Privacy Enhancing Identity Management: Protection Against Re-identification and Profiling

Sebastian Clauß
TU Dresden
Fakultät Informatik
D-01062 Dresden, Germany
sc2@inf.tu-dresden.de

Dogan Kesdogan
RWTH Aachen
Informatik IV
D-52056 Aachen, Germany
kesdogan@informatik.rwth-aachen.de

Tobias Kölsch
RWTH Aachen
Informatik IV
D-52056 Aachen, Germany
koelsch@i4.informatik.rwth-aachen.de

ABSTRACT

User centric identity management will be necessary to protect user's privacy in an electronic society. However, designing such systems is a complex task, as the expectations of the different parties involved in electronic transactions have to be met. In this work we give an overview on the actual situation in user centric identity management and point out problems encountered there. Especially we present the current state of research and mechanisms useful to protect the user's privacy. Additionally we show security problems that have to be borne in mind while designing such a system and point out possible solutions. Thereby, we concentrate on attacks on linkability and identifiability, and possible protection methods.

Categories and Subject Descriptors

K.4.1 [**Computers and Society**]: Public Policy Issues

General Terms

Management, Security

Keywords

Identity Management, Privacy

1. INTRODUCTION

Human beings want to control the privacy they have as well when visiting our cities with their different locations as when surfing the Internet with its various applications.

When visiting a bookshop in a city there is no need for showing a unique number at the entrance. But the Internet user does so when entering a virtual book store with his IP address. Here, anonymizing services can help to reach anonymity on the IP level. Users have the choice between simple anonymous proxies (e.g., Anonymizer [1]) or more secure services like Web mixes [1] or Tor [12], both more or less based on Chaum's Mixes [5]. The first ones only protect them against the Internet store while the latter additionally protect them against the provider of the anonymizing service.

When looking for a book and asking a bookshop assistant for advice one usually only reveals part of his interests, not a complete profile of past readings or even the name or similar personal data. But an Internet user often does so when logging in to a virtual bookstore's area. Here identity management (IDM) systems can help to control and reduce the amount of data transferred from the user to the server side to the minimum needed while reaching user's liability when the status of paying an item purchased is reached. So a user becomes known under different user profiles (or partial digital identities) to different applications. Identity management systems can be classified in user centric and server centric systems regarding the side (users or servers) who have control over personal data as we will outline in section 2.

As well on the IP layer as in concrete applications anonymity of users or unlinkability of user profiles is preferable but usually not achievable in a perfect way. Thus privacy enhancing technologies for both IP and application layer should inform a user about a resulting decreasing degree of anonymity and the increasing danger of identifiability.

Because a user can only be anonymous within a group of other users that might have the same user profile, a feedback about anonymity needs an estimation/calculation of the number of applications' users and a possible distribution of user profiles. Server centric types of identity management systems allow an easy calculation of these numbers, but usually do not reach anonymity of the user against the system providers while user centric approaches only allow to estimate the numbers.

Further the calculation/estimation of the usage of user profiles is needed because the linkability of user profiles might endanger a user's anonymity and lead to a re-identification. One specific user profile might be unlinkable to a specific user but a set of user profiles linkable to each other might build a comprehensive user profile for this user and reduces his anonymity (in the worst case to identifiability). When Clayton et al. studied technical attacks on an electronic student dating service [9] they found out that none

DIM'05, November 11, 2005, Fairfax, Virginia, USA.

[1] `http://www.anonymizer.com/`

of the typical technical attacks had been executed but some users tried to make 'social' attacks: They asked others for some of their habits or actions to build pseudonymous user profiles. With only small user profiles it was already easy to break some users' anonymity.

Recent anonymity research regarding protocols, design issues and attacks concentrated on the IP layer. An overview is given in [23], where Raymond discusses protocols, attacks, design issues, and open problems.

In section 3 we present a structural approach to classify attackers goals, attackers and attacks on identity management systems. Most techniques are already known from the IP layer and database security. This approach is usable for all known types of identity management systems.

2. OVERVIEW ON IDENTITY MANAGEMENT

In the digital world a person can be represented by sets of data (attributes) which can be managed by technical means, so-called digital identities.

Depending on the situation and the context only subsets of these attributes are needed to represent a person both in the physical and the digital world, so-called (digital) partial identities [18]. An identity management system provides the tools for managing these partial identities in the digital world.

A person typically uses different partial identities for work, others for leisure activities (e.g., doing sports, or with the family), or dealing with companies (e.g., a bank, a bookstore). Some partial identities containing the information which other communication partners typically know about a person, are shown in Figure 1. Some information is static (e.g., birthday) while other might change dynamically (e.g., interests).

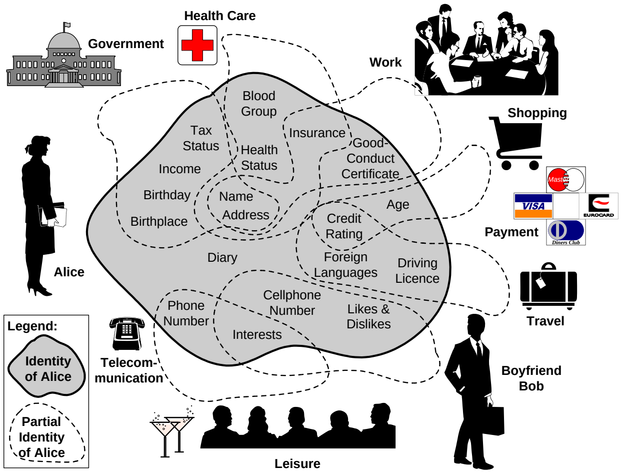

Figure 1: Partial Identities of Alice

Depending on the situational context and the communication partner each person might want to decide which partial identity to use in this relationship to the communication partner. Sometimes different names (nicknames, pseudonyms) are bound to the chosen partial identity.

Occasionally it is suitable to remain entirely anonymous, e.g. in the bookstore example when knowing exactly what to buy and paying anonymously. In other cases it is necessary to reveal identifying personal data, e.g. when being asked by a governmental representative for showing the identity card. Often, neither anonymity is acceptable to the communication partner nor identifiability of a person is needed, but only some (typically reliable and maybe certified) personal data is needed. The differentiated choices between the states of anonymity and identifiability depending on the user's wishes and the communication partner's prerequisite have to be supported by identity management systems.

Pseudonyms act as identifiers of subjects or sets of subjects (in the latter case called group pseudonyms). They comprise the entire field between and including anonymity and identifiability [18]. Therefore, pseudonyms serve as the core mechanism of an identity manager.

Wishing to use the same pseudonym more than once, the holder may take advantage of an established pseudonymous user account including e.g., presettings and reputation. Some kinds of pseudonyms enable dealing with claims in case of abuse of unlinkability to holders: Third parties may have the possibility to make the holder identifiable in order to provide the means for investigation or prosecution, or they may act as liability brokers of the holder to clear a debt or settle a claim. A pseudonym together with the data linked to it forms a partial identity.

This needs identity management systems to support and integrate both techniques for anonymity and authenticity to reach the following security goals:

Controlled pseudonymity of users: This consists of two aspects:

- **Unlinkability of pseudonyms and their holders (or holder anonymity):** The linkage of a pseudonym and its holder is not publicly known.
- **Unlinkability of pseudonyms:** The unlinkability results from their use in different contexts [20]: If the same pseudonym is used in many cases, the corresponding data about the holder, which is disclosed through each use, can be linked. In general, anonymity is the stronger, the less often and the less context-spanning the same pseudonyms are used. We distinguish between transaction pseudonyms, which are only used for one transaction, situation pseudonyms which are used in a specific context (e.g. according to the role of the holder or the relationship to the communication partner), and context-spanning person pseudonyms as substitutes for the holder's name respectively civil identity

(see also Figure 2).

Controlled liability of users: A pseudonym can be authenticated in a secure way and based on this authorized to use specific services. When necessary the holder of the pseudonym can be revealed and is liable for actions performed under this pseudonym.

The EU project FIDIS[2] distinguishes between three types of identity management systems (IMS) (see [19] for a detailed overview):

[2] Future of IDentity in the Information Society (`http://www.fidis.net/`)

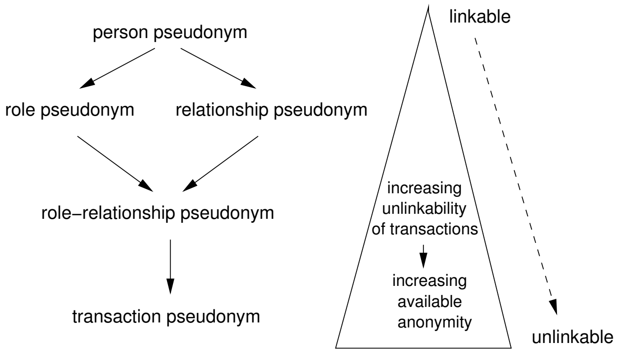
person pseudonym
role pseudonym
relationship pseudonym
role−relationship pseudonym
transaction pseudonym
linkable
increasing unlinkability of transactions
increasing available anonymity
unlinkable

Figure 2: Pseudonym types and related anonymity

1. Identity management systems for account management, especially implementing an AAA-infrastructure (authentication, authorization, and accounting),
2. Identity management systems for profiling of user data by an organization, e.g. by data warehouses which support personalized services or the analysis of customer behavior,
3. Identity management systems for user-controlled context-dependent role and pseudonym management.

Sections 2.1 and 2.2 refer to these types and describe them in more detail.

2.1 Server centric Identity management systems

Identity management systems of the first two types above are mainly implemented in a centralized way. The main goal of their usage is reliable identification of persons or reliable assignment of attributes to a person to reach liability of users while the second goal of identity management systems, the controlled pseudonymity, is neglected.

They store all personal data related to partial identities on the server side. The most simple form is a stand-alone system with only one partial-identity-database used at this server and for the applications provided to users.

Beneath this simple approach federated identity management became of great interest during the last years, it allows users to manage partial identities for different applications and with different communication partners. It has the following features:

- Identity provisioning: Based on one single registration at one service or so-called identity provider different services at different servers can create user accounts for partial identities of the identity this registration is associated with.
- Single-Sign-On: Based on the login to one user account at one service a user is able to use his user accounts at different services.
- Attribute exchange: The linkability of attributes to a partial identity at one service can be exchanged with other services.

A popular example of a specification for federated identity management is Liberty Alliance[3]. Although most current approaches conform to this specification are server centric identity management systems the specification would also allow user centric ones and is interoperable with such external identity management systems.

2.2 Decentralized Identity Management

Type 3 IMS are organized in a decentralized, user-oriented way and try to reach both aspects of identity management, controlled pseudonymity and reliability of users. This section gives an overview in basic principles and techniques used; for more details see [7, 8, 24].

Personal data is initially stored under the control of the user. Then, the user can decide, whether, to whom and for which purpose he wants to disclose personal data. This requires a network capable of keeping communication partners anonymous. Further, pseudonyms must be used in order to control linkability of personal data disclosed. Additional properties of pseudonyms, which are especially useful for privacy enhancing identity management, are explained in Section 2.2.1. In order to not only preserve privacy as much as possible, but also enable personal data to be certified by third parties, an additional infrastructure is needed. This is shown in Section 2.2.2.

2.2.1 Basic properties of pseudonyms

To reach both reliability and controlled pseudonymity of users pseudonyms can be created by the user owning them (allowing unlinkability of holder and pseudonym), or they can be generated and assigned by an application provider or by a third party (allowing controlled reliability of holders).

Digital pseudonyms could be realized as a public key to verify digital signatures where the holder of the pseudonym can prove holdership by forming a digital signature which is created using the corresponding private key. E.g., the public keys of PGP, bound to e-mail addresses, are digital pseudonyms.

Convertibility, i.e. transferability of attributes of one pseudonym to another is needed to reach unlinkability of pseudonyms. The user can obtain a convertible credential from one organization using one of her pseudonyms, but can demonstrate possession of the credential to another organization without revealing her first pseudonym. For this purpose, a credential can be converted into a credential for the currently used pseudonym. Therefore the use of different credentials is unlinkable. Chaum published the first credential system by [6]. Other systems have been proposed (e.g., Brands [2] and Camenisch/Lysyanskaya [4]).

Then, authorizations can be realized by credentials or attribute certificates bound to digital pseudonyms, but — in case of digital vouchers transferable to other people — by blind digital signatures or certificates as well.

2.2.2 Infrastructure

When a user wants to disclose personal data to a communication partner using a pseudonym, it can be linked to the pseudonym by a digital signature, which is issued on this pseudonym. To prevent the user from modifying the data beyond recognition, it must be certified by third parties. For instance, if a service can only be used by authorized users, but the users want to remain anonymous to the service, users need to show authorizations to the service which are issued by a third party and which are unlinkable to the users' pseudonyms.

[3] `http://www.projectliberty.org/`

Thus, for a comprehensive identity management, third parties must issue such authorizations (i.e., credentials) to the users. In the following, these third parties are called organizations. By issuing a credential, an organization certifies that the user owns a specific property or right. For instance, a governmental institution, such as a registration office, may issue credentials on the user's identification data like the name or the date of birth. One could also imagine credentials on the driving licence, age or rights of vote. A bank could certify that a user disposes of a specific amount of money.

When a user gets a credential, she can link it on demand to a pseudonym used during an action. The communication partner receiving the pseudonym verifies the credential to get the information certified by the credential issuing organizations.

If a communication partner wants to verify a credential, she needs some information of the credential issuing organization enabling her to perform the verification. For this purpose a PKI may be used. An organization publishes keys that can be used to verify the validity of organization's credentials. These keys must be certified by the CA and published on key servers, so that each potential verifier has access to them. The keys which are needed to verify the certificates of the organizations may be mutually certified and managed by using a PKI.

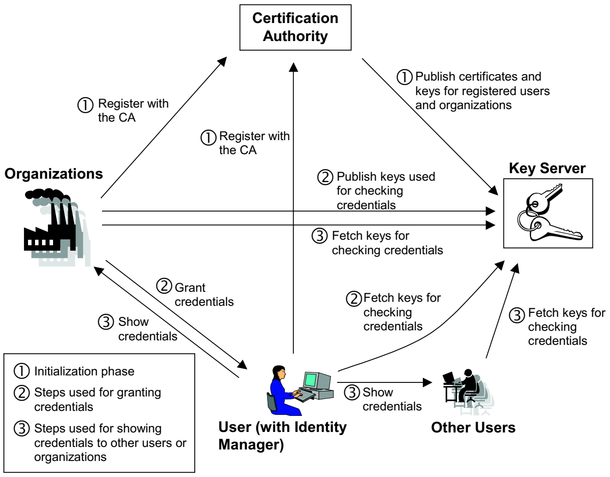
Certification Authority
① Register with the CA
① Register with the CA
① Publish certificates and keys for registered users and organizations
Organizations
② Publish keys used for checking credentials
Key Server
③ Fetch keys for checking credentials
② Grant credentials
③ Show credentials
② Fetch keys for checking credentials
③ Fetch keys for checking credentials
① Initialization phase
② Steps used for granting credentials
③ Steps used for showing credentials to other users or organizations
③ Show credentials
User (with Identity Manager)
Other Users

Figure 3: Infrastructure of an Identity Management System using Credentials

Thus, we need the following instances (see Figure 3):

- certification authorities where organizations and users can obtain certificates from,
- key servers where all published (certified) keys can be fetched from, especially the keys used to verify credentials.
- organizations which issue credentials to the users and publish keys to verify these credentials.

A credential system based on [4] is already implemented in the Idemix system [3] and will be used within the PRIME [22] project's approach for an comprehensive privacy enhancing identity management system.

Additional parties can support the user with her identity management and thus form a particular infrastructure. Such third parties comprise trustee services who may act as mediators like identity trustees, value trustees, or liability services. They may specialize on specific actions like payment or delivery services. Providing information about security and privacy risks with respect to deployed identity management systems is another important task which may be fulfilled by Privacy Information Services or Privacy Emergency Response Teams (PERT) analogous to today's Computer Emergency Response Teams (CERT).

2.2.3 Summary

Generally, Privacy-enhancing identity management systems are supposed to enable a user to control the nature and amount of personal information disclosed.

This can be realized with a comprehensive identity management system based on a communication network providing anonymity. The core components of such a system include digital pseudonyms built with various kinds of digital signatures. Thus, the possibility exists for modelling properties, specific to different application contexts.

The identity management system shall enable a user to minimize properties shown to a communication partner during her electronic communication. Therefore, it must be ensured that only this data is disclosed to a communication partner, which is intended to by the identity management. In this context, a measurement of her anonymity (or the opposite: identifiability) regarding some action is important for the user. It can inform her, to what extend a communication partner can identify her, i.e. link a digital identity to a real identity.

2.3 Anonymous Communication

In general, information about a user is not only transported on the application layer, but on the transport layer as well. Todays networks have no capabilities of hiding or pseudonymizing network addresses, which can contain quite sensible information.

Several solutions exist to solve this problem. Most solutions that became implemented in practice base on Mixes [5] or Onion Routing [12]. There, traffic relayed over a series of proxies, where the data is subject to cryptographic transformations to complicate traffic analysis.

An overview on these protocols and techniques is given in [23].

3. ATTACKS

In the context of identity management systems an attack is represented by a party that tries to find out information on a user this user does not want to disclose. The identity management system is to provide protection against re-identification and profiling even though the attacker can monitor and compromise certain parts of the system. It is assumed here that an individual Alice has several partial identities and various interests (i.e. interactions with other parties) and wants to control whom to give what information to achieve her personal and business goals. The attacker's goal is of course the opposite. He wants to learn the partial identities of Alice, the peer partners of Alice and the behavior of Alice (i.e. build a personal profile of Alice). To achieve his goal the attacker has two main resources:

Released Data The Attacker derives information from the data Alice releases. Either the released data is directly accessible to the attacker or he has to attack the database it is stored in.

Interaction The individual Alice has a daily behavior and interacts with her peer partners according to her interests (according to her personal profile).

IDM systems usually are no static systems. Since IDM systems are planned to be used for all kinds of user centric interactions. Several sources of information can be identified that help the attacker to achieve his goal by considering the following actions:

- System changes with time due to new information given by the user.
- Different partial identities of the same user are given to different organizations (bank, virtual store, distributor etc.) under different pseudonyms. These organizations have to interact with each other to provide a service.

Thus, the two main resources appear in all combinations. In its simplest form the attacker observes one database and the interaction of Alice using a pseudonym for this database (e.g. update of database, usage of time and frequency). Of course, the attacker is not bound to attack only one peer side (i.e. database). He can search for other peer sides and may derive more information about Alice by combining his knowledge from several databases. Finding peer sides can be very easy as stated above, since several peer sides can be highly connected to each other[4]. In the following we outline the goals an attacker might have (section 3.1), the capabilities he has (section 3.2) and several attacks (section 3.3) he actually might perform using these capabilities to reach (part of) his goal.

3.1 Attacker Goals

In the typical attacker scenario an opponent tries to discover information that is confidential, about a target to gain some kind of advantage. Some examples are:

- Blackmailing: force payment in exchange of non-disclosure of uncomfortable information.
- Revilement: publish private data to discredit someone.
- Insurance risk assessment: find out specific information on one person or a population to minimize the risk of a policy.

To obtain such information, the attacker is supposed to have some initial knowledge about his target (e.g. her birthday and her birthplace). An attacker stores all information acquired in the course of his target's transactions. Additionally he may have access to other collections of data, as publicly accessible or stolen databases or some observations made in the real world, or by observing the target otherwise. His goal is to combine these data sources to increase the knowledge he has about his target.

Even though it is hard to give a model about the exact goal a specific attacker might have we distinguish between the following kinds of success he might achieve:

- If the attacker can reach a *total success*, he can not only identify the individual but also has extensive knowledge about her profile.
- We call it a *simple success* if the attacker can identify the individual, but has incomplete information on her profile.
- If the above kinds of success are not achievable, then the attacker can still acquire a *partial success*. We distinguish here two types of partial success:
  - The attacker can successfully relate two independent pieces of information as originating from the same unknown individual.
  - By analyzing the given information the attacker downgrades the possible number of individuals to less than a given minimum.

The above goals can be restated using probabilities, e.g. the probability of linking of two partial identities should not be more than a given number x.

3.2 Attacker Model

Attackers can be distinguished in different ways: there are passive and active attackers. Some attackers have legitimate access to different databases (insiders) and some have only partial access or need to break into a database (outsiders).

	passive	active
insider (service provider)	the service provider gathers all information he gets through the transactions he is involved in	the service provider manipulates the communication and his database to get a better attack position
outsider (external attacker)	third person that observes transactions non-intrusively	third person that manipulates communication and databases to gain further information

Passive attackers try to receive the wanted information by passive data gathering, e.g. listening to the communication in a chat channel, or by sniffing web traffic. Other possibilities include collecting publicly available information, i.e. searching the world-wide-web for information on the target, or gaining access to databases with user data. While passive attackers are not as powerful as active attackers, they must not be underestimated because they act unnoticed by the target and do so from any distance. This lowers the attacker's risk of being discovered and thus raises the temptation to do so.

Active attackers try to conceive information by manipulating the user and his environment. The easiest way is by just asking the target about the wanted fact. A more subtle possibility is given, if the attacker has access to a data collection, in which pseudonymized users have the fact of interest disclosed along with other attributes. The attacker can now ask the target about apparently harmless attributes, not knowing that by disclosing these he will be uniquely identifiable on that data collection. If the user is known to release data differently in case disclosed data would make him reidentifiable, the attacker can manipulate the database by

[4] Usually to provide a complex service several parties are involved.

adding fake records in sparse regions of the feature space to make him believe disclosure of some data item is not critical.

An inside attacker is one that is involved in the user's transactions. He does not have to spend any effort in accessing the user profile, as it is gained as part of the transaction. This can be his chatting peer, the online shop he uses, a company (employee) with which the user stands in some kind of relation. This kind of attackers have full access on the presented data collection and are not bothered by the restrictions given e.g. by statistical databases.

Outsiders need to perform some special action to get access to user data. As presented earlier, this can be done e.g. by buying data collections, by sniffing the user's Internet connection, or by paying someone to provide him with information on the user. The accessed data sources can be by far inferior to those accessible to an insider. Note that a service provider that sniffs his users data is regarded as an outsider if the attack is based on the sniffing and as an insider if it is based on attacks on his customer databases.

To accomplish their goals the attacker may also take in account additional information, that is not directly given by the IDM or in the databases, as the traffic data of the user, and data on the user given on other protocol layers, the time of access, latency of responses, etc.

3.3 Review of different Attacks

This section gives an overview on the following kind of attacks:

Databases contain sensitive information about individuals or companies. (Restricted) access to some databases is given by different census agencies. Roughly four types of databases can be distinguished: person records, statistical databases, transaction databases, and unstructured knowledge bases.

Network Anonymity has the goal to protect the communication traffic of a user, i.e. hide all communication patterns. Security in anonymous communication is a well explored domain and a number of different theoretic and practical attacks on these systems have been identified. Even though there are some parallels with IDM systems and Network Anonymity we identify also major differences.

Interactivity Additionally, in identity management there are attacks based upon its interactive nature and coincidence of observed events.

3.3.1 Databases

In this type the attacker has some initial knowledge about his target to a point that he knows some specific attributes, e.g. *the owner of the red car that parks in front of my office.* He also has access to exactly one database (maybe that of the road traffic licensing department) and he is able to notice events that are related to the target by observing the person within some context or by observing the database.

Different types of databases can be distinguished based on the type of data store and the access methods they provide:

Person records These databases have a relatively static structure. They contain a well defined set of fields that hold the user data. Usually, each record can be associated to exactly one person and no person is represented more then once. These databases can be found as administration databases of enterprises, to which employees might have access. Alternatively many federal statistical offices provide access to anonymized microdata files. The level of anonymization of these files varies from case to case as described in Section 3.4. Also some companies create profit from creating user profiles based on some aspects of the users' life. They tend to have quite specific information on their clients. If an attacker has access to some of those databases and knows that his target is part of the database the attack can be performed quite straight forward. He compares the attributes he knows on his target and compares it to the records in the database, maybe using some appropriate distance measure. The record of his target should be that closest to the query record.

Statistical databases This kind of databases has the same inner structure as the previous type. However the access to these databases is restricted to statistical queries. They are also usually provided by statistical offices. As presented in [21] there are different methods to gain access to the relevant data in spite of these protections. However, e.g. the German federal statistical offices invest a lot of effort in protecting the databases from dangerous requests, as can be read in [25].

Using the database model we can find the same goal as in the IDM system: to not to give "too much" information, so the individual can be identified. However, the given information within the IDM system differs from statistical databases. Usually the attacker will not get access to statistics about groups of individuals, but exact information about individuals. E.g., the attacker controls the IDM system and has access to exact but incomplete information[5].

Attacks on statistical databases are sophisticated query techniques to gain exact information about a specific target, if its data is stored in a database. These attacks are applied, if the access to a database only allows statistical queries for privacy reasons, that means statistics calculated from small query-sets are not allowed. But in [11] Denning presents some techniques to subvert this limitation:

Tracker Attacks: The attacker blows up small query-sets with extra records to meet the security requirements of the database. This is done by queries which are more general than the intended one. With combination of the results one can construct the result of the target query.

Median Attacks: Suppose that the database permits queries that select a single value from a query-set, and a uniquely identifying formula is known. If the attacker is able to construct two queries in a way that the intersection of the query-sets contains only the victim, the values of the attacked attribute are all distinct and the result of both queries is the same. So, the query result is the victim's value of the attacked attribute.

Insertion and Deletion Attacks: Again a formula that identifies uniquely a victim is needed. The attacker inserts dummy records to blow up the query-set (i.e. he adds records that fulfil the victims identifying formula).

[5] Incomplete regarding identification of an individual.

After that he is able to query the statistic of the set characterized by the formula (that is the victims record plus all dummy records). Because the attacker knows all records but one, he is able to calculate back to the single unknown record.

A more profound elaboration on the possible attacks and their probabilities is presented in [15] or in [21].

Transaction databases These databases can be found in production systems, as accounting databases of a merchants software or the billing system of a phone company. Their usability for attacks heavily depends on their specific form. So the transaction logs of a cash register in a supermarket are usually not well suited for attacks in case of cash payment, as there is no link to the shoppers and the shopping profiles vary and overlap strongly. On the other hand the billing system of a phone company or the logs of a web server usually contain quite valuable information on specific persons, as the users is referenced to by a unique identifier. The attacks on transaction databases have the same form as those presented for the person records.

Unstructured knowledge bases Here a large unstructured collection of data exists and some search heuristics are applied to it so to retrieve information on the specific user of interest. One great example for this kind of databases is Google [6], where queries can be performed on a huge amount of publicly available homepages. An attack on this kind of data collections is performed by inquiring it using an appropriate query.

Note that the border between the different kinds of databases are not hard. So transaction databases can be seen as databases for person records if they contain a reference to the person that performed the transaction (as in the billing system example).

Server centric IDM uses central databases and allows attackers who succesfully attack databases the access to all user profiles stored there independent of the fact if these profiles have been already in use. In the contrast in user centric IDM user profiles only can be build and stored in central databases after the respective user used this profile with a peer partner. The IDM systems themselves could use the profiles available to inform users about their current state of anonymity/unlinkability. While in server centric IDM the number of profiles with the same attributes can be calculated in user centric IDM they can only be estimated.

3.3.2 Network Anonymity

Anonymity is the state of not being identifiable within a set, called the "anonymity set". General attacks downgrading the anonymity sets by observing the potential users are known in the area of network anonymity techniques [23, 17, 13].

These have as a common precondition that the attacker must be able to determine a sender anonymity set, that contains his victim with a high probability. This set is build on the network layer by passive observations, i.e. building sets of users contributing to a mix's batch [5]. Starting from the sender anonymity set, a recipient anonymity set is build. Any of those can be trivial, i.e. consist only of a single element, but they usually comprise multiple elements, although they might be weighted with different probabilities. A schematic picture of this can be seen in Figure 4.

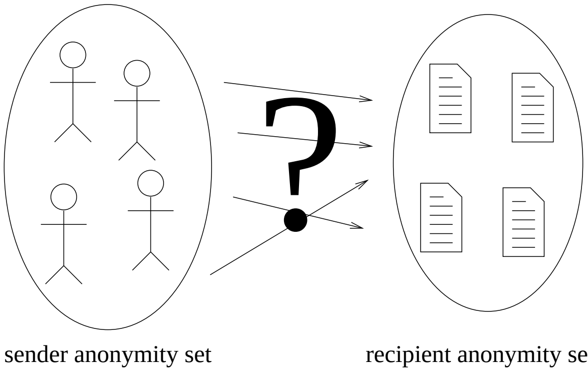

Figure 4: Model of Network Anonymity

The attacker's task is, to find the links between the elements in the sender anonymity set and the recipient anonymity sets. To this end, usually multiple observations are needed. The necessary number of observations depends on the attacker's knowledge of the system and the properties of the algorithm he's using. More details can be found in the according literature, like [17, 10, 13].

But, the same techniques can be used in IDM-contexts as well. Due to the abstract nature of the attacks, we can reidentify the same preconditions in IDM scenarios. Instead of linking two users in a network setting, we can adopt the attacks to link a user to his profile, or to certain parts of profiles.

Assume that the attacker is able to build a user anonymity set, and an anonymity set of profiles, partial identities, or interests. Thus the attacker is interested in linking all users on the one hand side to the profiles on the other. He can now use the existing attacks from network anonymity and apply their algorithms to his data.

In general, we can differ three cases:

- A record or profile can be linked directly to Alice by the attacker. In this case he is automatically successful and Alice has no protection at all.
- The attacker can build anonymity sets, and is enabled to roughly estimate a set of possible items that can be linked. The attacker can use the attacks described in this section to discover more about Alice.
- If the attacker can't link Alice or a profile to anything, Alice is perfectly protected. This is usually the case, if the supplied data is too unspecific to tell anything about it.

Server centric IDM allow attackers who succesfully attack databases the access to all user profiles stored there and thus calculate anonymity sets of profiles quite easily while user centric IDM only allow an estimation of anonymity sets based on past attacker observations. The IDM systems themselves could use this information on the anonymity set to inform users about their current state of anonymity / unlinkability.

[6] The well known Internet search engine is publicly accessible from `http://www.google.com`

3.3.3 Attacks Based on Interactivity

The attacks in the field of statistical databases and on anonymous communication can not directly be applied to the field of IDM. But they may form part of a such attack.

Now we will present an additional type of attacks. These are made possible by the interactive nature of identity management and by the fact that multiple access channels can be used by the offender. We derive the following (incomplete) list of attacks:

Timing Attack The attacker can be able to link timely events to the accessed data. If he can observe several databases, he can link changes to the records to some observations he made outside of the database. Say he observes the target's encrypted web traffic and notices that the user connects consecutively or simultaneously to two database and works with it. In a next step he checks all records that have changed at the time of the transaction and eventually this leads him to the wanted information on the user.

Wait and Seek Attack For example imagine a statistical service database to which users connect and enter an anonymous profile, which can later on be evaluated for statistical queries. An attacker that knows when his target connects to the service makes a query before he has entered his profile and one after the transaction. From the difference of the statistical results he can deduce the values entered by his target.

The opportunity to link timely events to the knowledge gained from the database is what makes the difference between the well known field of protection on databases and privacy enhancing identity management.

Linking Attack If the attacker has a single database, which contains a record of Alice, it might already put some of Alice's privacy at risk. Giving an attacker who has access to several databases is worse. If the databases' records contain data that sufficiently well identifies individuals this attack is trivial, of course, i.e. if the attacker is able to successfully identify her victim in each of these databases. But even otherwise could an attacker easily link individuals from all databases and join the information.

If the databases are pseudonymized or don't identify individuals sufficiently well there still remains a certain risk that records can be linked to broaden the attacker's knowledge. Even without learning Alice's real identity there might be enough information in the databases' records that an attacker might learn more about Alice, than a single database would allow him to do.

Selective Information Requests The attacker can alter the requests for information he pretends to require depending on the data transmitted so far. By this he can selectively ask for items that permit him to re-identify a user within some data collection he possesses.

3.4 Protection Methods

As we have seen there are many ways to compromise a user's anonymity. Now different techniques to protect the user privacy will be presented. Methods to thwart attacks on databases are mainly based on lying about attributes, roughening of data precision, and non disclosure of items. The other attacks can be aggravated by removing the timely correlation between events.

Many of the methods for databases have their roots in the protection of published micro data files as the needs there are quite similar to those encountered in IDM. However, it should be noted that there is one large difference. For the microdata files it is important that the correlations between different attributes are maintained, to maintain the utility of the data collection [24]. In the context of IDM this is not that important.

The formal anonymization is a quite simple way of anonymization in the field of statistics. It is done by simply removing all attributes that permit a direct contacting of a person (e.g. the full name, an address, a phone number) and typically wide spread identifiers as the social security number in the US. Before the presence of electronic data processing these means were sufficient to protect the subjects privacy. Now stronger means are necessary as motivated earlier.

A way to protect especially "outliers", in the sense of partial identities that are exposed, is by reducing the accuracy of the data. There are different methods to accomplish this. The simplest is by sampling every attribute range independently, e.g. by zeroing values after some digit or by creating value classes as "small", "medium", and "large". As outliers tend to be at the extremes of the scale, it is often useful to increase the size of the classes at the borders of the value range compared to those in the center. This method can easily be realized in the context of IDM. For discrete data, there is the possibility to take profit of semantical hierarchies, by merging some attributes to a super attribute. One problem with this method is that a consent between all users of an IDM on how this classing is performed is necessary. It does not lead to protection, if all users perform classing, but the classes are that different from user to user that they can be reidentified by their classing algorithm.

One problem with sampling is that usually outliers are not based upon one single attribute, but on a collection of attributes. A protection for outliers in this sense can be accomplished by looking for partial identities with a somewhat similar profile and forming a joint group and sharing the critical attributes. This method is usually referred to as aggregation and there are different techniques to perform such an aggregation, an overview can be found in [16]. It is obvious that this method cannot be applied locally, but includes some kind of communication between different individuals in which they agree upon common partial identities.

A similar result as with classing can also be realized by adding noise to an attribute or a group of attributes every time they are disclosed. The size of the noise has to be chosen, such that the service can reliably be provided (who would want to buy shoes that are 3 sizes to small just that no one can reidentify him), but sufficiently large that proper reidentification is effectively thwarted. Obviously if an attacker knows that his target may lie he will give more weight for the reidentification to the attributes that are necessary for the proper working of the service. Another problem with this method is to estimate the necessary noise level for the attributes. Here a defender needs information on the attribute distribution in the data collection. This is similar to the problem presented in the context of aggregation. If some (hopefully reliable) distribution is known a method, as described in [14] that estimates the utility of the different attributes and adds the noise according to this can be applied. As pointed out earlier, the correlation within the original data does not necessarily have to be kept. Contrarily, adding uncorrelated noise to correlated data can effectively confuses an opponent, as some reidentification metrics (e.g. some scaled Euclidean) assume correlated data.

An effective means to protect user data in statistical databases is by publishing only a representative sample of the original base [21, 16], especially in combination with data reduction or noisy data. By this an attacker cannot know if a specific target is present. As a result he is more heavily disturbed by the noise. In privacy enhancing IDM this method is analogous to using means for anonymous communication. Here the attacker cannot be sure that his target is the one from who he acquired his observations.

Protection against the attacks that are based on the timely correlation is quite difficult to achieve, as many of the aspects come from the server side. The user has no influence on when a database inserts a transaction, nor when one service forwards some personal information to another service (think of a book delivery with the online shop and the delivery service as involved parties, for example). One way of protection is by using anonymous communication, such that it is difficult to get a link to the timely correlations. Another possibility is that the service provider supports protection of his client at this level.

A non technical solution is by introducing strong data protection laws. However, the international nature of the Internet makes such a protection useless, or at least reduces its effect drastically. And even though this is an interesting approach, as computer scientists we are more interested in technical solutions, that prevent the mere possibility to snoop private data.

4. CONCLUSIONS

In this work we have given an overview of Identity Management Techniques. Identity management techniques are essential if we want to provide true anonymity and accountability at the same time. First we gave a survey of IDM in general and then concentrated on proposals for user centric IDM. Especially we have presented the approach of the EU founded project PRIME (Privacy and Identity Management for Europe). Although user centric identity management schemes are not new the security evaluation of IDM is a quite new area. To evaluate an IDM we have given a simple attack classification scheme. The main sources of the attacks are the released data (static IDM) and all the metadata during the interactions (dynamic IDM). We have identified three fields (i.e. statistical database, network anonymity, and interactivity) from which attacks on IDM can be derived. We have briefly discussed the protection methods against these attacks. Future work will show how risky the derived attacks are and how effective the protections methods can be.

5. ADDITIONAL AUTHORS

Additional authors: Lexi Pimenidis (RWTH Aachen, Informatik IV, D-52056 Aachen, Germany, email: lexi@i4.informatik.rwth-aachen.de),
Stefan Schiffner (Technische Universität Dresden, Fakultät Informatik, D-01062 Dresden, Germany, email: ss602038@inf.tu-dresden.de),
and Sandra Steinbrecher (Technische Universität Dresden, Fakultät Informatik, D-01062 Dresden, Germany, email: steinbrecher@acm.org)

6. REFERENCES

[1] O. Berthold, H. Federrath, and S. Köpsell. Web mixes: A system for anonymous and unobservable internet access. Designing Privacy Enhancing Technologies. Proc. Workshop on Design Issues in Anonymity and Unobservability, LNCS 2009, Springer-Verlag, Heidelberg 2001, pp. 115–129.

[2] S. A. Brands. *Rethinking Public Key Infrastructures and Digital Certificates - Building in Privacy.* PhD thesis, Netherlands, 1999. 2nd Edition: The MIT Press; August 2000.

[3] J. Camenisch and E. V. Herreweghen. Design and implementation of the idemix anonymous credential system. In *Proceedings of the 9th ACM conference on Computer and Communications Security*, Washington D.C., November 2002. ACM Press.

[4] J. Camenisch and A. Lysyanskaya. Efficient non-transferable anonymous multi-show credential system with optional anonymity revocation. Research Report RZ 3295 (# 93341), IBM Research, November 2000.

[5] D. Chaum. Untraceable electronic mail, return addresses and digital pseudonyms. Communications of the ACM, 24(2), pp. 84-88, 1981.

[6] D. Chaum. Showing credentials without identification: Signatures transferred between unconditionally unlinkable pseudonyms. In F. Pichler, editor, *Advances in Cryptology - EUROCRYPT '85, Workshop on the Theory and Application of of Cryptographic Techniques, Linz, Austria, April 1985, Proceedings*, volume 219 of *LNCS*, pages 241–244, Heidelberg, 1986. Springer Verlag.

[7] S. Clauß and M. Köhntopp. Identity management and its support of multilateral security. *Computer Networks*, 2001.

[8] S. Clauß, A. Pfitzmann, M. Hansen, and E. V. Herreweghen. Privacy-enhancing identity management. *The IPTS Report*, Special Issue: Identity and Privacy:8–16, 2002.

[9] R. Clayton, G. Danezis, and M. G. Kuhn. Real world patterns of failure in anonymity systems. Information Hiding 2001, LNCS 2137, pp. 230 245, Springer-Verlag Berlin 2001.

[10] G. Danezis and A. Serjantov. Statistical Disclosure or Intersection Attacks on Anonymity Systems. Proceedings of the 6th Information Hiding Workshop (IH2004), LNCS, Toronto, 2004.

[11] D. E. Denning. A security model for the statistical database problem. In *SSDBM*, pages 368–390, 1983.

[12] R. Dingledine, N. Mathewson, and P. Syverson. Tor: The second-generation onion router. In *Proceedings of the 13th USENIX Security Symposium*, August 2004.

[13] L. P. Dogan Kesdogan. The Hitting Set Attack on Anonymity Protocols. In *Proceedings of Information Hiding, 7th International Workshop*. Springer Verlag, 2004.

[14] G. Duncan, S. Keller-McNulty, and L. Stokes. Database security and confidentiality: Examining disclosure risk vs. data utility through the R-U confidetiality map. Technical Report 142, U.S. National Institute of Statistical Sciences, March 2004.

[15] U. W. Gerhard Paaß. *Datenzugang, Datenschutz und Anonymität*. Oldenbourg, München, 1985. (in german).

[16] J. Höhne. Methoden zur Anonymisierung wirtschaftsstatistischer Einzeldaten. *Forum der Bundesstatistik*, 42:69–94, 2003.

[17] D. Kesdogan, D. Agrawal, and S. Penz. Limits of Anonymity in Open Environments. In *Information Hiding, 5th International Workshop*. Springer Verlag, 2002.

[18] M. Köhntopp and A. Pfitzmann. Anonymity, unobservability, and pseudonymity - a proposal for terminology. Draft v0.12., June 2001.

[19] M. Bauer and M. Meints (Editors). Structured overview on prototypes and concepts of identity management systems; fidis del. 3.1. available from `http://www.fidis.net/fileadmin/fidis/deliverables/fidis-wp3-del3.1.over%view_on_IMS.pdf`.

[20] B. Pfitzmann, M. Waidner, and A. Pfitzmann. Rechtssicherheit trotz anonymität in offenen digitalen systemen. *Datenschutz und Datensicherung (DuD)*, 14(5-6):243–253, 305–315, 1990. Vieweg, Wiesbaden.

[21] K. Pommerening. *Datenschutz und Datensicherheit*. BI-Wissenschaftsverlag, Mannheim, Wien, Zürich, 1991. ISBN 3-411-15171-4 (in german).

[22] PRIME - Privacy and Identity Management for Europe. http://www.prime-project.eu.org.

[23] J.-F. Raymond. Traffic analysis: protocols, attacks, design issues, and open problems. In *International workshop on Designing privacy enhancing technologies*, pages 10–29, New York, NY, USA, 2001. Springer-Verlag New York, Inc.

[24] W. Winkler. Masking and re-identification methods for public-use microdata: Overview and research problems. Research Report # 2004-06, U.S. Bureau of the Census, October 2004.

[25] S. Zühlke, M. Zwick, S. Scharnhorst, and T. Wende. The research data centres of the federal statistical office and the statistical offices of the länder. FDZ-Arbeitspapier 3, Statistische Ämter des Bundes und der Länder, March 2005. `http://www.forschungsdatenzentrum.de/publikationen/arbeitspapiere/03.asp`.