Privacy Violation Detection

Tina Hermandsen Krekke

22nd June 2004

Preface

This master's thesis has been carried out at the Norwegian University of Science and Technology (NTNU) as a part of my Master of Science degree during the spring 2004. It looks into issues related to privacy and privacy violation detection, and the problem was initially suggested by the Norwegian Computing Centre.

I would like to thank my supervisors for excellent support and feedback, and for giving me your time:

- Jerker Danielsson, Research Scientist at the Norwegian Computing Centre, for input, information and reading through this report.
- Svein J. Knapskog, Professor at Department of Telematics, The Norwegian University of Science and Technology (NTNU).

Trondheim, 22nd June 2004

Tina Hermandsen Krekke

Summary

In information and communication technology (ICT) systems which stores and processes private, personal data, privacy violations may occur. A Privacy Violation Detector (PVD) aims at detecting such privacy violations.

Generally, an individual's privacy can be protected in two ways, either by minimizing the amount of personal data stored, or by enforcing privacy policies. The PVD may be a part of systems that enforce privacy policies.

This thesis describes and defines issues relating to privacy, such as treats, privacy policies and privacy violations, and outlines some privacy protection measures, e.g. privacy protective laws and Privacy Enhancing Technologies (PETs). Privacy violations in general are discussed and defined as events that breach a privacy policy or an agreement between a customer and the data collecting entity.

Motivations, in addition to the fact that the PVD will be able to detect privacy violations, are discussed. The PVD will be an reactive protection mechanisms, that is, able to detect privacy violations and cause some reaction after they occurred. This feature may be desirable, e.g. in medical information systems where access to data, regardless of authorisation, may be more important than enforcing a privacy policy, e.g. in an emergency situation. The privacy violation may then be investigated after it occurred to see whether it may be justified based on the present circumstances. Further, a PVD may increase the users and data subjects (i.e. the individual who's personal data is collected) trust in the information system collecting, processing and storing personal data and have a preventive function.

However, there are some issues relating to privacy violation detection. By monitoring events that occur in a system the privacy of the users of the system may be threatened. Therefore, measures must be taken in order to protect the privacy of the users such as pseudonymise and protect the logs.

Further, as a system that integrates a PVD may have similarities to current intrusion detection systems (IDSs) techniques, a brief outline of IDSs techniques are included. This constitutes the background for discussing a PVD, which may be based upon anomaly detection or policy-based detection.

A case study of applicability of a PVD in a medical information system context is provided. For the case, a privacy policy and data that needs to be logged in order to be able to detect a privacy violation is defined. Then, a privacy violation scenario is described, to show the applicability of a PVD in the scenario.

Contents

List of Figures

List of Tables

Chapter 1

Introduction

1.1 Background and Motivation

This master's thesis was suggested by the Norwegian Computing Center (Norsk Regnesentral, NR). NR has strategic institute program on privacy, entitled 'Personalised Internet-based Services and Privacy Protection', financed by the Norwegian Research Council. The main goal of the program is to study privacy issues in the context of customised and personalised mobile and Internet-based services and businesses [1].

Processing of personal data may be useful and necessary under some circumstances. In some cases, personal data must be collected due to legislation, or in order to provide some public service. Personal data may also personalise services to the user.

However, as increasing amounts of personal data is collected, stored and processed in information systems, an individual's privacy is increasingly threatened. It may also be difficult for individuals to exercise control over their own privacy, as it is difficult to get an overview of which types of system actually store personal information and what types of information they store.

Privacy can be protected in different ways, and privacy and data protection laws are important. However, privacy should also be enforced through technical means, by privacy-enhancing technologies (PETs), either by minimizing the amount of personal data collected, or by frameworks for enforcement of privacy policies.

[2] describes a framework for the enforcement of privacy developed, and partly implemented, by NR. The suggested framework consists of framework elements that together provide the functionality necessary for enforcement of applicable regulations and privacy agreements reached in connection with data collection. Further, each framework element is composed of components that support the implementation of its functionality. This thesis will further investigate the privacy violation detector (PVD), which is suggested to be a part of the monitoring element. The

[1] See http://snipsnap.nr.no:8668/privacy/space/Project+description

monitoring element monitors and analyses the audit trail generated by the other elements, and enable detection of a policy breach and cause some reaction after the breach happened. The PVD is an automated tool for detection of privacy violations. An employee browsing through customer records for email addresses is an example of a privacy violation.

1.2 Problem Description and Scope

The aim of this thesis is to outline a privacy violation detector (PVD). What constitutes a privacy violation is determined by a local privacy policy, and may vary from organisation to organisation. Individuals may also have different privacy preferences, which can be mirrored in individual privacy agreements with an organisation. Hence, what constitutes a privacy violation may also vary from individual to individual. Therefore, this thesis discusses and defines privacy in general, outlines potential threats to an individual's privacy due to e.g. increased use of information- and communication technology (ICT) systems. Further, privacy policies (and agreements) and privacy violations are discussed, as it is important to have an understanding of what constitutes a privacy violation in order to be able to detect them. Then, a study of privacy protective measures is included, as a PVD also may protect privacy.

A PVD is likely to have many similarities with current intrusion detection systems (IDSs), which are systems that automate the process of monitoring the events occurring in a computer system or network, and analyse them for signs of security problems [2]. The PVD aims at automating the process of monitoring the events occurring in a computer systems or network, and analyse them for signs of privacy violations. Therefore, a brief description of IDSs techniques is included, and the IDS terminology adopted when the PVD is outlined. The choice has been made to outline the different features of a potential PVD, rather than define a design for a PVD. Different detection methods and different potential features of a PVD are discussed.

However, as what constitutes a privacy violation varies from organisation to organisation, the privacy violation relevant data types are likely to be application dependent. Therefore, this thesis includes a case study where a (simplistic) privacy policy is defined, and data that needs to be logged in order to detect a privacy violation decided on. The case study includes a scenario to show how a PVD may work in that setting.

1.3 Outline of the report

The remainder of this thesis first looks into some privacy related issues in chapter 2. Chapter 2 also defines privacy, privacy policies and privacy violations, discusses

[2] Definition of IDS taken from [17].

threats to privacy and privacy protective measures. Chapter 3 outlines the motivation for a PVD, along with some goals and issues. Intrusion detection systems (IDSs) techniques are also introduced in chapter 3, as there are similarities between privacy violation detection and intrusion detection. Then, a PVD design discussion is provided. Chapter 4 describes a case study of medical information systems. Based on this case, a privacy policy is outlined and data that needs to be collected in the log files of such systems in order to determine a privacy violation is determined. Chapter 5 contains the conclusions of this work.

Chapter 2

Privacy

This chapter presents some privacy concepts. Privacy is defined, and subsequently threats towards privacy in our increasingly networked society are discussed. Then privacy policies and privacy violations are defined and discussed, before privacy protection measures, including privacy legislation and privacy enhancing technologies, are outlined.

2.1 Defining Privacy

Privacy is one of the fundamental human rights, as recognized by the United Nations Universal Declaration of Human Rights, see article 12, [3]. However, defining privacy is not straightforward. Privacy relates to our personal sphere and our ability to protect information about ourselves, and is very subjective. Privacy may be defined as:

> Privacy is the claim of individuals, groups and institutions to determine for themselves, when, how and to what extent information about them is communicated to others [1].

Other definitions of privacy include the right to a private life, the right to be left alone, the right to limit accessibility, the right to minimize intrusiveness, the right to expect confidentiality, to enjoy solitude, intimacy, anonymity, reserve, and the right to secrecy.

Further, privacy is about protection of personal data (also referred to as personal identifying data). Privacy protection is discussed in subchapter 2.5, but some essential privacy principles are stated next, as it may help gain a better understanding of privacy issues.

[1] Contributed by Dr. Alan F. Westin, Professor of Public Law and Government, Columbia University, see [4].

2.1.1 Privacy Principles

The following list, see [1], states the most essential requirements to privacy. The requirements are mainly based upon national laws, the EU directive and the OECD guidelines (see Appendix C for an introduction to these legal requirements). These requirements provide a common ground discussion of privacy issues:

Principle of lawfulness and fairness. Personal data should be collected and processed in a fair and lawful way.

Principle of the purpose specification and purpose binding. The *purposes* for which personal data is collected and processed should be specified and legitimate. The subsequent use of personal data is limited to those specified purposes, unless there is an informed consent by the data subject.

Principle of necessity of data collection and processing. The collection and processing of personal data should only be allowed, if it is *necessary* for the tasks falling within the responsibility of the data processing agency.

Information, notification and access rights of the data subjects. Data subjects have the right to information, to notification and the right to correction, erasure or blocking of incorrect or illegally stored data. These rights should not be excluded or restricted by a legal transaction.

Principle of security and accuracy. Security measures have to be taken to guarantee the *confidentiality*, *integrity*, and *availability* of personal data. Personal data also has to be kept *accurate*, *relevant* and *up to date*.

Supervision and sanctions. An independent data protection authority has to be designated and should be responsible for supervising the observance of privacy provisions. In the event of violation of the provisions of privacy legislation, criminal or other penalties should be envisaged.

2.1.2 Privacy vs. Security

Privacy and security are two different concepts that do not necessarily deal with the same issues, but they are related. Security deals with prevention and detection of unauthorised actions by users of a (computer) system, and a definition of security (computer security) is often stated based on three aspects [26]:

- **Confidentiality**, prevention of unauthorised disclosure of information.
- **Availability**, prevention of unauthorised withholding of information or resources.
- **Integrity**, prevention of unauthorised modification of information.

Confidentiality deals with the fact that unauthorised users should not be able to read, or learn, sensitive information. In such, it is related to privacy, as privacy may be defined as the protection of personal data [26]. Further, in accordance with the Directions to the personal Data Act [2] [7], actions must be taken

- against unauthorised access, or inspection of personal data were confidentiality is necessary.
- to provide access to personal data when availability is necessary.
- against unauthorised modification of personal data when integrity is necessary.

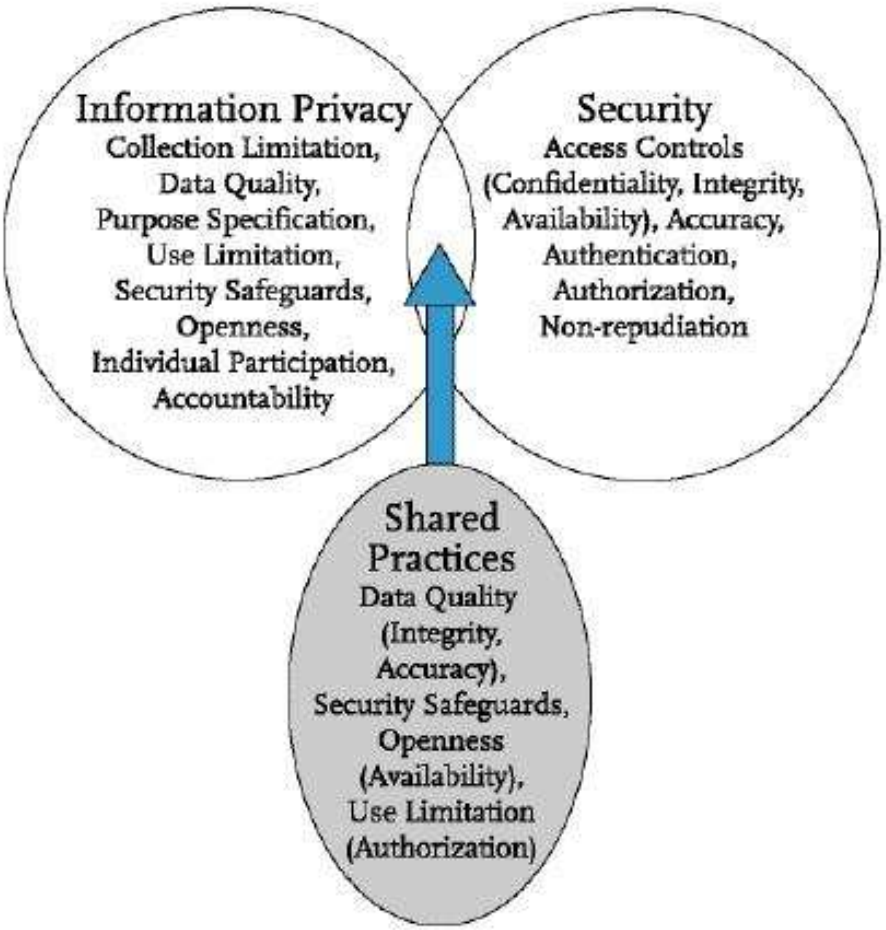

Figure 2.1: Elements of Information Security and Privacy [15].

[2] Forskrift til personopplysningsloven, in norwegian

Therefore, privacy and security are related, and privacy requires security (the principle of security and accuracy). Figure 2.1 illustrates that some components of privacy protection can be addressed by security safeguards, while others cannot. Some security functions may actually hinder or even threaten necessary privacy protection (i.e. intrusion detection systems and auditing). This is called the *security-privacy paradox* [15].

In figure 2.1 privacy is defined based upon the OECD guidelines [9], see appendix C, which includes the principles of collection limitation, data quality, purpose specification, security safeguards, openness, individual participation and accountability, and relates to the privacy principles stated above in subchapter 2.1.1. As shown, security measures does e.g. not cover the collection limitation principle which state that there should be limits to the collection of personal data, that the personal data should be obtained by lawful and fair means, if possible, with the concent of the data subject, and the purpose specification principle, which states that the purpose for which the personal data is collected must be specified upon collection. Further, the individual participation and accountability principle is not covered.

2.1.3 Aspects of Privacy

There are different aspects of privacy that can be pointed out, as recognized by [1]. These include:

- **Territorial privacy**, by protecting the close physical area surrounding a person, i.e. domestic and other environments such as the workplace or public space.
- **Privacy of the person**, by protecting a person against undue interference, such as physical searches, drug testing or information violation his/her moral sense,
- **Informational privacy**, by controlling whether and how personal data can be gathered, stored, processed or selectively disseminated.

For the remainder of this thesis, privacy should be understood as informational privacy, and personal data, for the remainder of this thesis also referred to as personal identifying data, will refer to any information that may be linked to a person, that is, the data subject.

Fully protecting a person's privacy may not be possible, as privacy is not an absolute right due to the fact that individuals cannot participate fully in society without revealing a certain amount of personal data. Disclosure of personal identifying information may be necessary, e.g. to provide personalised services. However, measures should be taken to give the individuals a level of protection that they require, and they should be given the ability to control whether and how personal data could be gathered, stored and processed, in a manner such as to preserve that level of protection.

2.2 Threats towards Privacy

With the advent of ICT systems, privacy has been getting increasing attention. ICT systems have the ability to store and process personal data, and a potential negative impact of this is invasion of privacy and misuse of personal data stored in the systems. Earlier, invasion of privacy was limited by the physical abilities to see, hear, and remember events. Over the years, new technologies have made the recording of events easier, potentially threatening privacy. The camera, the telephone, the microphone, the tape recorder, the television, the computer, the Internet and now, even mobile phones with connectivity and cameras, have all extended our abilities to communicate and made it more difficult to remain private.

Threats to an individual's privacy is mainly related to two reasons, see [5]. First, the amount of personal data collected and processed is increasing, e.g. due to the ability of ICT systems to store and process personal data. The trend in collecting increasing amount of information is expected to continue, and as a result, many details in the lives of people are being documented in databases somewhere. Secondly, it is difficult for individuals to control their own privacy. It may be hard for individuals to get an overview over which systems that actually collect and store their personal data, and what personal data they collect and store.

Some examples of threats are provided below [5]:

- Tax records and other public records made available on the Internet make efficient search and aggregation of information about individuals possible. Identity thefts and fraud are common uses of information gathered in this way.
- Outsourcing of services is increasingly popular. This gives opportunities to gather personal identifying information from different sources.
- There is also a development towards fewer and larger companies, which imply linking of information systems and databases. As this also happens across what used to be separate sectors, such as banking and insurance, a lot of power and information is gathered in few hands.
- Location-based services are assumed to become very popular, and this opens up new possibilities for tracking of individuals and building personal profiles.
- Threats from insiders. Even though the organisation responsible for stored personal information does not have any malicious intent, all its employees may not be that trustworthy, one such example is found in [25]. Hence, protective measures against disloyal employees are needed.
- The number of surveillance cameras is rapidly increasing, and particularly mobile phones with cameras and digital cameras may threaten individuals privacy, as it is easy to take photos, store them and disclose them to third parties.

- The growing amount of personal data that is collected and processed may also be communicated through networks across state borders [1]. It is a severe privacy threat if sensitive personal data, i.e. medical records, are communicated to and perhaps routed via different countries, which do not necessarily have appropriate privacy legislation.

- Data mining, which promise to efficiently discover valuable, non-obvious information from large databases, is vulnerable to misuse, and may compromise privacy [14]. The future of data mining relies on techniques that incorporate privacy concerns. Databases may also in the future incorporate privacy functionality, for a strawman design for such a database concept, see the Hippocratic database concept in [22].

Also, threats to privacy, and privacy violations, can be categorized on the type of intruder, or violator:

- Insiders, authorised users of the system who attempt to gain additional privileges for which they are not authorised, or who misuse the privileges given to them, i.e. due to accidental disclosure, insider curiosity and subornation.

- Outsiders, unauthorised users (attackers) accessing the system from the internet.

For the remainder of this thesis it will be assumed that the privacy threats mainly arise due to insider attacks. Outsider attacks relate to security issues, and should be prevented by implementing security functions such as intrusion detection systems (IDSs) (see subchapter 3.4), firewalls and access control. It is, however, assumed that ones the outsiders have successfully penetrated the system, they are "insiders".

Subchapter 2.5 includes an overview of privacy protective measures.

2.3 Privacy Policies, Promises and Agreements

A privacy policy demonstrates an organisations commitment to privacy, and describe how information collected about individuals should be processed and possibly disclosed. That is, a privacy policy may define which data is collected, for what purpose the data will be processed and accessed, whether the data may be disclosed to a third party or not and for how long the data will be retained. Clearly, the local privacy policy must be compliant with national laws and regulations, and may be defined as a superset of privacy policies mandatory to a particular ICT system by national laws and regulations. Organisations also have the option of adding more and stronger privacy requirements, to suit their, and their customers, privacy needs.

Here, a privacy policy will be defined as:

> A privacy policy is a set of general rules and principles that describes how information about an individual will be collected, processed and disclosed by a data collecting entity. The privacy policy must be compliant with applicable national laws, regulations and international guidelines.

Further, a privacy policy may be positive or negative. A positive policy explicitly state what is allowed, with respect to the personal data, to do in the system, whereas a negative policy explicitly state what is not allowed. A privacy policies may also be written in a formal language such as the Enterprise Privacy Authorisation Language (EPAL) [3]. EPAL is briefly introduced in Appendix B.

The privacy policy, which is local to a data collecting organisation, forms the basis for the data collecting organisation's privacy promise. A privacy promise is a set of agreements that that the data collector is willing to accept, and states the privacy principles the data collector promises to fulfil. Then, data subject can modify this based upon their privacy preferences, through opt-ins and opt-outs, if present, defined in the privacy promise. Figure 2.2 shows the relationship between an privacy policy, privacy promise and a privacy agreement.

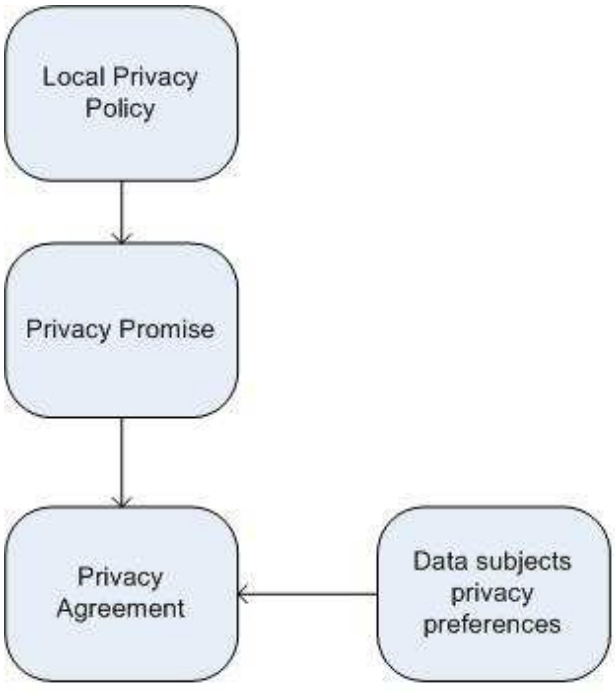

Figure 2.2: The making of an privacy agreement

An opt-in choice means that the organisation will not use the personal information unless the person specifically agrees to it and an opt-out choice means that the

[3] See e.g. http://www-1.ibm.com/services/security/epa.html. EPAL is also briefly outlined in appendix B.

organisation can use the information unless the person specifically says not to. Such individual agreements, here called privacy agreements, may be defined as a set of rules that determine how the personal data collected, about an individual, can and should be used. An agreement is derived from the privacy promise of the data collector and the privacy preferences of the data subject, and both parties consent to it.

The Word Wide Web Consortium's platform for privacy preferences (P3P) [23] allows users to be informed about a website's privacy policy. It allows web sites to present their data-collecting practices in a standardized, machine-readable, easy-to-locate manner. P3P user agents (P3P enabled browsers) read the published privacy policy and compare them with the user-specified privacy settings. P3P also enables web users to understand what data will be collected by sites they visit, how the data may be used, and what data they may opt-in or opt-out of. That is, a data collecting organisation may encode their privacy promise in P3P, and P3P may help the user opt-in and opt-out according to the privacy promise in order to agree on an acceptable agreement. The agreement could then e.g. be encoded in EPAL.

An understanding of the privacy policy is important in order to understand, or defined, what constitutes a privacy violation. Privacy violations are discussed next i subchapter 2.4.

2.4 Privacy Violations

Privacy refers to the right of individuals to control information about them, to keep it secret or to share it with others as they feel.

Defining a privacy violation, or an invasion of privacy, is not straightforward. What violates privacy is highly subjective, as privacy concerns our private life, and our right to control whether or not we want information about us to be communicated to the world.

Here, a privacy violation will be defined as follows:

> A privacy violation is an event that breaches the privacy policy or a privacy agreement between a data subject and a data collector.

Hence, what constitutes a privacy violation depend on the privacy policy defined by the data collecting organisation or on the individual agreements, see subchapter 2.3, and therefore, what constitutes a privacy violation may vary from organisation to organisation, and individual to individual. Due to the subjective nature of privacy, individuals are likely to establish different agreements with the data collector, if they are given the opportunity.

An example of a privacy violation include a drugstore chain that sold customer's medical information to a marketing company that sent consumers coupons for drugs related to their disorders. Obviously, customers trust such information to

be treated as confidential, and this is clearly a privacy violation. Generally, if the personal data of a data subject is accessed for a purpose that she/he has not consented to, it will be regarded as privacy violations, as it violate the agreement between a data subject and data collector.

Another example of a privacy violation is Lilly, a pharmaceutical company that disclosed e-mail addresses of their subscribers to its prozac remainder service [4].

Clearly, privacy violations must be avoided, and privacy protected. The next subchapter outlines some privacy protection measures.

2.5 Privacy Protection

Privacy is increasingly important in an increasingly networked society, and individuals' privacy should be protected by some means. Generally, privacy protection can be supported by, see [1]:

- Privacy and data protection laws promoted by government.
- Self-regulation by businesses.
- Privacy education of consumers and IT professionals.
- Privacy-enhancing technologies adopted by individuals or businesses.

2.5.1 Privacy and Data Protection Laws

To protect individuals' privacy, most nations have their own privacy protection laws. Sweden's Data Act from 1973 was the first national data protection act in the world. Norway got its data protection act a few years later, in 1978. However, privacy problems are not limited to one nation, privacy issues also arise due to communication across borders. This problem where recognized by the Organization for Economic Cooperation and Development (OECD) which in 1980 adopted its Guidelines on the Protection of Privacy and Transborder Flows of Personal Data. These guidelines were intended to help to harmonize the different national laws and enforce some minimum degree of privacy protection amongst the member countries, see [9]. The European Council also acknowledged the need for privacy protection and the EU Directive 95/46/EC on the protection of individuals with regard to the processing of personal data and on the free movement of such data was adapted in 1995, see [8].

Privacy and Data Protection Laws are described in Appendix C. The privacy principles described in subchapter 2.1.1 serves as an summary of the most essential privacy requirements formulated by the OECD Guidelines, the EU directive and other national laws.

[4] See http://www.ftc.gov/opa/2002/01/elililly.htm.

2.5.2 Self-regulation

Self-regulation bases itself upon the fact that organisations and the industry will restrict their practises according to fair information principles, such as the OECD Guidelines, see [9].

In order for enforcement of self-regulation to be effective, verification and monitoring should be in place. This could e.g. be undertaken by third party enforcement program or a seal program such as TRUSTe. TRUSTe [5] is a "seal of approval" that uses a standardised icon to link to an organisation's privacy practices and to indicate that these practices are monitored by outside auditors. Further, complaint resolution and education of consumer and businesses should be in place.

2.5.3 Privacy Education

Privacy issues should be an important part of the education of computer scientists, as they are responsible for a lawful ethically acceptable system design and system administration, see [1]. Privacy should be a design criterion, and emphasis should be on designing privacy-friendly and privacy-acceptable systems, and assessing privacy risks of ICT systems. Further, legal requirements of data protection acts have to be fulfilled, so computer scientist should be familiar with legal requirements of data protection and how to enforce them.

Also, in order for privacy-enhancing technologies (PETs), see subchapter 2.5.4, to play a more significant role in this networked society, it will be necessary to create more public awareness as well as consumer demand for them [10]. Therefore, privacy education will be important. If there is a demand for such technologies, providers will most likely respond to the market forces. PETs may also be assumed to provide a competitive advantage because they increase users' trust in the services and technologies involved.

In order for PETs to be effective means to technically enforce privacy aspects, users and consumers must have sufficient technical knowledge to apply them. Users and customers need information and education about their rights, about the value of their personal data, about privacy risks and the possibilities of self-protection by the use of PETs.

2.5.4 Privacy Enhancing Technologies (PETs)

Privacy cannot be sufficiently protected by privacy legislation alone [1], and enforcement of the legislation can be automated and made mandatory through technological measures. There are two approaches to the implementation of privacy-enhancing or privacy-assuring technologies and processes, as defined in [2]:

1. Minimize the amount of personally identifiable data through pseudonymisation and anonymisation, or by simply not collecting any data at all.

[5] See http://www.truste.org/

2. Ensure that the privacy agreement between the data subject and the data collector is enforced.

There is no conflict between these two approaches; they are both important means to protect and enhance our privacy. They can also be used in conjunction with each other. Generally, PETs that minimize the amount of identifiable data through pseudonymisation and anonymisation, can be adopted by both individuals and data collecting organisations, whereas PETs that ensure the enforcement of a a privacy agreement are more likely to be implemented by data collecting organisations, e.g. in order to increase the data subjects trust in their practises. A privacy violation detector (PVD), as described in this thesis, will be a part of frameworks that enforce a privacy policy and policy agreements.

Clearly, anonymity is an important privacy property. *Anonymity*, defined in [12] and [11], ensures that a user may use a resource or service without disclosing his or her identity. In cases where anonymity cannot be provided, e.g. in systems where accountability is required, *pseudonymity*, defined in [11], can protect the user's identity by using pseudonyms as IDs. Pseudonymity ensures that a user, acting under a pseudonym, may use a resource or service without unconditionally disclosing his or her identity. Pseudonyms may be personal (related to an individual) or role-based (related to the role the individual is currently performing). Additionally, cryptographic and one-way pseudonyms exist. Cryptographic pseudonyms are generated by encrypting the identity data, whereas one-way pseudonyms are implemented by a cryptographic one-way function. One-way pseudonyms can not be reversed, and therefore offer a higher degree of protection than cryptographic pseudonyms.

Unobservability and unlinkability are two other privacy properties that may be fulfilled by PETs. *Unobservability* of the user ensures that a user may use a resource or service without others, e.g. third parties, being able to observe that the resource or service is being used [12], whereas *unlinkability* ensures that a user may make multiple uses of resources or services without others being able to link these uses together [12]. For a formal definition on anonymity, unobservability, unlinkability and pseudonymity, see [1],[11] and [12].

Protecting the user and data subject identities

PETs can be used for protecting user identities by providing anonymity, pseudonymity, unlinkability and unobservability, i.e. minimize the amount of personal identifiable data stored, and this subchapter describes some techniques for protecting the user and data subject identities.

Personal data is often collected for research purposes, were the scientists rarely need to know the identities of the data subjects. The personal data should therefore be depersonalised. Depersonalisation can be reversible (pseudonymity) or non-reversible (anonymity). However, perfect depersonalisation, were the data subject is no longer identifiable, is practically impossible to achieve. Generally, data being collected for statistical purposes contain the following types of data attributes:

- **Identifying data** such as name, address and phone number.
- **Demographic data**, such as sex, date of birt, place, nationality, education, religion, marital status.
- **Analysis data**, such as diseases and habits. Generally, this is the data for which the statistical analysis is conducted.

Here, the records may be rendered anonymous or pseudonymous in order to depersonalise the it by removing the identity information from the record (anonymisation) or by replacing user identifying data with a pseudonym.

However, is it enough to simply remove (pseudonomise) the identifying data from the record in order to obtain anonymity (pseudonymity)? No. Even if an record is depersonalised an attacker with adequate supplementary knowledge about the demographic information of a data subject could use the knowledge to re-identify the data subject and disclose sensitive information. This risk of re-identification should therefore be taken into account.

There are efforts on minimizing this risk, such as computational disclosure control, see [13]. Computational disclosure control is motivated by the fact that depersonalisation does not equal anonymity [6]. Basically, computational disclosure control seeks disclosure of data such that interference about identity of people and organisations and about sensitive information contained in the released data cannot reliably be made, see [13].

The data subject's privacy may also be protected by making the personal identifying data less granular. That is, by replacing the address with e.g. the name of the municipality, and by replacing the exact date of birth with e.g. an interval containing that date of birth, it may be harder to re-identify the individual if the identity of the individual is anonymised or pseudonymised in the case of depersonalisation.

In [10] the concept of an identity protector was introduced. The identity protector (IP) provides anonymity, pseudonymity, unlinkability and/or unobservability for users and data subjects, by controlling the release of an individual's true identity to the various processes within the information system. It works by converting the user's actual identity into a pseudo-identity (a pseudonym), an digital identity that the user may adopt when using the system. The IP will generate pseudo-identities as needed, convert pseudo-identities into actual identities if needed (e.g. upon detection of a privacy violation to identify the violator), and combat fraud and misuse of the system (e.g. by preventing the user from using the anonymity as a shield to commit fraud).

An identity protector introduces two different domains to an information system, as shown in figure 2.3. One identity domain, in which the user's actual identity is known and accessible, and a pseudo domain in which the user's actual iden-

[6] In [13] it is shown that 53% of the US population can be uniquely identified on *place (e.g. city or town), sex* and *date of birth*. This demonstrates that by only removing name and phone number, the identity may be identified later, and is in fact not anonymised

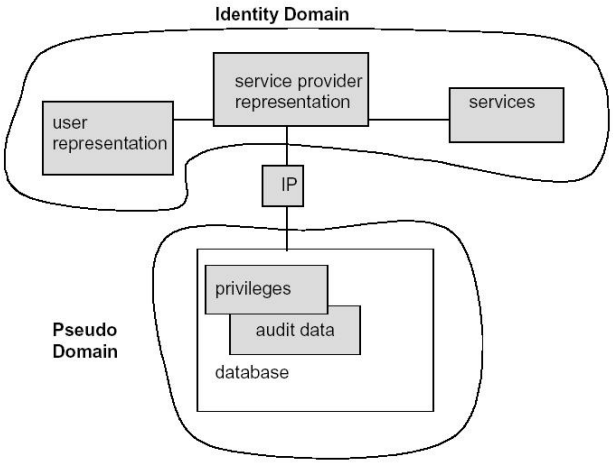

Figure 2.3: The Identity Protector ([1], [10]) prevents the registration of the users' real identity in the database.

tity is not known. In figure 2.3 the identity protector can be used to implement pseudonymous logging, in order to protect the users of the system.

The identity protector is implemented using techniques such as digital signatures, blind signatures, digital pseudonyms and trusted third parties, see appendix E.

Further, Mix-nets, described in appendix E, can provide both anononymity and unlinkability. A remailer technology, Mixmaster, provide anonymity, and is also briefly described in appendix E. Further, steganography can be used to provide unobservability. Steganography is the art and science of transmitting secret messages through innocuous carriers in such a way that the very existence of the embedded message is undetectable, see [1], and can be used to hide the existence of encrypted messages.

Enforcing Privacy Policies, Privacy Legislation and Privacy Agreements

By enforcing privacy policies, privacy legislation and privacy agreements, an individual's privacy will be enhanced. [2] outlines a framework for enforcement of privacy policies. This framework is also briefly described in appendix D, as this framework suggests the use of a privacy violation detector (PVD). Other efforts include IBM's Enterprise Privacy Architecture (EPA) [7]. The EPA helps organisations understand how privacy impacts business processes. The privacy architecture focus on meeting the needs of business use of data and the protection of individual privacy rights, and build a privacy policy from the bottom up, based on an analysis of business processes. In [1] Fishcer-Hübner also presents a formal task-based privacy model for enforcement of privacy policies. The idea is to control access to

[7] See e.g. http://www-1.ibm.com/services/security/epa.html

personal data through strict control of the tasks users perform, where a task consists of a set of allowed transformation procedures. Access to personal data is only allowed if it is necessary for the task, the user is authorised for the task, and the purpose of the task corresponds to the purpose stated when the information was collected, unless the user has consented to the new purpose. That is, the privacy model enforces the privacy requirements of purpose binding and necessity of data collection and processing.

Clearly, privacy enhancing technologies are valuable tools for privacy protection in addition to privacy legislation, as they will help complying with privacy legislations, and they are likely to become increasingly popular in the future. Some examples of PET that aim at minimizing the amount of personally identifiable data through pseudonymisation, anonymisation or not collecting any data at all are provided in Appendix E. An example of a framework for the enforcement of privacy policies is presented in Appendix D.

Chapter 3

Privacy Violation Detection

In an information and communication technology (ICT) system which stores and processes private information, identifiable down to the personal level, privacy violations may occur, either by intrusion into the system by outsiders, or by violation of access privileges by insiders. A Privacy Violation Detector (PVD) aims at detecting such privacy violations [1].

In [2] a PVD has been defined as:

> The privacy violation detector continually monitors access to personal data and detects misuse and/or anomaly behaviour.

A PVD system has many similarities with current Intrusion Detection Systems (IDS), and this chapter includes an brief overview of IDSs techniques in subchapter 3.4. Privacy violation detection is discussed in subchapter 3.5.

3.1 Enforcing privacy policies

As stated above, the PVD aims at detecting privacy violations. In subchapter 2.4 privacy violations were defined as events that breach a privacy policy, or some privacy agreement between a data subject and the data collector. The aim of the PVD is to detect privacy policy (or privacy agreement) breaches and in such try to enforce the privacy policy. For the remainder of this chapter, a privacy policy should be understood as the data collecting organisation's privacy policy, or the individual privacy agreements where such exist.

As stated in subchapter 2.3, a privacy policy must be compliant with applicable national legislation, regulations and international guidelines. The privacy principles stated in subchapter 2.1.1 defines some of the more important privacy principles that, most likely, are a part of a privacy policy. The PVD, generally, aims at

[1] As stated in subchapter 2.2, the focus is on insider attacks.

detecting privacy violations by comparing events occurring in a system to the privacy policy, and enforce the privacy policy. However, the PVD will not be able to enforce all parts of a privacy policy.

The PVD will be able to detect privacy violations related to processing of personal data, such as whether the processing of the personal data was according to the stated purpose. Related to this, the PVD will be able to detect privacy violations where the processing that occurred did not correspond to necessary tasks as defined in the privacy policy. Further, it will control whether personal data are being processed in a fair and lawful way. That is, the PVD is likely to enforce parts of the principle of purpose specification and purpose binding, the principle of necessity of data collection and processing and the principle of lawfulness and fairness stated in subchapter 2.1.1. However, these principles also refer to the handling of personal data upon collection, but the PVD will only be able to detect privacy violations during processing.

Further, the PVD will be able to detect violations relating to retention and possible disclosure of personal data, if such requirements are specified in a policy or an agreement.

3.2 Motivation

The obvious motivation for employing a privacy violation detector (PVD) is to be able to detect privacy violations and in such enforce the privacy policy, as discussed in subchapter 3.1. There are, however, other potential advantages as well.

Generally, access to personal data can be enforced using access control. Still, in many applications, e.g. medical information systems, it might be more important to be able to access information regardless of authorisation, e.g. in an emergency situation, than enforcing access rules. Then, using a privacy violation detection mechanism, such behaviour might be tolerated as violations will presumably be discovered, and proper actions taken against those violating the rules. Violations will be investigated after they occurred to see whether they could be justified given the presented circumstances. That is, a PVD will most likely be reactive, rather than proactive. Other security mechanisms, such as access control, are proactive mechanisms whose purpose is to prevent users from doing what they are not allowed to do. Reactive mechanisms, on the other hand, enable detection of a policy breach and cause some reaction after the breach happened, and are important, e.g. to build and maintain the users' trust in the system.

Clearly, establishing and maintaining trust is a challenge when dealing with personal data. It is important that the data subjects trust the data collection entity, because the data subject does not control the information once it is collected. However, the data subject should not have to trust all the employees of the data-collecting organisation. Enforcement systems, which the PVD will be a part of, help data-collecting organisations and data subjects to protect against malicious employees. Therefore, a privacy violation detection mechanism will help increase

the data subject's trust in the organisations handling of personal data.

A privacy violation detector might also detect weaknesses in the organisation, such as weaknesses in established guidelines, e.g. who is allowed to access what information, or weaknesses in system implementations.

A privacy violation detector may also have a preventive function. The fact that it exists, and that privacy violations most likely will be detected, may stop people from conducting privacy violating actions. Employees will have to think twice before commencing a privacy violation.

In addition, other administrative systems in an organisation may be very general, and not designed to cope with privacy legislations and privacy policies. Hence, the PVD could function as an add-on to existing systems helping them to enhance their privacy.

A PVD may be easier to add-on to a system compared to access control. This is illustrated by figure 3.1, which shows that access control needs to be integrated with the application, whereas the PVD could seamlessly be added to the system. The only issue related to the PVD is if the format of the logs generated by the application is different from the format required by the PVD, then the PVD will need some kind of adapter in order to be able to read the logs.

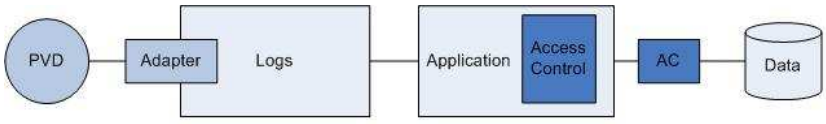

Figure 3.1: PVD vs. Access Control

Also, monitoring systems, such as a Privacy Violation Detector, are important parts of an internal control system, which are mandated by legislation, such as the Norwegian legislation, see §14 [2][6].

IBM also recognises the need for a privacy violation detection process [3], and defines it as an off-line check whether privacy has been violated.

3.3 Goals and Issues

Clearly, the main goal of the PVD is to detect privacy violations. Other goals include:

- The PVD should be easy to configure and update, as i.e. false alarms may trigger a need to do so.
- The PVD should provide minimal overhead for the monitored system.

[2] Internal Control
[3] See http://www.zurich.ibm.com/pdf/privacysummit/Karjoth.pdf.

- The PVD should not generate too many false alarms.
- The PVD should be informative. It should provide information so that an alert can, effectively, be further investigated by a privacy officer.
- The PVD should protect the privacy of the users using the monitored systems.
- The PVD should be easy to integrate with current systems.
- The PVD itself should be secured, that is, it should resist subversion and run continually.

It is important to agree upon a certain ambition level, that is, how certain we need to be that a detected privacy violation in fact is a privacy violation. How many false positives [4] are acceptable? When automating detection of privacy violations, a goal of no false positives is too strong. However, the number of false negatives [5] should be kept as low as possible. That is, all possible privacy violations should be detected, but the fact that some non-privacy violating events are marked as privacy violations is acceptable.

The aim is to provide an indication of possible breaches so that a human operator, a privacy officer, can further investigate these events. The PVD will function as a filter that analyse all events that occur in the monitored system, and provide the privacy officer with possible privacy violations and information relating to the violation.

Figure 3.2 shows the filtering aspect of the PVD. Here, three different outputs from the PVD are shown. The first output from the left illustrates a perfect output. There are no false positives and no false negatives. However, this is not very likely. To illustrate two more realistic cases, the second output from the left illustrates a situation where 75% of the actual privacy violations is detected. However, in this care, 50% of the non-privacy violating events are marked as potential privacy violations as well. Hence, the second output from the left outputs two false positives, and there is one false negative (PV3 is not detected). Clearly, this is not an ideal situation, as all privacy violations should be detected. The last output from the left illustrates a situation where all the actual privacy violations are detected, but there are still false positives. This is an acceptable situation, and represents the aim of the PVD. The most important goal is to reduce the number of false negatives while keeping the degree of false positives on a manageable level, that is, a certain degree of false positives may still be acceptable. This is because false positives only represent more work for the privacy officer, who will have to investigate events that turn out to be not privacy violating at all. False negatives may actually, if not detected, compromise a user's privacy.

[4] False positives are ordinary non-privacy violating events that are flagged as potential privacy violations by the PVD.

[5] False negatives are actual privacy violations that are not flagged as potential privacy violations by the PVD

There are, however, a potential conflict related to the ratio of false negatives and false positives. By aiming at less false negatives, more false positives may have to be accepted. As stated above, the number of false positives accepted should be kept on a manageable level. Generally, the more general the filtering (detection) rules are, the more false alarms are likely to occur. That is, the larger the size of the holes of the filter is, the more likely it is that all privacy violations will be detected, and that the number of false positives will increase. Therefore, the hole of the filter, or the generality of the rules, must be configured so that all privacy violations will go through and only a minimum of false positives occurs. If the rules are too specific, less false positives will occur, but it will be more likely that false negatives occur.

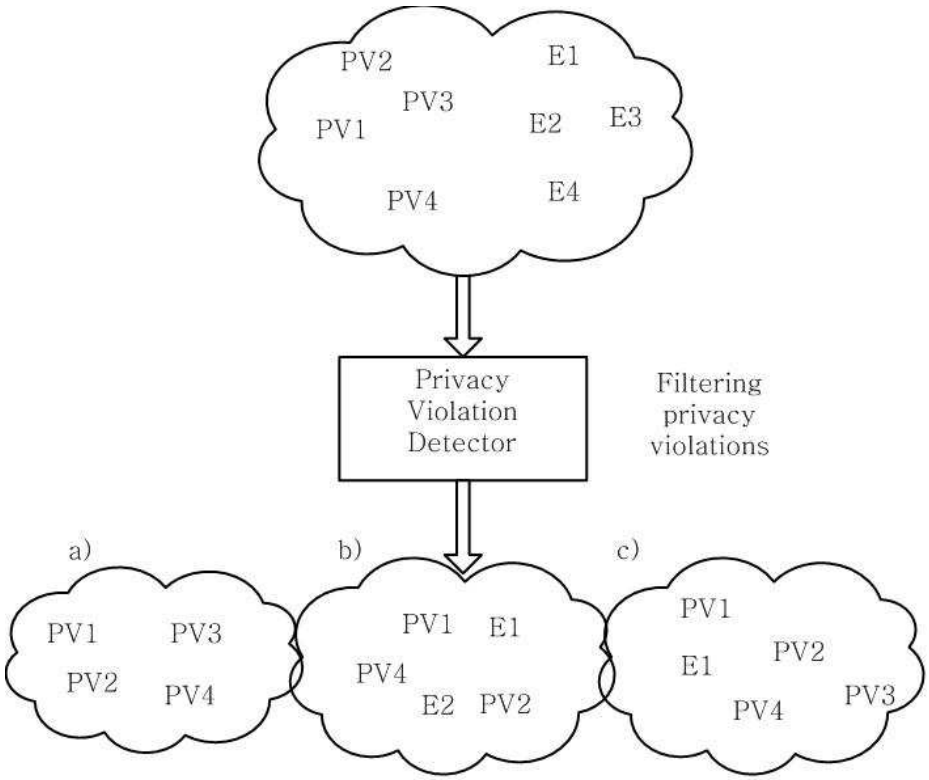

Figure 3.2: Filtering function of the PVD. Output a: Complete, no false negatives or positives. Output b: Not complete. False positive (PV3 not detected), and false negative (E1 and E2). Output c: Not complete, but no false positives make this output more acceptable than output b.

As briefly mentioned above, the PVD should output precise and consistent information relating to the potential privacy violation so that the privacy officer efficiently can investigate whether a privacy violation did in fact happen, or not. Clearly, the informativeness of the PVD is an essential feature, as the more high-quality (that is, relevant) information the privacy officer has about a potential privacy violation, the less time the privacy officer is likely to spend on the investigation. As one of the

goals associated with the PVD is to reduce the time spent by the privacy officer to detect privacy violations (compared to manually going through the whole log file), information should be provided along with the alarm to assist the privacy officer as much as possible. Different detection methods will provide different amounts (and type) of information, so the detection methods should be considered in relation to the need for information. This issue will briefly be discussed in subchapter 3.5.1.

By monitoring and logging events in a system in order to detect privacy violation, another issue rises. Is logging of events in order to detect privacy violations itself a privacy violation [21]? Too much fine-grained logging in order to protect a data subject's privacy can actually lead to violation of the user's privacy. By gaining access to the log, a person may be able to generate profiles about the users of the system. Therefore it is important that the privacy of the users is considered as well as the privacy of the data subjects. As logging of events may enhance the data subject's privacy, the user's privacy may be accomplished by using pseudonymous logging, see [21], introducing an identity protector, see [10], or simply regulating access to the log files. It is also important to be sure about the facts and discrete when investigating privacy violations, as being innocently accused of a privacy violation itself may feel as a privacy violation. Users of the systems must also be informed about the fact that their actions on the systems will be logged, and the logs analysed for signs of privacy violations.

Integration with current systems is also important, e.g. for its usability. Generally, the PVD should be easy to add-on to existing systems, enhancing their privacy protection. Practically, the integration with current systems may prove to be hard, as different systems and applications have different log formats and data types. Further, different application may log different things. It is not certain that all applications e.g. log the purpose of access to personal data, which is relevant in the case of privacy violation detection (the principle of purpose specification). Therefore, a PVD may need a adapter function, e.g. as shown in figure 3.1, according to its "target" system.

The PVD itself should be secured, that is, unauthorised or authorised users intentionally violating privacy, should not be able to "cover their tracks" by e.g. deleting entries in the log file.

It is also important to note that the PVD cannot:

- compensate for significant deficiencies in your organisations privacy policy
- compensate for privacy weaknesses in network protocols
- substitute for other types of security mechanisms such as authentication, encryption, firewalls or access control.

It is also important to note that when using a PVD it is important that users of the system are notified about the use of the PVD, and that they are informed about the privacy policy and the fact that their activities may be logged.

3.4 Intrusion Detection Systems

As this thesis aims at discussing whether or not current Intrusion Detection Systems (IDSs) technologies can be used for detecting privacy violations, an overview of IDSs is provided.

Intrusion Detection Systems (IDSs) are used to detect intrusions into computer systems and networks by attackers, and are in that way analogous [6] to burglar alarms protecting homes, buildings and other physical things. A definition of IDSs are provided below:

> Intrusion Detection Systems (IDS) are software or hardware systems that automate the process of monitoring the events occurring in a computer system or network, analyzing them for signs of security problems [17].

As the number of attacks on computers and networks has increased over the past few years, intrusion detection systems have become a increasingly important part of the security architecture of many organisations. Reasons for the growing number of attacks include increased connectivity and complexity and dependence on distributed network services. The increasingly sophisticated attacks towards computer systems and networks also motivate the need for IDSs. This trend is shown by figure 3.3. The figure shows that while the sophistication of attacks are increasing, the knowledge of novice intruders are decreasing because today anyone can attack a network due to widespread and available intrusion tools and attack scripts via the Internet.

Organisations may also choose to implement intrusion detection systems to document existing threats to the organisation, and to obtain useful information of intrusions that actually do take place.

Earlier, intrusions where detected based on manual inspection of the *audit trail*[7] generated by the operating system. However, the large size of such data makes manual inspection of these logs impractical. Intrusion detection systems automate the process of inspecting the audit log and other information.

Clearly, the aim of IDSs is to detect attacks as soon as possible to make it possible for the system operators to take appropriate action. In turn, information based on detected intrusions and their alarms may provide useful information that can help prevent such attacks from happening again, e.g. by tuning the IDS.

Figure 3.4 illustrates a possible organisation of a intrusion detection system. Generally, an intrusion detection system contain at least the following elements, see [19]:

- **Audit collection**. Audit data from which to make intrusion detection de-

[6] IDSs are almost analogous to burglar alarms as not all IDSes provide an alarm, it may only detect and log the suspected intrusion.

[7] The audit trail is a record of all activities on a system logged to a file in chronologically sorted order.

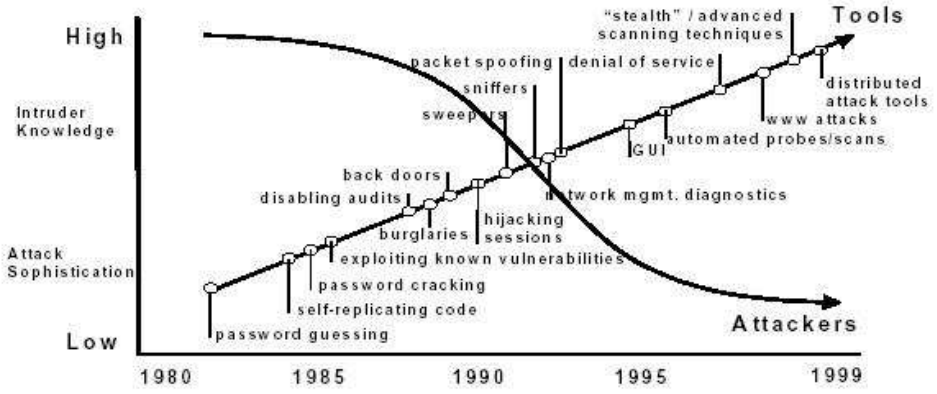

Figure 3.3: Attack Sophistication vs. Intruder Knowledge [18]. The sophistication of attacks are increasing, whereas the knowledge of novice intruders are decreasing.

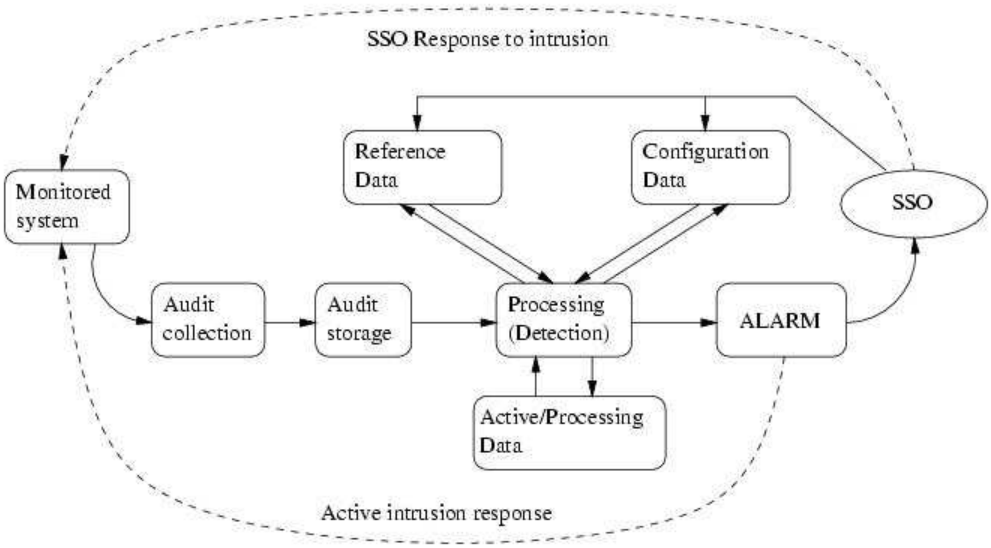

Figure 3.4: Organisation of a generalised intrusion detection system [19].

cisions must be collected. Different parts of the monitored system can be used as sources of data. Here, a network-based and host-based (application or operating system) approach will be described.

- **Audit Storage**. The audit data is stored, either for a long time or temporarily awaiting processing.
- **Processing**. This is the heart of intrusion detection systems, here different algorithms are executed to find evidence of suspicious behaviour in the collected audit data. Two detection techniques commonly used in IDSs - anomaly detection and misuse detection - are described later.
- **Configuration Data**. This includes information relating to how and where to collect audit data and how to respond to intrusions.
- **Reference Data**. This element stores information about known intrusion signatures and/or profiles of normal behaviour that are used by the processing element to detect intrusions.
- **Active/Processing Data**. As the processing element may need to store intermediate results and state, this element is needed.
- **Alarm**. This part of the system handles all output from the system, whether it is an automated response to the suspicious activity, or it involves notifying some site security officer.

The remainder of this subchapter gives an overview of different aspects of intrusion detection systems available today, based on the way they analyse and monitor events. Readers familiar with IDSs may skip to subchapter 3.5.

3.4.1 Detection

There are many different ways of analysing events on a system in order to detect whether or not an attack in fact happened. Two methods, namely anomaly detection and misuse detection, are discussed below.

Anomaly Detection

Anomaly detection assumes that all detected abnormal behaviour is intrusions or intrusion attempts. Such intrusion detection systems maintain a profile of normal activity for a system user, and then flag all activity varying from this profile as intrusion attempts.

Unfortunately, anomaly-based IDSs often produce a large number of false alarm (false positives), which are related to the fact that profiles of users system use and actual system behaviour can vary wildly. In order to generate profiles of normal behaviour, such IDSs also require extensive training sets of system events. They are, on the other side, able to detect new attacks as they are able to detect possible

attacks without specific knowledge of the attack type. Other issues include selecting threshold levels (how many file accesses per day is normal?), and the selection of features to monitor. Additionally, IDSs based on anomaly detection may be computationally expensive due to the overhead of keeping track of, and updating, several system profile metrics. Further, the informativeness of an anomaly-based IDSs may be limited, as it may be harder for the security operators to figure out exactly what triggered the alarm. However, if the profile is explicit and unambiguous, an anomaly-based IDS may be able to provide useful information.

Techniques used in anomaly detection may be statistical measures or threshold detection, see [17].

Misuse Detection

Misuse detectors analyse system activity, and look for events (or sets of events) that corresponds to predefined patterns (also called signatures) of events that describe known attacks.

An advantage of misuse detectors is that they can be tuned not to generate an overwhelming number of false alarms by removing signatures that are to general or do fit the target system, but they can only detect attacks they already know about. Therefore, intruders only need to slightly modify the attack method in order to bypass the detector. Further, the informativeness of misuse-based IDSs may be greater than anomaly-based IDSs, as the misuse detector knows which signature that matched, and can provide information relating to what occurred.

3.4.2 Information Sources

Intrusion detection systems may also be classified on how they monitor the systems and network, that is, what information sources they based their analysis on. Generally, there are two kinds; network-based and host-based (application-based) intrusion detection systems.

Network-based

Network-based intrusion detection (NIDS) systems detect attacks by capturing and analysing network packets [17]. By listening on a network segment or switch, one NIDS can monitor the network traffic affecting multiple hosts that are connected to the network segment, thereby protecting the hosts connected to the network segment. Therefore, one NIDS are able to monitor a large network (many hosts). Further, a NIDS consists of passive devices that listen to the network without interfering with the operation, and can be made reasonably secure against attacks and invisible for the attacker. Network protocols are standardized, so network-based IDSs do not have to be customised to the same extent as host-based IDSs.

However, in a large network the NIDS may fail to recognise an attack as it may be difficult to process all packets. Also, in many cases it cannot analyse encrypted

information, or tell whether an attack was successful or not. It can only tell that an attack was initiated. Additionally, NIDS may have problems with attacks that involve fragmented packets.

Host-based

A host-based intrusion detection system (HIDS) operate on information collected from within a particular computer system, and by analysing system logs, operative system logs and application logs (then also called application-based). As opposed to a NIDS, a HIDS can easier see the outcome of an attempted attack, and may therefore be able to provide the security officer with more information relating to the intrusion than NIDS. Further, HIDS can detect attacks that cannot be seen by a NIDS, as they have the ability to monitor events local to a host. HIDSs can also operate in environments where network traffic is encrypted, because it may gain access to the information after decryption.

The disadvantages include the fact that they are harder to manage than NIDS, as information must be configured and managed for every host monitored. Also, the IDS may be attacked and disabled as a part of the attack. Therefore, it is important that the logs are protected, in particular the application log that is not so protected as e.g. operative system logs. Further, HIDSs may not be suited for detecting network scans or other surveillance that target the entire network, and they use computing resources on the host they monitor.

An advantage of application-based IDSs is that they have the ability to detect suspicious behaviour due to authorised users exceeding their authorisation. Problems are most likely to appear in the interaction between the user, the data and the application, and application specific knowledge can contribute to the detection of such violations (insider intrusions).

3.4.3 Other features

Other features of IDSs include response options, timing and control strategy options. These features are, however, only briefly described, as they are not so important for the further discussion on privacy violation detection.

Response

Once IDS have obtained information and analysed it for symptoms of attacks, they generate responses, or alarms. Generally, responses may be passive or active.

Passive responses provide information to a system operator, e.g. the security officer, relying on him to take subsequent action based on that information. Active responses are automated actions taken when intrusions are detected. A passive response may include recording the results of the analysis in a log file, which in turn will be studied by a system operator. The system could also immediately

trigger an alarm, i.e. an alarm flag, message or e-mail. Then, upon the alarm, the security officer will take action.

Generally, the alarm should provide the security officer with information that may shorten the intrusion investigation time. As briefly mentioned earlier, misuse-based IDSs may provide more information related to a detection than anomaly-based, and a HIDS may be able to provide more information than a NIDS, as it can see the outcome of an attack.

Timing

Timing, referring to the time between the events that are monitored and the analysis of those events, can be either interval-based (batch mode) or in real-time (continuous) [17].

In an interval-based intrusion system the information flow from the monitoring points to the detection engine is not continuously. Real-time based intrusion detection systems, on the other hand, transfer the information from the monitoring points to the detection engine immediately. Therefore, only real-time based intrusion detection systems make active responses possible, and allow the IDSs to take action that affects the progress of the detected attack.

Control Strategy

IDS also differ in the way that they are controlled, how the elements of the IDS are controlled and how the input and output of the IDS is managed.

Generally, there are three options, see [17] for illustrations:

- centralized
- partially distributed
- fully distributed

With a centralized control strategy, all monitoring, detection and reporting is controlled from a central location. With a partially distributed approach, monitoring and detection is controlled from a local node, with hierarchical reporting to one or more central location(s), whereas with a fully distributed approach, monitoring and detection is done using an agent-based approach, where response decisions are made at the point of analysis.

3.5 The Privacy Violation Detector

As previously stated, the aim of a privacy violation detector (PVD) is to detect privacy violations in ICT systems that store and process personal identifying information.

Currently there is no good method to automate detection of privacy violations. It might, however, be possible to draw lines to current intrusion detection systems (IDS) which are used to detect intrusions into computer systems and networks by attackers. This subchapter aims at outlining a PVD and discusses whether the detection methods described in subchapter 3.4, which gave a brief introduction to IDSs, are applicable to detecting privacy violations. The other aspects of intrusion detection systems (IDSs), such as information sources, will also be considered in the PVD context.

The process of detecting privacy violations can be separated in to, at least, two different parts:

1. collection and organisation of data that, potentially, document privacy violations.
2. analyse the collected data to look for potential privacy violations

The first part of the process may be referred to as monitoring, or information sources, discussed in subchapter 3.5.2, whereas the last part may be referred to as detection, discussed in 3.5.1. With regards to the PVD, the detection is of most interest. However, the detector will not be able to detect privacy violations without information sources that feed the PVD with information. Information sources will therefore also be discussed here.

Figure 3.5 illustrates the context which the PVD will reside in, here called a privacy violation detection system (PVDS). The illustration borrows from the terminology and illustration provided in [19], and shown in figure 3.4.

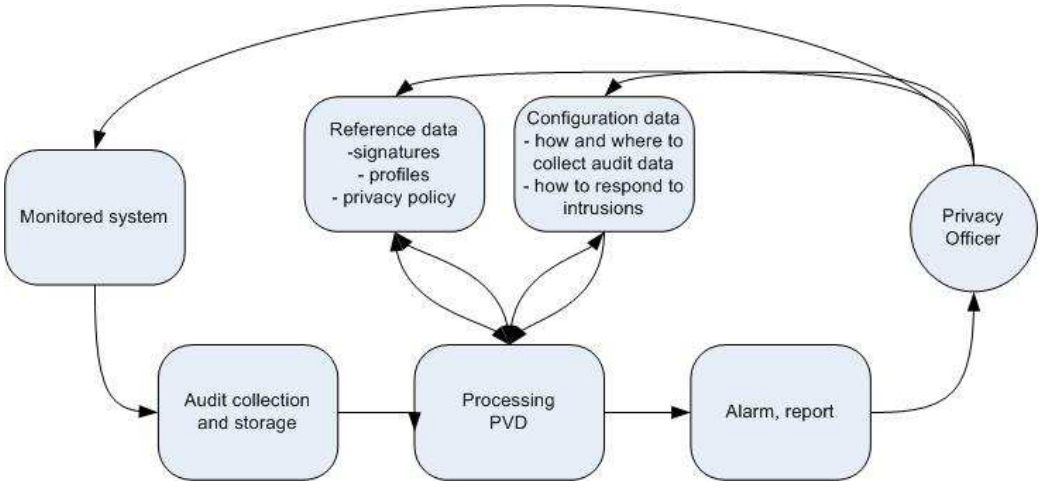

Figure 3.5: Organisation of a Privacy Violation Detection System

The figure shows the PVD, being the heart of the PVDS and performing the actual detection of potential privacy violations based upon processing of privacy relevant information from the monitored system. Upon an detection of a privacy violation,

an alarm is raised, and information relating to the potential privacy violation given to the privacy officer. Needed actions will then be taken by the privacy officer, and potential changes made to the reference data or configuration data, e.g. if a false alarm occurs.

Then, in order to make the investigation of privacy violations more efficient, the informativeness of the detection and the alarm is important. As the different detection methods may differ with respect to informativeness, this will be discussed later.

Next, detection methods are discussed followed by a discussion on different information sources that the PVD can rely on. Then, some other features, such as response and timing, are briefly discussed.

3.5.1 Detection

For detecting privacy violations, different detection approaches exist, including:

- anomaly-based detection.
- policy-based detection.

Anomaly detection requires that a profile of a normal behaviour is generated, and that a certain threshold or a statistical deviation from this normal profile is defined as abnormal and possibly a privacy violation. This corresponds to anomaly detection in IDSs; the difference is that for an anomaly-based PVD typical abnormal behaviour that corresponds to a privacy violation has to be defined, whereas for an anomaly-based IDS abnormal behaviour that corresponds to a intrusion have to be defined. Also, the profiles used as a benchmark may not be the same, as a profile for a normal non-privacy violating behaviour may focus on other things than a profile for detecting intrusions. A profile for a PVD may for example state that it is normal for nurses to access the nursing data for patients in their ward, and that if a nurse accesses the personal identifying data (contact data) for the patient, it is abnormal, and considered a privacy violation from a PVDs point of view.

A policy-based detection approach requires that audit data are compared to a defined privacy policy (or a privacy agreement). However, this requires that the privacy policy exist in a machine-readable format, such as XML, e.g. by encoding the privacy policy in EPAL [8]. A privacy policy may, as mentioned in subchapter 2.3, be positive or negative. A positive privacy policy explicitly state what is allowed in the system, whereas a negative privacy policy explicitly state what is *not* allowed. If the privacy policy is negative, then this approach converges to the misuse-based approach mentioned in subchapter 3.4. Misuse-based detection requires a signature, or a pattern, for each known privacy violation that audit

[8] Enterprise Authorization Language: A formal language for writing enterprise privacy policies to govern data handling practices in IT systems according to positive and negative authorisation rights, see http://www-1.ibm.com/services/security/epa.html.

data from the monitored system can be examined against. If the privacy policy is positive, all accesses may be compared to the machine-readable privacy policy, and if the access (log entry) does not correspond to any of the rules in the policy, a possible privacy violation has been detected.

An example of a privacy violation that could have been detected based upon anomaly detection, is a case where a doctor residing in Oslo normally only treating patients from Oslo and the surrounding area, suddenly makes a large amount of accesses to the patient database, say, in the northern region of Norway, Tromsø, requesting sensitive information about patients. Normally, the doctor will be allowed to access this information, but the fact that he makes a large amount of accesses, perhaps over a short period of time, makes these actions suspicious. Here, all the remote accesses deviate from what constitutes a normal behaviour for a doctor residing in Oslo, and could be investigated as a privacy violation. Other typical anomaly scenarios include receptionists accessing personal information outside business hours, a particular data subjects personal information being accessed abnormally many times during a short period of time, or a user accessing a lot of information (from different data subjects) during a short period of time.

In [22], which suggests that future database systems must include responsibility for the privacy of the data they manage, a *query intrusion detector* is suggested. The query intrusion detector will spot queries that differ from normal queries made to a database, and correspond to anomaly (privacy violation) detection. Further, [22] suggests a data retention manager to delete data items that have outlived their purposes. This could also be a task for the PVD, e.g. by deleting, or reporting, personal data which have been stored longer than their retention time.

To recapitulate, [2] suggests that the privacy violation detector will continually monitor access to personal data and detect misuse and/or anomaly behaviour. This motivates the use of, possibly, both misuse and anomaly detection, as outlined in 3.4, and corresponds to the suggested detection approaches discussed above.

Based upon the previous discussion, it is clear that the PVD need, at least, the following input:

- a audit log (log of events in the host, application, database or in the network)
- the privacy policy (and privacy agreements), the user profiles and/or signatures

The PVD needs an audit log which contains a list of privacy relevant actions that has occurred in the system. The PVD will then either compare the log entries to a privacy policy, a user profile or look for matching signatures, in order to detect privacy violations. Therefore, it may become necessary to define the format of the log and decide upon which events that needs to be logged, see subchapter 3.5.2.

The PVD will analyse the audit data and upon detection of a possible privacy violation, it will issue an alarm. Most likely, the alarm will be provided to the privacy officer so he can further investigate the potential privacy violation. As

previously stated, the informativeness of the PVD is important, as the more information relating to a privacy violation that a PVD can provide to the privacy officer, the more effective the actual investigation is likely to be. As with IDSs, the informativeness of a policy-based (misuse-based) PVD is likely to be better than the informativeness of an anomaly-based PVD. This is due to the fact that a policy-based PVD can explicitly state which policy rule, or signature that matched (detected the privacy violation). If the profile is explicit and unambiguous, however, the informativeness of a anomaly-based may increase.

3.5.2 Information sources

But what information sources will the PVD base its analysis on? As previously mentioned in subchapter 3.5.1, the PVD bases it detection on the audit log, and the privacy policy, the user profiles and/or signatures. The audit log is a ordered entry of events that have occurred in the system, and it is important that all possible privacy violating events are logged in order to detect them.

Therefore, privacy related events occurring on the network, in the host or application and in databases must be logged. This corresponds to a network-based or host-based approach, as for IDSs.

A network-based approach might be useful in the context of investigating unauthorised disclosure of personal data. That is, if an employee is transferring personal data to a third party and this action are not in compliance with the data subject's privacy preferences.

A host-based approach bases itself upon analysis of activity on a particular host due to access to personal data through an application (and may therefore be called application-based), or by analysing access to personal data through the database where the personal data is stored. Both the application and the database may generate logs of access history to personal data, which in turn may be used by the PVD to detect privacy violations. Furthermore, the approaches are not exclusive, and they could both be used by the PVD.

Therefore, this thesis suggests that the PVD will base its analysis on at least one of the following three sensors, or information sources:

- networks possibly conveying personal data
- applications accessing the personal data
- databases containing the personal data

The sensors log the activity on the network, the application and the database to e.g. a log server or feed it to the PVD. But in order to be able to detect privacy violations, what must be logged? Clearly the events in the monitored system that are necessary to log will to some extent be application dependent. However, all accesses to personal data must be logged, and the log record may consist of the following entries:

- The **user** who wants to access the data. The user identity could be pseudomised to protect the privacy of the users.
- The **action** performed on the data, e.g. store, write or read.
- The **purpose** of the access, i.e. task carried out by the user.
- The **data subject**, the owner of the data.
- The **type** of data about the data subject
- The **time** of access.

An important fact relating to logs is that the logs should be thorough enough for the privacy officer, upon detection of a potential privacy violation, to get information about that event in order to learn more about it and further investigate it. Therefore, the log should, at least, contain information relating to the user accessing the personal data, what data was accessed, who is the owner of the personal data (the data subject) and what was the purpose of the access. The most important [9] log field is, probably, the *purpose* field, as it is important to state the purpose for which the personal information will be used when dealing with personal information (the principle of purpose specification).

Further, as the information that is conveyed to the privacy officer relating to the privacy violation is important for efficient investigation, the information provided by the logs is important. As with IDSs, a host-based approach (i.e. information provided from applications and databases) might be able to provide the privacy officer with more detailed information relating to a privacy violation than a network-based approach, as it may be easier to see the outcome of the violation with a host-based approach .

In addition to the sensors stated above, feedback from IDS systems and firewalls in the system may e.g. inform the PVD that suspicious activities or intrusive activities are taking place, and that this could, in turn, lead to privacy violations. Hence, such information could "awaken" the PVD, taking it to a higher privacy protection level. It is, however, important that the PVD do not have too much information to search through, as this possibly can increase the numbers of false alarms and lead to an information overflow and a situation where real privacy violations (false negatives) are overseen.

3.5.3 Other features

Other features relating to the PVD are response options, timing and control strategies. These are briefly discussed here.

[9] or what makes privacy violation detection different from intrusion detection.

Response

The responses provided by the PVD will most likely be passive, that is, provide information about potential privacy violations to the system users, or a privacy officer, and rely on them to take subsequent action based on that information.

An important issue relating to responses is the informativeness of the alarm, as previously mentioned. Upon an alarm, the privacy officer will proceed to investigate whether the potential privacy violation was in fact a privacy violation or not. Therefore, it is essential that the privacy officer is provided with as much information relating to the incident as possible. As previously discussed in subchapters 3.5.1 and 3.5.2, the privacy officer will most likely get most information from a misuse-based, host-based PVD.

Timing

The PVD will most likely be interval-based. That is, at certain times the detector will run through the log, or the database, looking for potential privacy violations. This is due to the fact that the aim of the PVD is not real-time detection, but rather to off-line check at regular intervals to see whether privacy violations has occurred or not.

If the PVD is based upon a network-based approach, however, a real-time timing approach might be desirable, as the aim is to avoid disclosure of personal data, not discover that personal data was in fact disclosed to unauthorised parties after it occurred. However, real-time detection, and potentially prevention, of privacy violations is not necessarily the aim of this PVD.

Control Strategy

Control strategy refers to how the, possibly, different parts of the PVD, e.g. the sensors (monitors), are controlled and how input and output are managed. As previously mentioned in subchapter 3.4.3, possible strategies include centralized, partially distributed and fully distributed control.

With a centralized control, all sensors monitoring the system will provide information back to centralized location, that is, a PVD. Sensors include database sensors, network sensors and application sensors, as outlined in subchapter 3.5.2. A partially distributed and a fully distributed control strategy is more complicated to implement, but could be considered in large systems that generate a lot of information relating to privacy violations.

3.6 Privacy Violation Detection vs. Intrusion Detection

Clearly, there are a lot of similarities between detecting privacy violations and detecting intrusions, as shown by subchapter 3.4 and 3.5. But is it feasible to incorporate privacy violation detection into already existing intrusion detection systems? Or is it desirable to integrate these two different functionalities into one system? As IDSs already exist, is a PVD unnecessary?

Integrating these two similar, but yet very different, concepts may actually clutter up the system, more than it will do well. As current IDSs still struggle with large amounts of false alarms, adding features for detecting privacy violations will probably not better this fact. Therefore, the detection of privacy violations and intrusions should take place in different components, or subsystems. It is a need for a PVD.

The main difference between privacy violation detection and intrusion detection is the fact that they aim at detecting different things (privacy violations and intrusions are not the same, even though it might be argued that privacy violations are a subset of intrusions). Therefore, the detectors will need different input in order to be able to detect privacy violations or intrusions, e.g. the profiles, signatures and policies will not look the same. Also, the PVD aims at detecting breach of the privacy policy, whereas the aim of IDSs is to detect exploitation of security holes in the system and security policy violations. Furthermore, some organisations may not incorporate IDSs but recognize the need for a PVD. However, as shown earlier, privacy violation detection may base it self on the basic paradigms of IDSs.

Chapter 4

A Case Study: Detecting Privacy Violations in a Medical Information System

As the aim of this thesis is to define a privacy policy and determine which data that needs to be collected in a system in order to identify and detect privacy violations, a case has been chosen. For this thesis, a medical information system context has been selected due to the fact that personal information stored in medical databases is highly sensitive. Medical records include sensitive personal information that reveals some of the most intimate aspects of an individual's life. The medical record includes diagnostic and medical test results, details of a person's family history, history of diseases and treatments, history of drug use, sexual orientation and practices and possibly other sensitive information.

Clearly, the privacy of an individual's medical record is important to the individual, and should be protected. The individual's perception of privacy will also reflect upon the individual's willingness to share sensitive, and important, health information. It is important that health care information is accurate and comprehensive so that the quality of health care delivery is not compromised. Consequently, it is important that the privacy of medical records is maintained, in order for patients to trust their health care provider.

This chapter outlines a hospital case in order to show the applicability of a PVD in a hospital setting with a scenario. A privacy policy and data that needs to be logged in order to detect privacy violations is defined.

4.1 Outlining the case

For the remainder of this chapter a case will be outlined. It has been chosen to keep it simple as the goal is to show the applicability of a privacy violation detector (PVD) in a medical information system context.

The case environment consists of:

- A hospital treating patients while maintaining their privacy.
- Personnel working at the hospital, using the medical information system deployed on the hospital for access to personal identifiable information about patients for purposes such as diagnosis and treatment. There are different kinds of personnel, also called actors. This case assumes the following set of actors:
  - doctors, treating and diagnosing patients.
  - nurses, treating patients.
  - receptionists, registrating and localisating patients.
  - testers, performing and analyzing tests ordered by doctors.
- Medical journals stored in a database local to the hospital which contain both patient identifying data, including name, birth date and address, and medical data, such as diagnoses and medical history.
- The hospital deploys a privacy violation detector (PVD), which is able to detect whether privacy violations has occurred. The PVD is policy-based and takes a host-based approach, that is, it bases its analysis upon logs generated by the application that provide the actors with access to the personal data.
- All accesses to personal data are logged to a log server, in order to be checked by the PVD.
- A privacy policy.

Hospital personnel use the medical information system deployed on the hospital to access the personal identifiable information stored in the system in order to e.g. treat and diagnose patients. For simplicity this case assumes that access control is not deployed in the system. All accesses to the personal identifying information are logged to the log server. The health personnel carry out different roles with respect to the system, and their tasks can be illustrated by a use case diagram as shown in figure 4.1.

As the use case shows, the different actors (doctor, nurse, testers and receptionists) all carry out different use cases. Doctors are responsible for carrying out treatment and diagnosing patients, whereas nurses assist in treating the patients. Receptionists register patients upon arrival to the hospital, and are also able to locate patients so that visitors can get in touch with the patient. Testers are responsible for carrying out tests, such as pathology, radiology and laboratory tests and then analysing them. The results of the tests provides means for doctors to diagnose patients (therefore the array marked with "uses" in the use case diagram shown in figure 4.1).

The overview provided by figure 4.1 are simplified. However, it is chosen to keep it simple, as the main goal of this case example is to show potential use of a PVD.

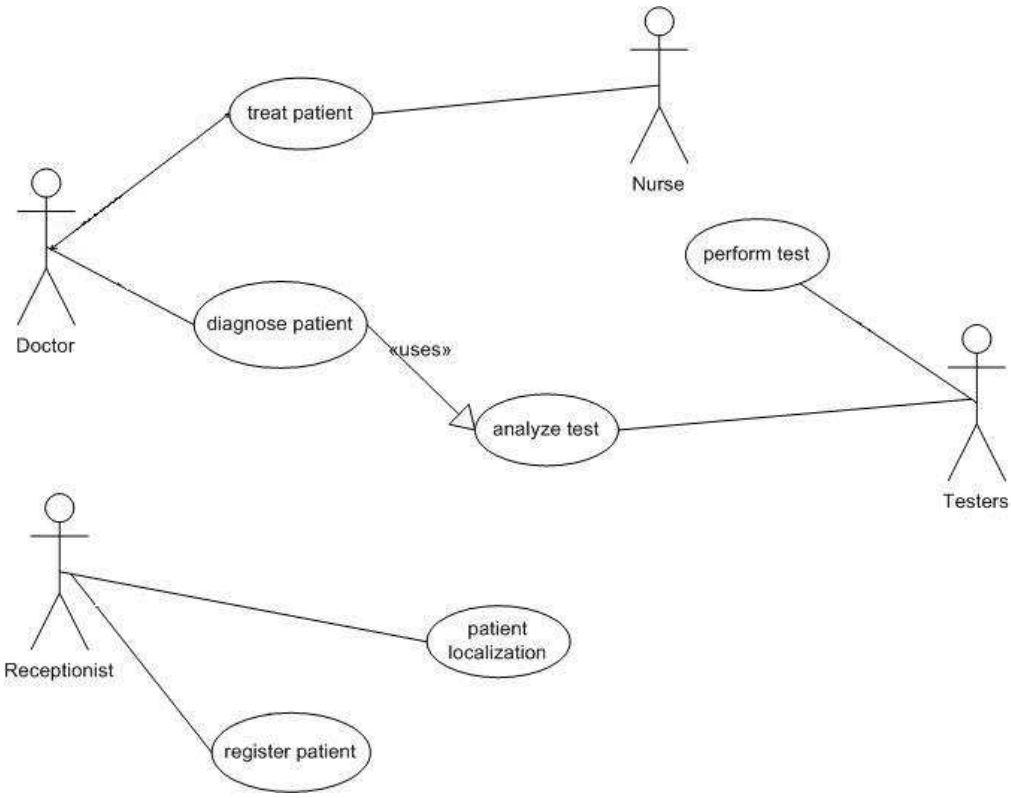

Figure 4.1: A use case diagram describing the simple hospital scenario.

Further, it is assumed that a medical journal maintained at the hospital contains the following data categories [1]:

- **Patient identifying data**, which include information such as name, address and date of birth.
- **Contact data**, which include information such as the patients name, room number and telephone number.
- **Medical history**, which includes a description of complaint, present illness, medical history, psychosocial history, family history and physical examinations.
- **Physicians order**, which includes test and medications ordered.
- **Progress note**, which describes the course of the patient illness, response to treatment and the status at discharge.
- **Departmental reports**, which includes contributions from pathology, radiology, laboratory and physical therapy.
- **Nursing data**, which includes the nurses notes and observations of the patient such as vital signs, fluid intake and output and dosage of medication.
- **Operative reports**,which include anesthesia report, description of surgical events and a recovery room record.
- **Discharge summary**, which recapitulates the patients treatment in the hospital and results.

The different actors should have access to different parts of this medical journal, as they should only be able to see what they really need to know in order to treat the patient. Table 4.1 shows a authorisation matrix which gives indications of which actors are allowed to access which parts of the medical journal, and what actions they are allowed to perform (such as read, write and create).

Access rights to a patient's data depend upon the purpose of the access. The purpose is incorporated in to the authorisation table in brackets, where w(t,d) means write access is allowed for the purpose of treatment and diagnosis. The abbreviations for the purposes are shown below:

- t - treatment
- d - diagnosis
- te - testing
- r - registration

[1] Content of patient records, outlined on http://mason.gmu.edu/ falemi/IT/paprecord.htm

- l - localisation

As table 4.1 shows, the doctors have access (read and write) to all fields of the medical journal except the patient identifying data. Further, nurses have read access to the physicians order field for the purpose of treatment, and write and read access to the nursing data field for the purpose of treatment. For this case, it is chosen to let the doctors have read and write access for the purpose of treatment and diagnose to the contact data field, e.g. in order to be able to register that a patient is relocated to another room. Nurses are given read access to the contact data field for the purpose of treatment. Testers are given create, read and write access to the departmental report field, whereas the receptionists are given read and write access to the patient identity data field for the purpose of registration. Further, they are given write access for the purpose of registration and read access for the purpose of localisation to the contact data field.

Clearly, the categorization of user categories, or actors, is a bit too rough. Doctor, for example, is not necessarily provided access to all patients full medical journal. Rather, they should be authorised on a "need to know" basis. Related to a patient, the doctors will perform different roles such as treating doctor, assisting doctors, examination specialist, surgeon and other doctors, which do not have a relationship to that patient. The treating doctor is responsible for the treatment of the patient and the advising doctor are called in to assist the treating doctor. The examination specialist is in charge of diagnosing the patient temporality. The surgeon is responsible for the patient during surgery.

A patient's treating doctor must be given access to all fields in the medical journal, except the patient identifying data such as home address and telephone number. The advising doctor only needs access to the fields in the medical journal that is necessary for diagnosing or treating the patient. For this case, the following approach is suggested: Upon arrival to a hospital, the patient is registered, and examined by an examination specialist. The examination specialists has read and write access to the patient's contact data, medical history, physicians order, progress notes, departmental reports and nursing data for the purpose of diagnosis. Then a treating doctor is appointed to the patient. This relationship must be maintained, i.e. in a database. Further, the treating doctor is allowed (due to professional secrecy) to appoint one of more advising doctors. The advising doctor is also given access to the patient's full medical journal, except the patient identifying data and the discharge summary. The surgeon is authorised to access (read and write) the patient's operative report for the purpose of treatment. All other doctors should be denied access to the patient's medical journal, and accesses done by doctors other that a treating doctor or an advising doctor should be investigated as a privacy violation. Therefore, the doctor field is marked with an *, indicating that not all doctors are allowed that extensive access. A more extensive authorisation table for doctor is provided in table 4.2.

Data category	Doctor*	Nurse	Tester	Receptionist
Patient id. data				r(r), w(r)
Contact data	r(d,t), w(d,t)	r(t)		r(l), w(r)
Medical history	r(d,t),w(d,t)			
Physicians order	r(d,t), w(d,t)	r(t)		
Progress notes	r(d,t),w(d,t)			
Departmental reports	r(d,t),w(d,t)		c(te),w(d),r(d)	
Nursing data	r(d,t),w(d,t)	w(t),r(t)		
Operative reports	w(d,t), r(d,t)			
Discharge summary	r(d,t),w(d,t)			

Table 4.1: Authorisations of the different actors to the different data categories in the simple hospital case.

Data category	Treating Doctor	Advising Doctor	Examination specialist	Surgeon
Patient id. data				
Contact data	r(d,t) , w(d,t)	r(d,t)		
Medical history	r(d,t),w(d,t)	r(d,t)	r(d), w(d)	
Physicians order	r(d,t), w(d,t)	r(d,t)	r(d), w(d)	
Progress notes	r(d,t), w(d,t)	r(d,t)	r(d), w(d)	
Departmental reports	r(d,t),w(d,t)	r(d,t)	r(d), w(d)	
Nursing data	r(d,t),w(d,t)	r(d,t)	r(d), w(d)	
Operative reports	r(d,t), w(d,t)	r(d,t)		r(t), w(t)
Discharge summary	r(d,t),w(d,t)			

Table 4.2: Authorisations for different categories of doctors to the different data categories in the simple hospital case.

The two tables, 4.1 and 4.2, basically outlines the privacy policy for the hospital.

4.2 A privacy policy

This subchapter defines a informal hospital policy based upon the simplified case outlined above.

4.2.1 A Hospital Policy

1. A treating doctor have write and read access to the <`contact-data, medical-history, physicians-order, progress-notes, departmental-report, nursing-data, operative-reports, discharge-summary`> fields of the patient's medical journal for the purpose of diagnosis and treatment.
2. A advising doctor have read access to the <`contact-data, medical-history, physicians-order, progress-notes, departmental-reports, nursing-data, operative-report`> fields of the patient's medical journal for the purpose of diagnosis and treatment.
3. A examination specialist have write and read access to the <`medical-history, physicians-order, progress-notes, departmental-reports, nursing-data`> fields of the patient's medical journal for the purpose of diagnose.
4. A surgeon have read and write to the <`operative-report`> field of the patient's medical journal for the purpose of treatment.
5. A nurse have read access to the <`contact-data,physicians-order`> fields of the patient's medical journal for the purpose of treatment, and read and write access to the <`nursing data`> field of the patient's medical journal for the purpose of treatment.
6. A receptionist have read and write access to the <`patient-id-data`> field of the patient's medical journal for the purpose of registration (once), read access to the <`contact-data`> field of the patient's medical journal for the purpose of localisation, and write access to the <`contact-data`> field of the patient's medical journal for the purpose of registation.
7. A tester have create access to the <`departmental-report`> field of the patient's medical journal for the purpose of performing a test, and read and write access to the <`departmental-report`> field of the patient's medical journal for the purpose of diagnosis.

4.2.2 Patient Policies

Individual patients may have their own policies as long as it is compliant with the hospital policy. For this case it is assumed that all patients have agreed upon the hospital policy in order to keep it simple.

4.2.3 Mapping of the hospital policy to EPAL

Based upon the hospital policy stated above, a more thorough informal policy is provided below and corresponds to authorisation rules defined by the Enterprise Policy Authorization Language (EPAL) [2].

1. Allow write and read on <`contact-data, medical-history,physicians -order, progress-notes,departmental-reports, nursing-data, operative-report, discharge-summary`> by doctors for the purpose of treatment and diagnosis, if the doctor is the patient's treating doctor.

2. Allow write and read on <`medical-history, physicians-order, progress -notes, departmental-reports, nursing-data, operative -report`> by doctors for the purpose of treatment and diagnosis, if the doctor is the patient's advising doctor.

3. Allow read access to <`contact-data`> by doctors for the purpose of treatment and diagnosis, if the doctor is the patient's advising doctor

4. Allow write and read on <`medical-history, physicians-order, progress-notes, departmental-reports, nursing-data`> by doctors for the purpose of diagnosis, if the doctor is the patient's examination specialist.

5. Allow write and read access to <`operative-reports`> by doctors for the purpose of treatment, if the doctor is the patient's surgeon.

6. Allow read on <`contact-data,physicians-order`> by nurses for the purpose of treatment, if the nurse is on duty on the patient's ward.

7. Allow read and write on <`nursing-data`> by nurses for the purpose of treatment, if the nurse is on duty on the patient's ward.

8. Allow read and write on <`patient-id-data`> receptionists for the purpose of registration. Notify the patient.

9. Allow read on <`contact-data`> by receptionists for the purpose of localization.

10. Allow write on <`contact-data`> by receptionists for the purpose of registration.

11. Allow create on <`departmental-report`> by testers for the purpose of performing a test.

12. Allow read and write on <`departmental-reports`> by testers for the purpose of diagnosis.

[2] EPAL authorisation rules take the form of "allow or deny actions on data-categories by user-categories for certain purposes under certain conditions while mandating certain obligations", see appendix B.

In subchapter 3.5.1 it was stated that a policy-based detection approach requires that audit data are compared to a defined privacy policy, and that in order to be able to do so, the privacy policy must exist in a machine-readable form, e.g. by encoding the privacy policy in EPAL. EPAL is described in appendix B. An EPAL encoding of parts of the informal privacy policy described above is shown below, an EPAL encoding of the whole policy are shown in appendix B.

```
<rule id="Rule1" ruling="allow">
<short-description language="en" />
<long-description language="en" />
<user-category refid="doctor" />
<data-category refid="contact-data" />
<data-category refid="medical-history" />
<data-category refid="physicians-order" />
<data-category refid="progress-notes" />
<data-category refid="departmental-reports" />
<data-category refid="nursing-data" />
<data-category refid="operative-reports" />
<data-category refid="discharge-summary" />
<purpose refid="treatment" />
<purpose refid="diagnosis" />
<action refid="write" />
<action refid="read" />
<condition refid="isTreatingDoctor">
<obligation refid="log">
</obligation>
</rule>

<rule id="Rule7" ruling="allow">
<short-description language="en" />
<long-description language="en" />
<user-category refid="nurse" />
<data-category refid="nursing-data" />
<purpose refid="treatment" />
<action refid="read" />
<action refid="write" />
<condition refid="isOnDutyOnPatientWard">
<obligation refid="log">
</obligation>
</rule>
```

As shown, the rules are identified by a rule number. Further, all rules are of the form "allow", meaning that this policy is positive, i.e. it states what is allowed to do. The user-category field refers to the actors in the system. The data-category field corresponds to the data categories defined earlier. Further, the purposes relate

to the purpose of access, and refer to the purposes already defined. The actions refer to the operation performed on the data. The condition puts limitations on the rule. The "isOnDutyOnPatientWard" refers to the fact that a nurse can only treat a patient when she is on duty at the same ward as the patient whose personal data she is accessing. The rules have also been encoded with the obligation to log all events, that is, all accesses to personal data.

4.3 The Log

This subchapter describes data that need to be logged in order for the PVD to be able to identify a potential privacy violation. Generally, all accesses to personal data must be logged, and at least the following information recorded:

- The **time** of access.
- The **user-category**, e.g. doctor or nurse.
- The **identity of the user**, e.g. by a unique doctor id, the full name of the doctor or a pseudonym.
- The **data-category**, e.g. medical-history or nursing-data.
- The **identity of the patient** (the data subject), either by a unique id number (e.g. social security number), the full name of the patient or a pseudonym.
- The **purpose** of the access.
- The **action** performed, such as create, write or read.

The log entries consisting of these fields will be compared to the privacy policy encoded in a machine-readable language, such as EPAL, in order to see if any privacy violation might have occurred. If a log entry does not correspond to any of the (positive) rules defined in the privacy policy, an alarm is raised. The log format shown here corresponds to the EPAL language, which also defines user-category, data-category, purpose and action. In addition to these fields, the identity of the user and the identity of the data subject (i.e. the patient) must be recorded (possibly pseudomised), in order to be able to confront a user with his or hers actions. Further, time should also be recorded.

4.4 Scenarios

This subchapter describes two scenarios that show how a privacy violation detector (PVD) might work in a hospital context.

First, a scenarios, called scenario one, which does not contain any privacy violating events is described. This is mainly to show how the hospital environment outlaid

above may operate. It should also be noted that this scenario may not be complete or correct according to how hospitals actually runs. Then, a scenarios describing detection of privacy violations occurring in this environment is described.

4.4.1 Scenario 1: No privacy violating events

A person feels pain, e.g. in the stomach region, and rushes to the hospital to have it checked out. Upon arrival to the hospital, the receptionist registers the patient with the correct contact data. Then he is diagnosed by an examination specialist, which is in the user-category doctor, and appointed to a treating doctor, which also is in the user-category doctor. In order to diagnose, the examination specialist may have to access the medical history, and add to it descriptions of the complaint and present illness. The treating doctor is in charge of treating (and further diagnosing) the patient, with assistance from the nurses. Test results from testers are also used in order to diagnose the patient.

Further, as the examination specialist has ordered tests in order to diagnose the patient, a tester will, upon the arrival of the patient, create a departmental report. Once the testing is completed, the tester writes the results to the departmental report. Then, a tester will analyse (read) the departmental report in order to diagnose the patient. After the examination specialist has made a diagnosis, i.e. based on the departmental report, a treating doctor is appointed. The treating doctor accesses e.g. the medical history, physicians order, departmental reports and progress notes in order to learn about the patient's illness. Further, the treating doctor may appoint an advising doctor if advises are necessary, which also are allowed access to the patients medical record. Once a correct diagnosis has been set, surgery is undertaken, and the surgery doctor writes to the operative-report information such as description of surgical events and recovery room. Whilst the patient is in the recovery room, he is being treated by nurses, which read and write to the nursing data field.

Finally, after successful treatment the patient is released from the hospital. The receptionist make discharges the patient, and the treating doctor writes to the discharge summary field.

For this scenario, the PVD will run through the log fields, and see that all entries are in compliance with the privacy policy. The PVD will have to check that the identity of the doctor, i.e. the treating doctor, corresponds to the registered treating doctor, as not all doctors are allowed an extensive access to all patients' records. For now, it is assumed that this information is maintained somewhere, e.g. in a local database.

4.4.2 Scenario 2: Privacy violating events

This scenario aims at describing a detection of a privacy violation. Again, as with scenario 1, the scenario is not complete, and possibly not realistic, but aims at illustrating the concept of privacy violation detection.

A person feels pain in the stomach region, as above, and rushes to the hospital to have it checked out. Upon arrival to the hospital, the receptionist registers the patient with the correct contact data, before the patient is diagnosed by an examination specialist. All the events, that is, the accesses to this patient's personal data are logged. A possible log is shown in table 4.5. The time of occurrence is for the sake of convenience only labelled time1, time2 etc. The patient's identity is pseudonomised in the system, and this patient is referred to as pat10. Further, the log entries corresponds to the log as described in subchapter 4.3. The user category corresponds to e.g. doctors and nurses; whereas the identity of the doctor and nurse are registered in the user-id field. This identity is also for privacy purposes pseudonomised.

As table 4.5 shows, first the receptionist (res4Y) accesses the contact-data field for pat10 for reading at time1 for the purpose of registration, assuming that it is an existing record on him. Then the receptionist checks that the information there is correct, and adds any information that is required, see the write access at time2. The examination specialist (docY, see table 4.4.2) opens (reads) the medical history and the physicians orders of pat10 for the purpose of diagnosis at time3 and time4. Further, docY writes a description of the complaint to the medical history field for the purpose of diagnosis at time5, and a list of tests which he orders to the physicians order at time6, also for the purpose of diagnosis.

examination specialist	docY
treating doctor	docX
advising doctor	docZ
surgeon	docW

Table 4.3: Doctors with access right to pat10's medical record.

Next, a tester (tC3) creates a departmental report at time7, and the test results are recorded there at time8. Further, another tester (tD5) reads the departmental report at time9, in order to analyse the tests, and writes a conclusion (possible diagnosis) to the departmental report at time10. Then the examination specialist reads the departmental reports at time11, and writes a diagnosis to the medical history at time11. Then the patient is released to a appointed treating doctor, see table 4.4.2, which is in charge of treating (and further diagnosing) the patient, with assistance from the nurses Then, the treating doctor updates himself, by accessing e.g. the medical history, physicians order, departmental reports and progress notes. Here, the treating doctor, docX, reads the medical-history at time13, and writes to the progress notes at time14, for the purpose of diagnosing and treatment respectively. At time15, however, there is an access to pat10's medical history by a docQ, not enlisted in table 4.4.2. This may correspond to a privacy violation. Further, the treating doctor appoints an advising doctor, as e.g. the patient does not seem to respond to the treatment given. The advising doctor, docZ, reads the medical-history at time16, the progress note at time18 and the departmental report at time19, in order to understand the patient's illness.

Together, the two doctors decide upon surgery. Once the surgery is undertaken, the surgery doctor, docW, writes, at time20, to the operative report information such as description of the surgical events and recovery note. Whilst the patient is in the recovery room, he is treated by nurses who write information such dosage of medication to the nursing data field, e.g. at time21. Also, at time17 there is an access to the patient's medical history field by a nurse, which according to the privacy policy is not allowed.

Finally, after successful treatment the patient is released from the hospital. The treating doctor writes to the discharge summary field, and the receptionist discharges the patient (time 24 and time25).

time15	doctor	medical-history	docQ*	pat10	treatment	read
time17	nurse	medical-history	nA*	pat10	treatment	read

Table 4.4: Potential privacy violations

For this scenario, the PVD will run through the log fields, and see that all entries is in compliance with the privacy policy except for the events occurring at time15 and time17. These events reported as potential privacy violations for further investigation. The potential privacy violations are shown in table 4.4.

For the potential privacy violation that occurred at time15, the identity of the doctor must be checked. docQ is not enlisted as one of the patient's doctors, see table 4.4.2. However, docQ might have been called in to help in e.g. treating or diagnosing the patient in a case of emergence where the treating doctor e.g. did not have the time to register the doctor in the patient-doctor database. If this was the case, there was no privacy violation at time15.

Then the potential privacy violation at time17 is further investigated. Here, nurses are not under any circumstances allowed to access the medical history field of a patient's medical journal; hence, this is a privacy violation.

This scenario bases its detection upon the rules, or policy, stated earlier. It is assumed that all obligations are fulfilled, that is, all accesses are logged. Further, conditions like "isTreatingDoctor", "isAdvisingDoctor", "isExaminationSpecialist" and "isSurgeon" has been checked, such as with the potential privacy violation at time15. The scenario does, however, assume that "isOnDutyOnWard" and "isOnDuty" conditions were fulfilled. In order to detect privacy violations were those conditions are not fulfilled, there is a need for a database, or a table, that lists the nurse and receptionist identity (i.e. pseudonym), and their working hours.

Further, the scenario has demonstrated policy-based detection of a privacy violation, based upon information provided from logs generated by e.g. the applications used to access personal data or the database where the personal data is stored. Scenarios could also be included to show detection of privacy violations based upon information provided by network traffic (e.g. privacy violations related to disclosure of personal data) and information provided by searching the database that stores the personal data (e.g. privacy violations related to the retention time).

time	user-category	data-category	user-id	patient-id	purpose	action
time1	receptionist	contact-data	res4Y	pat10	registration	read
time2	receptionist	contact-data	res4Y	pat10	registration	write
time3	doctor	medical-history	docY	pat10	diagnosis	read
time4	doctor	physicians-order	docY	pat10	diagnosis	read
time5	doctor	medical-history	docY	pat10	diagnosis	write
time6	doctor	physicians-order	docY	pat10	diagnosis	write
time7	tester	departmental-report	tC3	pat10	testing	create
time8	tester	departmental-report	tC3	pat10	testing	write
time9	tester	departmental-report	tD5	pat10	analysing	read
time10	tester	departmental-report	tD5	pat10	analysing	write
time11	doctor	departmental-report	docY	pat10	diagnosis	read
time12	doctor	medical-history	docY	pat10	diagnosis	write
time13	doctor	medical-history	docX	pat10	diagnosis	read
time14	doctor	progress-note	docX	pat10	treatment/diagnosis	write
time15	doctor	medical-history	docQ*	pat10	treatment	read
time16	doctor	medical-history	docZ	pat10	diagnosis	read
time17	nurse	medical-history	nA*	pat10	treatment	read
time18	doctor	progress-note	docZ	pat10	diagnosis	read
time19	doctor	departmental-report	docZ	pat10	diagnosis	read
time20	doctor	operative-report	docW	pat10	treatment	write
time21	nurse	nursing-data	nB	pat10	treatment	write
time22	doctor	progress-note	docX	pat10	treatment	write
time23	doctor	discharge-summary	docX	pat10	discharge	write
time24	receptionist	contact-data	res4P	pat10	discharge	write

Table 4.5: Logging a privacy violation.

Chapter 5

Conclusion

The aim of a PVD is to detect privacy violations, i.e. events that breach a privacy policy or individual privacy agreements.

A motivation for implementing a PVD, in addition to the fact that it will detect privacy violations, is the fact that the PVD will be a reactive mechanism. Privacy violations will be detected after they occurred. Further, it may have a preventive function, and help increase the users' trust in the system's handling of personal data. Monitoring mechanisms are also mandated by legislation.

There are some issues related to privacy violation detection that should be considered. It is particularly important that the users' privacy is maintained when protecting the data subjects' privacy. This could be accomplished by pseudonymising logs or protecting access to logs.

Two different privacy violation detection approaches, anomaly and policy-based detection, has been identified and discussed. Anomaly detection requires a profile of normal behaviour (that is, non-privacy violating behaviour) to be defined. Then, deviations from this profile are investigated as potential privacy violations. Policy-based detection compares the audit data from the monitored system against a defined privacy policy (or agreement). Therefore, the PVD needs an audit log, and the privacy policy (agreement), the user profiles and/or the privacy violation signatures in order to be able to detect privacy violations.

All accesses to personal data must be logged and later analysed by the PVD. As a minimum, a log entry should reveal the identity of the user accessing the personal data, the action the user performed, the purpose of the access, the identity of the data subject, the type of data accessed and the time. Further, the PVD is likely to base its analysis upon information from the networks (network-based), the application accessing the personal data or the database storing the personal data (host-based).

The applicability of a policy-based, host-based PVD has been shown in a simplified hospital case. Based upon a log and a predefined hospital policy, privacy violations were detected.

References

[1] Fischer-Hübner, S., *IT-Security and Privacy, Design and Use of Privacy-Enhancing Security Mechanisms*, Springer -Verlag 2001 ISBN 3-540-42142-4

[2] Danielsson, J., Arnesen, R.R., *A Framework for the Enforcement of Privacy Policies*, Proceedings of NORDSEC 2003, p. 13-23, Gjøvik, Norway, ISBN 82-993980-4-5.

[3] United Nations Universal Declarations of Human Rights, available at http://www.un.org/Overview/rights.html

[4] Westin, A.F., Privacy and Freedom, Atheneum, NY, 1967.

[5] Arnesen, R. R., Danielsson,J.,*Høringsuttalelse fra Norsk Regnesentral: Teknologirådets høring om IKT & Personvern*, available at (in norwegian) http://www.teknologiradet.no/files/norsk_regnesentral_copy.pdf

[6] LOV 2000-04-14 nr 31: Lov om behandling av personopplysninger (personopplysningsloven). Available from http://www.lovdata.no, in norwegian. An unofficial translation of the Personal Data Act can be found at http://www.datatilsynet.no/lov/loven/poleng.html.

[7] Forskrift til personopplysningsloven (personopplysningsforskriften), http://odin.dep.no/jd/norsk/regelverk/lover/012001-200005/hov004-bu.html

[8] Directive 95/46/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. Official Journal L281, 23/11/1995, available at http://europa.eu.int/eur-lex/en/index.html

[9] OECD Guidelines on Transborder flow of Personal Data, available at http://www1.oecd.org/publications/e-book/9302011E.pdf

[10] Information and Privacy Commissioner/Ontario, Registratiekamer, *Privacy-enhancing technologies: The path to anonymity, Revised Edition*, available at: http://www.cbpweb.nl/downloads_av/AV11.PDF

[11] Pfitzmann, A., Köhntopp, M., *Anonymity, Unobservability and Pseudonymity - A Proposal for Terminology*, Draft version 0.14, July 2002, available at http://freehaven.net/anonbib/papers/Anon_Terminology_v0.14.pdf

[12] ISO IS 15408, 1999, http://www.commoncriteria.org

[13] Sweeney, L., *Computational Disclosure Control, A Primer on Data Privacy Protaction*, August 2001, available at http://citeseer.ist.psu.edu/499190.html

[14] Agrawal, R., Srikant, R., *Privacy-Preserving Data Mining*, ACM SIGMOD Record , Proceedings of the 2000 ACM SIGMOD international conference on Management of data, Volume 29 Issue 2.

[15] The Information and Privacy Commissioner/Ontario, Deloitte & Touche *The Security-Privacy Paradox: Issues, Misconceptions, and Strategies*, August 2003, available at http://www.ipc.on.ca/userfiles/page_attachments/sec-priv.pdf

[16] Stadler, M., Piveteau, J., Camenisch, J., *Fair Blind Signatures*, Advances in Cryptology - Eurocrypt'95, volume 921 of Lecture Notes in Computer Science, pp. 209-219, Springer-Verlag, 1995.

[17] Bace, R., Mell, P., *NIST Special Publication on Intrusion Detection Systems*, SP 800-31, November 2001, available at http://csrc.nist.gov/publications/nistpubs/800-31/sp800-31.pdf

[18] Allen, J., Christie, A., Fithen, W., McHugh, J., Pickel, J., Stoner, E., *State of the Practice of Intrusion Detection Technologies*, Technical Report, CMU/SEI-99-TR-028, January 2000.

[19] Axelsson, S., *Research in Intrusion-Detection Systems: A Survey*, Technical report 99-15, Department of Computer Engineering, Chalmers University, Aug 1999.

[20] Sundaram, A., *An Introduction to Intrusion Detection*, ACM Crossroads, Computer Security, Issue 2.4, April 1996.

[21] Lundin, E., Jonsson, E.,*Privacy vs Intrusion Detection Analysis*, Second international workshop on the Recent Advantages in Intrusion Detection, September 7-9., 1999,

[22] Agrawal, R., Kiernan, J., Srikant, R., Xu, R., *Hippocratic Databases*, Proceedings of the 28th VLDB Conference, Hong Kong, China, 2002.

[23] The platform for privacy preferences 1.0 (P3P1.0) Specification, W3C Recommendation 16 April 2002, http://www.w3.org/TR/P3P

[24] Agrawal, R., *Why is P3P Not a PET?*, W3W Workshop on the Future of P3P, 12-13 November 2002, Dulles, Virginia, USA.

[25] Ansatte snoker i pasientjournaler, http://www.aftenposten.no/nyhter/iriks/article787554.ece, in Norwegian.

[26] Gollmann, D.,*Computer Security*, J. Wiley & Sons, 1999

[27] Macaulay, L , *Privacy Enhancing Technologies State of the Art Review*, Technical Report Series 1, February 2002 (TRS-2002-001), Version 1, available at http://www.co.umist.ac.uk/research/tech_reports/trs_2002_001_lam.pdf

Appendix A

Terminology

Anonymity Anonymity ensures that a user may use a resource or service without disclosing the user's identity.

Consent A consent is any freely given, specific and informed declaration by the data subject to the effect that he or she agrees to the processing of personal data relating to him or her [6].

Data collecting organisation or entity A data collecting organisation (also called entity) is the organisation that collects the personal data for a specific purpose.

Data controller The data controller is the person who determines the purpose of the processing of personal data and which means are to be used [6].

Data subject The data subject is the person to whom personal data may be linked [6].

False alarm A false alarm corresponds to a false positive or a false negative.

False negatives False negatives are actual privacy violations that are not flagged as potential privacy violations by the PVD.

False positives False positives are ordinary non-privacy violating events that are flagged as potential privacy violations by the PVD.

Intrusion detection system (IDS) Intrusion detection systems are software or hardware systems that automate the process of monitoring the events occurring in a computer or network, analysing them for signs of security problems [17].

Personal data Personal data refers to any information and assessments that may be linked to a natural person [6].

Personally identifiable data See Personal data.

Privacy agreement A privacy agreement is a set of rules that determine how the personal data the agreement pertains to can and should be used, and that both the data subject and the data collector has consented to [2].

Privacy officer A privacy officer is a person who is responsible for further investigating possible privacy policy breaching events, that is, privacy violations.

Privacy policy A privacy policy demonstrates an organisations commitment to privacy, and describe how information collected about an individual (that is, personally identifiable information) should be processed (e.g. accessed) and possibly disclosed. The privacy policy must be in compliance with national laws and regulations.

Privacy promise A privacy promise is a set of rules, or principles, relating to an organisation (e.g. data collector) handling of personal data that the organisation promises to fulfil.

Privacy violation A privacy violation is an event that breach the privacy policy defined by the data collector, or an event that breaches a privacy agreement between a data subject and a data collector.

Profile A profile is a definition of a users normal behaviour in a system.

Pseudonymity Pseudonymity ensures that a user may use a resource or service without disclosing its user identity, but can still be accountable for that use.

Sensitive personal data Sensitive personal data includes information such as racial or ethnic origin, or political opinions, philosophical or religious beliefs, the fact that a personal has been suspected of, charged with or convicted of an criminal act, information concerning health issues, information about a person's sex life and information on trade-union membership.

Unobservability Unobservability ensures that a user may use a resource or service without others, especially third parties, being able to observe that the resource or service is being used.

Unlinkability Unlinkability ensures that a user may make multiple uses of resources or services without other being able to link these uses together.

User The user is the person who processes personal data.

Appendix B

Enterprise Privacy Authorization Language

B.1 Introduction to EPAL

The Enterprise Privacy Authorization Language (EPAL) is a formal language to specify fine-grained enterprise privacy policies [1]. EPAL defines a policy terminology (vocabulary) and authorisation rules. The rules allow or deny actions depending on the purpose of the access.

An EPAL policy defines lists of hierarchies of data-categories, user-categories, and purposes, and sets of (privacy) actions, obligations, and conditions. These elements may be called the vocabulary. The data-categories define different categories of collected data that are handled differently from a privacy perspective (e.g. medical-record vs. contact-data). User-categories are the entities that use collected data (e.g., doctor or receptionist). Purposes model the intended service for which data is used (e.g. medical treatment or patient registration). Further, actions model how the data is used (e.g. read or write). Obligations define actions that must be taken, e.g. delete after 30 days, whereas conditions are boolean expressions that evaluate the context, e.g. the user-category must be the primary care physician of the data-subject.

A privacy policy written in EPAL is made up of privacy authorisation rules that:

> allow or deny **actions** on **data-categories** by **user-categories** for certain **purposes** under certain **conditions** while mandating certain **obligations**.

Further, the EPAL rules are sorted by descending precedence, in such, it allows for exceptions. Exceptions to a rule may be implemented by putting the exception first, e.g. a rule about a particular employee, for example, can be inserted before the rule about the department in order to implement an exception.

[1] The EPAL specification is outlined on http://www-1.ibm.com/services/security/epa.html.

An example of an informal privacy policy is the following:

> Allow a sales agent or a sales supervisor to collect a customer's data for order entry if the customer is older than 13 years of age and the customer has been notified of the privacy policy. Delete the data 3 years from now.

Translated to EPAL, this informal privacy policy may look like:

```
 <rule id="Rule1" ruling="allow">
<short-description language="en" />
<long-description language="en" />
<user-category refid="Sales Department" />
<data-category refid="customer-record" />
<purpose refid="order-processing" />
<action refid="store" />
<condition refid="the customer is older than 13 years of age" />
<obligation refid="delete the data 3 years from now">
</obligation>
</rule>
```

B.2 A hospital policy

In chapter 4 a case was chosen to show the applicability of a privacy violation detector (PVD). In subsection 4.2 and 4.2.3, the privacy policy for the case were outlined. The encoding of the privacy policy to EPAL is shown here:

```
 <rule id="Rule1" ruling="allow">
<short-description language="en" />
<long-description language="en" />
<user-category refid="doctor" />
<data-category refid="contact-data" />
<data-category refid="medical-history" />
<data-category refid="physicians-order" />
<data-category refid="progress-notes" />
<data-category refid="departmental-reports" />
<data-category refid="nursing-data" />
<data-category refid="operative-reports" />
<data-category refid="discharge-summary" />
<purpose refid="treatment" />
<purpose refid="diagnosis" />
<action refid="write" />
<action refid="read" />
<condition refid="isTreatingDoctor">
<obligation refid="log">
```

```
</obligation>
</rule>

<rule id="Rule2" ruling="allow">
<short-description language="en" />
<long-description language="en" />
<user-category refid="doctor" />
<data-category refid="contact-data" />
<data-category refid="medical-history" />
<data-category refid="physicians-order" />
<data-category refid="progress-notes" />
<data-category refid="departmental-reports" />
<data-category refid="nursing-data" />
<data-category refid="operative-reports" />
<purpose refid="treatment" />
<purpose refid="diagnosis" />
<action refid="write" />
<action refid="read" />
<condition refid="isAdvisingDoctor">
<obligation refid="log">
</obligation>
</rule>

<rule id="Rule3" ruling="allow">
<short-description language="en" />
<long-description language="en" />
<user-category refid="doctor" />
<data-category refid="contact-data" />
<purpose refid="treatment" />
<purpose refid="diagnosis" />
<action refid="read" />
<condition refid="isAdvisingDoctor">
<obligation refid="log">
</obligation>
</rule>

<rule id="Rule4" ruling="allow">
<short-description language="en" />
<long-description language="en" />
<user-category refid="doctor" />
<data-category refid="medical-history" />
<data-category refid="physicians-order" />
<data-category refid="progress-notes" />
<data-category refid="departmental-reports" />
<data-category refid="nursing-data" />
<data-category refid="operative-reports" />
```

```
<purpose refid="diagnosis" />
<action refid="write" />
<action refid="read" />
<condition refid="isExaminationSpecialist">
<obligation refid="log">
</obligation>
</rule>

<rule id="Rule5" ruling="allow">
<short-description language="en" />
<long-description language="en" />
<user-category refid="doctor" />
<data-category refid="operative-reports" />
<purpose refid="treatment" />
<action refid="write" />
<action refid="read" />
<condition refid="isSurgeon">
<obligation refid="log">
</obligation>
</rule>

<rule id="Rule6" ruling="allow">
<short-description language="en" />
<long-description language="en" />
<user-category refid="nurse" />
<data-category refid="physisians-order" />
<data-category refid="contact-data" />
<purpose refid="treatment" />
<action refid="read" />
<condition refid="isOnDutyOnPatientWard">
<obligation refid="log">
</obligation>
</rule>

<rule id="Rule7" ruling="allow">
<short-description language="en" />
<long-description language="en" />
<user-category refid="nurse" />
<data-category refid="nursing-data" />
<purpose refid="treatment" />
<action refid="read" />
<action refid="write" />
<condition refid="isOnDutyOnPatientWard">
<obligation refid="log">
</obligation>
</rule>
```

```
<rule id="Rule8" ruling="allow">
<short-description language="en" />
<long-description language="en" />
<user-category refid="receptionist" />
<data-category refid="patient-id-data" />
<purpose refid="registration" />
<action refid="read" />
<action refid="write" />
<condition refid="isOnDuty">
<obligation refid="log">
</obligation>
</rule>

<rule id="Rule9" ruling="allow">
<short-description language="en" />
<long-description language="en" />
<user-category refid="receptionist" />
<data-category refid="contact-data" />
<purpose refid="localisation" />
<action refid="read" />
<condition refid="isOnDuty">
<obligation refid="log">
</obligation>
</rule>

<rule id="Rule10" ruling="allow">
<short-description language="en" />
<long-description language="en" />
<user-category refid="receptionist" />
<data-category refid="contact-data" />
<purpose refid="registation />
<action refid="write" />
<condition refid="isOnDuty">
<obligation refid="log">
</obligation>
</rule>

<rule id="Rule11" ruling="allow">
<short-description language="en" />
<long-description language="en" />
<user-category refid="tester" />
<data-category refid="departmental-report" />
<purpose refid="perform-test" />
<action refid="create" />
<obligation refid="log">
```

```
</obligation>
</rule>

<rule id="Rule12" ruling="allow">
<short-description language="en" />
<long-description language="en" />
<user-category refid="tester" />
<data-category refid="departmental-report" />
<purpose refid="diagnosis" />
<action refid="read" />
<action refid="write" />
<obligation refid="log">
</obligation>
</rule>
```

Appendix C

Privacy Laws

This appendix gives an overview of the OECD Guidelines, the EU Directive and the Norwegian Personal Data Act, which lay out the basis for protection of privacy and e.g. defining privacy policies.

C.1 OECD Guidelines on the Protection of Privacy and Transborder Flows of Personal Data

The Organization for Economic Cooperation and Development (OECD) has recognized the need for a minimum adequate privacy standard for transmission of personal data across national boarders, and aims at harmonizing the different national laws and enforce some minimum degree of privacy protection amongst their member countries by their Guidelines on the Protection of Privacy and Transborder Flows of Personal Data. Another objective is to help raise international awareness of the importance of data protection.

The principles stated in [9] are the following:

Collection Limitation Principle There should be limits to the collection of personal data and any such data should be obtained by lawful and fair means and, where appropriate, with the knowledge or consent of the data subject.

Data Quality Principle Personal data should be relevant for the purpose for which they are to be used, and, to the extent necessary for those purposes, should be accurate, complete and kept up-to-date.

Purpose Specification Principle The purposes for which personal data are collected should be specified no later than at the time of data collection and the subsequent use limited to the fulfilment of those purposes or other that are not incompatible with those purposes and that are specified on each occasion of change of purpose.

Usage Limitation Principle Personal data should not be disclosed, made available or otherwise used for purposes other than those specified in accordance

with the 'Purpose Specification Principle' except with the consent of the data subject, or by the authority of law.

Security Safeguards Principle states that personal data should be protected by reasonable security safeguards against such risks as loss or unauthorized access, destruction, use, modification or disclosure of data.

Openness Principle There should be a general policy of openness about developments, practices and policies with respect to personal data. Means should be readily available of establishing the existence and nature of personal data, and the main purposes of their use, as well as the identity and usual residence of the data controller.

Individual Participation Principle An individual should have the right:

- to obtain from a data controller, or otherwise, confirmation of whether or not the data controller has data relating to him;
- to have communicated to him, data relating to him within a reasonable time; at a charge, if any, that is not excessive; in a reasonable manner; and in a form that is readily intelligible to him;
- to be given reasons if a request is denied, and to be able to challenge such denial; and
- to challenge data relating to him and, if the challenge is successful to have the data erased, rectified, completed or amended.

Accountability Principle A data controller should be accountable for complying with measures which give effect to the principles stated above.

C.2 The EU Directive on the protection of individuals with regard to processing of personal data and on the free movement of such data

As of 1995[1], the EU Directive 95/46/EC on the protection of individuals with regard to processing of personal data and on the free movement of such data requires all member states to implement legislation to protect the right to privacy with respect to the collection, processing, storage and transmission of personal data. The objective of the EU Directive is to help the EU member states to protect the fundamental rights and freedoms of natural persons and in particular their right to privacy with respect to the processing of personal data. Generally, the Directive aims at developing a high level of data privacy in the European Union. Free flow of personal data between the EU member states is also promoted, as the Directive seeks to harmonize the privacy laws in the different nations. Moreover, transfer of personal data to non-EU countries that do not meet the specified minimal European standard for data protection is prohibited.

[1] The European Commission's Directive on Data Protection went into effect in October 1998.

The Directive [8], Article 6 ("General rules on the lawfulness of the processing of personal data"), specifies that personal data must be:

- processed fairly and lawfully,
- collected for specified, explicit and legitimate purposes and not further processed in a way incompatible with those purposes,
- adequate, relevant and not excessive in relation to the purposes for which they are collected,
- accurate and kept up-to-date as necessary,
- kept in a form that permits identification of individuals for no longer than is necessary.

That is, Article 6 corresponds to the Collection Limitation Principle, the Purpose Specification Principle, the Usage Limitation Principle and the Data Quality Principle as stated by the OECD.

Article 7 ("Criteria of making data processing legitimate") states that personal data may be processed if the data subject has consented, as according to the Collection Limitation Principle, and additionally if the processing is necessary for e.g. the performance of a contract with the data subject, for compliance with a legal obligation or for the protection of vital interests of the data subject.

Processing of sensitive personal data [2] is according to Article 8 generally prohibited. However, there are some exceptions, such as if the data subject has consented to such processing or if it is necessary for the vital interests of the data subject.

Article 11 states that where data has not been obtained from the data subject, the controller must provide the data subject with information relating to e.g. the identity of the controller and the purposes of processing. Also, the EU Directive Article 12 ("Right of access"), which corresponds to the Individual Notification Principle of the OECD, states the data subject's right to:

- information, such as the identity of the data controller and the purposes of the processing for which the data are intended
- conformation as to whether or not data relating to him are being processed
- information such as the purposes of processing
- object to the processing of their personal data, see Article 14 [8].

Further, the confidentiality and security of processing is outlined in Article 16 and 17, and corresponds to the Security Safeguards principle stated by the OECD. The controller's obligation to notify the supervisory authority before carrying out

[2] e.g. data revealing racial or ethnic origin, political opinions, information concerning health issues and sex life, see Appendix A for a definition

processing is stated in Article 18, and the content of such notifications in Article 19. The EU Directive requires all member states to establish an independent supervisory authority to oversee the regulation of the usage of personal data, as stated in Article 28.

Article 25 states the principles for transfer of personal data to third countries outside EU. Generally, the export of personal data to third countries, which do not provide an adequate level of protection, is prohibited. However, Article 26 outlines exemptions to this general principle including if the data subject has given it consent to the transfer and if the transfer is necessary for the performance of a contract between the data subject and the controller.

In the United States there is no national comprehensive privacy law for the private sector and there is no independent supervisory authority designated to monitor the observance of privacy provisions. Hence, the United States do not provide an adequate level of data protection according to the EU Directive. Therefore, to bridge the different privacy approaches in the European Union and the United States, and to provide means for U.S organisations to comply with the EU Directive, a "safe harbour" framework has been developed. Basically, in order for companies to participate in the framework they must adhere to the Safe Harbour Principles [3]:

1. **Notice.** The customers should be told why their personal information is collected, how it will be used, where they can direct inquiries and complaints, to which third parties the business intends to disclose their personal information, and what choices they have to restrict the use and disclosure of their information.

2. **Choice.** If a business wants to use obtained personal information in a way not previously agreed upon or disclose such information to a third party, the business must obtain authorisation from the customer.

3. **Onward transfer.** To disclose personal information to a third party, the business must comply with the notice and choice principles. It must also limit its disclosure to third parties that follow the Safe Harbour Principles or that are subject to the EU Directive or any other "adequate" agreement.

4. **Security.** The business must take precautions to protect personal information from loss, misuse, unauthorized access, disclosure, alteration and destruction.

5. **Data Integrity.** All personal information collected must be relevant for the purposes it is to be used.

6. **Access.** The customer should have reasonable access to stored information about him or her and opportunities to correct and delete any inaccuracies.

[3] See e.g. http://www.export.gov/safeharbor, April 2004.

7. **Enforcement.** The business must provide an independent and available dispute resolution mechanism for investigating and resolving customers' complaints and disputes. The business must also have a procedure for verifying its compliance with the Safe Harbour Principles

Like the EU Directive, the Safe Harbour Principles allow for several exceptions, including national security, public interest and law enforcement.

C.3 The Norwegian Data Protection Act

Norway got its own Personal Data Act in 1978. However, the 1978 version is now made obsolete by the Personal Data Act of 2000. To ensure enforcement of the Personal Data Act, the Data Inspectorate, an independent administrative body under the Norwegian Ministry of Labour and Government Administration was set up in 1980. The purpose of this Act is to protect persons from violation of their right to privacy through the processing of personal data. The Act is intended to help to ensure that personal data are processed in accordance with fundamental respect for the right to privacy, including the need to protect personal information and private life and ensure that personal data are of adequate quality [6].

The need for a new Act in 2000 was partly due to the technological development seen and partly due to Norway's obligation to implement the 95/46 EU Directive on the protection of individuals with regard to automatic processing of personal data. Here, a comparison to the EU Directive will be made when appropriate.

In section 8, and according to Article 7 in the EU Directive, it is stated that personal data may only be processed if the data subject has given her consent [4], or the processing is necessary, i.e. due to contractual duty or in order to fulfil some legal obligation. Personal data may also be processed when it is deemed necessary for carrying out work due to public interests, such as processing of data for statistic, historic or scientific work.

Further, personal data must, according to section 11 and in compliance with Article 6, be:

- used only for explicitly stated purposes, and not used subsequently for purposes that are incompatible with the original purpose of the collection, without the consent of the data subject.
- adequate, relevant and not excessive in relation to the purpose of processing.
- accurate and up-to-date, and not stored longer than is necessary for the purpose of the processing.

[4] In the Act of 2000 a consent is defined as a freely given, specific and informed, indication of the wishes of the data subject, by which the data subject signifies his agreement to personal data relating to him being processed

Also, according to section 13 it should be ensured that there is a <satisfactory level of data security with regard to confidentiality, integrity and accessibility in connection with the processing of personal data, which corresponds to Article 16 and 17 of the Directive.

When personal data is collected from the data subject himself, the controller shall, as stated in section 19, inform the data subject of e.g. the name and address of the controller and the purposes of< processing. If data is collected from other persons than the registered person, such as credit and financial institutions, section 20 states the controller duty to notify the data subject when the data is gathered. This corresponds to Article 11 and 12 of the Directive.

When it comes to processing of sensitive personal data, the regulations are more rigid. The processing of sensitive information require a licence form the Data Inspectorate. Section 34 of the Act sets out the requirements that the Inspectorate might consider when deciding whether or not to give license. According to section 9, sensitive information may be processed if the data subject consents to such processing or if the data subject is unable to consent to such processing for reasons such as serious illness, processing may be necessary in order to protect his or her vital interests. That is, Section 9 corresponds to Article 8 of the EU Directive.

When it comes to transfer of personal data to other countries section 29 of the Act states that personal data may only be transferred to countries which ensure an adequate level of protection of the data. Countries which have implemented the EU Directive meet the requirement with regards to adequate level of protection. In such, section 29 corresponds to Article 25 of the EU Directive. However, personal data may also be transferred to countries which do not ensure an adequate level of protection if one of the exceptions in section 30 (the EU Directive Article 26) is fulfilled. Exceptions include if the data subject consents to such transfer or if there is an obligation to transfer the data due to i.e. membership in an international organisation. Further, according to the EU Directive, Article 28, Norway has got an independent supervisory authority to oversee the regulation of the usage of personal data.

Appendix D

A framework for the enforcement of privacy policies

Processing of personal data may be useful and necessary under some circumstances. In some cases, personal data must be collected due to legislation, or in order to provide some public service. Personal data may also be lawfully used to personalize services.

Usually, the data subject, being the person whose identity is connected to the data, has little control over the information collected and stored. Hence, the data subject must trust the data collection entity when personal data is collected and processed. To establish trust, systems for mandatory and automated enforcement should exist. The Norwegian Computing Center suggests a open framework for enforcement of privacy regulations and privacy commitments made by data collectors. The framework is described in [2] and will briefly be described here.

The framework is intended to function as a layer of control between personal data on one hand and services accessing and collecting data on the other hand. The suggested framework consists of framework elements that together provide the functionality necessary for enforcement of applicable regulations and privacy agreements reached in connection with data collection. In turn, each framework element is composed of components that support the implementation of its functionality. An overview of the different framework elements are given in figure D.1.

The **Access** element controls the flow of personal data between the personal data bundle repository and the other components of the framework. It is responsible for keeping track of all personal data that is held by the organisation and for regulating the access to this data in accordance with the configuration and the Agreements pertaining to the data.

The **Personal Data Bundle Repository** contains Personal Data Bundles. A Personal Data Bundle contains personal data and the Agreement regulation how the personal data can and should be used. An audit trail containing the access history of the personal data is also included in the Personal Data Bundle. An Agreement is here defined as a set of rules that determine how the personal data the

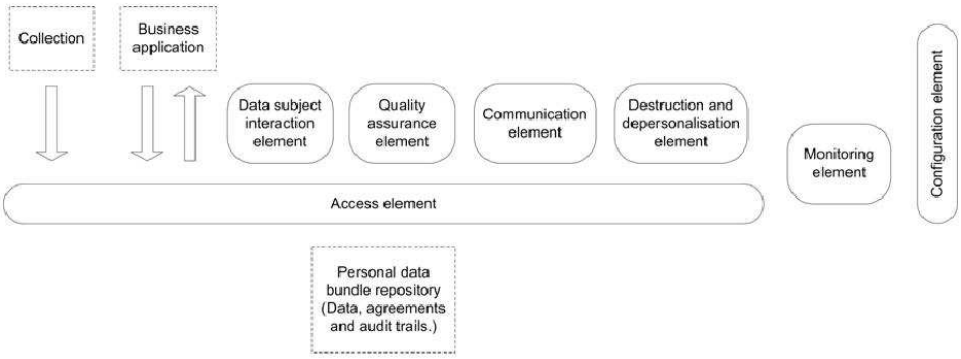

Figure D.1: A framework for enforcement of privacy policies [2].

Agreement pertains to can and should be used, and that both the data subject and data collector have consented to [2]. The agreement will be derived from the privacy promise made by the data collector and the data subject's preferences. A privacy commitment is based on an analysis of the need for personal data for business processes. Then, the data subject's privacy preference defines the Agreements that the data subject is willing to consent to. The agreement will typically state the purpose for the collection of the data (collected data must only be used for the stated purpose), how long the data will retained, if the personal data will be disclosed to third parties and whether or not the data subject is allowed to access its personal data.

The **Data Subject Interaction** element provides access to personal data, to the usage history and Agreements to data subjects. It also includes mechanisms for data subjects to submit complaints, and support for resolving such complaints.

The **Quality Assurance** element encompasses functionality that aims at upholding the correctness of the stored data. Norwegian legislation demands that the data controller ensures that personal data processed are up-to-date and accurate, and also adequate, relevant and not excessive in relation to the purpose of the processing [6].

The **Communication** element provides functionality for importing and exporting personal data in and out of the domain controlled by an instance of the framework.

The **Destruction and Depersonalisation** element is responsible for the last step of the life cycle of personal data (collection - processing - destruction). After this step the data should no longer be considered personal data. This element fulfills another important principle in Norwegian legislation stating that personal data may not be stored longer than necessary for the purpose [6].

The **Configuring** element is responsible for assuring that the configuration of the framework elements complies with applicable privacy regulations and that the framework is consistent with the local privacy policy. It includes functionality for generation of the other framework elements' configuration and functionality

for generation of privacy commitments. This functionality is automated or semi-automated, and based on the local privacy policy and applicable regulations. An overview of the Configuration process is shown in figure D.2.

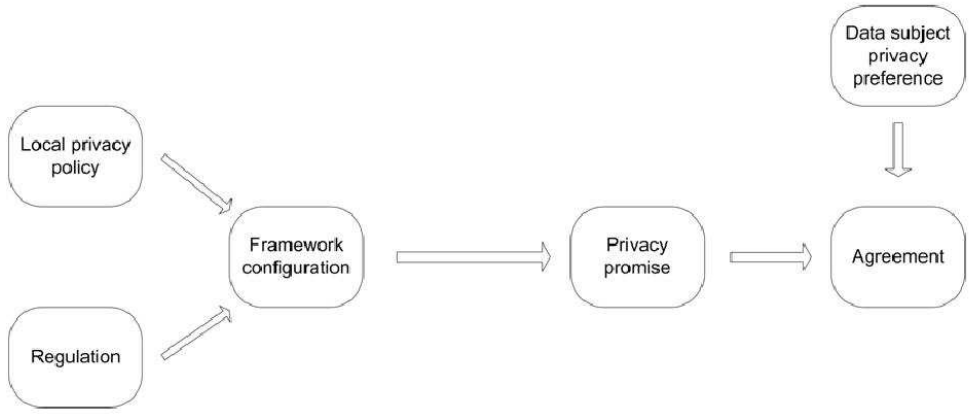

Figure D.2: The configuration process and the making of an agreement [2].

The **Monitoring** element monitors and analyses the audit trail generated by the other elements. The monitoring element is reactive, that is, they may enable detection of a policy breach and cause some reaction after the breach happened. Such security mechanisms are important, in order for users's to trust the systems. Monitoring elements can also be considered a part of internal control systems, which is mandated by Norwegian legislation [6].

In particular, the Monitoring element consist of:

- a Audit Manager, which support auditing and implements functionality for searching and reviewing the audit trail.
- a Privacy Violation Detector (PVD), which continually monitors access to personal data and detects misuse and/or anomaly behaviour. A PVD is outlined in chapter 3 and 4 of this thesis.
- a Remote Privacy Audit Manger, which provide e.g the Data Inspectorate with the possibility to remotely monitor and review the site.

Appendix E

PET Examples

Blind signatures Blind signatures can be considered an extension of the digital signature. A digital signature is the electronic equivalent to a handwritten signature, and provides a proof of authenticity. Public key encryption [1] forms the basis for digital signatures. When the sender encrypts a document with his or her private key, it is equivalent to signing a document with a hand written signature, as the private key is unique to that individual. The recipient may decrypt the document with the senders public key, and if it is successfully decrypted, the recipient may be sure about the documents authenticity. Basically, digital signatures ensure authentication of individuals, while blind signatures ensure authentication of individuals without identification, that is, blind signatures ensures the anonymity of the sender. One application involving blind signatures is the use of digital cash, which can be used as an form of electronic payment that is transmitted over computer networks. Clearly, this is a PET as blind signatures eliminate the collection of personally identifiable information. See [10] and [16].

Digital pseudonyms Digital pseudonyms build upon the blind signature technique. It is a method of identifying an individual through a pseudo-identity, created for a particular purpose, and it permits users to preserve their anonymity.

Trusted Third Party A trusted third party is an independent third party who is trusted by both the user and service provider. This trusted third party can be trusted to keeping things such as the master key linking digital pseudonyms with the true identities of their users, and keeps the relationship between a user's true identity and his/her pseudo-identity completely secret. However, if certain conditions require it, the trusted party will be permitted to reveal the user's identity (under previously agreed upon terms) to a service provider. The conditions under which an individuals identity would be revealed must be known to both user and service provider prior to entering into an agreement with the trusted party. See [10].

[1] In public-key systems two keys are created for each individual; one public, one private. The private key is kept secret, whereas the public key is made publicly available.

Anonymizers Anonymizers [2] enable users to surf the web without being tracked, monitored, profiled or exposed. Users visit web sites through a anonymizing software, rather than a standard browser, which enable user anonymity.

Mix-Nets The technique of Mix networks realises unlinkability of sender and receiver, sender anonymity against the recipient and optionally recipient anonymity. A Mix is a special network station, which collects and stores incoming messages, discards repeats, changes their appearance by encryption and outputs them in a different order. The relation between sender and recipient is hidden by the Mix. A Mix-net consists of a chain of independent mixes, which improve the security. The Mix-net is described in e.g. [1].

Remailers/Mixmaster Allows users to anonymously send email and post to newsgroups, basically by stripping email headers of personally identifiable information. The Mixmaster remailer, see [27], also provide protection against eavesdropping attacks. It works as follows:

1. chaining, with each link of the chain being encrypted
2. constant-length messages to prevent passive correlation attacks where the eavesdropper matches up incoming and outgoing messages by size.
3. message reordering code to stop passive correlation attacks based on timing coincidences.

As the safety for eavesdropping relies on "safety in number" (where the target message cannot be distinguished from any of the other messages in the remailer net), there is also a need for continuously-generated random cover traffic to hide the real messages among the random noise.

[2] See e.g. http://www.december.com/cmc/mag/1997/sep/boyan.html