatliches Gewaltmonopol und privates Sicherheitsgewerbe – Plädoyer für eine Police-Private-Partnership, Neue Juristische Wochenschrift 1997, pp. 889-896.

[St05] v. Stechow, C.: Datenschutz durch Technik. Rechtliche Förderungsmöglichkeiten von privacy enhancing technologies am Beispiel der Videoüberwachung, DUV Verlag, Wiesbaden 2005.

[TSS96] Tomko G. J.; Soutar C.; Schmidt G. J.: Fingerprint controlled public key cryptographic system. U.S. Patent 5541994, July 30, 1996 (Priority date: Sept. 7, 1994).

[WHO 03] Woodward, J. D., Jr.; Orlans, N. M.; Higgins, P. T.: Biometrics. Identity Assurance in the Information Age, McGraw-Hill Osborne Media, New York 2003.

[WP03] Article 29 Data Protection Working Party: Working document on biometrics, 12168/02/EN, WP 80, available at http://ec.europa.eu/justice_home/fsj/privacy/docs/wpdocs/2003/wp80_en.pdf, Brussels 2003.

[WP09] Article 29 Data Protection Working Party: The Future of Privacy. Joint contribution to the Consultation of the European Commission on the legal framework for the fundamental right to protection of personal data, 02356/09/EN, WP168, available at http://ec.europa.eu/justice_home/fsj/privacy/docs/wpdocs/2009/wp168_en.pdf Brussels 2009.

Applikation des DBSCAN Clustering-Verfahrens zur Generierung von Ground-Truth Fingerabdruck-Minutien

Sebastian Abt[1], Christoph Busch[2], Claudia Nickel[1]

[1] Hochschule Darmstadt – Fachbereich Informatik
[2] Hochschule Darmstadt – CASED

{sebastian.abt,christoph.busch,claudia.nickel}@h-da.de

Abstract: Biometrische Verfahren werden häufig im Rahmen der Authentifizierung zum Schutz wichtiger Daten und Systeme eingesetzt. Hierzu werden biometrische Referenzen gespeichert, die Informationen über biometrische Charakteristika der betroffenen Individuen enthalten. Um diese biometrischen Referenzen in offenen und verteilten biometrischen Systemen nutzen zu können, ist eine Standardisierung der Datenaustauschformate notwendig. Eine gute Erkennungsleistung ist nur erzielbar, wenn ausgetauschte Datensätze auch standardkonform erstellt werden. Das Einhalten dieser Standards muss mittels geeigneter syntaktischer und semantischer Konformitätstests sichergestellt werden. Im Rahmen eines semantischen Konformitätstests wird geprüft, ob die gespeicherten Referenzen ein wahres Abbild der biometrischen Charakteristika darstellen. Zur Durchführung semantischer Konformitätstests ist jedoch die Existenz eines entsprechenden Referenzdatensatzes unbedingte Voraussetzung. Derartige Referenzdatensätze werden zur Zeit jedoch erst erstellt. Diese Arbeit leistet einen Beitrag zur Erzeugung eines solchen Referenzdatensatzes am Beispiel von Fingerabdruck-Minutien. Hierzu wird das DBSCAN Clustering-Verfahren auf Datenpunkte angewendet, die von daktyloskopischen Experten durch manuelle Analyse von Fingerabdruck-Bildern bestimmt wurden. Mit Hilfe des Clustering-Verfahrens werden Ground-Truth Fingerabdruck-Minutien erzeugt, die als Referenzdatensatz zum Durchführen semantischer Konformitätstests von automatischen Fingerabdruckidentifikationssystemen genutzt werden können.

Keywords: Fingerabdruck-Minutien, Interoperabilität, Ground-Truth, Referenzdaten, Konformitätstest, Clustering, DBSCAN.

1 Einleitung

Der Schutz digitaler und physikalischer Systeme und Daten stellt eine immer bedeutender werdende Aufgabe in unserer heutigen Gesellschaft dar. Zur Autorisierung eines Zugriffsversuchs ist eine erfolgreiche Authentifizierung notwendig. Die Authentifizierung einer Identität kann entweder im Rahmen einer automatischen *Identifikation* einer betroffenen Identität oder im Rahmen einer *Verifikation* erfolgen, bei der während der Authentisierung eine Identitätsbehauptung getroffen wird.

Beispielsweise werden zum Schutz persönlicher, in einem Computer oder Mobiltelefon

gespeicherter Daten heutzutage vornehmlich so genannte *wissensbasierte Verfahren* eingesetzt (z.B. die Eingabe einer persönlichen Identifikationsnummer (PIN) oder die Eingabe einer Benutzername/Passwort-Kombination). In der Grenzkontrolle kommen zur Zeit in der Regel *besitzbasierte Verfahren* (das Vorlegen eines amtlichen Personaldokuments) zur Anwendung. In der Gebäudesicherung wird häufig eine Kombination aus besitzbasierten und wissensbasierten Verfahren im Kontext der Zutrittskontrolle eingesetzt (z.B. eine Kombination aus elektronischem Transponder und PIN). Besitzbasierte und wissenbasierte Verfahren erliegen jedoch der Gefahr des Verlierens bzw. Stehlens oder Vergessens von PIN/Passwort oder Token und eröffnen zusätzlich die Möglichkeit des unerlaubten Weitergebens der eigenen PIN bzw. des persönlichen Tokens. Die Rechtmäßigkeit eines Zugriffs kann daher selbst nach erfolgreicher Identifikation bzw. Verifikation nicht garantiert werden. Zur Minimierung dieser Risiken lässt sich die Authentifizierung durch den Einsatz von Biometrie erweitern. Nach ISO/IEC 24745 [ISO09b] dienen biometrische Systeme zur Erkennung von Individuen auf Basis physiologischer (z.B. Fingerabdrücke, Gesicht, Iris) oder verhaltensbasierter (z.B. Gang, Stimme, Handschrift) Charakterisitka. Auf Grund ihrer physikalischen Zugehörigkeit zum betroffenen Individuum lassen sich biometrische Charakteristika nur schwer verteilen und vergessen und können somit wissensbasierte oder besitzbasierte Verfahren im Sinne einer Zwei- oder Dreifaktor-Authentifizierung ergänzen.

Um die *Interoperabilität* verschiedener biometrischer Systeme gewährleisten zu können, ist eine Standardisierung der von den Systemen verwendeten digitalen Repräsentationen biometrischer Charakteristika (biometrische Referenzen) notwendig. Im Bereich der (automatischen) Fingerabdruckidentifikation ist dies im Rahmen des ISO Minutien Interoperabilitätsstandards IS 19794-2 [ISO05] geschehen. Analog hierzu definiert ISO/IEC 29109-2 [ISO09a] drei Level zur *Konformitätsprüfung* der von automatischen Fingerabdruckidentifikationssystemen (AFIS) erzeugten Daten. *Level 1* Konformitätstests befassen sich mit der Existenz und korrekten Kodierung aller notwendiger Datenfelder. *Level 2* Konformitätstests beschäftigen sich mit den korrekten Inhalten der Datenfelder sowie der Konsistenz der Werte in Beziehung stehender Datenfelder. *Level 3* Konformitätstests sollen Aussagen über die semantische Konformität der Daten treffen, d.h. z.B. über die korrekte Detektion von Minutien innerhalb spezifischer Toleranzen [BLT+09].

Zur Durchführung *semantischer (Level 3) Konformitätstests* ist die Existenz eines Ground-Truth Referenzdatensatzes notwendig. Erste Arbeiten zur Generierung eines derartigen Referenzdatensatzes wurden von Busch et al. [BLT+09] durchgeführt. Hierbei wurden 5000 Fingerabdruck-Bilder aus den National Institute of Standards and Technology (NIST) Spezialdatenbanken SD14 und SD29 zusammengestellt, die von mehreren daktyloskopischen Experten des Bundeskriminalamts (BKA) manuell und ohne Unterstützung von AFIS-Systemen untersucht werden. Die auf diese Weise untersuchten Fingerabdruck-Bilder resultieren, abhängig von der Anzahl untersuchender Experten, in einer Vielzahl subjektiver, sich in ihren Ausprägungen (Detektion, Platzierung und Typisierung von Minutien) unterscheidenden Klassifikationsergebnissen pro Bild (für weitere Details siehe [BLT+09]). Zur Generierung eines für semantische Konformitätstests geeigneten Ground-Truth Referenzdatensatzes ist die Fusion dieser unterschiedlichen Einzelergebnisse notwendig. Diese Notwendigkeit stellt die Motivation dieser Arbeit dar, in der die Applikation des DBSCAN Clustering-Verfahrens auf Fingerabdruck-Minutien untersucht wird.

Der Rest dieser Arbeit gliedert sich wie folgt: nachdem Kapitel 1 in das Themenfeld einführte und die Arbeit motivierte, erläutert Kapitel 2 die notwendigen Grundlagen zur Datenbasis und zum DBSCAN Clustering Verfahren. Verwandte Arbeiten werden in Kapitel 3 diskutiert. Kapitel 4 beschreibt die Anwendung des DBSCAN Clustering-Verfahrens zur Erzeugung von Ground-Truth Fingerabdruck-Minutien. Kapitel 5 diskutiert die Ergebnisse der Arbeit und Kapitel 6 schließt die Arbeit ab.

2 Hintergrund

Dieses Kapitel erläutert die zu Grunde liegende Datenbasis (Kapitel 2.1) sowie das DBSCAN Clustering Verfahren (Kapitel 2.2).

2.1 Datenbasis

Bei den im Rahmen dieser Arbeit zur Verfügung stehenden Daten handelt es sich um einen Auszug aus der in [BLT+09] beschriebenen Fingerabdruck-Bilder. Der Datensatz umfasst $npic = 17$ Bilder P_j, $j = 1, \ldots, npic$, die von $nexp = 9$ daktyloskopischen Experten des Bundeskriminalamts analysiert wurden (Experten-IDs 11 bis 19)[1]. Die Ergebnisse der von den Experten durchgeführten manuellen Untersuchungen stehen in einem proprietären GTM-Dateiformat (siehe [BLT+09]) zur Verfügung. Da sich diese Arbeit zunächst lediglich mit dem Erstellen eines Ground-Truth Datensatzes für Fingerabdruck-Minutien befasst, werden die in den GTM-Dateien enthaltene Informationen über Deltas und Core eines Fingerabdrucks sowie weitere Metainformationen (Bildqualität, Fingerabdrucktyp, etc.) nicht weiter betrachtet. Es handelt sich bei den während des Clusterings verwendeten Daten folglich um Fingerabdruck-Minutien $m_{j,k} \in P_j$ repräsentierende 6-Tupel $m_{j,k} = (ExpId, Type, XPos, YPos, Angle, Quality)$, wobei

- $ExpId \in \{11, 12, 13, \ldots, 19\}$ die eindeutige Identifikationsnummer des die Minutie definierenden Experten bezeichnet,
- $Type \in \{0, 1, 2\}$ den Typ der Minutie bezeichnet; der Wert 0 kodiert hierbei eine Minutie unbestimmbaren Typs, Wert 1 bezeichnet eine Minutie vom Typ *Ridge Ending* und Wert 2 eine Minutie des Typs *Ridge Bifurcation* [ISO05],
- $XPos, YPos \in \mathbb{R}$ die X- bzw. Y-Koordinate der Minutie im Fingerabdruck-Bild beschreiben [ISO05],
- $Angle \in [0, 255]$ den in einem Byte kodierten Richtungswinkel der Minutie beschreibt [ISO05],
- $Quality \in [0, 100]$ die Güte der Minutie bezeichnet; hierbei handelt es sich um einen rein subjektiven Wert, der vom Experten festgelegt wird und als Maß für die Vertrauenswürdigkeit seiner Bewertung dienen kann.

[1] Zum Zeitpunkt der Untersuchung stand kein umfangreicheres Datenmaterial zur Verfügung.

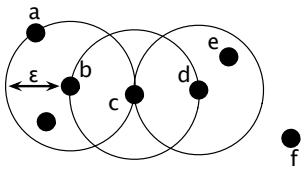

Abbildung 1: Verdeutlichung des DBSCAN Clustering-Verfahrens. Die Punkte b, c, d stellen *Kernpunkte* bzgl. $\varepsilon, MinPts$ dar. Der Punkt a ist *direkt dichte-erreichbar* von Punkt b, Punkt e ist *direkt dichte-erreichbar* von von Punkt d und *dichte-erreichbar* von Punkt b. Die Punkte a, e sind *dichte-verbunden* über b, c, d. Punkt f stellt Rauschen dar.

2.2 DBSCAN Clustering-Verfahren

Beim *Density Based Spatial Clustering of Applications with Noise (DBSCAN)* [EKSX96] Clustering-Verfahren handelt es sich um ein von Ester et al. entwickeltes dichtebasiertes Clustering-Verfahren. Dichtebasierte Clustering Verfahren sind Methoden, die zur Entdeckung von Clustern beliebiger Form geeignet und an die räumlichen Clustering-Fähigkeiten des Menschen angelehnt sind. Derartige Verfahren definieren Cluster in der Regel als punktdichte Regionen in einem gegebenen Raum D, die untereinander durch Regionen geringerer Dichte separiert werden. Diese separierenden Regionen geringerer Dichte – bzw. die darin liegenden Punkte – werden als Rauschen betrachtet.

DBSCAN definiert einen Cluster als eine in Bezug auf *Dichte-Erreichbarkeit* maximale *dichte-verbundene* Punktmenge. Zur Steuerung des DBSCAN Clustering-Verfahrens dienen die Parameter

1. $\varepsilon \in \mathbb{R}$ zur Definition der Größe der ε-Nachbarschaft $N_\varepsilon(p)$ eines Punktes $p \in D$, sowie
2. $MinPts \in \mathbb{N}$ zur Definition der minimalen Größe einer Punktmkenge $C \subseteq D$ um sie als Cluster-Kandidaten zu betrachten.

Die ε-Nachbarschaft $N_\varepsilon(p) \subseteq D$ eines Punktes p ist definiert als diejenigen Punkte $q \in D$, für die gilt $\delta(p, q) \leq \varepsilon$, wobei $\delta : D \times D \rightarrow \mathbb{R}$ eine beliebige auf D definierte Distanzfunktion bezeichnet. Außer der Angabe von ε und $MinPts$ verwendet DBSCAN keine weiteren externen Informationen über die zu Grunde liegenden Daten zur Clusterbildung. Insbesondere ist es nicht notwendig die Anzahl der zu erzeugenden Cluster zu spezifizieren.

DBSCAN definiert zwei Punkte $p, q \in D$ als *dichte-verbunden* in Bezug auf $\varepsilon, MinPts$ wenn es einen Punkt $o \in D$ gibt, für den gilt: p und q sind *dichte-erreichbar* von o unter Berücksichtigung von $\varepsilon, MinPts$ (vgl. Abbildung 1).

Ein Punkt $p \in D$ ist *dichte-erreichbar* von einem Punkt $q \in D$ in Bezug auf $\varepsilon, MinPts$, wenn eine Menge von Punkten $p_1, p_2, \ldots, p_n \in D$ existiert, mit $p_1 = q$ und $p_n = p$ so dass für alle $i = 2, 3, \ldots, n$ gilt: p_i ist *direkt dichte-erreichbar* von p_{i-1}. Wir schreiben hierfür $q \rightsquigarrow_{\varepsilon, MinPts} p$.

Ein Punkt $p \in D$ ist *direkt dichte-erreichbar* von einem Punkt $q \in D$ in Bezug auf $\varepsilon, MinPts$ wenn $p \in N_\varepsilon(q)$ und $|N_\varepsilon(q)| \geq MinPts$ gelten. Ein Punkt q, für den $|N_\varepsilon| \geq MinPts$ gilt bezeichnet man *Kernpunkt*.

Unter Verwendung dieser Definitionen definiert sich ein Cluster $C \subseteq D$ bzgl. $\varepsilon, MinPts$ wie folgt:

1. $\forall p, q \in D : (p \in C \wedge p \rightsquigarrow_{\varepsilon, MinPts} q) \Rightarrow (q \in C)$
2. $\forall p, q \in C : p$ und q sind dichte-verbunden bzgl. $\varepsilon, MinPts$.

Für weitere Details zu DBSCAN und Implementierungsaspekte siehe [EKSX96].

3 Verwandte Arbeiten

Henniger et al. [HS07] unterstreichen die Notwendigkeit semantischer Konformitätstests. In Ermangelung eines Ground-Truth Referenzdatensatzes führen sie semantische Konformitätsprüfungen zwischen Systemen unterschiedlicher Hersteller manuell durch.

In [BLT+09] beschreiben Busch et al. die Notwendigkeit semantischer Konformitätstests sowie Ansätze zur Ermittlung von Konformitätsraten. Ferner wird die Kompilation von Fingerabdruck-Bildern aus NIST Spezialdatenbanken beschrieben und mit der daktyloskopischen Untersuchung begonnen.

Darauf aufbauend geben Lodrova et al. [LBT+09] einen Überblick über Fehler, die während einer automatischen Minutien-Detektion auftreten können und skizzieren ein Clustering-Verfahren zur Erzeugung von Ground-Truth Minutien auf der in Kapitel 2.1 beschriebenen Datenbasis. Bei dem von Lodrova et al. beschriebenen Clustering-Verfahren handelt es sich um ein hierarchisches Verfahren, das zwei Cluster $C_{j,1}, C_{j,2} \subseteq P_j$ der Größe n, d.h. $n = |C_{j,1}| = |C_{j,2}|$, zusammenführt, genau dann wenn

1. alle Minutien in $C_{j,1}$ und $C_{j,2}$ von unterschiedlichen Experten erzeugt wurden, d.h. wenn gilt: $\forall m_{j,1,k} \in C_{j,1} \nexists m_{j,2,k} \in C_{j,2} : ExpId(m_{j,1,k}) = ExpId(m_{j,2,k})$, wobei $ExpId : D \rightarrow \mathbb{N}$ die Identifikationsnummer des zur Fingerabdruck-Minutie $m_{j,i,k}$ gehörenden Experten bestimmt.
2. alle Minutien in $C_{j,1}$ und $C_{j,2}$ innerhalb eines Kreises mit Radius $\frac{W}{4}$ liegen; W bezeichnet hierbei den durchschnittlichen Abstand zwischen zwei benachbarten Papillarleisten.
3. $C_{j,1}$ und $C_{j,2}$ über $(n-1)$ identische Minutien $m_{j,i,k}$ verfügen, d.h. wenn gilt: $|C_1 \cap C_2| = n - 1$

Dieser Schritt wiederholt sich für alle $n = 1, \ldots, nexp_j$, wobei $nexp_j$ die Anzahl der ein Fingerabdruck-Bild P_j untersuchenden daktyloskopischen Experten bezeichnet. Nachdem auf diese Weise Minutien-Cluster erzeugt wurden, werden Ground-Truth Minutien durch das Bestimmen geeigneter Cluster-Zentren definiert.

4 Erzeugen eines Ground-Truth Referenzdatensatzes

Dieses Kapitel beschreibt das Erzeugen von Ground-Truth Fingerabdruck-Minutien mittels Clustering. Hierzu werden in Kapitel 4.1 zunächst die Herausforderungen beim Erzeugen von Ground-Truth Fingerabdruck-Minutien sowie unsere Vorgehensweise beschrieben. Kapitel 4.2 erläutert anschließend die Applikation des DBSCAN Clustering-Verfahrens auf die in Kapitel 2.1 beschriebenen Daten. Kapitel 4.3 thematisiert das Erstellen von Ground-Truth Fingerabdruck-Minutien unter Verwendung der Clustering Ergebnisse.

4.1 Vorgehensweise und Herausforderung

Das Vorgehen zur Erzeugung von Ground-Truth Fingerabdruck-Minutien ist zweigeteilt. Zunächst werden durch Anwendung des DBSCAN Clustering-Verfahrens Fingerabdruck-Minutien Cluster $C_{j,i} \subseteq P_j$ pro Fingerabdruck-Bild P_j erzeugt. In einem zweiten Schritt wird auf Basis der Clustering Ergebnisse pro Cluster $C_{j,i}$ ein geeigneter Repräsentant $gtm_{j,i}$ erzeugt. Ein Repräsentant $gtm_{j,i}$ stellt eine Ground-Truth Minutie dar.

Die in Kapitel 2.1 beschriebenen von daktyloskopischen Experten generierten Fingerabdruck-Minuten $m_{j,k} = (ExpId, Type, XPos, YPos, Angle, Quality) \in P_j$ stellen die Datenbasis für die Erzeugung von Ground-Truth Fingerabdruck-Minutien dar. Die Herausforderungen, die sich durch das Verwenden dieser Daten ergeben sind:

- *Die Anzahl vorhandener Fingerabdruck-Minutien ist unbekannt:* Die Fingerabdruck-Minutien werden von daktyloskopischen Experten individuell und ohne Verwendung von AFIS-Systemen detektiert. Dies hat zur Folge, dass die Anzahl der pro Fingerabdruck-Bild detektierten Fingerabdruck-Minutien pro daktyloskopischem Experten variieren können.
- *Die Anzahl der daktyloskopischen Experten ist variabel:* Fingerabdruck-Bilder können von einer unterschiedlichen Anzahl von Experten analysiert worden sein. Je nach Bild können somit pro potenzieller Ground-Truth Fingerabdruck-Minutie unterschiedlich viele Einzelergebnisse vorliegen.
- *Fingerabdruck-Bilder sind von unterschiedlicher Qualität:* Bei den gegebenen Fingerabdruck-Bildern kann es sich um gerollte Bilder (SD29) oder flache Abdrücke handeln (SD14), die wahlweise als Live-Scans oder Ink-Bilder vorliegen. Diese Variation hat Einfluss auf die Güte der Einzelbilder und damit auf die Güte und Zuverlässigkeit der von daktyloskopischen Experten detektierten Fingerabdruck-Minutien.

Auf Grund dieser Rahmenbedingungen stehen zur Clusterbildung nur wenige Zusatzinformationen zur Verfügung. So lässt sich auf Basis der maximalen Anzahl am Minutien-Detektionsprozess beteiligter daktyloskopischer Experten $nexp$ sowie auf Basis der ein einzelnes Fingerabdruck-Bild P_j untersuchenden daktyloskopischer Experten $nexp_j$ auf die maximale Anzahl in einem Cluster $C_{j,i}$ erlaubter Fingerabdruck-Minutien $m_{j,i,k} \in$

$C_{j,i}$ schließen, d.h. $\forall j \in \{1,\ldots,npic\} \forall i \in \{1,\ldots,ncl_j\} : nclmin_{j,i} = |C_{j,i}| \leq nexp_j \leq nexp$, wobei ncl_j die Anzahl der in einem Fingerabdruck P_j gebildeten Cluster $C_{j,i}$ bezeichnet. Ferner lässt sich fordern, dass einem Cluster $C_{j,i}$ pro daktyloskopischem Experten maximal eine Fingerabdruck-Minutie zugeordnet wird, d.h. $\forall j \in \{1,\ldots,npic\}$ $\forall i \in \{1,\ldots,ncl_j\} \nexists (m_{j,i,k}, m_{j,i,l}) \in C_{j,i} : ExpId(m_{j,i,k}) = ExpId(m_{j,i,l})$.

4.2 Clusterbildung mittels DBSCAN

Wie in Kapitel 2.2 beschrieben, handelt es sich bei DBSCAN um ein dichtebasiertes Clustering-Verfahren, das nur wenige Eingabeparemeter ($\varepsilon, MinPts$) benötigt und auf einer verrauschten Datenbasis arbeiten kann. Zur Anwendung von DBSCAN bedarf es lediglich der Definition einer geeigneten Distanzfunktion $\delta : D \times D \rightarrow \mathbb{R}$. In unserem Kontext bezeichne $D = P_j = \{m_{j,1}, m_{j,2}, \ldots\}$ die Menge aller von daktyloskopischen Experten erzeugten Fingerabdruck-Minutien $m_{j,k}$ eines vorgegebenen Fingerabdruck-Bildes P_j. Ferner ist es notwendig die Eingabeparameter $\varepsilon, MinPts$ geeignet zu interpretieren und zu bestimmen.

4.2.1 Definition der Distanzfunktion

Zur Anwendung des DBSCAN Clustering-Verfahrens ist die Definition einer geeigneten Distanzfunktion notwendig. Im Kontext des Fingerabdruck-Minutien Clustering stehen hierfür die in Kapitel 2.1 beschriebenen Attribute eines Minutien 6-Tupels $m_{j,k}$ zur Verfügung. Die unmittelbare Verwendung einer gängigen Distanzfunktion, wie z.B. die euklidische Distanz, Manhattendistanz oder Mahalanobisdistanz, scheint hierfür nicht ausreichend bzw. auf Grund der unterschiedlichen Attributklassen (numerisch, nominal) nicht möglich. Im Rahmen dieser Arbeit wurde stattdessen die folgende Distanzfunktion $d : D \times D \rightarrow \mathbb{R}$ verwendet:

$$d(m_{j,k}, m_{j,l}) = \beta\gamma\sqrt{(XPos(m_{j,k}) - XPos(m_{j,l}))^2 + (YPos(m_{j,k}) - YPos(m_{j,l}))^2},$$

wobei $XPos : D \rightarrow \mathbb{R}$ die zu einer Fingerabdruck-Minutie $m_{j,k}$ gehörende X-Koordinate des Fingerabdruck-Bilds und analog $YPos : D \rightarrow \mathbb{R}$ die Y-Koordinate des Fingerabdruck-Bilds bestimmt. Die Distanz zweier Fingerabdruck-Minutien berechnet sich somit durch Multiplikation der Koeffizienten β und γ mit der euklidischen Distanz der Minutien-Koordinaten im zweidimensionalen Raum. Dabei erzwingt der Koeffizient β die Forderung, dass alle Fingerabdruck-Minutien eines Clusters $C_{j,i}$ von unterschiedlichen daktyloskopischen Experten erzeugt wurden, d.h.

$$\beta = \begin{cases} +\infty, & \text{falls } ExpId(m_{j,k}) = ExpId(m_{j,l}) \\ 1, & \text{falls } ExpId(m_{j,k}) \neq ExpId(m_{j,l}) \end{cases}.$$

Der Koeffizient γ dient zur Bestrafung unterschiedlicher Richtungswinkel zweier Fingerabdruck-Minutien $m_{j,k}, m_{j,l}$ und berechnet sich wie folgt: sei φ der innere Winkel (gemes-

sen in Grad) zwischen zwei Fingerabdruck-Minutien $m_{j,k}, m_{j,l}$ in Bezug auf deren Richtungswinkel, so ist der Koeffizient γ definiert als

$$\gamma = \begin{cases} +\infty, & \text{wenn } \varphi > \theta_2 \\ \varphi/\theta_1, & \text{wenn } \theta_1 < \varphi \leq \theta_2 \\ 1, & \text{wenn } \varphi \leq \theta_1 \end{cases} .$$

Hierbei bezeichnet θ_2 eine obere Grenze für einen akzeptablen Winkel φ und θ_1 eine untere Grenze, bei deren Überschreitung – innerhalb der oberen Grenze – eine lineare Bestrafung anhand des Ausmaßes der Winkeldifferenz φ und der unteren Grenzen θ_1 vorgenommen wird.

4.2.2 Interpretation der Eingabeparameter

Der Eingabeparameter ε dient zur Definition der ε-Nachbarschaft $N_\varepsilon(m_{j,k})$ einer Fingerabdruck-Minutie $m_{j,k}$ und ist eng verbunden mit der Distanzfunktion $\delta(\cdot,\cdot)$. ε dient somit als Grenzwert für die maximale Distanz zweier Fingerabdruck-Minutien $m_{j,k}, m_{j,l}$. Auf Basis der weiter oben vorgestellten Distanzfunktion $d(\cdot,\cdot)$ lässt sich ε in etwa als die Anzahl der Pixel interpretieren, die maximal zwischen zwei Minutien-Punkten $m_{j,k}, m_{j,l}$ liegen dürfen, sodass sie noch als zusammengehörig aufgefasst werden können. Eine Heuristik für die Wahl von ε stellt hierbei analog zu [LBT+09] der Wert $\varepsilon = \frac{W}{4}$ dar, wobei W den durchschnittlichen Abstand zweier benachbarter Papillarleisten beschreibt.

Der Parameter $MinPts$ dient ebenfalls zur Definition der ε-Nachbarschaft $N_\varepsilon(m_{j,k})$ einer Fingerabdruck-Minutie $m_{j,k}$ und spezifiziert die geforderte Dichte. Die Wahl des Parameters $MinPts$ kann zur Steuerung der Zuverlässigkeit bzw. Aussagekraft eines Clusters $C_{j,i}$ bzw. einer daraus abgeleiteten Ground-Truth Fingerabdruck-Minutie $gtm_{j,i}$ verwendet werden und sollte im Intervall $[1, nexp_j]$ liegen. Eine Wahl von $MinPts$ nahe bei 1 sorgt für potenziell weniger aussagekräftigere Cluster, da zwei hinreichend nahe beieinanderliegende Minutien $m_{j,k}, m_{j,l}$ zur Generierung eines Clusters ausreichen. Eine Wahl von $MinPts$ nahe $nexp_j$ sorgt sicherlich für stabile Cluster $C_{j,i}$ und damit Ground-Truth Fingerabdruck-Minutien $gtm_{j,i}$, resultiert jedoch in potenziell weniger Cluster und Ground-Truth Fingerabdruck-Minutien.

4.3 Bestimmung der Ground-Truth Fingerabdruck-Minutien

Nach erfolgter Clusterbildung unter Verwendung des DBSCAN Clustering-Verfahrens und der weiter oben eingeführten Distanzfunktion $d(\cdot,\cdot)$ sowie passender Parameter $\varepsilon, MinPts$ ist es notwendig, Ground-Truth Fingerabdruck-Minutien $gtm_{j,i} = (Type, XPos, YPos, Angle, Quality, Support)$ eines Fingerabdruck-Bildes P_j zu ermitteln, die als Repräsentanten der erzeugten Cluster $C_{j,i}$ angesehen werden können. Auf Basis der zu einem Cluster $C_{j,i}$ gehörenden Fingerabdruck-Minutien $m_{j,i,k} \in C_{j,i}$ $(k = 1, \ldots, nclmin_{j,i} = |C_{j,i}|)$ sowie der Anzahl ein Fingerabdruck-Bild P_j untersuchenden daktyloskopischen Experten $nexp_j$ lassen sich die Attribute der Ground-Truth Fingerabdruck-Minutien $gtm_{j,i}$ wie folgt bestimmen:

- $Type$ wird auf denjenigen Wert gesetzt, der eine 2/3 Mehrheit in der Minutienmenge $C_{j,i}$ konstituiert. Sollte es keine 2/3 Mehrheit geben, gilt $Type = 0$ (Unknown).
- $XPos = \sum_{k=1}^{nclmin_{j,i}} XPos(m_{j,i,k})/nclmin_{j,i}$, d.h. das arithmetische Mittel der X-Koordinaten.
- $YPos = \sum_{k=1}^{nclmin_{j,i}} YPos(m_{j,i,k})/nclmin_{j,i}$, analog zu XPos.
- $Angle$ wird als der Median aller in einem Cluster $C_{j,i}$ enthaltenen Richtungswinkel bestimmt.
- $Quality = \sum_{k=1}^{nclmin_{j,i}} Quality(m_{j,i,k})/nclmin_{j,i}$.
- $Support = nclmin_{j,i}$.

Hierbei beschreibt $Support$ die Anzahl der einem Cluster $C_{j,i}$ zu Grunde liegenden Fingerabdruck-Minutien und kann in Kombination mit dem Attribut $Quality$ als Kennzeichen für Güte und Zuverlässigkeit der Ground-Truth Minutie $gtm_{j,i}$ verwendet werden.

5 Evaluierung

Das automatisierte Evaluieren der erzeugten Ground-Truth Fingerabdruck-Minutien $gtm_{j,i}$ stellt eine große Herausforderung dar, da per Definition kein Referenzdatensatz zur Verfügung steht, der zur Evaluierung verwendet werden kann. Nichtsdestotrotz soll im Folgenden eine erste Bewertung des Verfahrens vorgenommen werden. Zur Evaluierung der Ergebnisse wurde ein Prototyp entwickelt, der das DBSCAN Clustering-Verfahren unter Verwendung der in Kapitel 4.2.1 beschriebenen Distanzfunktion sowie variierenden Eingabeparametern $\varepsilon \in [0, 5]$ und $MinPts \in [1, 4]$ auf die in Kapitel 2.1 beschriebenen Daten ($npic = 17$ Fingerabdruck-Bilder $P_1, \ldots, P_{npic}$) anwendet. Die übrigen Eingabeparameter wurden in diesem ersten Schritt als $\theta_1 = 20$ Grad und $\theta_2 = 90$ Grad fixiert. Tabelle 1 zeigt die Ergebnisse der Evaluierung. In den Spalten werden folgende Daten abgebildet:

Nr – Fortlaufende Nummerierung der Testläufe mit variierenden Eingabeparametern $Parm$.

$Parm$ – Die Eingabeparameter ε und $MinPts$.

$\bar{x}_{dist}$ – Die mittlere Distanz zwischen den Fingerabdruck-Minutien $m_{j,i,k}$ und Ground-Truth Minutien $gtm_{j,i}$ aller Cluster $C_{j,i}$ aller Fingerabdruck-Bilder P_j. D.h. $\bar{x}_{dist} = \frac{1}{npic} \sum_{j=1}^{npic} AvgDist_j$, mit $AvgDist_j = \frac{1}{ncl_j} \sum_{i=1}^{ncl_j} AvgDist_{j,i}$, wobei $AvgDist_{j,i}$ definiert ist als $AvgDist_{j,i} = \frac{1}{nclmin_{j,i}} \sum_{k=1}^{nclmin_{j,i}} d(m_{j,i,k}, gtm_{j,i})$.

$\bar{\sigma}_{dist}$ – Die zu $\bar{x}_{dist}$ gehörende mittlere Standardabweichung berechnet nach $\bar{\sigma}_{dist} = \frac{1}{npic} \sum_{j=1}^{npic} \sqrt{\frac{1}{ncl_j - 1} \sum_{i=1}^{ncl_j} (AvgDist_{j,i} - AvgDist_j)^2}$.

Nr	$Parm$	$\bar{x}_{dist}$	$\bar{\sigma}_{dist}$	$\bar{x}_{supp}$	$\bar{\sigma}_{supp}$	$\bar{x}_{ncl}$	$\bar{x}_{pnc}$	r_{miss}
1	$\varepsilon = 1, MinPts = 1$	0.000	0.000	2.154	0.393	43.941	0.756	0
2	$\varepsilon = 1, MinPts = 2$	0.000	0.000	3.154	0.281	6.500	0.946	1
3	$\varepsilon = 1, MinPts = 3$	0.000	0.000	4.170	0.097	1.889	0.980	8
4	$\varepsilon = 1, MinPts = 4$	0.000	0.000	5.000	0.000	1.000	0.984	14
5	$\varepsilon = 2, MinPts = 1$	0.970	0.876	3.744	1.803	72.588	0.308	0
6	$\varepsilon = 2, MinPts = 2$	1.114	0.830	4.683	1.731	44.000	0.473	0
7	$\varepsilon = 2, MinPts = 3$	1.111	0.796	5.197	1.455	27.647	0.618	0
8	$\varepsilon = 2, MinPts = 4$	1.108	0.808	5.662	1.070	16.824	0.744	0
9	$\varepsilon = 3, MinPts = 1$	1.613	1.246	5.243	2.310	62.941	0.143	0
10	$\varepsilon = 3, MinPts = 2$	1.693	1.223	6.164	1.867	48.294	0.229	0
11	$\varepsilon = 3, MinPts = 3$	1.648	1.008	6.511	1.559	41.294	0.308	0
12	$\varepsilon = 3, MinPts = 4$	1.642	1.077	6.807	1.267	34.118	0.410	0
13	$\varepsilon = 4, MinPts = 1$	1.971	1.442	6.146	2.259	56.294	0.089	0
14	$\varepsilon = 4, MinPts = 2$	2.029	1.456	6.751	1.841	48.588	0.137	0
15	$\varepsilon = 4, MinPts = 3$	2.009	1.435	7.057	1.487	44.294	0.185	0
16	$\varepsilon = 4, MinPts = 4$	1.976	1.379	7.341	1.152	39.471	0.263	0
17	$\varepsilon = 5, MinPts = 1$	2.195	1.646	6.599	2.165	53.471	0.065	0
18	$\varepsilon = 5, MinPts = 2$	2.237	1.659	7.108	1.724	48.294	0.097	0
19	$\varepsilon = 5, MinPts = 3$	2.228	1.662	7.326	1.519	45.471	0.132	0
20	$\varepsilon = 5, MinPts = 4$	2.182	1.550	7.514	1.298	42.235	0.183	0

Tabelle 1: Ergebnisse der Evaluierung der erzeugten Cluster und Ground-Truth Minutien unter Anwendung des DBSCAN Clustering-Verfahrens mit unterschiedlichen Parametern $\varepsilon, MinPts$ und festen Parametern $\theta_1 = 20, \theta_2 = 90$.

$\bar{x}_{supp}$ – Der mittlere Support pro Cluster $C_{j,i}$ pro Fingerabdruck-Bild P_j berechnet nach: $\bar{x}_{supp} = \frac{1}{npic} \sum_{j=1}^{npic} AvgSupp_j$, mit $AvgSupp_j = \frac{1}{ncl_j} \sum_{i=1}^{ncl_j} nclmin_{j,i}$.

$\bar{\sigma}_{supp}$ – Die zu $\bar{x}_{supp}$ gehörende Standardabweichung $\bar{\sigma}_{supp}$ berechnet nach $\bar{\sigma}_{supp} = \frac{1}{npic} \sum_{j=1}^{npic} \sqrt{\frac{1}{ncl_j - 1} \sum_{i=1}^{ncl_j} (nclmin_{j,i} - AvgSupp_j)^2}$.

$\bar{x}_{ncl}$ – Die mittlere Anzahl detektierter Cluster $C_{j,i}$ pro Fingerabdruck-Bild P_j. D.h. $\bar{x}_{ncl} = \frac{1}{npic} \sum_{j=1}^{npic} ncl_j$.

$\bar{x}_{pnc}$ – Der mittlere Anteil der keinem Cluster zugeordneten Minutien, d.h. der als Rauschen interpretierten Minutien, $m_{j,k}$ aller Fingerabdruck-Bilder P_j berechnet nach $\bar{x}_{pnc} = \frac{1}{npic} \sum_{j=1}^{npic} |N_j|/|P_j|$, wobei $|P_j|$ die Anzahl aller von daktyloskopischen Experten detektierten Minutien $m_{j,k}$ eines Fingerabdruck-Bilds P_j und $|N_j|$ die Anzahl der keinem Cluster $C_{j,i}$ zugeordneten Minutien $m_{j,k}$ eines Fingerabdruck-Bildes P_j bezeichnen. D.h. $N_j = \{m_{j,k} \in P_j | \forall_{i=1}^{ncl_j} m_{j,k} \notin C_{j,i}\}$.

r_{miss} – Die absolute Anzahl der Fingerabdruck-Bilder P_j für die bei gegebenen Parame-

tern $\varepsilon, MinPts$ kein einziger Cluster $C_{j,i}$ gefunden werden konnte, d.h. $r_{miss} = |\{P_j | ncl_j = 0\}|$.

Die Durchführung der vorläufigen Evaluierung zeigt, dass alle generierten Cluster $C_{j,i}$ valide sind, d.h. dass alle Minutien $m_{j,i,k} \in C_{j,i}$ von unterschiedlichen daktyloskopischen Experten generiert wurden. Lediglich bei der Wahl von $\varepsilon = 1$ und $MinPts \in \{2, 3, 4\}$ konnten für einige Bilder P_j keine Cluster gebildet werden, was eine Wahl von $\varepsilon > 1$ nahelegt. Aus Tabelle 1 lässt sich entnehmen, dass ε und $\bar{x}_{dist}$ positiv korreliert sind (was auf Grund der Definition des Clustering-Verfahrens leicht einzusehen ist). Hieraus lässt sich schließen, dass die Wahl von ε zum Erreichen guter Ergebnisse nach oben beschränkt werden sollte. Ferner zeigt die Evaluationsreihe in Tabelle 1, dass bei Fixierung des Eingabeparameters ε der Parameter $MinPts$ mit $\bar{x}_{ncl}$ negativ und $\bar{x}_{pnc}$ positiv korreliert ist. D.h. bei zunehmendem $MinPts$ sinkt die Anzahl der bzgl. $\varepsilon, MinPts$ gefundenen Cluster $C_{j,i}$, gleichzeitig steigt der mittlere Anteil der keinem Cluster zugeordneten Fingerabdruck-Minutien. Diese Beobachtung spricht für die Wahl eines kleinen Parameters $MinPts$. Hierbei muss jedoch sichergestellt werden, dass die Qualität bzw. der Support der gefundenen Cluster $C_{j,i}$ hinreichend hoch ist.

Beschränkt man zur Wahl der geeigneten Parameter $\varepsilon, MinPts$ die zulässige mittlere Distanz $\bar{x}_{dist} < 2.0$ sowie den mittleren Anteil keinem Cluster zugeordneter Fingerabdruck-Minutien $\bar{x}_{npc} < 0.25$, so ergeben sich die zulässigen $\varepsilon, MinPts$-Kombinationen: $9, 10$ und 13 (siehe Tabelle 1). Auf Grund der großen Anzahl durchschnittlich gefundener Cluster ($\bar{x}_{ncl} = 56.294$) mit durchschnittlich hohem Support ($\bar{x}_{supp} = 6.146, \bar{\sigma}_{supp} = 2.259$) sowie des sehr niedrigen mittleren Anteils keinem Cluster zugeordneter Fingerabdruck-Minutien ($\bar{x}_{pnc} = 0.089$) scheint die Wahl von $\varepsilon = 4, MinPts = 1$ in Kombination mit $\theta_1 = 20, \theta_2 = 90$ das beste Ergebnis zu liefern.

6 Zusammenfassung und Ausblick

Diese Arbeit skizzierte einen Ansatz zur Erzeugung von für die Durchführung semantischer Konformitätstests benötigter Ground-Truth Minutien unter Verwendung des DBSCAN Clustering-Verfahrens. Die beschriebene Methode verfügt über vier frei konfigurierbare Parameter $\varepsilon, MinPts, \theta_1, \theta_2$. Durch bestimmen geeigneter Eingabeparameter lässt sich das Verfahren je nach Bedarf zur Erzeugung vieler Ground-Truth Minutien mit potenziell geringem Support oder weniger Ground-Truth Minutien mit hohem Support und hoher Qualität optimieren. Das hier beschriebene Verfahren wird auf von daktyloskopischen Experten des Bundeskriminalamts manuell detektierten Fingerabdruck-Minutien angewandt.

Die Verarbeitung dieser Daten teilt sich in zwei Schritte: (1) das Erzeugen von Fingerabdruck-Minutien Clustern unter Verwendung von DBSCAN. Hierzu wurde eine geeignete Distanzfunktion definiert sowie die notwendigen Eingabeparameter auf den Anwendungskontext projiziert. (2) Das Erzeugen geeigneter Repräsentanten der in Schritt (1) gewonnenen Cluster. Diese Repräsentanten dienen als Ground-Truth Fingerabdruck-Minutien.

Erste Tests auf Basis der vorliegenden Daten ergeben vielversprechende Ergebnisse bei

der Wahl der Parameter $\varepsilon = 4, MinPts = 1, \theta_1 = 20$ Grad und $\theta_2 = 90$ Grad. Nichtsdestotrotz bleibt die Frage der plausiblen und aussagekräftigen Bewertung der gewonnenen Ground-Truth Daten offen, da diese Daten auf Grund Ihrer Definition als Referenzdatensatz nicht automatisiert auf Plausibilität bzw. semantische Konformität geprüft werden können. Aus diesem Grund sollte in einem weiteren Schritt neben der Durchführung weiterer Tests – insbesondere einem Vergleich mit dem in [LBT+09] beschriebenen Verfahren, der zum Zeitpunkt des Einreichens dieses Papiers nicht möglich war – auf größeren Datenbeständen der Frage der plausiblen und aussagekräftigen (automatisierten) Bewertung der von unterschiedlichen Verfahren erzeugten Ground-Truth Minutien-Kandidaten nachgegangen werden, um vergleichbare Ergebnisse zu erreichen und einen für die Standardisierung verwertbaren Referenzdatensatz erzeugen zu können. Ferner sollte die Wahl Fingerabdruck-Bild spezifischer Parameter $\varepsilon, MinPts, \theta_1, \theta_2$ evaluiert und Heuristiken zur automatisierten Bestimmung der Parameter erforscht werden. Eine erste Heuristik für die Wahl von ε wurde in Kapitel 4.2.2 gegeben.

Literatur

[BLT+09] Christoph Busch, Dana Lodrova, Elham Tabassi, Wolfgang Krodel und Martin Drahansky. Semantic Conformance Testing for Finger Minutiae Data. In *Proceedings of IEEE IWSCN 2009*, Seiten 17–23, 2009.

[EKSX96] Martin Ester, Hans-Peter Kriegel, Jörg Sander und Xiaowei Xu. A Density-Based Algorithm for Discovering Clusters in Large Spatial Databases with Noise. In *Proceedings of 2nd International Conference on Knowledge Discovery and Data Mining (KDD-96)*, Seiten 226–231, 1996.

[HS07] Olaf Henniger und Dirk Scheuermann. Minutiae Template Conformance and Interoperability Issues. In *Proceedings of the Special Interest Group on Biometrics and Electronic Signatures (BIOSIG) 2007*, Seiten 25–32, 2007.

[ISO05] ISO/IEC JTC1 SC37 Biometrics. ISO/IEC 19794-2 Information technology – Biometric data interchange formats – Part 2: Finger minutiae data. International Organization for Standardization, 2005.

[ISO09a] ISO/IEC JTC1 SC27 Biometrics. ISO/IEC 29109-2 Information technology – Conformance testing methodology for biometric data interchange formats defined in ISO/IEC 19794 – Part 2: Finger minutiae data. International Organization for Standardization, 2009.

[ISO09b] ISO/IEC JTC1 SC27 Biometrics. ISO/IEC FCD 24745 Information technology – Biometric information protection. International Organization for Standardization, 2009.

[LBT+09] Dana Lodrova, Christoph Busch, Elham Tabassi, Wolfgang Krodel und Martin Drahansky. Semantic Conformance Testing Methodology for Finger Minutiae Data. In *Proceedings of the Special Interest Group on Biometrics and Electronic Signatures (BIOSIG) 2009*, Seiten 31–42, 2009.

BIOSIG 2010

Further Conference Contributions

Fingerprint Recognition with Cellular Partitioning and Co-Sinusoidal Triplets

Jan Hirzel
Email: janhirzel@gmail.com

Daniel Hartung
Email: daniel.hartung@hig.no

Christoph Busch
Email: christoph.busch@hig.no

Abstract: In this fingerprint verification approach, a fingerprint image is divided into equally sized cells and the pattern is represented by a substitute resulting in a feature vector of fixed length. A related ISO standard recommends three different approaches for the selection of these. It suggests a cell size of approximately two ridges per cell. For the co-sinusoidal triplet approach for the retrieval of the spectral component this assumption was investigated. The influence of the cell size on the biometric performance was supported and additionally, a sound comparison method was implemented. To maintain a comprehensible evaluation the open Fingerprint Verification Competition (FVC) databases FVC2000 and FVC2002 were used.

1 Introduction

The biometric characteristic of fingerprints is widely used for verification and identification purposes. Fingerprint recognition became more and more popular through high distribution of fingerprint sensors and the convenience in use. Traditional approaches are based on the extraction of a few stable points (minutiae) that uniquely describe a fingerprint. There are alternative approaches to this method which do not rely on minutiae, like fingerprint correlation [BVG+00] or finger pattern comparison [Hup07]. This paper is focused on fingerprint ridge pattern comparison. The algorithm described is based on the information of the fingerprint ridges and not just singular unique points. There are numerous ways how the pattern is examined. One of the most prominent ones is described in the ISO standard of 2009 [fS06] and will be discussed herein. The approach investigated in this paper has the advantage that features are - unlike finger minutiae data - already in the form of a fixed length feature vector, which is required for further processing such as template protection schemes.

2 Quantized Co-Sinusoidal Triplets

The first step in the generation of a fingerprint template is the conversion of the fingerprint to spectral data. The steps suggested in the ISO standard [fS06] are as follows: Image Preprocessing, Cellular Partitioning and Spectral Component Selection. The image preprocessing step is optional but can have a great effect on the results, since poor image quality leads to missing and spurious features.

2.1 Image Preprocessing

A basic improvement for the problem of strongly varying image qualities that does not consume much processing power is the histogram normalization of the fingerprint image (as for example proposed by Alparslan and Fuatince [AF81]) which already leads to a more consistent image quality. Further enhancements include Gabor filtering [YLJF03], image segmentation [HJ04] or binarization and thinning [Tha03].

2.2 Cellular Partitioning

Figure 1: The tessellation of a fingerprint with an optional X- and Y-Offset.

Figure 2: Left: A cell taken from the fingerprint. Right: The approximation through the co-sinusoidal triplets. The δ parameter is defined as $\delta = d/\lambda \cdot 360$. Source: [fS06]

After image preprocessing, a tessellation of the fingerprint image has to be done. Therefore, the fingerprint will be divided by a grid into cells of size $w \times w$. There is no specification of the optimal size of the cells but in [fS06] it is suggested to have a maximum of approximately two ridges per cell. Therefore, the size depends on the resolution of the sensor. Vannfält and Åström [Vs06] suggested in their thesis a cell size of 5×5 pixels for a sensor resolution of 250ppi. In case of margins at the borders, an offset can be chosen to get maximum coverage of the fingerprint area. Cells that cannot be fully filled with image information (the rightmost and bottommost cells) will be discarded.

2.3 Spectral Component Selection

There are three different, already established methods to retrieve the spectral information from a cell: Quantized co-sinusoidal triplets (QCT), Discrete Fourier Transformation (DFT) and Gabor filters. There are numerous publications about the use of the DFT (like [Bra89], [Nus82] or [Win76]) as well as the use of Gabor filters (like Yang [YLJF03] or Huppmann [Hup07]) but only the master's thesis of Vannfält and Åström [Vs06] was concerned with QCT, even though an algorithm using QCT won the FVC of 2000.
QCT are based on the approximation of each finger pattern cell through a cosine triplet (θ, λ, and δ). As seen in Figure 2, the three parameters describe the angle of propagation, the wavelength and the phase. The range of the parameters can be restricted to a minimum and maximum to get the highest variance without repeating structures.
Angle of propagation θ: represents the directional information of a cell. The angle of propagation is measured perpendicular to the crest of the co-sinusoidal function. If the crest is parallel to the vertical axis x, the angle is 0 and it increases with counter-clockwise rotation. The interval $[0, \pi[$ describes all possible orientations of the function.
Wavelength λ: describes the quantity of ridges and the distance between them for one cell. The frequency f is directly related to this parameter since λ is defined as $\lambda = \frac{1}{f}$. The range of the frequency is $[0, \text{maximal spatial frequency}[$ where the maximal spatial frequency is the Nyquist frequency [Gre59]. For 2D signal processing the Nyquist frequency is equal to the length of the image diagonal divided by two.
Phase δ: describes the distance of the first crest to origin of the cell. It is specified in angular coordinates and therefore is in the interval of $[0, 360]$. It is defined as $\delta = \frac{d}{\lambda} \cdot 360°$, where d is the distance. The 2D co-sinusoidal function to approximate the cell is defined as follows:

$$\text{Cell}_{\theta,\lambda,\delta}(s,t) = \cos(P \cdot 2\pi \cdot f + \delta), \text{where } P = s \cdot \cos(\theta) - t \cdot \sin(\theta), \text{and } f = \frac{1}{\lambda} \quad (1)$$

The parameters s and t of the function describe the position of each pixel inside the cell. The valid interval for s and t is $s = [1, \text{image width}]$ and $t = [1, \text{image height}]$ where $s, t \in \mathbb{N}$. The resulting values of the function will be quantized and are accurate enough to reconstruct the ridges of the fingerprint depending on the precision for each parameter. The amount of possible values of the quantization (bit-depth) depends highly on the cell size. The smaller the cells, the less information is necessary to approximate a cell adequately. Therefore, it is required to find a suitable bit-depth depending on the resolution of the cells (see 3 for an example). The substitution through the function automatically leads to a tolerance for errors in each cell (noise in the fingerprint image) and therefore it is very important to correctly estimate the bit depth.

To select the most suitable triplet to approximate a cell, the following approach is chosen. First a normalization of the fingerprint values to the range [-1, 1] is done. Then, the distance between the fingerprint cell and all possible synthetic cells (candidates) is calculated and the synthetic cell structure with the minimum distance is used to represent the information of the cell. Here, the Euclidean distance function is used to determine the resemblance of two cells, even though different distance functions, like hamming distance can also be used.In the case that there are more than one cell with the same distance, the

Figure 3: On the left is the original fingerprint image, on the right is the synthetic resemblance using the previously gathered triplets (cell size: 14). Source: FVC2000 DB_1a

following prioritization shall be employed. The triplet with the lowest frequency (δ) has the highest priority, then the triplet with the highest wavelength and finally the triplet with the lowest angle of propagation.
The number of the possible candidates depends on the bit-depth for the parameters and is 2^{l+m+n} where l, m and n are the the bit depths for θ, λ, and δ. The specific values are defined through an equidistant distribution between zero and the respective maximum values. For θ the maximum is 180, for λ the maximum is the Nyquist frequency and for δ the maximum is 360.

3 Experiments

In order to investigate the influence of the cell size, the False Acceptance Rate (FAR), the False Reject Rate (FRR) and the Equal Error Rate (EER) for cell sizes in the range of 5×5 pixels up to 18×18 pixels were studied. The comparison algorithm used is based on the similarity of the cell triplets in the reference and the probe. Each cell triplet of the probe will be compared to the corresponding cell triplet in the reference. A comparison score will be calculated depending on the similarity of all cell pairs. With an increasing score the probe resembles the reference better, therefore it is a similarity score.

In order to allow a positive comparison when noise is present in the images, a certain difference is acceptable. This is taken into account by matching cells only if their similarity is above a certain threshold. One problem that occurs is that occasionally, some cells score high enough to be above the threshold even though the surrounding cells do not match. In order to reduce the errors introduced by these outliers, the neighborhood around the current cell is taken into account by giving it a higher score if the surrounding cells match as well.

All possible combinations of the three parameters were considered as well during the tests but did not lead to better results. The test images were preprocessed by the VeriFinger SDK 6.0. Different image enhancement filters followed by a binarization were applied

and the orientations and positions of the cores were extracted. The images were then aligned on the cores to overcome translation and orientation problems. Images with no cores present and images that VeriFinger could not process were excluded from the test. Thus, the fingerprint set was reduced by 18% to a total of 656 prints from originally 800 for FVC2000 DB2_b and by 8.75% to a total of 730 of 800 prints for FVC2002 DB2_b. The chosen bit-depth for the angle of propagation θ was 5, for the wavelength λ 4 and for the phase δ 5.

4 Results

After conducting an evaluation on the dimensions for each parameter of the quantized co-sinusoidal triplet, the following results were discovered. The acquisition is based on the FVC2000 DB1_A and FVC2002 DB2_A. The results show the Equal Error Rates (EER) for each cell size. As can be seen in Table 1, the cell size that results in the least average error rate and therefore the optimal cell size for the FVC dataset DB1_a is 16.

Cell Size (in Pixels)	θ	λ	δ	Mean (for θ, λ and δ)
5	0.2254	0.4274	0.2082	0.2870
6	0.2145	0.4535	0.1798	0.2826
7	0.2211	0.4450	0.1595	0.2752
8	0.2255	0.4312	0.1503	0.2690
9	0.2420	0.3970	0.1482	0.2624
10	0.2262	0.3553	0.1461	0.2546
11	0.2176	0.3917	0.1545	0.2166
12	0.2108	0.2783	0.1606	0.2136
13	0.2019	0.2627	0.1766	0.2137
14	0.2056	0.2457	0.1869	0.2127
15	0.2059	0.2483	0.2460	0.2334
16	0.2101	0.2520	0.2703	0.2441
17	0.2205	0.2574	0.2780	0.2520
18	0.2231	0.2558	0.2849	0.2546

Table 1: The EER for θ, λ, δ and the mean for all three parameters

5 Conclusion

After conducting the evaluation of the different cell sizes using the quantized co-sinusoidal triplet approach, the assertion of the ISO standard [fS06] - to have approximately two ridges per cell - could be experimentally validated. The minimum possible ridge frequency is zero, which is present when a cell is of homogeneous intensity. This was observed very frequently with small cell sizes (around 5×5 to 6×6 pixels). It implies that the suggested average of two ridges per cell is violated. Large cell sizes lead to a rougher approximation

of the actual content of the fingerprint image and make the method less error prone. When looking at the results in Table 1, the cell sizes between 11×11 and 14×14 produce stable equal error rates at a stable level. The maximum ridge frequency for those cells sizes is 3 (defined by the Nyquist frequency). The optimal solution resulting in the best performance considering the EER was achieved with a size of 14×14 for the combination of all three parameters. A cross-check with the database FVC2000 DB_1a and FVC2002 DB_2a with different sensor properties shows, that the cell size has to be adapted specifically for the available image data.

References

[AF81] E. Alparslan and M. Fuatince. Image enhancement by local histogram stretching. *IEEE Transactions on Systems Man and Cybernetics*, 11:376–385, 1981.

[Bra89] R.N. Bracewell. The fourier transform. *Scientific American*, 260(6):86–95, 1989.

[BVG+00] A. M. Bazen, G. T. B. Verwaaijen, S. H. Gerez, L. P. J. Veelenturf, and B. J. van der Zwaag. A correlation-based fingerprint verification system. In *11th Annual Workshop on Circuits Systems and Signal Processing (ProRISC), Veldhoven, the Netherlands*, pages 205–213, Netherlands, November 2000. STW Technology Foundation.

[fS06] International Organization for Standardization. Biometric Data Interchange Formats - Part 3: Finger Pattern Spectral Data, 2006.

[Gre59] Ulf Grenander. *Probability and statistics: the Harald Cramér volume*. Almqvist and Wiksell, 1959.

[HJ04] L. Hong and A. Jain. Fingerprint enhancement. *Automatic Fingerprint Recognition Systems*, pages 127–143, 2004.

[Hup07] Markus Huppmann. *Fingerprint Recognition by Matching of Gabor Filter-based Patterns*. Technische Universität München, 2007.

[Nus82] HJ Nussbaumer. Fast Fourier transform and convolution algorithms. *Berlin and New York, Springer-Verlag(Springer Series in Information Sciences.*, 2, 1982.

[Tha03] R. Thai. Fingerprint image enhancement and minutiae extraction. *The University of Western Australia*, 2003.

[Vs06] A. Vannfält and A. Åström. *Fingerprint Spectral Matching - Quantized Co-Sinusoidal Triplets*. Chalmers University of Technology, 2006.

[Win76] S. Winograd. On computing the discrete Fourier transform. *Proceedings of the National Academy of Sciences of the United States of America*, 73(4):1005, 1976.

[YLJF03] J. Yang, L. Liu, T. Jiang, and Y. Fan. A modified Gabor filter design method for fingerprint image enhancement. *Pattern Recognition Letters*, 24(12):1805–1817, 2003.

Modular Biometric Authentication Service System (MBASSy)

Heiko Witte, Claudia Nickel
Hochschule Darmstadt*
heiko.witte@cased.de, c.nickel@fbi.h-da.de

Abstract: In this paper we present the design of a modular authentication system, which enables users to select an authentication procedure by preference.
A survey, carried out by N.L. Clarke and S.M. Furnell [CF05], proved the classical PIN-authentication to be inconvenient for many users. Passwords and PINs are either secure or easy to remember. Since humans tend to forget complex permutations of characters and numbers, the chosen secrets are often insecure.
The purpose of the system is to provide an infrastructure which allows the implementation and usage of alternative authentication procedures. This approach could lead to an increased acceptance of authentication on smartphones and thus an increased security of the devices. A prototype based on the concepts presented in this paper was implemented for the android operating system and a gait recognition module is being actively developed. Further modules like face recognition, voice recognition or graphical authentication schemes can be integrated which depicts the flexibility of the system.

1 Introduction

With the growing amount of smartphones in use, an increased demand in information security arises. The classical approach of user authentication relies on knowledge-based methods, with the PIN being the common implementation among these. Many users tend to forget passwords and PINs, which leads to increased efforts that the majority of users avoid. This decreases the security of the devices.
In this paper, we present a prototype of a modular authentication system for the android operating system. Authentication algorithms are outsourced into distinct application packages, which either require interaction with the user or run in the background. The system was developed to facilitate the deployment of biometric authentication algorithms. Biometrics provide alternative ways to authenticate a user, which may reduce user effort and thus increase the acceptance of authentication on mobile devices.
The modular design of the system allows the user to select the kind of authentication he prefers, thus encouraging the development of alternatives to the classical PIN authentication. The ability to activate multiple modules enables further extensions of the system.

*This work was supported by CASED (www.cased.de)

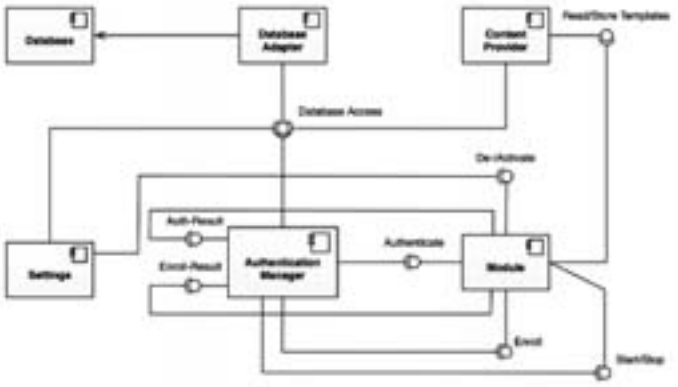

Figure 1: System overview

2 Components of MBASSy

The Modular Biometric Authentication Service System (MBASSy) was developed for the android operating system. Information on the android application fundamentals may be obtained from the extensive online documentation [Goo10]. All android specific terms in this paper are in *italics*. The system design follows most of the guidelines described in a publication by Saltzer and Schroeder [SS75]. The system consists of the following four main components: Database, Modules, User Interface, Background Service.

2.1 Modules

Modules contain a comparison function and the methods to obtain and process the user input. Two distinct types of modules are supported by MBASSy:

Background Modules run in a background process and do not require interaction with the user during the authentication process.

Foreground Modules require interaction with the user in all cases.

An interface shall be implemented to provide accessor methods for MBASSy. The interface is necessary to perform the following actions: ACTIVATE_MODULE, DEACTIVATE_MODULE, AUTHENTICATE, ENROLL, START (Background Modules), STOP (Background Modules).
When a module receives the activation action, a token is passed in the received *Intent*. This token must be used in every access request to the database. It shall be saved in the application package for later reference. If the token is lost, the module will lose access to its data.

When the enrollment action is invoked in the module interface, an authentication token is passed to the module in a data structure. The authentication token must be used in combination with the module token to gain write access in the *ContentProvider*.

2.2 Database

A database adapter class provides accessor methods for internal usage within the application package. A *ContentProvider* exposes a restricted version of those accessor methods to the modules. Read access is possible at any time, while write access is only granted in case of a pending enrollment request. This restriction was imposed on the modules to stay in control of the database size. In both cases, modules must supply their token and, in case of requesting write access, the authentication token which was supplied by the authentication service.
Read access is always available through a standard Java *InputStream* by constructing a URI of the following form:

```
content://com.cased.biometrics.provider.mbassy/ \
biometric_data/MODULE_TOKEN/TEMPLATE_ID
```

The template identifiers are part of the result set of a *ContentResolver* query. When requesting write access, the above URI is used without the appended template identifier. The required authentication token, as well as the user identifier, module token and user data are passed to the insert method of the *ContentResolver* in a *ContentValues* object. The user data shall be provided as a byte array. The *ContentProvider* takes care of writing the data to a file and storing the appropriate URI in the database.
All data is stored in the form provided by the module. The system does not encrypt or otherwise ensure the security or integrity of the data. Thus, modules shall provide ways to protect the data. The applicable protection algorithm may vary depending on the type of authentication algorithm. In the case of biometrics, biometric template protection is an appropriate way to ensure privacy [BBGK08].

2.3 User Interface

The user interface consists of the system preferences, account settings, module settings and system information overview. The module settings are an integral part of the system and will be discussed in detail.
When opening the module settings, a broadcast is sent to the operating system. The *Action* of the *Broadcast Intent* is **GET_MODULE_INFO**.
The *BroadcastReceiver* of a module is activated by the system and a response is generated by setting up an *Explicit Intent* which is sent back to the module settings. The following information of a module is passed in this *Intent*: Module name, Module type, Full package name, Full package path to the class implementing the required interface, Module token

(if available). The action of the *Intent* is **SET_MODULE_INFO**. If a module token is found in the database, it is displayed in the "Active Modules" section, otherwise in the "Available Modules" section. Modules can be activated, deactivated and prioritized in the module settings. The information on activated modules and their priorities is stored in the database. The priority is used by the database adapter to provide an ordered list of modules to the authentication service, according to the users preference.

2.4 Background Service

The background service constitutes the core of MBASSy. The service is responsible for managing authentication and enrollment requests. A broadcast receiver is used to take action upon receiving system broadcasts for the events "screen off", "screen on" and "battery low".
When the service receives a "screen off" event, a timer is started, which locks the device after a user defined period of time. When MBASSy is in a locked state and receives the system event "screen on", the background service fetches a list of active modules from the database and sequentially sends authentication requests to them.
MBASSy supports two authentication modes. The "Single Module Mode", requires a positive authentication result of a single module to successfully complete the authentication process. The "All Modules Mode" requires positive authentication results of all active modules in the system. In this early stage of development, modules return match decisions rather than comparison scores. The systems capabilities will be extended in the future by providing an authentication mode which accepts comparison scores as input and makes authentication decisions based on score level integration [BCP+04].
The background service stores log messages for every authentication or enrollment process in the database. This data may be used to learn about the system and user behaviour or the performance of modules. The background service performs the following tasks:

Authentication Requests An authentication request is performed by setting the respective *Action* of an *Explicit Intent*. The background service generates a unique authentication token, which shall be returned by the module when completing the request. The result is discarded if the returned authentication token does not match the one issued by the service. The data of an authentication result is also stored in an *Explicit Intent*, which contains at least the result as a boolean value, the authentication token and, in the case of a match, an identifier of the matched reference template as *Extra* values. Future versions of MBASSy will accept a comparison score as *Extra* value. The authentication process is depicted as a sequence diagram in figure 2.

Enrollment Requests Enrollment requests are issued automatically by the system when a new user account is created or when the user manually enrolls for a specific module in the account settings. The communication between MBASSy and the module is similar to the authentication process.
MBASSy stores the authentication token using the *SharedPreferences* of android, which are also accessible by the *ContentProvider*. A module shall pass the authenti-

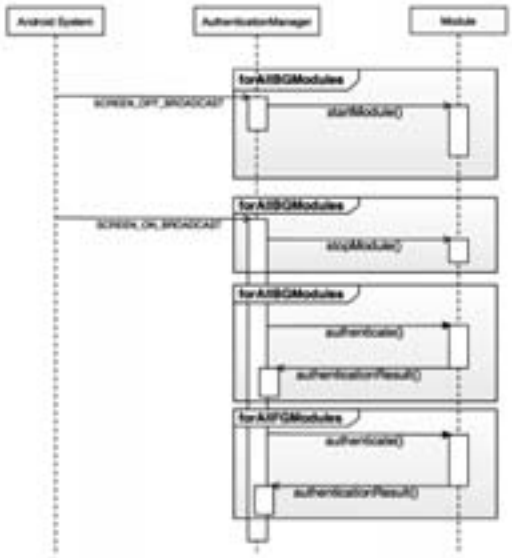

Figure 2: The authentication process

cation token as described in section 2.2. Write permission is denied without a valid combination of module token and authentication token. The enrollment process is depicted as a sequence diagram in figure 3.

Lock Screen Android provides no function for third party developers to securely lock a device. Due to this lack of functionality, a workaround was implemented. The workaround disables keys on the smartphone and replaces the home screen of android with a lock screen. As soon as Google releases a SDK which provides an official function to lock a device, it will be adopted in MBASSy.

3 Conclusion

Due to the flexibility of MBASSy, a combination of different authentication types is possible like e.g. knowledge-based authentication with biometric authentication. This could be a PassShape [WL08] module in combination with a gait recognition module. Fusion of multiple biometric modalities can also be accomplished by activating the respective authentication mode. Other use cases include the development of modules that collect biometric samples to create databases for testing purposes.
Authentication systems like the Local Authentication Subsystem (LASS) of Windows Mobile [Mic10] do not provide the functionality of MBASSy. Even though LASS supports the deployment of alternative authentication algorithms in a DLL, it does not support the activation of multiple modules.

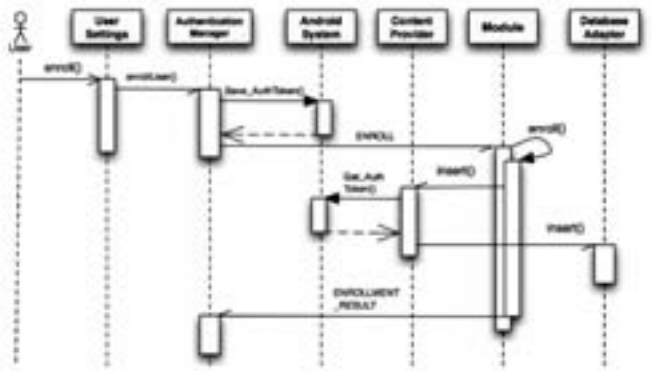

Figure 3: The enrollment process

References

[BBGK08] Jeroen Breebaart, Christoph Busch, Justine Grave, and Els Kindt. A Reference Architecture for Biometric Template Protection based on Pseudo Identities. *BIOSIG 2008 - Proceedings of the Special Interest Group on Biometrics and Electronic Signatures*, pages 25–38, 2008.

[BCP+04] Ruud M. Bolle, Jonathan H. Connell, Sharath Pankanti, Nalini K. Ratha, and Andrew W. Senior. *Guide to Biometrics*. Springer, 2004.

[CF05] N.L. Clarke and S.M. Furnell. Authentication of users on mobile telephones - A survey of attitudes and practices. *Computers Security*, pages 519–527, 2005.

[Goo10] Google. Android Application Fundamentals. `http://developer.android.com/guide/topics/fundamentals.html`, March 2010. 14.03.2010.

[Mic10] Microsoft. Local Authentication Subsystem (LASS). `http://msdn.microsoft.com/en-us/library/aa923670.aspx`, April 2010. 11.05.2010.

[SS75] J.H. Saltzer and M.D. Schroeder. The Protection of Information in Computer Systems. *Proceedings of the IEEE*, 63(9):pages 1278–1308, March 1975.

[WL08] Roman Weiss and Alexander De Luca. PassShapes: utilizing stroke based authentication to increase password memorability. *ACM International Conference Proceeding Series. Proceedings of the 5th Nordic conference on Human-computer interaction: building bridges*, 358:383–392, 2008.

Towards Best Practices for Biometric Visa Enrolment

Fares Rahmun[1], Dr. Sibylle Hick[2]

[1]German Federal Office of Administration (BVA)
Government Information Technology Solutions
Barbarastr.1, 50735 Köln
Fares.Rahmun@bva.bund.de

[2]secunet Security Networks AG
Biometrie und hoheitliche Dokumente
Kronprinzenstrasse 30, 45128 Essen
sibylle.hick@secunet.com

Abstract:
Public sector applications - e.g. the process of application of a biometric visa - are connected to different kinds of technical, organisational, and legal requirements. But how can those requirements and recommendations be retrieved, analysed and described? The German federal government has taken part in the European pilot project BioDEV II in order to gain comprehensive experiences regarding the handling and processing of biometric visa. Furthermore, the Federal Office for Information Security (BSI) publishes technical guidelines that address different topics in information security. In this context, the Technical Guideline Biometrics for Public Sector (TR Biometrics) has been published in order to describe technical and organisational requirements and recommendations in the context of biometrics for electronic identity documents. In this paper the authors give an overview of the pilot project BioDEV II and show how the results have been introduced in the TR Biometrics.

1 Introduction

Based on the European Regulation No 767/2008 [EC-767-2008] a central Visa Information System (VIS) is prepared in Europe which will introduce a new generation of short-stay visas. The application for a visa will be connected with biometrics in particular the acquisition of fingerprints. Thereby, this biometric data is transmitted and stored in the central VIS which is connected to a Biometric Matching System (BMS). Hence, different processes have to be considered and established starting from the secure acquisition of the biometric data to the later verification of the biometric features at border control. Within the scope of verification and identification not only questions regarding the different requirements and the quality of the captured biometric data have arisen but it must also be answered how the usability can be realised in an adequate way. By taking part in the European pilot project BioDEV II the German federal government has gained comprehensive experiences regarding the complexity of the processes not only on the organisational level but also regarding technical aspects.

The results of the BioDEV II project as well as the experiences made in Germany with the introduction of the new ePassports starting in 2005 have been gathered and analysed in several projects. In this paper the approach of BioDEV II will be. Afterwards the concept of Technical Guidelines for governmental identity documents in Germany and in particular the Technical Guideline "Biometrics for Public Sector Applications" (TR Biometrics) published by the German Federal Office for Information Security (BSI) will be introduced briefly. It will be shown how the results of the BioDEV II project have found their way as requirements, recommendations and best practices for the course of actions into this technical guideline that shall be undertaken if the afore described processes are realised.

2 Pilot Project BioDEV II

With the pilot project BioDEV II (Biometric Data Experimented in Visas) the European Commission has offered the possibility to gain experience with handling of visa by representing the complete and comprehensive process chain for the issuance of visa and the adjacent usage. Eight member states of the Schengen area took part in this pilot project. Those were Austria, Belgium, France, Germany, Luxembourg, Portugal, United Kingdom and Spain. The acquisition of fingerprint data and the decision if the captured biometric data has adequate quality have to be performed based on consistent rules and conditions. These are beside other time, usability and quality assurance. Thereby, the interaction between the different components and the involved roles has to be considered. As a result, the foundation to check an identity at border control is enabled. In the pilot project - from October 2007 to August 2009 - in particular the interactions of the software and hardware components have been observed by logging the results of the different components. In order to perform these tests a national VIS was installed and hosted by the German Federal Office of Administration (BVA). This national VIS was designed in respect to the central VIS/BMS. During the evaluation the processes for visa enrolment and border control were tested and analysed.

For the enrolment process two consular posts one in Syria (Damascus) the other in Mongolia (Ulan Bator) have been selected and these offices were equipped with a basic enrolment client software from October 2007 to December 2008. The acquisition of fingerprints with this software provided basic quality assurance mechanisms and additionally allowed the operators to make decisions regarding the fingerprint quality. The logging results showed that the quality of the captured fingerprints led to high rejection rates up to 82%. Therefore, the deployment of an enhanced enrolment client was decided. In order to evaluate a greater amount of properties two solutions of enrolment clients were provided. The properties comprised besides other hardware auto-capture within the fingerprint scanner vs. auto-capture as part of the software. Furthermore, the acquisition of multiple slaps of fingerprints has been supported and thereby composite records have been generated. The decision which fingerprint had to be included in the final set has been made based on independent quality assurance mechanisms as well as based on the mechanism of cross-matching. If applicable, the repetition of fingerprint acquisition has been necessary. In doing so, the acquisition of multiple slaps has been considered as well as the approach of optional single-fingerprint acquisition. By taking several slaps the possibility to lift the fingerprints after each capture or to permanently apply the fingerprints to the fingerprint scanner has been observed, too. Appliance of an open quality assurance software (NIST QA) together with segmentation on the one hand and the use of vendor specific quality assurance software together with segmentation on the other hand have also been considered.

The mechanisms that were applied in the enhanced enrolment clients had a great influence to the rejection rates. Accordingly, the rejection rates based on the Sagem Kit 4 went down in Damascus from 69% (phase 1) to 25% (phase 2) while the results changed in Ulan Bator from 82% (phase 1) to 43% (phase 2). Additionally, improvements within the scope of the quality assurance algorithm of Sagem Kit 4 lead to even better rejection results down to 3%. Nevertheless, the better results in phase 2 of BioDEV II have only been possible for a trade: the processes took 1 ½ to 3 times longer time than before. This fact shall not be underestimated since usability has shown to be a key element within the complete process.

Besides technical mechanisms furthermore organisational safeguards have been introduced and their impacts have been observed. This included acquisition guides (i.e. posters) that visualise how the fingers shall be placed on the sensor as information for the operators and a separate information poster for the applicants during enrolment. The same is true for training videos. Additional tools to improve image quality e.g. if a person has dry fingers has been provided. The displaying of the captured fingerprints through feedback monitors has also been conducted.

After the performance of the enrolment process the analysis of the verification process at border control was necessary to complete the test system solution. Here, in accordance with the enrolment two different German airports were selected to collect analysis data. Those were the airports in Schoenefeld and Tegel both located in Berlin. Verification for travellers holding a biometric visa can be divided into two positions. In the first position the properties time and usability play an important role in order to allow high throughput. This means that the live captured fingerprint images are only matched against the biometric data that is stored in the VIS for this person (1:1 verification). The second position shall be entered by a traveller if the verification has failed at first position. This means that the comparison of the live captured fingerprint images of the visa holder did not match against the corresponding biometric data stored in the VIS whereas the connection is made over the visa sticker number. In this case 1:1 verification as well as 1:n identification are provided and can be executed on request. In the second position up to ten fingerprints may be captured (acquisition of slaps of the left and right hand as well as the thumbs). In any case, the border control process anticipates that at first the visa sticker is read and the according biometric features are captured. This data for verification is than coded and sent to the national VIS. The comparison result is sent as response from the VIS. Again, the BioDEV II project had the agenda to evaluate different properties. In particular the verification time and the use of different quality assurance algorithms were analysed and different fingerprint sensors types were applied.

3 Provision of a Technical Guideline for Biometrics

The Federal Office for Information Security (BSI) has published several technical guidelines in different application areas in order to describe adequate security requirements and/or safeguards. With the introduction of electronic Passports in Germany commencing on November 2005 the specification of new processes regarding the application, processing and transmission of electronic and biometric data became necessary. The normative requirements were described in the Technical Guideline “zur Produktionsdatenerfassung, -qualitätsprüfung und -übermittlung für Pässe” [TRPDÜ].

Very fast after the introduction of the ePassports it became clear that the experiences gained within this scope were important not only for other governmental electronic identity documents such as the national identity card or visa but also for different kind of public sector applications (i.e. application profiles) like application of a new document or verification and identification processes at border control. This is in particular true for the hardware and software components that were already put in operation at the municipalities and shall be used again for further applications or with other governmental documents. Therefore, several technical guidelines were published which can be traced back or connected to the TR PDÜ and represent different parts. Those are:

- TR PDÜ hD [TRPDÜ]
  description of the business processes and the complete application process
- TR Biometrics [TRBIO]
  encapsulation of requirements and recommendations regarding biometrics
- TR XhD [TRXhD]
  XML representation for the exchange of application data
- TR SiSKo hD [TRSiSKohD]
  requirements regarding secure communication processes for governmental documents.

3.1 Technical Guideline Biometrics for Public Sector Applications

The processing of visa or the new national identity cards will be connected to biometric data of the document holder as described in section 2 or as it is already applied for ePassports. Therefore, it has been proven advantageous to specify uniform quality requirements and recommendations as well as interoperability requirements in order to be able to use the same hardware and software components in different public sector applications and to establish uniform conditions.

As a consequence three parts of the TR Biometrics (TR-03121) have been published (compare [TRBIO]). In the first part (TR-03121-1) the framework and concept of the guideline and how to apply it are introduced and described. In part two (TR-03121-2) the software architecture and Application Profiles depending on the governmental document are specified. Thereby, the software architecture is based on the BioAPI concept [ISO06] which allows an easy and flexible way to integrate further biometric modalities or capture devices. The Application Profiles give more information about the processes in relation to the biometric feature, the target groups and the relevant documents that shall be referenced (e.g. regulations, standards, technical descriptions, etc.). The requirements, recommendations, and best practices that are valid for a specific Application Profile are encapsulated in so called Function Modules. This has the effect that it allows a specific target group to select only those Function Modules that are relevant for their business processes. E.g. a provider for quality assurance software is interested in the requirements regarding fingerprint comparison while a provider of hardware components (e.g. a fingerprint sensor) needs to know the specific requirements regarding acquisition and biometric image processing. Thus, a mapping table is part of every Application Profile so that an overview of all Function Modules (i.e. requirements and recommendations) is given from which an instance of the target groups can choose from. The actual Function Modules are listed and

specified in part three of the technical guideline (TR-03121-3) depending of the biometric feature (fingerprints or facial image).

The guideline is designed in a way that it is easily possible to add new Application Profiles and it's connected Function Modules for a governmental document. With the results of the BioDEV II project extensive experiences regarding the enrolment and later usage (i.e. border control) have been made. In the following section it will be shown how these results have found their way into the TR Biometrics.

3.2 New Application Profiles in the TR Biometrics for Biometric Visas

Two main processes have been analysed in the BioDEV II project and have been depicted in section 2 of this paper:

- Enrolment for biometric visa and
- Performance of border control with biometric visa.

In order to transform the results regarding biometrics of the pilot to requirements and recommendations for the TR Biometrics at first, new Application Profiles "Application for Biometric Visa", "Basic Identity Check Biometric Visa", and "Extended Identity Check Biometric Visa" have been added in part two. The first listed Application Profile starts with a short overview of the current situation of the Schengen Member States regarding biometric visa. Afterwards it is shown how the instances of the biometric visa application process correlate to each other. Biometric data for a new visa is captured in the visa application office and then sent to the National Central Authority (NCA) which forwards the data to the central VIS respectively the BMS. The BioDEV II project has shown that quality assurance is very important and has a large effect on the overall process. Significant results can only be provided if adequate logging information is available. Therefore, a description which information shall be logged needs to be given. Within the Application Profile this is highlighted by the Biometric Evaluation Authority (BEA) and furthermore described in part 3 in a separate Function Module (FM Coding and FM Logging). Furthermore a short process overview how the fingerprint images are captured and a facial image is obtained is given. While the process for fingerprint acquisition is based on the BioDEV II results the description how a facial image can be provided is based on the experiences made with the application of a German ePassport. Finally, the mapping table shows which Function Modules are important for this specific Application Profile. Thereby, Function Modules are "reused" from former defined Application Profiles if the requirements and recommendations are true for the situation in hand and additional (new) Function Modules are added if the requirements are specific to the Application Profile "Application for a Biometric Visa".

The transformation of the requirements and recommendations - which are highlighted as Functions Modules in the afore mentioned mapping table - can be found in part 3 of the TR Biometrics. The Function Module Process (abbreviated with P-FP-VAPP) shows the balance that has been figured out between quality assurance and usability. Here as an example, the concept of composite records based on cross-matching has been adopted. Requirements for the acquisition hardware and software are chosen based on the requirements and recommendations that have been found in the context of the national identity card. Biometric image processing and thereby applying segmentation is based on the experiences made in BioDEV II. The Function Module

Coding (referenced with COD-FP-VAPP) describes the structure how the biometric data and the corresponding additional quality information are provided whereas the focus of the technical guideline is set on the quality values and logging information. While the purpose of logging is to collect specific data to understand the processes it is the objective of the evaluation to analyse the data based on specific evaluation methods. The requirements for evaluation are described in a separate Function Module (FM EVA-FP-VAPP). Safeguards for quality assurance have been tested with acquisition guides that visualise the correct positioning of the hands. This approach has been proven to be advantageous since it supports the official at the application counter. Therefore, these examples and the specification of the operation of devices and process requirements have been added besides further recommendations in the Function Module Operation (FM O-FP-VAA). The same is true for the user interface (FM UI-FP-APP).

The technical guideline containing the three named Application Profiles for biometric visa as well as the assigned Function Modules has been published in a developer version 2.1 on the web page of the Federal Office for Information Security.

4 Conclusions

The afore described derivation of requirements and recommendations from the pilot project BioDEV II into the Technical Guideline Biometrics for Public Sector Applications has shown that not only technical but also organisational safeguards have to be taken into account. Thereby, the different properties which are possible had to be evaluated against each other in order to describe processes that show the best balance between high quality and usability. Here, in particular the time factor is very important because high quality is connected to a longer enrolment time while a good usability requires fast and easy usage. But an adequate decision can only be made if comprehensive logging data is available that allows evaluation of the situation also in the case that errors might occur.

Bibliography

[EC-767-2008] Regulation (EC) No 767/2008 of the European Parliament and of the council of 9 July 2008 concerning the Visa Information System (VIS) and the exchange of data between Member States on short-stay visas (VIS Regulation).

[TRPDÜ] Federal Office for Information Security (BSI): Technische Richtlinie zur Produktionsdatenerfassung, -qualitätsprüfung und -übermittlung für Pässe (TR PDUE), Version 2.1.3, 2009.

[TRBIO] Federal Office for Information Security (BSI): Technical Guideline TR-03121, Biometrics for Public Sector Applications (TR Biometrics), Version 2.1, 2010.

[TRXhD] Federal Office for Information Security (BSI): Technische Richtlinie – XML-Datenaustauschformat für hoheitliche Dokumente (TR XhD).

[TRSiSKohD] Federal Office for Information Security (BSI): Sichere Szenarien für Kommunikationsprozesse im Bereichhoheitlicher Dokumente (TR SiSKohD), Version 1.3, 2009.

[ISO06] ISO/IEC 19784-1:2006 “Information technology – Biometric application programming interface – Part 1: BioAPI specification”.

Activity Related Biometrics based on motion trajectories

Anastasios Drosou[1,2], Konstantinos Moustakas[2], Dimos Ioannidis[2], Dimitrios Tzovaras[2]

[1]Imperial College London, SW7 2AZ London, UK {a.drosou09@imperial.ac.uk}

[2]CE.R.T.H. - I.T.I., P.O. Box 361 57001 Thermi-Thessaloniki, Greece

Abstract: The current paper contributes to the concept of activity-related biometric authentication in ambient Intelligence environments. The motivation behind the proposed approach derives from activity-related biometrics and is mainly focusing on everyday activities. The activity sequence is captured by a stereoscopic camera and the resulting 2.5D data are processed to extract valuable unobtrusive activity-related features. The novel contribution of the current work lies in the warping of the extracted movements trajectories, so as to compensate for different environmental settings. Authentication is performed utilizing both HMM and GMMs. The authentication results performed on a database with 32 subjects show that the current work outperforms existing approaches especially in the case of non-interaction restricting scenarios.

1 INTRODUCTION

Biometrics have recently gained significant attention from researchers, while they have been rapidly developed for various commercial applications ranging from surveillance and access control against potential impostors to smart interfaces. These systems require reliable personal recognition schemes to either confirm or determine the identity of an individual requesting their services. A number of approaches have been described in the past to satisfy the different requirements of each application such as unobtrusiveness, reliability, permanence, etc. Biometric methods are categorized to physiological and behavioral [JRP04], depending on the type of used features.

Behavioral biometrics, are related to specific actions and the way that each person executes them. They can potentially allow the non-stop (on-the-move) authentication or even identification in an unobtrusive and transparent manner to the subject and become part of an ambient intelligence environment. Behavioral biometrics are the newest technology in biometrics and they have yet to be researched in detail. They are supposed to be less reliable than physiological biometrics, however they are less obtrusive and simpler to implement [JRP04].

Recent work and efforts on human recognition have shown that the human behavior (e.g. extraction of facial dynamics features[HPL07]) and motion (e.g. human body shape dynamics during gait [ITD$^+$07]), when considering activity-related signals, provide the potential of continuous authentication for discriminating people .

Moreover, prehension biometrics belong to the general category of behavioral biometrics and can also been thought as a specialization of activity related biometrics [KCC02]. Ac-

tivity related signals have exhibited the potential to accurately discriminate between subjects, while they remain stable for the same subject over time [KCC02]. Moreover, they are targeting to the user convenience (unobtrusiveness) as well as to an optimal performance in various realistic environments (invariance).

In this concept, an interesting biometric characteristic can be the user's response to specific stimuli within the framework of an ambient intelligence (AmI) environment. The present paper extends the applicability of activity-related biometric traits [DMIT10] and investigates their feasibility in user authentication applications. Specifically, it deals with the major problem of small variances in the interaction setting, which are introduced by the arbitrary positioning of the environmental objects, in respect to the user, at each trial. Thus, a generic approach for coping with this issue is attempted through the utilization of a warping algorithm, whereby the the behavioural information of the movement is not affected at all.

The overall workflow of the system follows: The user is expected to act with no constraints in an ambient intelligence environment. Meanwhile, some events, such as the ringing of the telephone, the need for typing of a password to a panel or even an instant message for online chatting, trigger specific reaction from the user. The users movements are recorded by one stereo camera and the raw captured images are processed, in order to track the users head and hands.

Then, the feature extraction step follows, where the user-specific trajectories are processed. The proposed biometric system supports two modes: a)The enrollment mode, whereby a user is registered through the training of a Hidden Markov Model or a Gaussian Mixture Model. b)The authentication mode, where the HMMs or the GMMs evaluate the claimed ID request by the user, as valid or void.

The proposed algorithm has been tested and evaluated in a large proprietary database and considerable improvements in recognition performance are seen in comparison to the state-of-the-art methods. The effectiveness of the proposed system is demonstrated by two experiments, where the potential of a person authentication using the biometric signature of just one activity as well as the combination of two separate activities activities will be examined. Additionally, the comparison results between the two proposed statistical methods (HMM - GMM) are presented.

2 ACTIVITY RELATED FEATURE EXTRACTION

Lacquaniti et. al. has proved in [LS82] that the motion pattern for a given movement is considered consistent from trial to trial and independent of the movement speed assumption. Thus, we can claim that the trajectories of each body part for a given activity can be seen as a biometric pattern.

Just like the approach suggested in [DMIT10], the estimated positions of the head and the palms on each frame are used to describe the movement performed by the user (Figure 1a). Before the actual feature extraction, a series of normalization operations are applied to the trajectories. Given the depth information provided by the the disparity image, it is easy to acquire the 3D data from the 2D image. The following trimming that is performed on these signals is a 3-step procedure [DMIT10]. Additionally, the homogenization of the

extracted trajectories is further improved by resizing them to a fixed, a priori set vector length.

The acquired set of smooth trajectories (head & two palms) is shown in Figure 1b. These trajectories represent with high accuracy the movement of the corresponding body parts in XY axis, while the Z-axis (depth) is represented by the diameter of the circles.

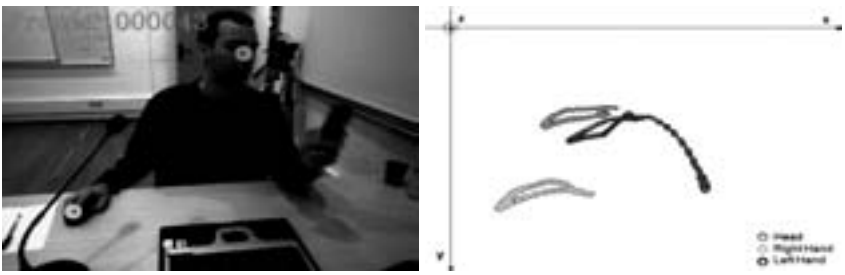

Figure 1: Set of trajectories describing a combined activity: Interaction with an office Panel during a phone conversation.

Spatial Warping A major issue is the invariance of the extracted trajectories even under small variations in the interaction setting between separate trials (different positions of the interaction objects) Normally, an increased False Rejection Rate (FRR) would apppear, due to variances in the interaction setting (object's positions, etc.) and not because the user not a genuine client for the system. In order to provide enhanced invariability to the extracted trajectories, the concept of *spatial warping* is introduced, inspired from the Dynamic Time Warping (DTW) method [SC90].

Without loss of generality regarding the environmental object, the case of a user answering successive phone calls will be studied. In the sequel, the relative distance between the user and the phone is not expected to remain fixed, either due to a shift of the user's body or due to small displacements of the phone. The same method can be applied to any environmental object with which the user is expected to interact (i.e. mouse, keyboard, pencil, book, etc.).

In a regular short phone conversation, there are two "extreme" positions of the hand that holds the telephone. Specifically, these can be seen in Figure 2 at point P_{Phone}, when the user has just grasped the phone just before he picks it up and at point P_{Head}, when the phone has touched the user's ear. The distance between these two "extreme" spots may vary even between the same user from trial to trial, since it depends on the slight variations of the environmental setting. Nevertheless, since we are mainly interested in the motion pattern of the trajectories and not its size, we apply the warping method on the hand trajectories.

Specifically, the exact location of these two points in the 3D space is automatically stored in the database for each user during the enrollment procedure. At the authentication step is warped according to the environmental characteristics of the enrollment moment.

In other words, the head-to-phone distance d is used for the deformation of an incoming set of trajectories, according to the claimed ID. Specifically, the blue line in Figure 2 indicates

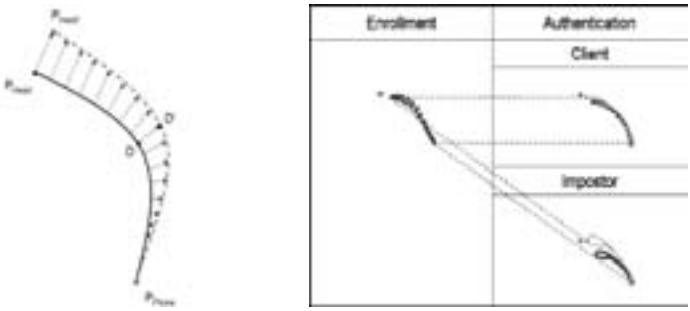

Figure 2: Normalization method for adjusting small variances in the head-to-phone distance.

the actual extracted trajectory in the authentication stage. $P_{Head'}$ and P_{Phone} are the stored locations of the user's head and the phone respectively obtained in the enrollment phase. The suggested method indicates that P_{Head} and $P_{Head'}$ as well as $P_{Phone'}$ and P_{Phone} are aligned, while all other points P_D of the XYZ signature in between are linearly transformed to the new point $P_{D'}$ as following:

$$P_{D'} = s(d)P(D) \tag{1}$$

where $s(d) = (P_{Head} - P_{Head'})d$ and $d = \sqrt{P_{Head}^2 - P_{Phone}^2}$.

In Figure 2b it can be noticed that the application of this method will the deformation (stretching/compression) of the trajectories to a common length. The advances of the suggested method towards enhanced invariancy can clearly be noticed in the ROC-curve diagrams, presented in Section 4.

3 Clustering & Classification

The training and verification step of our method has been tested under two statistical models. Namely, the Hidden Markov models and the Gaussian Mixture models have been mobilized and investigated, so as to perform authentication in the context of the proposed framework.

At the enrollment stage, both Hidden Markov Models (HMMs) and Gaussian Mixture Models (GMMs) are capable of clustering several sets $x_k(l_{head}, l_{rHand}, l_{lHand})$ of trajectories into a user specific signature. The Baum-Welch algorithm is utilized in the HMM case, while a weighted linear combination of $M = 5$ unimodal Gaussian densities and the Expectation-Maximization (EM) form the GMM. At the classification stage, on the other hand, the extracted set of trajectories is compared to the corresponding claimed signature, utilizing of the maximum likelihood criterion (HMM case) or Equation 2 (GMM case) and is finally classified as an impostor/client trajectory.

$$L(x|GMM_{n,k}) = \sum_{i=1}^{M} \log \sum_{j=1}^{L} \alpha_j \phi(j|\mu_j \Sigma_j) \tag{2}$$

whereby $M = 5$ denotes the five clusters of each GMM, α_j the weight factor of the j-th cluster, μ_j and Σ_j are the mean value and the variation of the distribution in the j-th cluster.

A user is only then authenticated, if the signature likelihood, returned by corresponding classifier, is bigger than an experimentally selected threshold.

4 EXPERIMENTAL RESULTS

The proposed methods were evaluated on the proprietary ACTIBIO-dataset [DMIT10], which consists of 29 regular subjects, performing a series of everyday office activities (i.e. a phone conversation, typing, talking to a microphone panel, drinking water, etc.) with no special protocol in 8 repetitions in total, equally split in two sessions.

Considerable improvements in the potential of recognition performance has been seen in comparison, after the application of the warping algorithm. Additionally, important outcomes have been extracted in terms of the applicability of the two utilized classifiers (Section 3) to the proposed system.

The proposed framework has been evaluated in the context of three verification scenarios. Specifically, the potential of the verification of a user has been tested, based on his a)activity-related signature during a *phone conversation*, b) activity-related signature during the *interaction with an office panel* and c) the *fusion at the score level* of the latter two activities. Specifically for scenario *c* the results from the activity *Phone Conversation* contributed with a factor of 0.2, while the the weight factor for the activity *Interaction with an Office Panel* is 0.8.

The evaluation of the proposed approach in an authentication scenario utilizes ROC-Curves and the corresponding equal error rates (EER) scores as shown in Table 1, whereby a noticeable improvement compared to original authentication capabilities of the system proposed in [DMIT10] can be seen.

Table 1: Authentication Performance - EERs.

EER	non-Warped			Warped	
	Phone	Panel	Fused	Phone	Fused
HMM	15%	9.8%	7.4%	12%	6.5%
GMM	25.2%	16.2%	15.3%	21.1%	14.9%

Given an HMM classifier, the original system could achieve an overall lowest EER of 7.4% in case there was a score level fusion in the results of two activities, namely a phone conversation and the user's interaction with the office panel. Each partial activity alone exhibited a much lower EER of about 15% and 9.8% respectively.

The evaluation of the warping method (Section 2) takes place in respect with the *phone conversation* activity, whereby the corresponding trajectory is warped for each user, according to the stored features of the claimed ID. There is a noticeable improvement with our approach, since there is a decreasing in the EER, of about 3%, compared to the simple approach. Given that the EER rate scores of the second activity remains untouched, the overall recognition performance of the system can achieve even lower EER scores, after tje score level fusion from these two separate activities.

On the other hand, the evaluation of the system, given a GMM classifier, can be characterized as rather worse in both cases (warped - nonWarped). Specifically, the overall EER lies

at 15.3%, while each partial activities exhibit even lower EERs (phone conversation:25.2%, office panel:16.2%).

The reason for this deterioration in the authentication capabilities of the system, when a GMM classifier is used can be found in the fact that GMMs fail to calculate the transition probability within the given signal. A GMM can be viewed as a single-state HMM with a Gaussian mixture density. This proves to be useful in some cases, in terms of complexity, processing load and control over the statistical model, however, in this case the overall recognition performance deteriorates. Still, the proposed warping method has improved the results even with the GMM classifier, compared to the simple approach.

5 CONCLUSION & FUTURE WORK

In this paper we presented an extension to an unobtrusive authentication method that is related to activity-related biometrics and includes the dynamic characteristics derived when performing everyday activities, as a response to specific stimuli. The trajectories extracted from each user are warped towards more invariant activity related features, which are less dependent on the environmental setting but still retain the behavioural information. Moreover, the comparison between hidden Markov models and Gaussian mixture models towards user recognition exhibited the superiority of the first ones. This can be explained by the GMMs failure to take into account the transition probability between different states within the same model. Obviously, proposed method can achieve very high rates of authentication performance and therefore comprises a very interesting approach for further research in activity-related biometrics.

References

[DMIT10] A. Drosou, K. Moustakas, D. Ioannidis, and D. Tzovaras. On the potential ofactivity-related recognition. In *The International Joint Conference on Computer Vision, Imaging and Computer Graphics Theory and Applications (VISAPP 2010)*, 2010.

[HPL07] Abdenour Hadid, Matti Pietikäinen, and Stan Z. Li2. *Learning Personal Specific Facial Dynamics for Face Recognition from Videos*, pages 1–15. Springer Berlin / Heidelberg, 2007.

[ITD+07] D. Ioannidis, D. Tzovaras, I. G. Damousis, S. Argyropoulos, and K. Moustakas. Gait Recognition Using Compact Feature Extraction Transforms and Depth Information. *IEEE Trans.Inf. Forensics Security.*, 2(3):623–630, 2007.

[JRP04] A. K. Jain, A. Ross, and S. Prabhakar. An Introduction to Biometric Recognition. *IEEE Trans. Circuits Syst. Video Technol.*, 14(1):4–20, 2004.

[KCC02] A. Kale, N. Cuntoor, and R. Chellappa. A framework for activity-specific human identification. In *IEEE Proc. International Conference on Acoustics, Speech, and Signal Processing (ICASSP)*, volume 4, pages 3660–3663, 2002.

[LS82] F Lacquaniti and J F Soechting. Coordination of arm and wrist motion during a reaching task. *The Journal of neuroscience : the official journal of the Society for Neuroscience*, 2(4):399–408, April 1982.

[SC90] H. Sakoe and S. Chiba. Dynamic programming algorithm optimization for spoken word recognition. *Readings in speech recognition*, 1990.

Increasing the Reliability of Biometric Verification by using 3D Face Information and Palm Vein Patterns

Olegs Nikisins, Modris Greitans, Rihards Fuksis, Mihails Pudzs, Zanda Serzane

Institute of Electronics and Computer Science
14 Dzerbenes Str., Riga, Latvia, LV-1006
olegs.nikisins@gmail.com
modris_greitans@edi.lv

Abstract: The improvement of the reliability for biometric image-based identification systems is discussed in this paper. Accuracy rising of the verification process is based on two main principles: implementation of multidimensionality and advanced image acquisition. The multidimensionality principle is based on two factors: the system structure (multimodality of biometric system) and the acquired biometric data (visual and spatial information about the object). The advanced image acquisition is based on multi-view principle and photography that is sensitive in infrared spectrum.

1 Introduction

Biometrical recognition process is mainly based on one parameter, such as face, fingerprint, hand, palm or iris, captured with single image sensor. However, mono-modal systems can be fooled by using fake parameters. Two methods of increasing the security level of the system are implementation of the multidimensionality and advanced acquisition of biometric data. Multidimensionality includes the combination of two or more biometric parameters in one system (multimodality) and diversity of single biometric parameter representation (intensity and shape descriptive arrays), but it is very important to choose the right parameters to achieve higher precision and easy enrolment. The advanced image acquisition makes unauthorized capturing of biometric data problematic and decreases identity stealing possibilities. The advanced imaging is based on two main approaches: multi-view photography and imaging in invisible light. In this paper we introduce the improvements of multimodal biometric system [BPM10] by implementing 3D facial information in addition with palm blood vessel pattern. This maintains previously described advantages and increases the security level of the system.

2. Basic idea of the proposed approach

The main idea of the proposed system focuses on reliability increasing of biometric verification process by introduction of multimodal approach and reliable biometric

parameters. These parameters include 3D facial information over the eye-line and palm vein patterns. Proposed system consists of passive stereo camera module [BBK05] (multi-view imaging principle) for 3D face information extraction, palm vein image acquisition block (infrared imaging) and the PC for cameras output data storage and processing. This system configuration is suitable both for biometric parameters extraction and for the recognition tasks, however only the first stage is described in this paper. Passive stereo camera system consists of two calibrated cameras [SHB08]. We describe the methodology of face contour extraction over the line between eye pupil centres, what can be used to separate face photos from real faces. This approach allows the extraction of 3D face information without the usage of active light sources (projector), what is comfortable for the user. Palm vein image acquisition system consists of infrared camera with additional optical filter and an IR light source. The filter and the light source with the wavelength of 850nm are selected, because the visibility of palm veins in our system is the best at this wavelength, what is established empirically. An effective algorithm for palm vein detection is described in later sections.

3. Eye pupil detection based on modified Hough transform

Accurate eye detection is an important biometrical feature extraction task [PZV05]. This section of the paper proposes the method for eye pupil detection. Determined eye pupils location is used for 3D facial information extraction over the eye line, what is important for the increasing of biometrical system's fake resistance. Proposed eye pupil detection method is based on modified Hough transform [Ba81]. Eye pupil has circle shape and the concept of Hough transform for circle detection provides good eye pupils segmentation results. Algorithm for eye pupil detection consists of following stages: captured image Gaussian low-pass filtering; edge detection with Sobel operator; Hough transform of the image for circle detection; resulting (transformed) pattern Gaussian low-pass filtering; detection of two maximums (coordinates of eye pupils) in the pattern.
The central idea of Hough transform for circle detection is to determine all potential circle points (image edges) and to transform them into parametric space. The parametric space is three-dimensional *(a,b,r)*, where coordinates *(a, b)* defines the centre of the circle and *r* is radius.

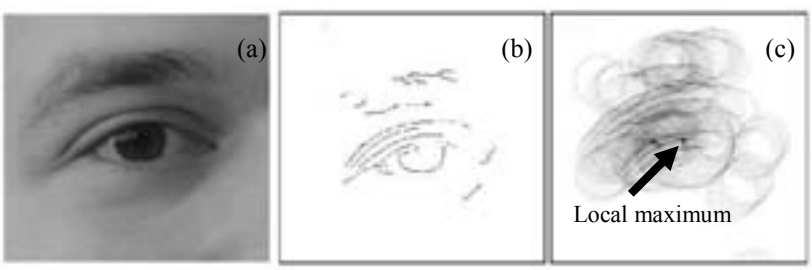

Figure.1: Hough transform for eye pupil detection. (a) Original image. (b) Edges detected with Sobel operator after low-pass filtering (c) Parametric space.

The main idea of our modified Hough transform is to exclude accumulator and to process parameter space (Figure.1 (c)) directly. This processing is performed with low-

pass filtering of the parametric space pattern, what results in space blurriness. The operation of local maximums detection is performed next to the blurred parametric space, what is the resulting stage of eye pupil detection.

4. 3D information extraction over corresponding line segments

One of the methods to increase reliability of face recognition system is to implement the use of 3D face information. However the main problem is to create precise and task - appropriate 3D image acquisition system. Offered solution is based on the "passive" stereo system [BBK05]. The operation of passive stereo imaging system is the most comfortable for the user, but this approach commonly requires complicated algorithms to resolve the *correspondence problem*, so we propose to use well – observable facial features for the calculation of 3D information. In details, points of face contour along the eye-line are selected as a parameter for 3D face information evaluation, what will let to segregate real faces from spoof objects (face photos).

Once eye pupil detection task is performed an equation of the line between eye pupil centres can be found. This allows to find brightness values along the eye-line (Figure.1).

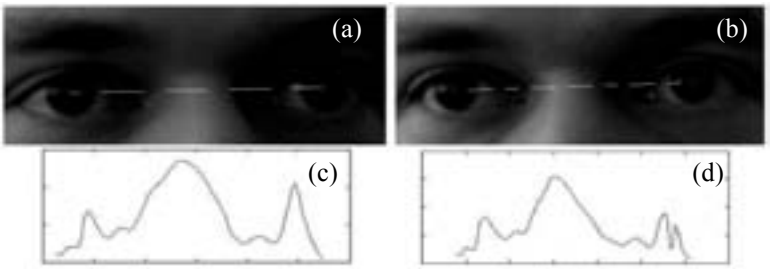

Figure.2: images form stereo system. (a) Image from left camera with detected eye-pupils and corresponding eye-line. (b) Image from right camera with detected eye-pupils and corresponding eye-line. (c), (d) Intensity along the eye-line for left and right camera images

Intensity signals are used to resolve a correspondence problem. Algorithm for corresponding point detection consists of following stages: Gaussian low-pass filtering of the signal (Figure.2, (c) and (d) signals); signal rationing to same maximum and minimum values; extremum calculation for both signals; locating corresponding extremums in both signals. Once corresponding image points are located we can compute the scene points using the **triangulation** method [SHB08]. The solution of the triangulation task for corresponding points gives their 3D coordinates, however only 2 coordinates (Figure.3) are used in our approach: horizontal position of the point (location along x axis) and the distance from points to the stereo system (location along y axis). Information about locations of the eye - line points (Figure.3) is next used for object shape estimation. We use distance **d** from point to line between eye centres, as an estimation parameter. Obviously, that distances **d** are small, if face photo is placed in front of biometric recognition system, so the implementation of this parameter into recognition process will increase fake resistance of the system.

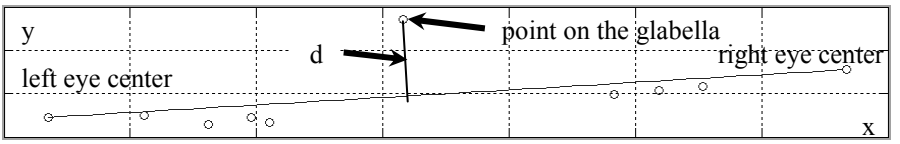

Figure.3: location of points detected along the eye-line in x, y coordinates

5. Palm vein image acquisition in infrared light

The idea of palm vein imaging is based on absorption of infrared (IR) light [LL06]. Haemoglobins in blood have different absorption properties. Absorption maximum within the near-infrared (NIR) area for deoxygenated haemoglobin is 760 nm, but for oxygenated is 850nm. Thus, palm veins appear darker than the surrounding tissue when captured in IR light.

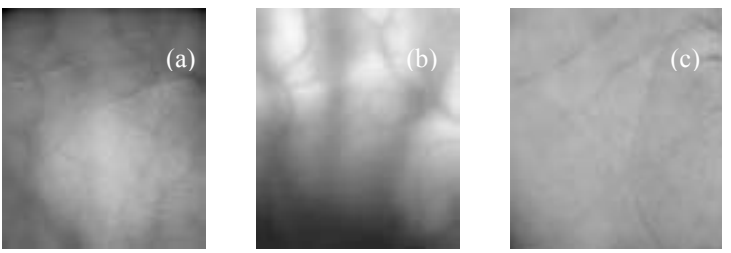

Figure.4: Infrared palm images captured with reflection (a) and transmission (b) methods. Palm image in visible light (c).

Infrared palm images can be obtained by two approaches – reflection and transmission method [FGNP10]. Based on previous experiments [FGNP10] it is effectively to choose reflection method, because it requires less power for the light source and reduces system dimensions in comparison with transmission setup. Blood vessel images captured with described methods are displayed in Figure.4. In the image captured with reflection method palm vein layout is visible and could be considered as an additional biometric parameter for person recognition; in image acquired with transmission approach two parameters are visible: bone structure and palm veins. Both parameters detected with transmission method are useful for reliability increasing, because vein layout is excellent uniqueness identifier, but bone structure could be used for aliveness detection. An infrared light source and IR filter with central wavelength of 850 nm has been selected for our system.

6. Image pre-processing. Complex matched filtering

To perform palm blood vessel extraction task complex matched filtering (CMF) approach was developed. Method is based on two-dimensional matched filtering (MF) [CSKNG89]. CMF was introduced in [GPF09] for feature detection in images with structured objects. This method improves computational efficiency and provides

additional information about the object, such as vessel orientation. CMF filters an image with one complex mask, which incorporates all necessary angles and scales, what is an advantage in comparison to MF approach. The kernel of complex matched filter is defined by following equation:

$$CMF(x,y) = \sum_{m=1}^{M} \sum_{l=0}^{L-1} \exp(j2\varphi_n) G(x,y,\varphi_l,c_m),$$

where c_m is scaling factor, M – total number of used scales, L – total number of used angles, $\varphi_l = l\pi / L$ and $G(x,y,\varphi_l,c_m)$ - Gaussian mask for matched filtering.

CMF output is a matrix of vectors. Vector magnitudes illustrate matching intensity, while angles characterize the orientation of blood vessel. These vectors might be used for person identification and is a good biometrical parameter for highly reliable biometric system design, due their advanced acquisition.

6. Experimental results

We use an average distance from points to line between eye centres (Figure.3), as an estimation parameter. Usage of 7M pixel digital cameras for stereo system setup provides following experimental results: average distance is less than 1 (mm) for face photo images, and more than 2.5 (mm) for real faces (notes: distance from stereo system to the object is less than 70 cm; distance between optical centres of two cameras is 15 cm). Difference in limits of average distances is not significant, what could be explained with "relatively flat" face contour over the eye-line.

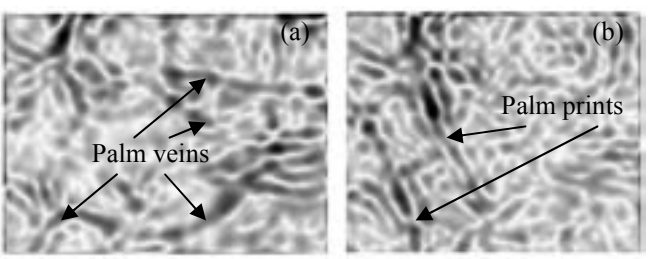

Figure.5: Palm vein image complex matched filtering. (a) Image is acquired in infrared light. (b) Image is acquired in visible light

Infrared palm vein imaging system with CMF algorithm has been tested on a database of 400 images from 50 persons (8 from each person). Figure 5 shows typical complex matched filtering result of palm images acquired in infrared and visible light for the same person. The segmentation of biometric features will obviously provide different results for visible light image (Figure.5, a) and IR image (Figure.5, b), what is excellent parameter for aliveness detection. When real hand is replaced with palm photo, the result of complex matched filtering for images acquired in visible and infrared light will be the same. This fact is useful for the purpose of fake object rejection. The combination of visible and infrared palm imaging techniques is a powerful tool for aliveness detection and reliability increasing.

7. Conclusions

The implementation of advanced imaging and multidimensionality factors into biometric recognition system benefits in increased reliability of the result. Information about the face shape and palm vein layout adds to the system such decidable factors as fake resistivity and ability of aliveness detection. Proposed methodology for 3D data extraction could be extended for the reconstruction of the whole face, by simple line transfer along the face in parallel to the eye-line, what is an effective solution of "correspondence problem". The complex matched filtering of the palm vein image decreases computation complexity and provides extra information for the recognition stage. Our future research will be focused on the explication of 3D face reconstruction idea, its combination with IR palm imaging and the development of robust recognition algorithms in order to develop reliable multimodal biometric authentication system.

8. Acknowledgments

This research is supported by the Latvian State research program in innovative materials and technologies and ESF project Nr. 1DP/1.1.1.2.0/09/APIA/VIAA/020.

Bibliography

[Ba81] Ballard, D.H.: Generalizing the Hough Transform to Detect Arbitrary Shapes. Pattern Recognition, Vol.13, No.2, 1981; pp. 111-122

[BBK05] Bronstein, A.; Bronstein, M.; Kimmel, R.: Three-Dimensional Face Recognition. International Journal of Computer Vision 64(1), 2005; pp. 5-30

[BPM10] Boulgouris, N.; Plataniotis, K.; Micheli-Tzanakou, E.: Biometrics, theory, methods and applications. Wiley, USA, 2010

[CSKNG89] Chaudhuri, S.; Shatterjee, S.; Katz, N.; Neleson, M.; Goldbaum, M.: Detection of Blood Vessels in Retinal Images Using Two-Dimensional Matched Filters. IEEE Transactions on medical imaging. Vol. 8, No. 3, 1989; pp. 263-269

[FGNP10] Fuksis, R.; Greitans, M.; Nikisins, O.; Pudzs. M.: Infrared imaging system for analysis of blood vessel structure. Electronics and Electrical Engineering, 1(97), 2010; pp. 45-48

[GPF09] Greitans, M.; Pudzs, M.; Fuksis, R.: Object Analysis in Images using Complex 2D Matched Filters. Proceedings of the IEEE Region 8 Conference EUROCON 2009; pp. 1392-1397

[L96] Lutkepohl, h.: Handbook of Matrices. John Wiley & Sons, Chichester, England, 1996

[LL06] Lingyu, W.; Leedham, G.: Near- and Far- Infrared Imaging for Vein Pattern Biometrics. Proceedings of the IEEE international Conference on Video and Signal Based Surveillance (AVSS'06), 2006

[PZV05] Park, K.; Zhang, H.; Veznevets, V.: Image-Based 3D Face Modelling System. EURASIP Journal of Applied Signal Processing, 13, 2005; pp. 2072-2090

[Ro06] Roberts, C.: Biometric Technologies – Palm and Hand. Link: http://www.ccip.govt.nz/newsroom/information-notes/2006/biometrics-technologies-palmhand.pdf, 2006

[SHB08] Sonka, M.; Hlavac, V.; Boyle, R.: Image Processing Analysis, and Machine Vision. Cengage Learning, USA, 2008

User Survey on Phone Security and Usage

Frank Breitinger, Claudia Nickel

Hochschule Darmstadt*
frank.breitinger@stud.h-da.de, c.nickel@fbi.h-da.de

Abstract: Mobile phones are widely used nowadays and during the last years developed from simple phones to small computers with an increasing number of features. These result in a wide variety of data stored on the devices which could be a high security risk in case of unauthorized access. A comprehensive user survey was conducted to get information about what data is really stored on the mobile devices, how it is currently protected and if biometric authentication methods could improve the current state. This paper states the results from about 550 users of mobile devices. The analysis revealed a very low securtiy level of the devices. This is partly due to a low security awareness of their owners and partly due to the low acceptance of the offered authentication method based on PIN. Further results like the experiences with mobile thefts and the willingness to use biometric authentication methods as alternative to PIN authentication are also stated.

1 Introduction

The number of mobile phone users worldwide exceeded the mark of 4 billion last year for the first time; this means that two-thirds of the world's population use mobile phones. Especially in industrialized countries the trend is strongly towards increased usage of mobile data services [Bit09]. With growing availability of data tariffs and new functionalities, mobile internet and e-mail on mobile phones have become accessible to the masses. Modern mobile phones also allow photography, have integrated calendars or can be used as a notepad – with consistently small size. These factors have led to an increase in different kinds of sensitive data stored on the mobile phone which makes them even more attractive to thieves. [Fla06] states that 800,000 inhabitants of England and Wales have been victim of mobile phone theft between mid of 2005 and mid of 2006. This documents the need for protecting the stored information by applying secure and user-friendly authentication methods which could e.g. be provided by using biometrics.
This paper summarizes the results of a comprehensive survey about the security and usage of mobile phones. In order to reach a large and diverse group of people, the survey was realized as a printed survey and as an online questionnaire which was available for about six weeks in April and May 2010. Promotion has been made on social networks like facebook[1], different mailing lists, a gym and a physiotherapy.

*This work was supported by CASED (www.cased.de).

[1] www.facebook.com

2 Analysis

A total of 548 people took part in the survey, of which about 11% filled out the paper questionnaires. The survey has been in German. Age and gender distribution are given in table 1.
It can be seen that the majority of participants (55%) was between 18 and 30 years old. There has been a significant difference between the age of the online participants (60% were between 18 and 30 years) and the people handing in the paper questionnaires, where nearly 75% of the respondents were older than 41 years. There is no significant difference between the number of males and females through all age groups.

	< 18	18-23	24-30	31-40	41-50	51-60	> 60	unknown	sum
male	26	73	88	27	27	20	12	0	273
female	35	62	76	28	30	18	21	1	271
unknown	1	1	0	0	2	0	0	0	4
total	62	136	164	55	59	38	33	1	548

Table 1: Age and gender distribution of the participants.

2.1 Familiarity and usage

The first four questions have been about the usage of the phone in general and the familiarity to the phone's features and IT security. In each case the participant could choose a value between 1 and 4, the results are given in table 2. It reveals that approximately 65% of the mobile phone owners are private users[2]. Mobile phones are important to their owners and most of them are mainly familiar with the features. Less familiar is the topic IT security. Only 19% answered, that they are familiar with this topic, nearly 15% even said they are not interested in IT security.

2.2 Using additional features

In one question participants should state which additional features they are using. They could select between SMS, e-mail, internet, camera, calendar or add further ones. In the following question the kind of stored data should be stated. Again one could choose between the proposed data (phone numbers, addresses, e-mails, appointments, birthdays and

[2]This shows the same trend as the comprehensive online survey from 2008 (see [Ifa08]).

Question	1	2	3	4	no answ.
Usage of mobile device (1 = private, 4 = business)	65.3	19.7	9.9	4.9	0.2
Importance of mobile device (1 = very important, 4 = unimportant)	36.3	39.4	20.6	3.3	0.4
Knowledge of the features of the device (1 = very good, 4 = no interest)	40.0	40.1	14.2	5.3	0.4
Knowledge of IT security (1 = very good, 4 = no interest)	19.3	30.8	33.8	14.6	1.5

Table 2: General questions about familiarity and usage of mobile devices (results in percent).

additional functionalities - Top 5	**stored data - Top 5**
1. SMS (92%)	1. phone numbers (99%)
2. camera (65%)	2. appointments (45%)
3. calendar (53%)	3. birthdays (40%)
4. internet (25%)	4. addresses (34%)
5. e-mail (17%)	5. e-mails (24%)

Table 3: Top 5 of additional functionalities and stored data on mobile phones.

passwords/PINs) or add further ones. In both cases participants were allowed to choose several answers. The top 5 of used additional features besides phoning and stored data are given in table 3. On the 6th place with 13% are *passwords/PINs*, which will be in most cases freely available in case the mobile phone is lost (see section 2.4), as only 8% of the phones containing stored PINs of passwords are sufficiently secured.

2.3 Carrying and attending the mobile phone

The answers to the question "I carry my mobile phone mainly...", show fundamental differences between men and women. One of the seven possibilities (back trousers pocket, front trousers pocket, breast pocket, pocket at the belt, back pack, hand bag, other) could be chosen. Two-thirds of men answered to carry their mobile phone in the front trousers pocket while 63% of women carry their mobile phone in their purse. See figure 1 for more details.

All respondents were asked how they take care of their mobile phone and nearly two-thirds answered their phone is *always within reach*. See figure 2 for the proposed values and the distribution of given answers.

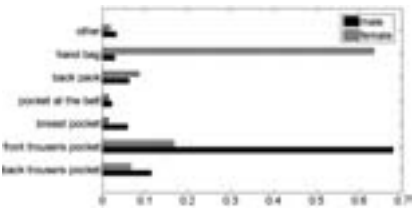

Figure 1: I carry my mobile phone mainly in my ...

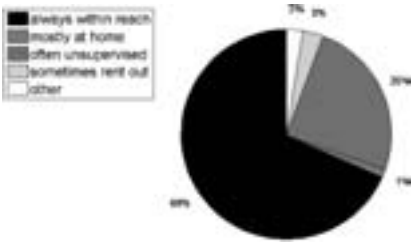

Figure 2: My mobile phone is....

2.4 Security settings

Section 2.2 showed that almost every mobile phone user saves personal data and 13% of the respondents even save passwords or PINs on their mobile devices. To see how this data is protected, the participants were asked to state which kind of action/input is necessary when using the phone after a standby phase.
Figure 3 shows the possible answers to this multiple choice question. The most common setting when reactivating the mobile phone from standby mode is keylock. Only 13% of the phones are sufficiently protected by PIN or visual code[3] which is in half of the cases combined with keylock. Comparing these answers to the ones regarding the familiarness with IT security there are some accordances. Most participants using a PIN or visual code said they are familiar with IT security (32% chose option 1, 37% option 2, 25% option 3, 5% option 4, 1 didn't answer at that question).
When asked for the reason for the low level of security (immediately usable or only key-

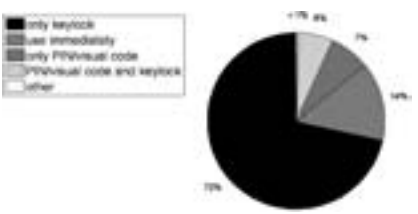

Figure 3: Security level when reactivating from standby mode.

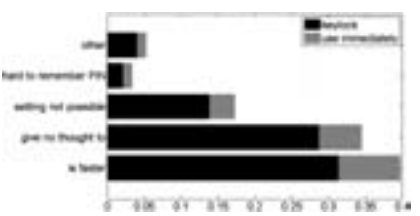

Figure 4: Reasons for the chosen low security setting.

lock) 40% answered it was chosen because *it is faster* and a further 34% answered they *did not think about it*. 17% stated that their phone does not have the function *enter PIN after standby mode*. For 3% of the participants, the PIN is too difficult to memorize (see figure 4).
Table 4 shows the willingness of the participants to use biometric authentication in total and depending on the so far chosen security setting. More than 50% are interested in an

[3]A visual code is a technique mostly used from the android operation system in which a pattern must be drawn; similar to a PIN.

would choose biometrics	total	only keylock	use immediately	only PIN/ visual code	PIN/VC and keylock	other
yes	54.4%	55.1%	51.3%	41.5%	67.7%	66.7%
no	37.8%	37.8%	38.5%	43.9%	29.4%	33.3%
not specified	7.9%	7.1%	10.3%	14.6%	2.9%	0.0%

Table 4: Willingness to use biometric authentication methods instead of the so far chosen security setting.

alternative biometric authentication process, 73% of those just use the key lock so far. If this option would be available, 55% of the participants currently using keylock and 51% of the ones which can directly use their phone after standby phase would use biometric authentication instead. This indicates that offering biometric authentication methods on mobile devices would highly increase the security of the data stored in mobile devices.
Participants which are willing to use biometric authentication, were asked which biometric modality/ies they would use. The favourite modality is fingerprint (87%), followed by speaker recognition (20%), face recognition (19%) and gait recognition (9%).

2.5 Mobile phone theft

With increasing popularity of mobile phones, also the number of thefts increases. The survey from [Fla06] showed that "4% of households owning a mobile phone have experienced a mobile phone theft". The german TNS-Emnid are talking about more than 7% in year 2008 [NM08] and our survey shows a further growth as approximately 10% of the participants have experienced a mobile phone theft, 22% of those more than once. Analysis of this survey showed that the theft rate for men is 10% higher than the theft rate for women. Four out of five participants who have had their mobile phone stolen are younger than 30. Of the participants which experienced a mobile phone theft, 75% answered their mobile phone is always within reach and 21% said its unsupervised most of the time. The locations in which the thefts occurred (see figure 5) are similar to the ones reported in [Fla06, p15-17].

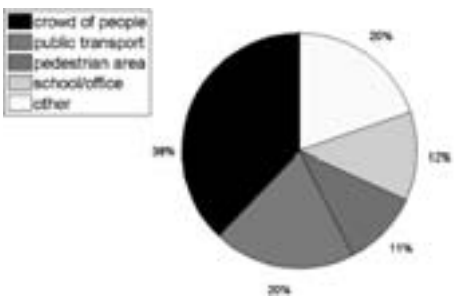

Figure 5: Locations at which mobile phones have been stolen.

3 Conclusion

As the number of mobile phones, their functionalities and application scenarios increases and hence also the amount of data stored on mobile devices, it is interesting and important to analyse the security awareness of the users which is mirrored by their chosen security settings. This paper states the result of a comprehensive survey with 548 participants. It is shown that after a stand by period only 13% of the mobile devices are secured with a PIN or visual code. This means that in 87% all data is freely available in case the phone is stolen or lost. The reason for unsecured phones is in 74% of the cases that it is faster or people did not even think about securing it. Offering biometric authentication methods on mobile phones would increase the number of secured phones as these methods would be used by about 54% of the participants. One reason for this might be that the problem of memorization and speed (see section 2.4) could be solved with biometric authentication.
Comparing the results of this survey to the ones from 2006 in [Fla06], there are still a lot of parallels. In general it is necessary to increase the user's security awareness such that he chooses sufficient security settings. This could for example be achieved by publications like the *Guidelines on mobile Phone and PDA Security* (2008, see [JS08]) by the National Institute of Standards and Technology (NIST). On the other hand configuring a PIN or biometric authentication as default setting when reactivating the mobile phone would probably also increase the number of secured phones as many people did not even think about changing the settings. In addition to the possibility of data loss because of stolen of lost phones, also attacks on mobile devices (see e.g. [HJO08]) should be considered.

References

[Bit09] Bitkom. Mehr als vier Milliarden Handy-Nutzer weltweit. http://www.bitkom.org/de/presse/62013_60608.aspx, August 2009. last checked: 2010-07-28 [german].

[Fla06] John Flatley. Mobile phone theft, plastic card and identity fraud. *http://rds.homeoffice.gov.uk/rds/pdfs07/hosb1007.pdf*, www.homeoffice.gov.uk/rds, 2005/06.

[HJO08] S.M. Habib, C. Jacob, and T. Olovsson. A practical analysis of the robustness and stability of the network stack in smartphones. In *11th International Conference on Computer and Information Technology (ICCIT 2008).*, pages 393 –398, 24-27 2008.

[Ifa08] Ifak Institut, Media Markt Analysen. Typologie der Wünsche 2009. http://de.statista.com/statistik/diagramm/studie/103676/umfrage/private-oder-dienstliche-handynutzung/, October 2008. last checked: 2010-05-19 [german].

[JS08] W. Jansen and K. Scarfone. Guidelines on Cell Phone and PDA Security. *NIST Special Publication 800-124*, http://csrc.nist.gov/publications/nistpubs/800-124/SP800-124.pdf, October 2008.

[NM08] Netzzeitung.de and Felix Magin. Das sollte man beim Handy-Verlust machen. http://www.netzeitung.de/wirtschaft/ratgeber/1099657.html, July 2008. last checked: 2010-07-28 [german].

The meaningful, safe and reliable use of biometrics [1]

Prof. Dr J. (Jan) H.A.M. Grijpink

Ministry of Justice / Utrecht University
P.O. Box 20301
NL-2500 EH The Hague
j.grijpink@minjus.nl

Abstract: *This paper introduces the current results of the debates within the Netherlands Biometrics Forum (NBF) on the principles that lead to a meaningful safe and reliable use of biometrics. Biometrics is becoming an important element of our information society, but new technology is often initially used incorrectly. This is partly due to so-called fallacies of the wrong level. In practice, large-scale systems tend to work out differently compared with small-scale applications, thus presenting additional problems at that larger scale that should have been taken into account during design and development. That implies that the first major applications can confront us with worrying social risks for which effective solutions have yet to be found. This paper, therefore, proposes to explicitly use the concept of identity fraud (identity theft) as litmus test of any biometrics application. The point is made that excessive concern about privacy inadvertently exacerbates these social risks. It is made clear that the assessment criterion 'safety' implies the protection of privacy, but that this does not necessarily apply the other way round. Because current large-scale applications seem to have neglected major privacy and security risks, this paper is primarily meant to stimulate this debate.*

1 The Netherlands Biometrics Forum (NBF) [2]

The Netherlands Biometrics Forum (NBF) is a foundation that advocates the meaningful, safe and reliable use of biometrics. Having the international character of biometrics in mind it focuses on the Dutch situation highlighting both the interests of the public at large and of the professionals involved. The NBF attempts to create more awareness regarding what can and cannot be done with biometrics, and a clearer understanding of the opportunities and risks. It is desirable to ultimately achieve social acceptance of this technology. That calls for trust. The NBF is convinced that this trust cannot be enforced

---

[1] Jan Grijpink is Principal Adviser at the Dutch Ministry of Justice (Information Strategy) and professor of Information Science (Chain-computerisation) at the University of Utrecht. He is chairman of the Netherlands Biometrics Forum (NBF).

[2] www.biometrieforum.nl

but must rather be earned. During the past two years many professionals from the public and private sector and the scientific community in the Netherlands have worked on formulating key principles for the meaningful, safe and reliable use of biometrics in what is known as a position paper. This document is periodically updated on the basis of new experiences and insights. In this phase of development of the NBF's position, the NBF wants these principles shared and challenged.

2 Biometrics

The term 'biometrics' is taken to mean: automated recognition of individuals based on their behavioural and biological characteristics. These days, information technology makes it possible to quickly digitise behavioural and biological characteristics so that we can either depict them or subject them to calculations. This can not only be done with unalterable characteristics such as the contour of a hand or a finger, a fingerprint or the pattern of an iris but with alterable characteristics as well, such as a voice, the way somebody moves his hand when writing his signature, or the rhythm of typing certain words on a keyboard. Biometric verification involves comparing a previously measured characteristic against the result of a new measurement at the time and place of the check. The result of the previous measurement can be registered in the verifying authority's information system, or on a chip card or another electronic document held by the person being checked.

Many people find it difficult to fathom the technology needed for biometric person recognition because it is based on the laws of probability and thus necessarily leads to a number of erroneous acceptances and rejections (the extent of which depends on the tolerances set by the operator himself). For that reason biometrics never offers complete certainty (100%) that someone is the right person. That way, biometrics also makes erroneous connections between people and their documents or data. The fact that biometrics cannot make any statements about the integrity of these documents and data or about the accuracy of the link itself implies that biometrics is unable to conclusively establish who somebody is. Contrary to what many people think, biometrics can only calculate the probability that somebody is the right person! This makes biometrics vulnerable to privacy and security concerns. [Gr01; Gr08].

3 The future of biometrics

The importance of computerised person recognition is becoming increasingly important in an anonymous information society characterised by increasing global mobility. As compared to administrative verification methods such as a PIN code, password or key, only biometrics is based on a person-related behavioural or biological characteristic as the point of recognition. Biometrics will ultimately become indispensable for sensitive work processes in the public and private sectors. Biometrics is especially useful when we need to know for sure that the person we are dealing with is the right person, or when someone wants to prevent his identity from being stolen and misused by somebody else.

That constantly sets different requirements for computerised person recognition, depending on the risks in a given context.

Two examples [3]

1. A swimming pool organisation wanted to use fingerprint verification to exclude a certain group of boys that was repeatedly harassing girls. A worthy aim, but the devil is in the detail. All visitors (both male and female) were asked to register their fingerprints in the swimming pool's computer system. This application threatens the bright future of biometrics. First, if you have the fingerprints of the boys you want to exclude from swimming, it is sufficient to check the fingerprints of male visitors belonging to the relevant age group. [4] Second, if someone's fingerprint is included in the blacklist, he can be sent on his way. Therefore, there is no need to store fingerprints at all. There is no point whatsoever in checking and storing the fingerprints of girls. And the story gets even worse. A woman of 82 refuses to cooperate with having her fingerprints checked and is therefore banned from the swimming pool.

The NBF's position is this: biometrics must be necessary and the purpose is the deciding factor regarding the rights and wrongs of how biometrics is to be implemented and used.

2. A car rental company was having a lot of difficulties with cars being returned. Many rented cars were not returned or were taken to the wrong place. Biometrics looked promising, but must not be too expensive. A creative employee came up with a solution without the need for expensive electronics: the fingerprint was placed on the paper rental contract with gel, with the assurance that the paper containing the fingerprint would be returned when the car was brought back. This experiment proved to be a resounding success: during the experiment no stolen or incorrectly returned vehicles! All well and good. But watch out! This simple biometrics system was introduced elsewhere by the same company, too. A few months later this site's administration proved to be full of copies of rental contracts with fingerprints without there being any need for them!

For that reason the NBF calls for attention to be paid to a biometrics application as a whole, the development of an application in the course of time being just as important as practical details such as the contracts' administration.

These examples illustrate how easy it is to use biometrics incorrectly. This engenders unnecessary resistance among the public and undermines social acceptance. Because

---

[3] The examples in this paper are mostly taken from public sources, but some stem from private practice of the NBF's participants. They have all been used to underpin and test the NBF's position during the development of the NBF's position paper.

[4] Such a blacklist may be constructed and maintained under the European Data Protection Directive and Dutch national law if the culprit's fingerprints are taken after a case of misbehaviour and used during a limited period of time and if the list's purpose is clearly explained to the public and the boys involved.

biometrics will ultimately become indispensable to our information society, the NBF regards this as a problem. Both examples also highlight the importance of providing information to the public and organisations wishing to use biometrics. We must guard our biometric details jealously, certainly those which are derived from unalterable biometric characteristics such as our fingerprints. Once compromised, the problem will remain for a long time without the possibility of defending ourselves by altering that biometric characteristic.

4 Fallacies of the wrong level

In information science, in common with other social sciences, we often gain insights from small-scale applications, such as at the level of a person or an organisation. We then translate those insights – usually without a second thought – into large scale applications, such as at the level of a chain or a social sector. In doing so we are likely – often without noticing it – to make what is known as a fallacy of the wrong level, for insights are related to the level at which they are gained and are generally invalid at other levels (higher or lower)! [Gr05; Gr06a, Gr10]. That results in all sorts of assumptions and principles in large-scale systems being incorrect, so that these systems contain more shortcomings and risks than we think or expect.

Two examples: the biometric passport and the biometric visa

1. Our first example concerns the new biometric passport. This is based on the notion that somebody can accurately be verified by his fingerprint. This essentially small-scale notion should not automatically be extended to the national or international scale of border control. Otherwise it is uncertain whether the biometric passport delivers what is expected of it. Large-scale systems function differently from small-scale ones because on this scale there is no coordinating or enforcing authority. Moreover, large-scale systems involve huge numbers of stakeholders (members of the public, travellers and patients) and cooperating autonomous organisations and professionals that causes large-scale processes to be barely manageable. Despite all good intentions much goes wrong.

Biometrics, too, can be supposed to work differently at large-scale level (chain, sector, country) than one might think from small-scale ideas, and can sometimes be counter-productive. Imitating or counterfeiting the fingerprint on the passport can enable someone to get through the check without it being possible to find out afterwards who it was because traces left inherently point to the official holder, not to the identity fraudster. Scaling up without taking a closer look at the risks of the large-scale situation is therefore a risky undertaking. And even then it is advisable to scale up gradually. For instance, by first having a fingerprint check carried out at the moments of the application and the delivery of a passport without the fingerprints being stored in the passport. Then, in a later stage, one could also store the fingerprints in the passport on voluntary basis, to begin with for those wishing to travel to the US. And so on.

2. The biometric visa, the second example, has already been introduced to keep out unwanted foreigners even before they come to the Netherlands. For that reason the fingerprints of the traveller are taken at the Dutch embassy in the country of origin during the visa application and sent to the Netherlands. If those fingerprints are included in the database of fingerprints of unwanted foreigners, the visa is refused.

Biometrics can in some cases be counterproductive at that large-scale level. Take a situation where a criminal network wants to send someone to the Netherlands for a criminal act. If the visa is refused, the network knows that it will either have to send someone else or choose a route where the checks are less well organised. That means that rather than the anticipated tighter grip on incoming passenger traffic, the target group of unwanted foreigners can imperceptibly become invisible!

It should be noted that for technical reasons it is not possible to place fingerprints on the visa as we are now doing with the passport. We therefore check visa applicants using the fingerprints in the database only. With these two variants of biometrics, if unforeseen problems arise in the future we can try out which of the two biometrics systems is the most flexible. It would of course have been better to carry out such an experiment beforehand, since once introduced it is barely possible to change a large-scale system such as this.

5 Identity fraud/theft as the touchstone for a biometrics application

By identity fraud we mean somebody with malicious intent deliberately contriving the appearance of an identity that does not belong to him, using the identity either of someone else or of a non-existent person. An identity fraudster has no need for a document or identity card: he can also use a personal number, a photo, an occurrence or a biometric detail because they all contain a suggestion on which people base their conclusion as to who they are dealing with. Identity fraud proves to be easy and does not involve too much risk. When carrying out identity checks we use barely any verification details other than those held by the person being checked. That reduces the chance of getting caught. And if someone gets caught, he has not (yet) done anything wrong! If the identity fraud succeeds nobody is the wiser, while the benefits can be substantial and of long duration. Official means of identifying people, such as an identity card, citizen service number or a biometric detail on the passport are of extra value to identity fraudsters because they must and can be used everywhere. Added to that is the fact that official verification procedures are known, uniform and predictable and can be inconspicuously observed in search of weak spots. Fallback procedures for situations in which the normal procedure cannot be followed (‘I've forgotten my passport...’ or equipment failure) are usually sloppy and improvised and can be triggered by the identity fraudster himself without the identity checking officer knowing, for instance by deliberately using a wrong or invalid token or ID document.

On the other hand, there is the weak position of the victim to consider. As the world becomes more digital identity fraud leaves more and more (technical) traces; but those

traces lead not to the perpetrator, but – inherent in the precise nature of identity fraud - to the victim, who is then faced with proving that he has *not* done something. For that reason the safe use of biometrics makes it necessary to substantially reduce the predictability of identity checks and sharply increase the quality of exception and emergency procedures. Indeed, biometrics should help to achieve that, too.

We therefore need to examine whether identity fraud is being prevented by biometric verification rather than made easier. [Gr04a; Gr04b; Gr06b]. This specific safety aspect of a biometrics application could be scrutinised with questions like the following. Can someone successfully pass through the identity check by imitating the biometric characteristic of the rightful holder? Can someone influence the check and get wrongly recognised as the rightful holder? Is it possible to obtain from the results of the check information that can be used with malicious intent (see the example of the biometric visa)? That is how the phenomenon of identity fraud / identity theft functions as the touchstone for *safe* biometrics. This safety assessment always relates to the biometrics application as a whole, including technology, organisation, procedures and not least the extent to which people cooperate or, conversely, have a vested interest in errors or misuse.

The NBF regards preventing identity fraud/theft as the touchstone for a safe biometrics application. Each biometric technology is in itself easy to mislead or to misuse. The NBF's position is therefore that it is necessary to make *simultaneous* use of several biometric details or technologies in combination with other data or resources since an identity fraudster will not be able to successfully make use of them all at the same time.

6 Privacy and safety

The traces left by identity fraud lead not to the perpetrator but to the victim. Identity fraud thus seriously violates the victim's privacy. That is especially true of identity fraud with an unalterable biometric characteristic since this form of identity fraud can continue to follow someone for a lengthy period without there being much he can do about it. Official bodies initially regard the victim as the perpetrator because all of the clues point in his direction. That often leaves someone having to prove that he is not the perpetrator, which is often hard to do and wrongly leaves the victim under a cloud of suspicion. In the case of biometrics privacy is thus closely related to safety and reputation, depending on the how it has been misused. The discussion about *privacy* in the context of biometrics is therefore unlikely to abate any time soon, but it will however remain abstract for as long as the relationship with someone's *safety* is not expressly made. The point frequently made in discussions about privacy along the lines of 'I don't mind what they know about me, I've got nothing to hide' is put forward by people who have not yet faced a wrongful accusation. If that accusation is based on a misused unalterable biometric detail, the chances of putting up a good defence are not good. This is not a hypothetical risk. At present, virtually all biometrics applications are not safe, certainly in cases where an unalterable biometric characteristic is used on a large scale. The NBF's position is that those concerned about privacy should at this stage focus more sharply on the safety of the large-scale use of biometric personal details, taking account of the

application as a whole and of target groups with other interests. This concept of *safety* goes far beyond the standard privacy discussions concerning the *protection* of these sensitive personal details. [Gr06b; Gr08].With a view to large-scale safety, the NBF's position paper contains various requirements that will need to be met and can help us to assess the social acceptability of a specific biometrics application.

Example

The biometric passport provides us with an interesting example. In Germany, the discussion on privacy has led to two fingerprints being placed on the German passport without the government keeping a copy of any kind. As a consequence, the fingerprints of the person being checked can only be compared with the fingerprints on the passport. That seems fair enough at small-scale level, but it is completely inadequate for large-scale application. The German government then faces a situation where, after issuing the passport, it is no longer able to independently verify whether the fingerprints on it are still the original ones or whether the person present really is the same as the passport's official holder. For the first purpose integer copies are needed of the two fingerprints that have been put on the passport. For the second purpose one or two additional fingerprints of the official holder are needed that have not been put on the passport, because the ones on the passport can be imitated or counterfeited. In The Netherlands, therefore, the government has opted to store the fingerprint of four fingers in a municipal database: the two fingerprints on the passport and two others. The first two make it possible to verify the integrity of the passport, the second pair to directly – i.e. independently of the passport – establish whether the person present is the same person as the legitimate passport holder. If used correctly, these databases enable to detect and to prevent identity fraud.

In The Netherlands, too, the privacy discussion is fighting the biometric passport system's underlying municipal databases and is threatening the safety of our large-scale biometric passport system by making these integrity and authenticity checks impossible. Thus, concerns about privacy could inadvertently hugely increase the social risks of large-scale biometrics applications. It must be made clear to those with serious concerns about privacy that the assessment criterion 'safety' implies the protection of privacy, but that this is not necessarily the case the other way round.

7 Ten principles of meaningful, safe and reliable use of biometrics [5]

1. Biometric person recognition alone is not conclusive. Biometrics is based on probability theory and therefore leads inherently to a number of wrong acceptances and wrong rejections (the precise number depends on the tolerance parameters configured by an operator). Moreover, biometric characteristics and biometric data derived from them can be imitated and counterfeited.

[5] The most recent integral text of the NBF position paper can be found on www.biometrieforum.nl

2. Biometrics can only recognise people, it cannot establish identities. Biometrics can link a person to a document or detail, but that says nothing about the integrity of that document or detail, or whether the link is itself accurate.

3. Safety assessments are indispensable to safe and reliable large-scale biometrics applications. Owing to uncontrollable organisational and human factors, a large-scale biometrics application can only be rendered safe and reliable with an enormous additional effort. In practice, it is still not possible to achieve that. Safety assessments must always relate to the biometrics application as a whole, including technology, organisation, procedures and the extent to which people cooperate or, conversely, have a vested interest in errors or misuse of the biometrics system or the biometric detail.

4. The principle of "*at least three matches*". Biometrics quickly gains reliability and safety if the biometric characteristic is used in combination with another biometric characteristic or detail and a non-biometric detail, such as a PIN code. In principle, the use of a *separate* (= *disconnected*) biometric detail which, with reasonable effort, can be linked to the person involved must therefore be discouraged.

5. It is important to actively discourage the trivial use of biometrics. The use of biometrics must be absolutely necessary for the envisaged purpose and not replaceable by other, less invasive or burdening measures.

6. A person subjected to a biometric verification has the right to be assured that a number of requirements have been met. The NBF operates a checklist of seven requirements that can be used as a test for the social acceptability of a biometrics application. [6]

7. It should be practically impossible to re-use biometric details in an application outside of it. Additionally, it must be possible to derive the originating application from the biometric detail itself.

8. The storage of biometric details should only be permitted if indispensable to that application in question. Biometric details should then be stored in distorted and encrypted form only.

9. It should only be permitted to link a file containing biometric details to external databases in situations provided for by law and on condition that there are no direct links to biographical details of the person involved.

---

[6] These rights are: the biometrics application is used only for its intended purpose, a simple and straightforward objections and complaints procedure, a fallback procedure proportionate to the risks involved, preventative measures against theft or misuse of biometric data, the operator's active support (compensation for damages and rehabilitation), disclosure as to who has had access to one's biometric details and explicit measures against a person who attempts to misuse or succeeds in misusing biometric details.

10. A mandatory register of large-scale use of biometrics. Large-scale biometrics applications should be registered, certified and monitored. The misuse of biometrically based identities should be reported to this register's manager who also has to guard against any unnecessary or non-secure storage of biometric details and to verify whether the operator of such a large-scale biometrics application has taken sufficient preventative measures against the theft and misuse of the biometric data he controls.

8 Concluding remarks

In this introductory stage of biometrics applications public acceptance and thrust are to be earned by the biometrics community. Two major aspects stand out: (1) our focus must be on avoiding teething troubles gaining experience with the use of biometric details at a smaller scale and (2) when biometrics applications are scaled up, more attention must be paid to assumptions and expectations that might not be valid at that level. Risk assessments must uncover the inherent security and safety problems and risk management must form an essential part of a biometrics application taking into account the extent to which people can be expected to co-operate or, conversely, have a vested interest in errors or misuse if they can get away with it. The NBF's position paper, therefore, highlights these aspects to stimulate social debate.

Bibliography

[Gr01] Grijpink, J.H.A.M., (2001). Biometrics and Privacy, in: *Computer Law and Security Report*, May/June 2001, vol. 17 (3) 2001, pp. 154-160. Oxford, UK: Elsevier Science Ltd

[Gr04a] Grijpink, J.H.A.M., (2004). Identity fraud as a challenge to the constitutional state, in: Computer Law and Security Report, vol. 20 (1) 2004, pp. 29-36. Oxford, UK: Elsevier Science Ltd

[Gr04b] Grijpink, J.H.A.M., (2004). Two barriers to realizing the benefits of biometrics: A chain perspective on biometrics, and identity fraud as biometrics' real challenge, in: Optical Security and Counterfeit Deterrence Techniques V, edited by Rudolf L. van Renesse, Proceedings of SPIE-IS&T Electronic Imaging, SPIE Vol. 5310, pp. 90-102

[Gr05] Grijpink, J.H.A.M., (2005). Our emerging information society: The challenge of large-scale information exchange in the constitutional state, in: *Computer Law and Security Report*, vol. 21 (4) 2005, pp. 328-337. Oxford, UK: Elsevier Science Ltd

[Gr06a] Grijpink, J.H.A.M., (2006). Criminal Records in the European Union: The challenge of large-scale information exchange, in: *European Journal of Crime, Criminal Law and Criminal Justice*, Volume 14 (2006) 1, pp. 1-19. Leiden: Brill Academic Publishers

[Gr06b] Grijpink, J.H.A.M., (2006). Identity fraud and biometrics: An assessment model for the use of biometrics, in: *Computer Law and Security Report*, vol. 22 (4) 2006, pp. 316-319. Oxford, UK: Elsevier Science Ltd

[Gr08] Grijpink, J.H.A.M., (2008). Biometrics and security: Trend report on biometrics: Some new insights, experiences and developments, in: *Computer Law and Security Report*, vol. 24 (3) 2008, pp. 261-264. Oxford, UK: Elsevier Science Ltd

[Gr10] Grijpink, J.H.A.M., (2010). Chain analysis for large-scale communication systems. *Journal of Chain-computerisation,* 1, 1-32

GI-Edition Lecture Notes in Informatics

P-1 Gregor Engels, Andreas Oberweis, Albert Zündorf (Hrsg.): Modellierung 2001.

P-2 Mikhail Godlevsky, Heinrich C. Mayr (Hrsg.): Information Systems Technology and its Applications, ISTA'2001.

P-3 Ana M. Moreno, Reind P. van de Riet (Hrsg.): Applications of Natural Language to Information Systems, NLDB'2001.

P-4 H. Wörn, J. Mühling, C. Vahl, H.-P. Meinzer (Hrsg.): Rechner- und sensorgestützte Chirurgie; Workshop des SFB 414.

P-5 Andy Schürr (Hg.): OMER – Object-Oriented Modeling of Embedded Real-Time Systems.

P-6 Hans-Jürgen Appelrath, Rolf Beyer, Uwe Marquardt, Heinrich C. Mayr, Claudia Steinberger (Hrsg.): Unternehmen Hochschule, UH'2001.

P-7 Andy Evans, Robert France, Ana Moreira, Bernhard Rumpe (Hrsg.): Practical UML-Based Rigorous Development Methods – Countering or Integrating the extremists, pUML'2001.

P-8 Reinhard Keil-Slawik, Johannes Magenheim (Hrsg.): Informatikunterricht und Medienbildung, INFOS'2001.

P-9 Jan von Knop, Wilhelm Haverkamp (Hrsg.): Innovative Anwendungen in Kommunikationsnetzen, 15. DFN Arbeitstagung.

P-10 Mirjam Minor, Steffen Staab (Hrsg.): 1st German Workshop on Experience Management: Sharing Experiences about the Sharing Experience.

P-11 Michael Weber, Frank Kargl (Hrsg.): Mobile Ad-Hoc Netzwerke, WMAN 2002.

P-12 Martin Glinz, Günther Müller-Luschnat (Hrsg.): Modellierung 2002.

P-13 Jan von Knop, Peter Schirmbacher and Viljan Mahni_ (Hrsg.): The Changing Universities – The Role of Technology.

P-14 Robert Tolksdorf, Rainer Eckstein (Hrsg.): XML-Technologien für das Semantic Web – XSW 2002.

P-15 Hans-Bernd Bludau, Andreas Koop (Hrsg.): Mobile Computing in Medicine.

P-16 J. Felix Hampe, Gerhard Schwabe (Hrsg.): Mobile and Collaborative Busi-ness 2002.

P-17 Jan von Knop, Wilhelm Haverkamp (Hrsg.): Zukunft der Netze –Die Verletzbarkeit meistern, 16. DFN Arbeitstagung.

P-18 Elmar J. Sinz, Markus Plaha (Hrsg.): Modellierung betrieblicher Informationssysteme – MobIS 2002.

P-19 Sigrid Schubert, Bernd Reusch, Norbert Jesse (Hrsg.): Informatik bewegt – Informatik 2002 – 32. Jahrestagung der Gesellschaft für Informatik e.V. (GI) 30.Sept.-3.Okt. 2002 in Dortmund.

P-20 Sigrid Schubert, Bernd Reusch, Norbert Jesse (Hrsg.): Informatik bewegt – Informatik 2002 – 32. Jahrestagung der Gesellschaft für Informatik e.V. (GI) 30.Sept.-3.Okt. 2002 in Dortmund (Ergänzungsband).

P-21 Jörg Desel, Mathias Weske (Hrsg.): Promise 2002: Prozessorientierte Methoden und Werkzeuge für die Entwicklung von Informationssystemen.

P-22 Sigrid Schubert, Johannes Magenheim, Peter Hubwieser, Torsten Brinda (Hrsg.): Forschungsbeiträge zur "Didaktik der Informatik" – Theorie, Praxis, Evaluation.

P-23 Thorsten Spitta, Jens Borchers, Harry M. Sneed (Hrsg.): Software Management 2002 – Fortschritt durch Beständigkeit

P-24 Rainer Eckstein, Robert Tolksdorf (Hrsg.): XMIDX 2003 – XML-Technologien für Middleware – Middleware für XML-Anwendungen

P-25 Key Pousttchi, Klaus Turowski (Hrsg.): Mobile Commerce – Anwendungen und Perspektiven – 3. Workshop Mobile Commerce, Universität Augsburg, 04.02.2003

P-26 Gerhard Weikum, Harald Schöning, Erhard Rahm (Hrsg.): BTW 2003: Datenbanksysteme für Business, Technologie und Web

P-27 Michael Kroll, Hans-Gerd Lipinski, Kay Melzer (Hrsg.): Mobiles Computing in der Medizin

P-28 Ulrich Reimer, Andreas Abecker, Steffen Staab, Gerd Stumme (Hrsg.): WM 2003: Professionelles Wissensmanagement – Erfahrungen und Visionen

P-29 Antje Düsterhöft, Bernhard Thalheim (Eds.): NLDB'2003: Natural Language Processing and Information Systems

P-30 Mikhail Godlevsky, Stephen Liddle, Heinrich C. Mayr (Eds.): Information Systems Technology and its Applications

P-31 Arslan Brömme, Christoph Busch (Eds.): BIOSIG 2003: Biometrics and Electronic Signatures

P-32 Peter Hubwieser (Hrsg.): Informatische Fachkonzepte im Unterricht – INFOS 2003

P-33 Andreas Geyer-Schulz, Alfred Taudes (Hrsg.): Informationswirtschaft: Ein Sektor mit Zukunft

P-34 Klaus Dittrich, Wolfgang König, Andreas Oberweis, Kai Rannenberg, Wolfgang Wahlster (Hrsg.): Informatik 2003 – Innovative Informatikanwendungen (Band 1)

P-35 Klaus Dittrich, Wolfgang König, Andreas Oberweis, Kai Rannenberg, Wolfgang Wahlster (Hrsg.): Informatik 2003 – Innovative Informatikanwendungen (Band 2)

P-36 Rüdiger Grimm, Hubert B. Keller, Kai Rannenberg (Hrsg.): Informatik 2003 – Mit Sicherheit Informatik

P-37 Arndt Bode, Jörg Desel, Sabine Rathmayer, Martin Wessner (Hrsg.): DeLFI 2003: e-Learning Fachtagung Informatik

P-38 E.J. Sinz, M. Plaha, P. Neckel (Hrsg.): Modellierung betrieblicher Informationssysteme – MobIS 2003

P-39 Jens Nedon, Sandra Frings, Oliver Göbel (Hrsg.): IT-Incident Management & IT-Forensics – IMF 2003

P-40 Michael Rebstock (Hrsg.): Modellierung betrieblicher Informationssysteme – MobIS 2004

P-41 Uwe Brinkschulte, Jürgen Becker, Dietmar Fey, Karl-Erwin Großpietsch, Christian Hochberger, Erik Maehle, Thomas Runkler (Edts.): ARCS 2004 – Organic and Pervasive Computing

P-42 Key Pousttchi, Klaus Turowski (Hrsg.): Mobile Economy – Transaktionen und Prozesse, Anwendungen und Dienste

P-43 Birgitta König-Ries, Michael Klein, Philipp Obreiter (Hrsg.): Persistance, Scalability, Transactions – Database Mechanisms for Mobile Applications

P-44 Jan von Knop, Wilhelm Haverkamp, Eike Jessen (Hrsg.): Security, E-Learning. E-Services

P-45 Bernhard Rumpe, Wofgang Hesse (Hrsg.): Modellierung 2004

P-46 Ulrich Flegel, Michael Meier (Hrsg.): Detection of Intrusions of Malware & Vulnerability Assessment

P-47 Alexander Prosser, Robert Krimmer (Hrsg.): Electronic Voting in Europe – Technology, Law, Politics and Society

P-48 Anatoly Doroshenko, Terry Halpin, Stephen W. Liddle, Heinrich C. Mayr (Hrsg.): Information Systems Technology and its Applications

P-49 G. Schiefer, P. Wagner, M. Morgenstern, U. Rickert (Hrsg.): Integration und Datensicherheit – Anforderungen, Konflikte und Perspektiven

P-50 Peter Dadam, Manfred Reichert (Hrsg.): INFORMATIK 2004 – Informatik verbindet (Band 1) Beiträge der 34. Jahrestagung der Gesellschaft für Informatik e.V. (GI), 20.-24. September 2004 in Ulm

P-51 Peter Dadam, Manfred Reichert (Hrsg.): INFORMATIK 2004 – Informatik verbindet (Band 2) Beiträge der 34. Jahrestagung der Gesellschaft für Informatik e.V. (GI), 20.-24. September 2004 in Ulm

P-52 Gregor Engels, Silke Seehusen (Hrsg.): DELFI 2004 – Tagungsband der 2. e-Learning Fachtagung Informatik

P-53 Robert Giegerich, Jens Stoye (Hrsg.): German Conference on Bioinformatics – GCB 2004

P-54 Jens Borchers, Ralf Kneuper (Hrsg.): Softwaremanagement 2004 – Outsourcing und Integration

P-55 Jan von Knop, Wilhelm Haverkamp, Eike Jessen (Hrsg.): E-Science und Grid Adhoc-Netze Medienintegration

P-56 Fernand Feltz, Andreas Oberweis, Benoit Otjacques (Hrsg.): EMISA 2004 – Informationssysteme im E-Business und E-Government

P-57 Klaus Turowski (Hrsg.): Architekturen, Komponenten, Anwendungen

P-58 Sami Beydeda, Volker Gruhn, Johannes Mayer, Ralf Reussner, Franz Schweiggert (Hrsg.): Testing of Component-Based Systems and Software Quality

P-59 J. Felix Hampe, Franz Lehner, Key Pousttchi, Kai Ranneberg, Klaus Turowski (Hrsg.): Mobile Business – Processes, Platforms, Payments

P-60 Steffen Friedrich (Hrsg.): Unterrichtskonzepte für inforrmatische Bildung

P-61 Paul Müller, Reinhard Gotzhein, Jens B. Schmitt (Hrsg.): Kommunikation in verteilten Systemen

P-62 Federrath, Hannes (Hrsg.): „Sicherheit 2005“ – Sicherheit – Schutz und Zuverlässigkeit

P-63 Roland Kaschek, Heinrich C. Mayr, Stephen Liddle (Hrsg.): Information Systems – Technology and ist Applications

P-64 Peter Liggesmeyer, Klaus Pohl, Michael Goedicke (Hrsg.): Software Engineering 2005

P-65 Gottfried Vossen, Frank Leymann, Peter Lockemann, Wolffried Stucky (Hrsg.): Datenbanksysteme in Business, Technologie und Web

P-66 Jörg M. Haake, Ulrike Lucke, Djamshid Tavangarian (Hrsg.): DeLFI 2005: 3. deutsche e-Learning Fachtagung Informatik

P-67 Armin B. Cremers, Rainer Manthey, Peter Martini, Volker Steinhage (Hrsg.): INFORMATIK 2005 – Informatik LIVE (Band 1)

P-68 Armin B. Cremers, Rainer Manthey, Peter Martini, Volker Steinhage (Hrsg.): INFORMATIK 2005 – Informatik LIVE (Band 2)

P-69 Robert Hirschfeld, Ryszard Kowalcyk, Andreas Polze, Matthias Weske (Hrsg.): NODe 2005, GSEM 2005

P-70 Klaus Turowski, Johannes-Maria Zaha (Hrsg.): Component-oriented Enterprise Application (COAE 2005)

P-71 Andrew Torda, Stefan Kurz, Matthias Rarey (Hrsg.): German Conference on Bioinformatics 2005

P-72 Klaus P. Jantke, Klaus-Peter Fähnrich, Wolfgang S. Wittig (Hrsg.): Marktplatz Internet: Von e-Learning bis e-Payment

P-73 Jan von Knop, Wilhelm Haverkamp, Eike Jessen (Hrsg.): “Heute schon das Morgen sehen“

P-74 Christopher Wolf, Stefan Lucks, Po-Wah Yau (Hrsg.): WEWoRC 2005 – Western European Workshop on Research in Cryptology

P-75 Jörg Desel, Ulrich Frank (Hrsg.): Enterprise Modelling and Information Systems Architecture

P-76 Thomas Kirste, Birgitta König-Riess, Key Pousttchi, Klaus Turowski (Hrsg.): Mobile Informationssysteme – Potentiale, Hindernisse, Einsatz

P-77 Jana Dittmann (Hrsg.): SICHERHEIT 2006

P-78 K.-O. Wenkel, P. Wagner, M. Morgenstern, K. Luzi, P. Eisermann (Hrsg.): Land- und Ernährungswirtschaft im Wandel

P-79 Bettina Biel, Matthias Book, Volker Gruhn (Hrsg.): Softwareengineering 2006

P-80 Mareike Schoop, Christian Huemer, Michael Rebstock, Martin Bichler (Hrsg.): Service-Oriented Electronic Commerce

P-81 Wolfgang Karl, Jürgen Becker, Karl-Erwin Großpietsch, Christian Hochberger, Erik Maehle (Hrsg.): ARCS´06

P-82 Heinrich C. Mayr, Ruth Breu (Hrsg.): Modellierung 2006

P-83 Daniel Huson, Oliver Kohlbacher, Andrei Lupas, Kay Nieselt and Andreas Zell (eds.): German Conference on Bioinformatics

P-84 Dimitris Karagiannis, Heinrich C. Mayr, (Hrsg.): Information Systems Technology and its Applications

P-85 Witold Abramowicz, Heinrich C. Mayr, (Hrsg.): Business Information Systems

P-86 Robert Krimmer (Ed.): Electronic Voting 2006

P-87 Max Mühlhäuser, Guido Rößling, Ralf Steinmetz (Hrsg.): DELFI 2006: 4. e-Learning Fachtagung Informatik

P-88 Robert Hirschfeld, Andreas Polze, Ryszard Kowalczyk (Hrsg.): NODe 2006, GSEM 2006

P-90 Joachim Schelp, Robert Winter, Ulrich Frank, Bodo Rieger, Klaus Turowski (Hrsg.): Integration, Informationslogistik und Architektur

P-91 Henrik Stormer, Andreas Meier, Michael Schumacher (Eds.): European Conference on eHealth 2006

P-92 Fernand Feltz, Benoît Otjacques, Andreas Oberweis, Nicolas Poussing (Eds.): AIM 2006

P-93 Christian Hochberger, Rüdiger Liskowsky (Eds.): INFORMATIK 2006 – Informatik für Menschen, Band 1

P-94 Christian Hochberger, Rüdiger Liskowsky (Eds.): INFORMATIK 2006 – Informatik für Menschen, Band 2

P-95 Matthias Weske, Markus Nüttgens (Eds.): EMISA 2005: Methoden, Konzepte und Technologien für die Entwicklung von dienstbasierten Informationssystemen

P-96 Saartje Brockmans, Jürgen Jung, York Sure (Eds.): Meta-Modelling and Ontologies

P-97 Oliver Göbel, Dirk Schadt, Sandra Frings, Hardo Hase, Detlef Günther, Jens Nedon (Eds.): IT-Incident Mangament & IT-Forensics – IMF 2006

P-98 Hans Brandt-Pook, Werner Simonsmeier und Thorsten Spitta (Hrsg.): Beratung in der Softwareentwicklung – Modelle, Methoden, Best Practices

P-99 Andreas Schwill, Carsten Schulte, Marco Thomas (Hrsg.): Didaktik der Informatik

P-100 Peter Forbrig, Günter Siegel, Markus Schneider (Hrsg.): HDI 2006: Hochschuldidaktik der Informatik

P-101 Stefan Böttinger, Ludwig Theuvsen, Susanne Rank, Marlies Morgenstern (Hrsg.): Agrarinformatik im Spannungsfeld zwischen Regionalisierung und globalen Wertschöpfungsketten

P-102 Otto Spaniol (Eds.): Mobile Services and Personalized Environments

P-103 Alfons Kemper, Harald Schöning, Thomas Rose, Matthias Jarke, Thomas Seidl, Christoph Quix, Christoph Brochhaus (Hrsg.): Datenbanksysteme in Business, Technologie und Web (BTW 2007)

P-104 Birgitta König-Ries, Franz Lehner, Rainer Malaka, Can Türker (Hrsg.) MMS 2007: Mobilität und mobile Informationssysteme

P-105 Wolf-Gideon Bleek, Jörg Raasch, Heinz Züllighoven (Hrsg.) Software Engineering 2007

P-106 Wolf-Gideon Bleek, Henning Schwentner, Heinz Züllighoven (Hrsg.) Software Engineering 2007 – Beiträge zu den Workshops

P-107 Heinrich C. Mayr, Dimitris Karagiannis (eds.) Information Systems Technology and its Applications

P-108 Arslan Brömme, Christoph Busch, Detlef Hühnlein (eds.) BIOSIG 2007: Biometrics and Electronic Signatures

P-109 Rainer Koschke, Otthein Herzog, Karl-Heinz Rödiger, Marc Ronthaler (Hrsg.) INFORMATIK 2007 Informatik trifft Logistik Band 1

P-110 Rainer Koschke, Otthein Herzog, Karl-Heinz Rödiger, Marc Ronthaler (Hrsg.) INFORMATIK 2007 Informatik trifft Logistik Band 2

P-111 Christian Eibl, Johannes Magenheim, Sigrid Schubert, Martin Wessner (Hrsg.) DeLFI 2007: 5. e-Learning Fachtagung Informatik

P-112 Sigrid Schubert (Hrsg.) Didaktik der Informatik in Theorie und Praxis

P-113 Sören Auer, Christian Bizer, Claudia Müller, Anna V. Zhdanova (Eds.) The Social Semantic Web 2007 Proceedings of the 1st Conference on Social Semantic Web (CSSW)

P-114 Sandra Frings, Oliver Göbel, Detlef Günther, Hardo G. Hase, Jens Nedon, Dirk Schadt, Arslan Brömme (Eds.) IMF2007 IT-incident management & IT-forensics Proceedings of the 3rd International Conference on IT-Incident Management & IT-Forensics

P-115 Claudia Falter, Alexander Schliep, Joachim Selbig, Martin Vingron and Dirk Walther (Eds.) German conference on bioinformatics GCB 2007

P-116 Witold Abramowicz, Leszek Maciszek (Eds.) Business Process and Services Computing 1st International Working Conference on Business Process and Services Computing BPSC 2007

P-117 Ryszard Kowalczyk (Ed.) Grid service engineering and manegement The 4th International Conference on Grid Service Engineering and Management GSEM 2007

P-118 Andreas Hein, Wilfried Thoben, Hans-Jürgen Appelrath, Peter Jensch (Eds.) European Conference on ehealth 2007

P-119 Manfred Reichert, Stefan Strecker, Klaus Turowski (Eds.) Enterprise Modelling and Information Systems Architectures Concepts and Applications

P-120 Adam Pawlak, Kurt Sandkuhl, Wojciech Cholewa, Leandro Soares Indrusiak (Eds.) Coordination of Collaborative Engineering - State of the Art and Future Challenges

P-121 Korbinian Herrmann, Bernd Bruegge (Hrsg.) Software Engineering 2008 Fachtagung des GI-Fachbereichs Softwaretechnik

P-122 Walid Maalej, Bernd Bruegge (Hrsg.) Software Engineering 2008 - Workshopband Fachtagung des GI-Fachbereichs Softwaretechnik

P-123 Michael H. Breitner, Martin Breunig, Elgar Fleisch, Ley Pousttchi, Klaus Turowski (Hrsg.)
Mobile und Ubiquitäre Informationssysteme – Technologien, Prozesse, Marktfähigkeit
Proceedings zur 3. Konferenz Mobile und Ubiquitäre Informationssysteme (MMS 2008)

P-124 Wolfgang E. Nagel, Rolf Hoffmann, Andreas Koch (Eds.)
9th Workshop on Parallel Systems and Algorithms (PASA)
Workshop of the GI/ITG Speciel Interest Groups PARS and PARVA

P-125 Rolf A.E. Müller, Hans-H. Sundermeier, Ludwig Theuvsen, Stephanie Schütze, Marlies Morgenstern (Hrsg.)
Unternehmens-IT: Führungsinstrument oder Verwaltungsbürde
Referate der 28. GIL Jahrestagung

P-126 Rainer Gimnich, Uwe Kaiser, Jochen Quante, Andreas Winter (Hrsg.)
10th Workshop Software Reengineering (WSR 2008)

P-127 Thomas Kühne, Wolfgang Reisig, Friedrich Steimann (Hrsg.)
Modellierung 2008

P-128 Ammar Alkassar, Jörg Siekmann (Hrsg.)
Sicherheit 2008
Sicherheit, Schutz und Zuverlässigkeit
Beiträge der 4. Jahrestagung des Fachbereichs Sicherheit der Gesellschaft für Informatik e.V. (GI)
2.-4. April 2008
Saarbrücken, Germany

P-129 Wolfgang Hesse, Andreas Oberweis (Eds.)
Sigsand-Europe 2008
Proceedings of the Third AIS SIGSAND European Symposium on Analysis, Design, Use and Societal Impact of Information Systems

P-130 Paul Müller, Bernhard Neumair, Gabi Dreo Rodosek (Hrsg.)
1. DFN-Forum Kommunikations-technologien Beiträge der Fachtagung

P-131 Robert Krimmer, Rüdiger Grimm (Eds.)
3rd International Conference on Electronic Voting 2008
Co-organized by Council of Europe, Gesellschaft für Informatik and E-Voting.CC

P-132 Silke Seehusen, Ulrike Lucke, Stefan Fischer (Hrsg.)
DeLFI 2008:
Die 6. e-Learning Fachtagung Informatik

P-133 Heinz-Gerd Hegering, Axel Lehmann, Hans Jürgen Ohlbach, Christian Scheideler (Hrsg.)
INFORMATIK 2008
Beherrschbare Systeme – dank Informatik
Band 1

P-134 Heinz-Gerd Hegering, Axel Lehmann, Hans Jürgen Ohlbach, Christian Scheideler (Hrsg.)
INFORMATIK 2008
Beherrschbare Systeme – dank Informatik
Band 2

P-135 Torsten Brinda, Michael Fothe, Peter Hubwieser, Kirsten Schlüter (Hrsg.)
Didaktik der Informatik – Aktuelle Forschungsergebnisse

P-136 Andreas Beyer, Michael Schroeder (Eds.)
German Conference on Bioinformatics GCB 2008

P-137 Arslan Brömme, Christoph Busch, Detlef Hühnlein (Eds.)
BIOSIG 2008: Biometrics and Electronic Signatures

P-138 Barbara Dinter, Robert Winter, Peter Chamoni, Norbert Gronau, Klaus Turowski (Hrsg.)
Synergien durch Integration und Informationslogistik
Proceedings zur DW2008

P-139 Georg Herzwurm, Martin Mikusz (Hrsg.)
Industrialisierung des Software-Managements
Fachtagung des GI-Fachausschusses Management der Anwendungsentwick-lung und -wartung im Fachbereich Wirtschaftsinformatik

P-140 Oliver Göbel, Sandra Frings, Detlef Günther, Jens Nedon, Dirk Schadt (Eds.)
IMF 2008 - IT Incident Management & IT Forensics

P-141 Peter Loos, Markus Nüttgens, Klaus Turowski, Dirk Werth (Hrsg.)
Modellierung betrieblicher Informations-systeme (MobIS 2008)
Modellierung zwischen SOA und Compliance Management

P-142 R. Bill, P. Korduan, L. Theuvsen, M. Morgenstern (Hrsg.)
Anforderungen an die Agrarinformatik durch Globalisierung und Klimaveränderung

P-143 Peter Liggesmeyer, Gregor Engels, Jürgen Münch, Jörg Dörr, Norman Riegel (Hrsg.)
Software Engineering 2009
Fachtagung des GI-Fachbereichs Softwaretechnik

P-144 Johann-Christoph Freytag, Thomas Ruf, Wolfgang Lehner, Gottfried Vossen (Hrsg.)
Datenbanksysteme in Business, Technologie und Web (BTW)

P-145 Knut Hinkelmann, Holger Wache (Eds.)
WM2009: 5th Conference on Professional Knowledge Management

P-146 Markus Bick, Martin Breunig, Hagen Höpfner (Hrsg.)
Mobile und Ubiquitäre Informationssysteme – Entwicklung, Implementierung und Anwendung
4. Konferenz Mobile und Ubiquitäre Informationssysteme (MMS 2009)

P-147 Witold Abramowicz, Leszek Maciaszek, Ryszard Kowalczyk, Andreas Speck (Eds.)
Business Process, Services Computing and Intelligent Service Management
BPSC 2009 · ISM 2009 · YRW-MBP 2009

P-148 Christian Erfurth, Gerald Eichler, Volkmar Schau (Eds.)
9th International Conference on Innovative Internet Community Systems
I^2CS 2009

P-149 Paul Müller, Bernhard Neumair, Gabi Dreo Rodosek (Hrsg.)
2. DFN-Forum Kommunikationstechnologien
Beiträge der Fachtagung

P-150 Jürgen Münch, Peter Liggesmeyer (Hrsg.)
Software Engineering
2009 - Workshopband

P-151 Armin Heinzl, Peter Dadam, Stefan Kirn, Peter Lockemann (Eds.)
PRIMIUM
Process Innovation for Enterprise Software

P-152 Jan Mendling, Stefanie Rinderle-Ma, Werner Esswein (Eds.)
Enterprise Modelling and Information Systems Architectures
Proceedings of the 3rd Int'l Workshop EMISA 2009

P-153 Andreas Schwill, Nicolas Apostolopoulos (Hrsg.)
Lernen im Digitalen Zeitalter
DeLFI 2009 – Die 7. E-Learning Fachtagung Informatik

P-154 Stefan Fischer, Erik Maehle Rüdiger Reischuk (Hrsg.)
INFORMATIK 2009
Im Focus das Leben

P-155 Arslan Brömme, Christoph Busch, Detlef Hühnlein (Eds.)
BIOSIG 2009:
Biometrics and Electronic Signatures
Proceedings of the Special Interest Group on Biometrics and Electronic Signatures

P-156 Bernhard Koerber (Hrsg.)
Zukunft braucht Herkunft
25 Jahre »INFOS – Informatik und Schule«

P-157 Ivo Grosse, Steffen Neumann, Stefan Posch, Falk Schreiber, Peter Stadler (Eds.)
German Conference on Bioinformatics 2009

P-158 W. Claupein, L. Theuvsen, A. Kämpf, M. Morgenstern (Hrsg.)
Precision Agriculture Reloaded – Informationsgestützte Landwirtschaft

P-159 Gregor Engels, Markus Luckey, Wilhelm Schäfer (Hrsg.)
Software Engineering 2010

P-160 Gregor Engels, Markus Luckey, Alexander Pretschner, Ralf Reussner (Hrsg.)
Software Engineering 2010 – Workshopband
(inkl. Doktorandensymposium)

P-161 Gregor Engels, Dimitris Karagiannis Heinrich C. Mayr (Hrsg.)
Modellierung 2010

P-162 Maria A. Wimmer, Uwe Brinkhoff, Siegfried Kaiser, Dagmar Lück-Schneider, Erich Schweighofer, Andreas Wiebe (Hrsg.)
Vernetzte IT für einen effektiven Staat
Gemeinsame Fachtagung Verwaltungsinformatik (FTVI) und Fachtagung Rechtsinformatik (FTRI) 2010

P-163 Markus Bick, Stefan Eulgem, Elgar Fleisch, J. Felix Hampe, Birgitta König-Ries, Franz Lehner, Key Pousttchi, Kai Rannenberg (Hrsg.)
Mobile und Ubiquitäre Informationssysteme
Technologien, Anwendungen und Dienste zur Unterstützung von mobiler Kollaboration

P-164 Arslan Brömme, Christoph Busch (Eds.)
BIOSIG 2010: Biometrics and Electronic Signatures Proceedings of the Special Interest Group on Biometrics and Electronic Signatures

P-165 Gerald Eichler, Peter Kropf,
Ulrike Lechner, Phayung Meesad,
Herwig Unger (Eds.)
10th International Conference on
Innovative Internet Community Systems
(I^2CS) – Jubilee Edition 2010 –

P-166 Paul Müller, Bernhard Neumair,
Gabi Dreo Rodosek (Hrsg.)
3. DFN-Forum Kommunikationstechnologien
Beiträge der Fachtagung

P-167 Robert Krimmer, Rüdiger Grimm (Eds.)
4th International Conference on
Electronic Voting 2010
co-organized by the Council of Europe,
Gesellschaft für Informatik and
E-Voting.CC

P-168 Ira Diethelm, Christina Dörge,
Claudia Hildebrandt,
Carsten Schulte (Hrsg.)
Didaktik der Informatik
Möglichkeiten empirischer
Forschungsmethoden und Perspektiven
der Fachdidaktik

P-169 Michael Kerres, Nadine Ojstersek
Ulrik Schroeder, Ulrich Hoppe (Hrsg.)
DeLFI 2010 - 8. Tagung
der Fachgruppe E-Learning
der Gesellschaft für Informatik e.V.

P-170 Felix C. Freiling (Hrsg.)
Sicherheit 2010
Sicherheit, Schutz und Zuverlässigkeit

P-171 Werner Esswein, Klaus Turowski,
Martin Juhrisch (Hrsg.)
Modellierung betrieblicher
Informationssysteme (MobIS 2010)
Modellgestütztes Management

P-174 Arslan Brömme, Torsten Eymann,
Detlef Hühnlein, Heiko Roßnagel,
Paul Schmücker (Hrsg.)
perspeGKtive 2010
Workshop „Innovative und sichere
Informationstechnologie für das
Gesundheitswesen von morgen“

The titles can be purchased at:

Köllen Druck + Verlag GmbH
Ernst-Robert-Curtius-Str. 14 · D-53117 Bonn
Fax: +49 (0)228/9898222
E-Mail: druckverlag@koellen.de