A Framework for Quantification of Linkability Within a Privacy-Enhancing Identity Management System*

Sebastian Clauß

TU Dresden, Germany
Sebastian.Clauss@tu-dresden.de

Abstract. Within a privacy-enhancing identity management system, among other sources of information, knowledge about current anonymity and about linkability of user's actions should be available, so that each user is enabled to make educated decisions about performing actions and disclosing PII (personal identifiable information).

In this paper I describe a framework for quantification of anonymity and linkability of a user's actions for use within a privacy-enhancing identity management system. Therefore, I define a model of user's PII and actions as well as an attacker model. Based thereon, I describe an approach to quantify anonymity and linkability of actions. Regarding practical applicability, a third party service for linkability quantification is discussed.

1 Introduction

A privacy-enhancing identity management system[1] shall assist a user in using services (on the Internet) in a least privacy invading way. A basic technique for this is to be initially anonymous, e.g. by using an anonymity service at the network layer. So, in principle a user can control that only information required to perform services is disclosed. Depending on the service, such required information restricts privacy of the user to some extent.

Within a privacy-enhancing identity management system, a user needs to get reasonable information about his privacy status in order to make educated decisions about performing actions and disclosing PII (personal identifiable information). Among other sources of information, knowledge about current anonymity or about linkability of certain actions can help a user to assess his privacy.

In this paper a framework for quantification of anonymity and linkability of a user's actions is described, with the perspective of using such quantification within a privacy-enhancing identity management. After outlining related work on anonymity and linkability measurements in Section 2, a model for users and actions with regard to transferring PII is introduced in Section 3. Further, an

* Parts of this work have been supported by the Project FIDIS, a Network of Excellence within the EU's 6th Framework Programme.

[1] See e.g. [1] for details on functionality of a privacy-enhancing identity management system.

G. Müller (Ed.): ETRICS 2006, LNCS 3995, pp. 191–205, 2006.

attacker model is defined in Section 4. Based on these models, an approach for quantifying user's anonymity and linkability between actions is described in Section 5. Therefore, the problem of getting information needed for such quantification is discussed in Section 5.1. The basic model does not incorporate the time aspect, i.e. that user's attributes may change over time. An enhancement regarding this issue is discussed in Section 5.5. As quantification of anonymity and linkability of actions is a highly resource consuming calculation, in Section 5.6 possibilities to make use of third parties for this task are discussed.

2 Related Work

Over the years aspects of anonymity and (un-)linkability evaluation have been researched mainly with respect to evaluation of anonymity providing services on the network layer, e.g., mixes [2].

Regarding the connection layer, methods for anonymity evaluation have been described by Díaz et al. [3] and Serjantov and Danezis [4]. They describe properties of anonymity services and derive methods for measuring anonymity. The scenario used there is not directly comparable with the one I use. Their approach differs in that they consider a known set of users, whereas in our model the number of entities (users) is only restricted by the possibility to distinguish entities by their attributes[2]. Another difference is the main objective: these authors define a measure of anonymity based on prior knowledge, whereas they do not define how such knowledge is gathered and organised. Similar to these papers, I use entropy based measures for anonymity.

Based on [3] and [4], Steinbrecher and Köpsell [5] describe a general information-theoretic model for (un-)linkability of similar items (e.g., subjects, messages, events, actions, etc.) within a system. This model is consistent with ours. In Section 5.4 I apply methods for linkability measurement described in this paper.

Regarding the application layer, Díaz et al. [6] describe how entropy can be used to measure anonymity, but similar to [3], they assume that the number of users is known. They also assume that the attacker does get more information about a message than just the data in it, e.g. he can also see which user sends at a given time. In our model, the attacker only gets to know properties of the user (entity). Such properties may also be used to model information gained on the connection level, but our system abstracts from this by only talking of entities' attributes which can have different values. Similarly to the other papers referenced above, they also do not describe how exactly the attacker gains information, and how this information is aggregated.

Besides the information theoretic approaches discussed above, Hughes and Shmatikov [7] describe the partial knowledge of a function based on a mathematical abstraction. They specify anonymity properties using a modular approach. Their approach can be used independent of the underlying algebra or logic. In contrast to our approach, their approach is not probabilistic, i.e. items of interest are considered either fully linkable or not linkable at all.

[2] See Definition 6, *observer state*.

3 Modelling Users and Actions

In this section I define a model with regards to the entities involved, actions, data flow and the modelling of the data.

Within the model there is a finite set $\mathcal{E}$ of *entities*[3] $e \in \mathcal{E}$. An entity represents a subject out of the real life, like a real person or an artificial person (e.g., legal person). Entities are considered to be able to communicate by means of computer networks, mainly the internet. Further, entities have properties, by which entities can be classified into different subsets of $\mathcal{E}$. For each entity $e \in \mathcal{E}$ there exists at least one set of properties by which it can be identified within $\mathcal{E}$. These properties are also called personal identifiable information (PII).

Actions take place in form of communication between entities. Thereby, a single action takes place between exactly two entities e_m and e_n. In an action at least one of the communicating entities transmits data to the other one. Data transmitted during an action is considered to *belong together*. Further, these data can contain properties of the originating entity.

3.1 Attributes and Digital Identities

In order to structure the properties of entities transmitted during actions, data can be modelled as *attributes* and their *values*. These terms are defined as follows:

Definition 1 (Attribute Value). *An* attribute value *is a property of an entity.*

Definition 2 (Attribute). *An* attribute $\mathcal{A}$ *is a finite set of values* $a \in \mathcal{A}$*, and a special value "not applicable"* $\ominus$*.*

For example, the attribute "gender" consists of two attribute values "male" and "female". In case also legal persons or machines are considered as entities, the attribute "gender" would get the value "not applicable".

In addition to information from the content of messages, knowledge gained at the connection layer can also be modelled by means of attributes.

Within the model, I assume a finite set of attributes. Based on the attributes, the *digital identity* can be defined corresponding to the definition in [8].

Definition 3 (Digital Identity). *A* digital identity *is a complete bundle of attribute values. Thereby "complete bundle" means, that a* digital identity *comprises one value of every attribute.*

An entity has at least one digital identity. In case multiple values of an attribute are properties of one entity e, this entity has multiple digital identities.

Example 1. In Figure 1 relations between attributes, their values and a digital identity are shown by a concrete example.

[3] For reasons of intuitive understanding, I often speak about *users* throughout this paper. Regarding the model defined here, a *user* is an *entity*.

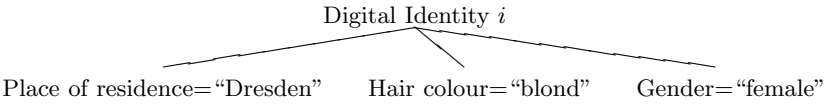

Fig. 1. Attributes with values assigned to a digital identity

The above definition is static, i.e. attribute values within a digital identity cannot change[4].

For describing a general framework for anonymity and linkability quantification within the next sections, this static model will be used. In Section 5.5 an approach for incorporating time properties will be described.

In this model values of attributes are assumed to be discrete. So the model does not directly consider attributes, where values can be continuous. But to use such values in the digital world, they must be measured and thereby (considering a certain fuzziness of the measurement) are transformed into discrete values, which then are covered by the model given[5].

4 Attacker Model

Generally, an entity's security goal is to find out, whether and to what extent the disclosure of a set of PII items (values of attributes) helps an attacker to link the current action to other actions of the entity, or to identify[6] the entity. To clarify this, the attacker and the success criterion of an attack, i.e. *linkability of actions* or *identification*, must be defined.

With respect to privacy-enhancing identity management, the attacker is assumed to be a set of service providers, with which users perform actions. More formally, the attacker is characterised by the following assumptions with respect to an entity e:

- The attacker controls one or more communication partners of e, i.e. gets to know data disclosed by e during actions with these communication partners.
- The attacker has general knowledge about attributes of entities, i.e. has access to public information services, e.g., phonebook entries, statistical offices.

Using this information, the attacker tries to identify entities, and tries to find out, whether different actions can be linked.

[4] I.e., with this definition a change of an attribute value of an entity means that the entity switches to another digital identity.

[5] An observer may nevertheless make a difference between attributes resulting from measurements and attributes defined in a discrete space, e.g., authorisation tokens. I discuss this issue in Section 5.2 with respect to matching functions for attribute values.

[6] Throughout this document, the term "identified" is used as the opposite of "anonymous".

Regarding the model defined in Section 3.1 two goals can be specified:

- Observing an action, an attacker wants to find out the digital identity this action belongs to.
- Observing two actions, an attacker wants to find out whether they originate from the same digital identity.

In order to quantify anonymity and unlinkability of *entities*, we need to specify the relation between entities and digital identities. In order to identify entities, the set of attributes considered within the system must contain a subset which is sufficient to identify entities (in the physical world[7]). Under this assumption, two cases need to be considered regarding the relation between entities and digital identities. Either, attributes are defined in a way that an entity has only one digital identity. In this case, identifying a digital identity means identifying an entity. In case an entity may have more than one digital identity, the attacker needs to group digital identities by data identifying the entity (in the physical world). (See Section 5.3 and 5.4 on how this influences measurements.)

5 Quantification of Anonymity and Linkability

A user wants to know, how his current privacy status is, or how it will be regarding actions planned given the current circumstances. The focus of this document is technical. So, privacy is seen here from a technical point of view, i.e. it is interpreted as degree of anonymity or linkability.

According to [8], unlinkability of actions can defined as follows:

Definition 4 (Unlinkability). Unlinkability *of two or more actions means that within the system, from the attacker's perspective, these actions are no more and no less related after his observation than they are related concerning his a-priori knowledge.*

In the scope of this section, "related" means the grade of certainty of the attacker, that these actions originate from the same entity. Measuring linkability means determining this certainty.

Anonymity can be defined as follows [8]:

Definition 5 (Anonymity). Anonymity *is the state of being not identifiable within a set of subjects, the anonymity set.*

For quantifying anonymity, the information gained from action c needs to be compared to the information needed to identify a certain entity out of a given set of entities (the anonymity set).

Within this section I first deal with anonymity and unlinkability with respect to digital identities as defined in Section 3.1. After that, I discuss anonymity and unlinkability with respect to entities under the assumption that a set of data sufficient for identification of entities in the physical world is known.

[7] This paper does not aim at defining, which data is sufficient for this purpose. Here, possibly legal definitions may suffice, but it may vary depending on the attackers context.

5.1 Sources of Information

Quantification of privacy is not possible without information from outside the user's domain. In other words, the user needs to know what an assumed attacker knows, in order to quantify information contained in disclosed data against this knowledge. The following example illustrates this:

Example 2. Let's assume a user disclosing his surname "Sebastian". Within the system, there may be many users named "Sebastian", but this name could also be unique. Only with knowledge about the other users it is possible to quantify, how much information is contained in this name.

Privacy quantification would be easy, if there was a "Big Brother"-like source of information [9], which has all knowledge available within the system. But such a source of information does not exist in today's internet[8].

So, multiple sources of information need to be taken into account, each having different partial knowledge about the system. Sources of information can be distinguished into the following categories:

- The user himself, i.e., the user who wants to quantify his privacy in a certain situation. The main source of information the user has itself is the history of data disclosed in past actions.
- Other users. Information from other users can be parts of their disclosure history.
- Public parties, e.g., public statistical offices, or special services supporting anonymity and linkability quantification by aggregating data about users in order to generate specific statistics regarding service providers considered as attackers.
- Service providers. Usually, it seems strange to assume service providers to be sources of information, because they are rather seen as the attackers on privacy, which a user wants to defend against. But the goal of a service provider can be seen differentiated. On the one hand, his goal is to find out as much as possible about profiles of users in order to optimise his services etc. On the other hand, some of the profiled information can be published for marketing or corporate image purposes, e.g., the total number of users may be a criterion to decide on the acceptance of a service for users, so it could be published by the service. There could also be privacy certificates for which service providers can apply, which are only issued in case certain information important for user's anonymity and linkability quantification is provided.

In order to utilise sources of information for evaluating the privacy situation, they need to be trusted regarding *correctness* of information. This also includes that information needs to be up to date.

In this paper, I will not go into detail how trust in sources of information may be established. Possibilities here are trust because of legal regulations, because of personal or third party evaluations, reputation systems, etc.

[8] Even that it would make privacy quantification easy, it would also be a perfect attacker on privacy.

Correctness of Information. In order to make use of information for evaluation of the privacy situation of a user, the information needs to be correct regarding the following aspects:

- the scope of the information, i.e., it must be clear, about which number, kind etc. of people a statistic contains information,
- the validity period of the information.

As noted above, there is no general master source of information. So, verification of correctness of data is not generally possible. The only possibility for verification is to have multiple sources of information about the same items of interest, which can be compared against each other. Then, techniques known from research on fault tolerance, e.g., majority voting, can be used to decide on the correct information.

Summary. The above sections make clear, that evaluation of a user's privacy exclusively bases on sources of information, which can be more or less trusted by the user. In this context, the evaluation can never be objective and universal. It will always be a subjective view, and under realistic circumstances, it will nearly never be possible to exactly get to the same results as an (assumed) attacker.

As a consequence of this, a technical evaluation of privacy should not be used as a automatic criterion to base decisions about actions on, but it can give hints.

5.2 A Method for Calculation

In this section, a method for calculation of anonymity and linkability measures is described in order to show, how information from the different sources could be aggregated. Further, this method forms the basis for enhancements described in Section 5.5.

Aggregating Input Data. The mathematical model sketched here enables to use entropy metrics for determining an anonymity set size for a given set of disclosed data items. Here, only the aspects of the model are described, which are needed to describe the calculation anonymity and linkability metrics. A more detailed description can be found in [10]. This model operates on static digital identities as defined in Section 3.

By observing actions an observer gets a limited insight into user's PII and into relations between PII items. The observer can collect this information, and conduct any desired statistical analysis on them. With a growing number of observations the information on probability distributions of the digital identities gets more exact[9]. I define the knowledge of an attacker which he gained by observations in form of the *observer state*:

Definition 6 (Observer State). *The* State $Z^{\mathcal{X}}$ *of an observer* $\mathcal{X}$ *is a triple* $(\mathcal{I}, h, g)$*, where:*

[9] "exact" here means exact with respect to the observation. Observations may nevertheless yield incorrect information (see Section 5.1).

– *$\mathcal{I}$ is the set of all digital identities possible.*

$$\mathcal{I} = \mathcal{A}_1 \times \mathcal{A}_2 \times \ldots \times \mathcal{A}_n$$

– *$h : \mathcal{I} \mapsto \mathbb{R}$ is a function, which assigns a probability to every digital identity, i.e.,* $(\forall i \in \mathcal{I}.0 \leqq h(i) \leqq 1)$
– *g is the number of observations leading to this* state.
– *the sum of all probabilities is 1.*

$$\sum_{(\mathcal{I})} h(i) = 1$$

$h(i)$ denotes the probability that within the set $\mathcal{I}$ of all possible identities the identity i is observed by the attacker. When the attacker observes an action of a user, the probability of the identities matching to the observation (i.e., the suspects with respect to the observation) is raised, whereas the probability of all other identities is lowered. After defining observations, I specify a method for matching identities and observations.

Definition 7 (Observation). *An* observation *is a (possibly incomplete) bundle of attribute values. Such a bundle contains at most one value per attribute. The set $\mathcal{B}$ of all possible* observations *is the cross product of all attributes with an additional element "not observed" $\perp$.*

$$\mathcal{B} = (\mathcal{A}_1 \cup \perp) \times (\mathcal{A}_2 \cup \perp) \times \ldots \times (\mathcal{A}_n \cup \perp)$$

Intuitively, this means that during actions a user discloses PII. The observer *observes* this PII and gets a more and more refined view on the digital identities and by that on the users.

Within the set of all possible digital identities an observer can separate suspect digital identities with respect to an observation from non-suspect digital identities. The set of *suspects* related to an observation can be defined as follows:

Definition 8 (Suspects). *The set of* suspects $\mathcal{V}_b$ *related to an observation $b = (x_1, .., x_n)$ contains all digital identities $i = (x'_1, .., x'_n)$, whose attribute values are either equal to attribute values of b or are not contained in b.*[10]

$$\mathcal{V}_b = \{i | x_k \in \{x'_k, \perp\}, k = 1, .., n\} \tag{1}$$

As stated above, the observer "learns" by observations. The following definition formalises this learning process:

[10] The matching function "equality" used here is a simple example. This makes only sense, if attribute values are discrete and not related to each other. If this is not the case, e.g., if measuring faults for originally continuous attribute values (see Section 3.1) need to be taken into account, other matching functions should be used which reflect such properties of attributes.

Definition 9 (Observer State Update). *Let $b \in \mathcal{B}$ be an observation and $\mathcal{Z}$ a set of observer states. An* observer state update *$\delta : \mathcal{Z} \times \mathcal{B} \to \mathcal{Z}$ constructs a new observer state from a given state and an observation.*

These definitions are a framework for formalising concrete observations and statistical analysis based on digital identities. In order to not restrict this model to passive (observing only) attackers, it is intentionally not defined how an observation is done. So, an attacker may observe messages, but may also actively insert or fake messages in order to observe users' reactions.

Based on the above definitions a statistical observer model is defined as follows:

Definition 10 (Statistical Observer Model). *A* statistical observer model *of an observer $\mathcal{X}$ comprises a set $\mathcal{I}$ of digital identities, a set of observations $\mathcal{B}$, a set $\mathcal{Z}^{\mathcal{X}}$ of observer states and a function δ, which derives new observer states from previous states and observations.*

The statistical observer model specifies the observer's knowledge in form of statistics about digital identities together with a method for aggregating newly gained knowledge. This is an abstract definition, as it leaves open how actually the aggregation of new observations influences the probabilities of digital identities.

In order to actually perform calculations within this framework model, a concrete model can be defined as follows[11]:

Let $\mathcal{I}$ be a set of digital identities and $\mathcal{B}$ the set of all observations possible. The set of states $\mathcal{Z}$ is defined inductively. First, I define the initial state, in which the attacker did not do any observations. For the initial state $Z_0 = (\mathcal{I}, h, g)$ it shall hold, that $g = 0$ and $(\forall i \in \mathcal{I}.h(i) = \frac{1}{|\mathcal{I}|})$.

Now I specify how an observation actually changes the probabilities of the digital identities. A function $\delta : \mathcal{Z} \times \mathcal{B} \to \mathcal{Z}$ derives a new state $Z_{k+1} = (\mathcal{I}, h_{k+1}, g_{k+1})$ from a previous state $Z_k = (\mathcal{I}, h_k, g_k)$ and an observation $b \in \mathcal{B}$ as follows:

$$h_{k+1} : i \mapsto \frac{h_k(i) * g_k + x}{g_k + 1} \tag{2}$$

$$x = \begin{cases} \frac{1}{|\mathcal{V}_b|} \text{ iff } i \in \mathcal{V}_b \\ 0 \quad \text{otherwise} \end{cases}$$

$$g_{k+1} = g_k + 1 \tag{3}$$

This intuitively means, that first each observation gets an equal "weight" 1. Then, this "weight" is divided by the number of suspects of this observation. By doing that, more significant observations (i.e., observations containing values of more attributes) get a bigger influence on the probability of the suspect identities

[11] The concrete model described here is an example, in order to show a possibility how observations can be aggregated in a meaningful way into a *statistical observer model*. There may exist other concrete models.

than less significant ones. Further, the "weight" of the observation is set into relation to the number of observations already aggregated, so that every observation already aggregated has the same overall influence on the probabilities.

In fact, the observer model defined above sums up relative frequencies. With a growing number of observations, it can be assumed that the relative frequencies converge to probabilities. By induction over g, it can be shown, that function h always has the properties of a probability distribution, i.e., $\sum_{(i \in \mathcal{I})} h(i) = 1$ and $h(i)$ is not negative[12].

A useful feature of this observer model is the fact, that two observer states can be aggregated without the need to add every single observation of one state to the other. So, observer states of different sources of information can be aggregated easily into a general state:

Definition 11 (State Aggregation). *Two states* $Z^A = (\mathcal{I}, h^A, g^A)$ *and* $Z^B = (\mathcal{I}, h^B, g^B)$ *based on the same set of digital identities are* aggregated *to a new state* $Z^A \cup Z^B = (\mathcal{I}, h^C, g^C)$ *as follows:*

$$g^C = g^A + g^B \tag{4}$$

$$h^C : i \mapsto \frac{g^A h^A(i) + g^B h^B(i)}{g^C} \tag{5}$$

For a proof the correctness of state aggregation see [10].

5.3 Quantifying Anonymity

As described in Section 2, Shannon entropy [11] is often used as a metric for anonymity. Given an observer state Z, the Shannon entropy $H_\varnothing$ of an information b can be computed.

Definition 12 (Shannon entropy). *Let* b *be an observation and* $\mathcal{V}_b$ *a set of suspects related to observation* b*. The* Shannon entropy *of* b *related to a state* Z *is the Shannon entropy of the suspects* $\mathcal{V}_b$*.*

$$H_\varnothing = -\sum_{(v \in \mathcal{V}_b)} p(v|b) \log_2 p(v|b) \tag{6}$$

$$p(v|b) = \frac{p(v \wedge (\bigvee_{(w \in \mathcal{V}_b)} w))}{p(\bigvee_{(w \in \mathcal{V}_b)} w)} \tag{7}$$

$$= \frac{h(v)}{\sum_{(i \in \mathcal{V}_b)} h(i)} \tag{8}$$

Thereby, $h(i)$ *denotes the probability of the identity* i *within the observer state* Z*.*

[12] See [10] for the proof.

Given a Shannon entropy $|\mathcal{S}| = 2^{H_\varnothing}$ denotes the equivalent size of a uniformly distributed anonymity set $\mathcal{S}$.

I first evaluate the case that a user has only one digital identity. The Shannon entropy $H_\varnothing$ specifies the average amount of information needed in addition to b in order to uniquely identify a digital identity. In case of a user evaluating his anonymity, he usually knows his digital identity. So, it may be more useful for him to compute the amount of information needed to identify *him*, i.e., his digital identity. This so called "individual anonymity" can be computed as follows:

$$H(i) = \log_2 p(i|b) \tag{9}$$

From the viewpoint of a single user, *individual anonymity* is the most accurate anonymity measure.

In case a user has multiple digital identities, this measure can also be used, but before calculating entropy all suspect digital identities belonging to the same user need to be grouped into one "personal" digital identity. This grouping is done by summing up their probabilities. This grouping needs to be done by information considered to be sufficient to identify users (in the physical world.)[13] The entropy is then calculated based on the "personal" digital identities.

5.4 Quantifying Linkability of Actions

Regarding linkability, it is interesting for a user, to what extent it can be determined that actions have been done by the same user. More formally, there are two actions c_1 and c_2 which have been observed in the form of observations b_1 and b_2.

According to [5], linkability of items of interest can be measured regarding equivalence classes, for which (after observations) an attacker has partial knowledge about which items of interest belong to which class.

Applied to the model used here, the equivalence classes are the digital identities. By an observation of an action, suspect digital identities can be determined corresponding to the observation of this action (see Definition 8), i.e., information about association of items of interest (actions) to equivalence classes (digital identities) is gained.

Regarding observations b_1 and b_2, the suspect sets are $\mathcal{V}_{b_1}$ resp. $\mathcal{V}_{b_2}$. Within a set of suspects, a digital identity has the probability $p(v|b)$, which is derived from the current observer state as shown in equations (7) and (8).

The probability p_r, that actions c_1 and c_2 belong to the same digital identities, can be computed as follows:

$$p_r = \sum_{(v \in V_{b_1 \cup b_2})} p(v|b_1) \cdot p(v|b_2)$$

Thereby, $V_{b_1 \cup b_2}$ denotes the set of digital identities, which are contained in both sets V_{b_1} and V_{b_2}. According to [5], the probability $p_{\neg r}$, that the actions c_1 and c_2 do not belong to the same digital identity is $1 - p_r$.

[13] See also Section 4.

From probabilities p_r and $p_{\neg r}$ a degree of linkability d can be computed by using the Shannon entropy [5]:

$$d := H(p_r, p_{\neg r}) = -p_r \cdot \log_2 p_r - p_{\neg r} \cdot \log_2 p_{\neg r}$$

The degree of linkability d specifies, how much an observer has learnt about the relation between c_1 and c_2 from observations V_{b_1} and V_{b_2}, taking also into account the a-priori knowledge about the digital identities derived from the current observer state.

If $p_r > p_{\neg r}$, the degree denotes the certainty of the observer, that actions c_1 and c_2 *belong to the same digital identity*, otherwise it denotes the certainty of the observer that the actions do *not belong to the same digital identity*.

In case a user has only one digital identity, linkability related to a digital identity is the same as linkability related to a user. In case a user may have more than one digital identity, before actually calculating linkability the suspect digital identities belonging to the same user first need to be grouped into "personal" digital identities, as described in Section 5.3 for the same purpose. Then, the calculation of linkability can be performed as shown above, but based on the "personal" digital identities.

5.5 Incorporating Time

As described in Section 3.1, the above model does not consider changes of attribute values of users. But for a system more closely modelling the real world this is an important feature, because many attributes of users can be subject to change over time, e.g., the family name may be changed by marriage.

In order to also consider timely changes of attributes within a digital identity, I define the *dynamic digital identity* as follows:

Definition 13 (Dynamic Digital Identity). *A* dynamic digital identity *is a bundle of functions $f_{\mathcal{A}}(t) : f(t) \rightarrow a$ for each attribute $\mathcal{A}$. This means, that for each attribute a function exists, which determines the value of the attribute at a given point in time t.*

A (static) digital identity can be seen as a snapshot at a point in time t of a dynamic digital identity. At a given point in time, the digital identity of an entity comprises all information, which can be transmitted by the entity during an action, so (regarding data to be possibly communicated to other entities) an entity can be seen as an incarnation of a particular digital identity.

For an observer, this means that observations "grow older", i.e., an observation matches a set of digital identities only at the time of the observation. For the knowledge base of the observer, the observer state, this means that probabilities of digital identities need to be adjusted according to time by using the functions $f_{\mathcal{A}}(t)$ of the digital identity. In most cases, the observer will not fully know this function, so he needs to estimate it. This leads to growing uncertainty with regards to older information.

As this "ageing" does not change the structure of the observer state, quantification of anonymity and linkability of actions can be performed in the same way as described above for the static model.

So, in general the observer state model described in Section 5.2 can be enhanced to incorporate time. For use for anonymity and linkability quantification within a real system, specific time dependent functions need to be defined for the single attributes.

5.6 Linkability Quantification Service

Despite getting enough and reliable information[14], the major problem with linkability quantification is resource usage. Detailed quantification needs a major amount of storage as well as computing power, depending on the number of attributes and attribute values to consider. Especially if considering e.g., mobile phones or PDAs as devices of the user, both storage and computing power are very much limited. Even if calculation within the general observer model described above could be optimised to some extent, this will usually exceed resources available on such smaller devices.

A usual way to address such a problem would be to introduce a third party linkability quantification service (LQS), which could compute anonymity and linkability quantification on behalf of users.

In the following I will go in some more detail about intended functionality of such a service, and especially on privacy and security risks introduced by an LQS and possibilities to solve them.

Basic Functionality. Basically, the service has two functions:

- Answering user requests for linkability computation.
- Gathering base data needed for linkability computation, i.e. the data described in Section 5.1. This data can be aggregated to an observer state as described in Section 5.2.

The first function is processed in the following way:

Input: The user inputs a request for (pre-)computation of measurements of his anonymity or linkability of actions. Such a request consists of one or more sets of data to be disclosed or already disclosed, relative to which anonymity or linkability can be quantified.

Processing: Measurements are computed using the base data aggregated from sources of information.

Output: The service outputs the measurement results, together with an arbitrary set of details regarding the computation, e.g., a description of the sources of information used.

In general, a LQS is just a linkability quantification done by a TTP. For the linkability quantification calculation, in terms of input and output, it is rather straight forward to compute it remote instead of local, but as the TTP running the LQS is not under the user's control, security and privacy issues arise.

[14] See Section 5.1 for details.

Security and Privacy Issues. Without countermeasures, the user needs to trust in the LQS for the following reasons:

1. Users need to trust for correct computation of linkability measurements by the LQS.
2. Users need to trust, that the LQS does not disclose their action data to third parties. This is a rather important issue, as the LQS essentially aggregates the user profiles, which should be prevented from being known to the attacker. Additionally the LQS gets to know data, which a user will potentially not disclose to the attacker. It is only input to LQS for computing linkability measures for the potential case of a disclosure. So, the LQS is a major goal for attackers which want to get user profiles.

The first problem can be solved by using redundant LQS' in order to detect wrong computation by techniques known from research on fault tolerance, e.g., majority voting.

For solving the second problem technically, an approach would be needed, so that the LQS can compute linkability metrics without getting knowledge of the input data. Basically, a secure-function-evaluation[15]-like approach using multiple instances (i.e., no single instance alone has enough knowledge to reconstruct the user profiles) could help, but this is a rather theoretic approach, as this adds a huge amount of extra resource usage to the service.

Besides this, legal regulations could help to restrict misuse, but on the other hand, to avoid the need to utilise such regulations is a goal of linkability computation at the user's side.

So, even if such a service would be desirable to save resources at the user's device, more research needs to be done on possibilities to implement desired security features to it.

6 Summary

In this paper I describe a framework to quantify anonymity and linkability of actions for use within a privacy-enhancing identity management. An appropriate attacker model is defined, and an approach for computing such quantification based on observations of user's actions is proposed. Further, I discuss an enhancement to the basic approach regarding time dependency of observations. The problem of getting enough information to do the quantification is analysed. Regarding the problem of high resource consumption for quantification computations, I analyse possibilities for utilising third party services especially with respect to privacy and security requirements.

Further research needs to be done regarding optimising computations with respect to resource consumption and regarding matching functions for attribute values depending on attribute characteristics. Another topic for further research is how to secure privacy of PII when using third party services for quantification of anonymity and linkability.

[15] e.g. [12].

Acknowledgement. I want to thank Sandra Steinbrecher and Stefan Schiffner for valuable discussions and hints during creation of this work.

References

1. Clauß, S., Köhntopp, M.: Identity management and its support of multilateral security. Computer Networks, Special Issue on Electronic Business Systems (37) (2001) 205–219 Elsevier, North-Holland.
2. Chaum, D.: Untraceable electronic mail, return addresses, and digital pseudonyms. Communications of the ACM **4**(2) (1981) 84–88
3. Díaz, C., Seys, S., Claessens, J., Preneel, B.: Towards measuring anonymity. In Dingledine, R., Syverson, P., eds.: Proceedings of Privacy Enhancing Technologies Workshop (PET 2002). Number 2482 in LNCS, Springer-Verlag (2002) 54–68
4. Serjantov, A., Danezis, G.: Towards an information theoretic metric for anonymity. In Dingledine, R., Syverson, P., eds.: Proceedings of Privacy Enhancing Technologies Workshop (PET 2002). Number 2482 in LNCS, Springer-Verlag (2002) 41–53
5. Steinbrecher, S., Köpsell, S.: Modelling unlinkability. In Dingledine, R., ed.: Proceedings of Privacy Enhancing Technologies workshop (PET 2003). Number 2760 in LNCS, Springer-Verlag (2003) 32–47
6. Díaz, C., Claessens, J., Seys, S., Preneel, B.: Information theory and anonymity. In: Proceedings of the 23rd Symposium on Information Theory in the Benelux, May 29-31, 2002, Louvain la Neuve, Belgium, Werkgemeenschap voor Informatie en Communicatietheorie (2002)
7. Hughes, D., Shmatikov, V.: Information hiding, anonymity and privacy: A modular approach. Journal of Computer Security **12**(1) (2004) 3–36
8. Pfitzmann, A., Hansen, M.: Anonymity, unlinkability, unobservability, pseudonymity and identity management - a consolidated proposal for terminology. Version 0.27 at `http://dud.inf.tu-dresden.de/literatur/Anon_Terminology_v0.27.pdf` (2005) Version 0.8 in: Hannes Federrath (Ed.): Designing Privacy Enhancing Technologies; Proc. Workshop on Design Issues in Anonymity and Unobservability; LNCS 2009; 2001; 1-9.
9. Orwell, G.: Nineteen Eighty-Four. Martin Secker & Warburg (1949)
10. Clauß, S., Schiffner, S.: Anonymität auf Anwendungsebene. In Dittmann, J., ed.: Proceedings of Sicherheit 2006. Volume P-77 of Lecture Notes in Informatics., Bonn, GI (2006) 171–182 (german).
11. Shannon, C.: A mathematical theory of communication. The Bell System Technical Journal **27** (1948) 379–423
12. Micali, S., Rogaway, P.: Secure computation. In: Crypto '91. Number 576 in LNCS, Springer Verlag (1992) 392–404