# Modelling of Pseudonymity under Probabilistic Linkability Attacks

Martin Neubauer
Institute of Communication Networks and Computer Engineering (IKR)
Universität Stuttgart
Pfaffenwaldring 47, 70569 Stuttgart, Germany
Email: martin.neubauer@ikr.uni-stuttgart.de

***Abstract*—This paper contributes to the field of measuring (un)linkability in communication systems; a subproblem of privacy protection. We propose an attacker state model for attacks on unlinkability of partial identities named linkability graph. It covers probabilistic linkability attacks based on heterogeneous and time-variant characteristics. From our model, we derive linkability measures and argue prospects for safeguard design. Our model reduces space and time complexity compared to other contributions in literature. This enables simulative privacy analysis of complex context-aware systems that employ multiple partial identities per user.**

## I. Introduction

Recent incidences in Germany, such as surveillance of supermarket cashiers [1] and public transportation employees [2], as well as acquisition of personal data for tax evader prosecution by the government [3], foster awareness of privacy issues. It becomes evident that benefits of digital data processing (ease of storage, transfer and reproduction) are disadvantages for privacy. For example, it is profitable to offer marginal discounts for disclosure of personal data and profiles, which are then subject to uncontrollable circulation.

With the advent of context-awareness [4], [5] and ubiquitous computing [6], [7], the situation becomes more critical. Context-awareness entails extensive data collection, often without any restrictions on the manifold purposes of use. For example, purposes of use may be only partially known or specified at the point in time data is collected.

One privacy safeguard is pseudonymisation. Instead of representing users by their fully-fledged true identity, they are represented by partial identities. This means that users own multiple identifiers (pseudonyms), each equipped with a subset of personal data [8]. The pseudonym refers to the user, the personal data subset and undertaken actions. This approach intentionally compartmentalises the user's true identity into multiple partial identities by splitting personal data into subsets and separation of action traces. However, this is only effective as long as this intentional separation is not reversed by an attacker. The reversal is known as the linkability threat.

Privacy-agnostic system design can enable linkability. For example, if globally unique identifiers, such as device addresses (e. g., MAC- or IP-addresses) are required. These allow, e. g. linking of web sessions originating from the same IP-address by both the addressed web server and outside attackers. Anonymity safeguards, e. g. MIX-networks [9], [10], the reduced latency version Web MIXes [11] and Onion-Routing [12] help against outside attackers. They prevent linkability due to globally unique identifiers and observable behaviour patterns by either sender, recipient or relationship anonymity/unobservability [8].

Linkability can also result from voluntary data disclosure, e. g. from data of overlapping Information Cards [13] and tracked behaviour patterns. Today, voluntarily contacted correspondent nodes already create user profiles from such data, e. g. for the purpose of targeted advertising. This challenges the partial identities approach. First, partial identities are prerequisites for privacy but are no guarantee in the presence of an attacking correspondent node. Second, linkability quantification depends on the attacker's knowledge, capabilities and on the characteristics of data. For example, on the range of values or how many users have some value in common.

We model linkability caused by voluntary data disclosure to correspondent nodes in a context-aware system. The correspondent nodes are attackers. Our model enables simulative analysis of privacy protection by partial identities under probabilistic linkability attacks. Therefore, it focuses on the attacked users. We take a viewpoint on the system that allows for comparison of probabilistically linked partial identities and the truth. This viewpoint enables quantification of what a particular attack means to users' privacy. Further, it enables identification of possibilities how users can influence or even control linkability. We neither aim for comparing attack strategies nor finding the best attack.

This paper is structured as follows. Section II introduces the attacker model and the linkability graph as our model for the attacker state. It shows the evaluation to linkability measures and concludes with a discussion of properties. Section III applies it to the multiple partial identities approach and exemplifies how linkability graphs can be obtained and evaluated in practice. Section IV discusses relations to previous work. Section V concludes the paper and exhibits open issues.

## II. Measuring Linkability

### A. Attacker Model

We adopt the term *item of interest* (IOI) from [8] to refer to observed elements. The characteristics of IOIs form the basis for any attack. Furthermore, we adopt the term *subject* [8] as a representative for a possibly acting entity, such as a user.

The attacker's aim is to cluster IOIs according to ownerships. This is, identify items owned by one subject and distinguish them from items of other subjects. We note that this does not imply identification of any subject as such, which would mean identification in real world. Instead, the aim is to find IOIs owned by the same, yet unidentified subject.

We assume that observed characteristics are not necessarily unique for some IOI or subject. Hence, any attack on unlinkability of IOIs is probabilistic, i. e. that equal characteristics do not guarantee that IOIs are owned by the same subject. This is addressed in detail in Section II-B.

It is useful to assume the hypothetical omnipresent attacker in a worst-case privacy analysis. Although this overestimates the attacker's abilities in realistic settings, we can assess whether safeguards are effective. If a safeguard is effective in the analysis, it is concluded that it is in realistic settings as well. However, if the analysis shows that a safeguard is ineffective, no conclusion is drawn for realistic settings.

We model linkability caused by voluntary data disclosure. We deny the unrealistic omnipresent outsider and postulate a local attacker who operates correspondent nodes. Subjects voluntarily interact with correspondent nodes, such as web servers of a service provider. The implications are:

1) If measures indicate an unlinkability breach, we know, as in the omnipresent attacker case, that the safeguard is ineffective.
2) If measures do not indicate an unlinkability breach, we cannot assess the maximum strength of an attacker against which the safeguard, in particular the multiple partial identities approach, protects. Hence, we yield a lower bound only.
3) Any local attacker may have partial knowledge about the system. For example, the attacker has no a priori knowledge about the total number of subjects and may observe a subset of IOIs only. We therefore adopt the open world assumption. This means that observations can argue for and against IOIs being linked or they do not allow any conclusion.

Without loss of generality, we henceforth focus on a single, passive correspondent node attacker without any collaboration, such as voluntarily visited web servers of one service provider. However, our model can be extended to collaborating and active attackers.

We assume that the attacker has "unlimited" computing power and storage capacity, i. e. he can execute complex algorithms in short periods of time and can store huge amounts of data. Thus, our approach does not depend on any restrictions regarding these capabilities.

We assume that state-of-the-art cryptography with correct treatment of key material is secure. Use of cryptography influences the amount of available cleartext characteristics. However, the attack principle, the attacker state model and the derived linkability measures are not influenced. Hence, in the following cryptography is not considered.

### B. Linkability Graph

In the following, we look at an attacker $\alpha$ who mounts a particular linkability attack $\mathcal{A}$ in a communication system. The users of the communication system perform actions, e. g. they interact with other users, consume services by means of partial identities or send and retrieve messages. They are the considered subjects. Dependent on the attacker's focus, the IOIs may include received messages, service consumption records or partial identities. For sake of clarity, we assume that only one type of IOIs is considered in any analysis, e. g. only partial identities.

Characteristics of IOIs may include sender's IP-address, email-address or gender. We assume that the universe of characteristics is constant and known. We do not mandate that all characteristics apply to all IOIs. For example, one message may contain the sender's email-address while another one does not. Consequently, we regard the attack algorithm to conclude about an IOI-pair as specific to the subset of common characteristics of these IOIs. We assume that the attacker cannot conclude anything from an empty set of common characteristics, i. e. he is forced to ignore such observations.

Furthermore, we assume that each IOI has one owner. For example, if the attacker aims for identifying messages of the same sender, the owner of a message is its sender. Although the following analysis relies on unambiguous enumeration of subjects, it otherwise imposes no further requirements on the actual owner representation.

Despite uniqueness of true ownerships, the attacker can be uncertain about any individual conclusion, e. g. due to ambiguity of observed characteristics. As a result, he obtains probabilistically linked IOIs. We do not postulate transitivity of these probabilistic links. We model the attacker's guess about existence of a link and the imposed ignorance by continuous variables. This is inspired by approaches to uncertainty in [14].

The attacker concludes about links over time. Hence, his a posteriori state is the accumulation of educated guesses about linked and unlinked IOIs, and expresses right and wrong conclusions, as well as ignored observations.

In the following, we focus on attacker $\alpha$ and model his state after attack $\mathcal{A}$. We assume that the nonempty universe of IOIs $I$ and an index set $S$ for the nonempty universe of subjects are known. This is required for quantification of linkability, but we do not assume that an attacker has this knowledge. We require that each subject of $S$ has at least one IOI in $I$. If this is not fulfilled in real settings, subjects without IOI must be removed from $S$ before applying our model. Furthermore, we define links to be pairs of IOIs. The set of all possible undirected[1] links is

$$L = \{l \in \mathbf{N} \mid l \leftrightarrow (p,q) = (q,p) \ : \ p,q \in I\} \quad . \qquad (1)$$

Let $G^\alpha = (I^\alpha, L^\alpha)$ be a loop-free undirected graph with vertices $I^\alpha \subseteq I$ and links $L^\alpha \subseteq L$. Graph $G^\alpha$ is the linkability graph for attacker $\alpha$ and attack $\mathcal{A}$.

[1] linkability is symmetric, i. e. $x$ linked to $y$ if and only if $y$ linked to $x$

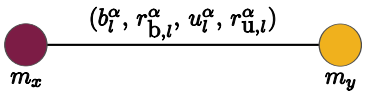



Fig. 1. Graphical representation of the attacker state for one link

The attacker observes $I^\alpha$ and aims for identifying ownerships as introduced in Section II-A. The attacker's speculation about ownerships is denoted by the mapping

$$\alpha : L^\alpha \mapsto (b_l^\alpha, r_{\mathrm{b},l}^\alpha, u_l^\alpha, r_{\mathrm{u},l}^\alpha) \quad \text{with} \quad l \in L^\alpha \ . \tag{2}$$

Thereby, the four variables represent the attacker's a posteriori state. The variables are specific to the attacker and attack. Their values vary over time due to made observations and drawn conclusions.

Each tuple covers two aspects. First, $b_l^\alpha$ and $r_{\mathrm{b},l}^\alpha$ capture the attacker's state with respect to drawn conclusions. This is, the attacker's belief $b_l^\alpha$ that the IOIs of link $l$ have the same owner and the obtained reliance[2] $r_{\mathrm{b},l}^\alpha$ in this belief. Second, $u_l^\alpha$ and $r_{\mathrm{u},l}^\alpha$ capture the imposed ignorance. They address observations which the attacker cannot exploit by attack $\mathcal{A}$. The attacker's ignorance for link $l$ is $u_l^\alpha$ and the obtained reliance is $r_{\mathrm{u},l}^\alpha$. The codomains are $b_l^\alpha, u_l^\alpha \in [0,1]$, $r_{\mathrm{b},l}^\alpha \in [0, \widehat{r}_{\mathrm{b}}^\alpha]$ and $r_{\mathrm{u},l}^\alpha \in [0, \widehat{r}_{\mathrm{u}}^\alpha]$ with $\widehat{r}_{\mathrm{b}}^\alpha, \widehat{r}_{\mathrm{u}}^\alpha \in (0, \infty)$.

The greater the attacker's belief in existence of link $l$, the greater is $b_l^\alpha$. The more evidence gathered and conclusions drawn supporting the belief, the greater is the reliance $r_{\mathrm{b},l}^\alpha$. Contrary, the more observations ignored, the greater is $u_l^\alpha$. Nevertheless, the more observations available, the greater is the reliance $r_{\mathrm{u},l}^\alpha$.

Figure 1 shows the graphical representation of the attacker's state for one link. In a realistic setting, the depicted vertices $m_x, m_y \in I$ may represent two sent messages. The true ownerships are represented by the vertex fillings. Identical vertex filling refers to identical owner, i. e. the two IOIs in Fig. 1 have different owners. However, we remind that this information is not part of the attacker state.

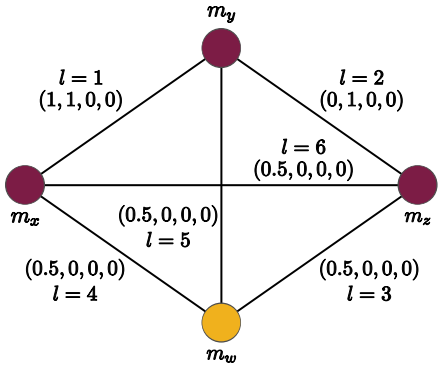
Any a priori knowledge about ownerships an attacker has can be taken into account by an appropriate initialisation of belief and reliance values on a per link basis. For an attacker without any a priori knowledge about ownerships we define the initial state to

$$\begin{aligned} &L^\alpha = \emptyset \qquad \text{and} \\ &(b_l^\alpha = 0.5,\ r_{\mathrm{b},l}^\alpha = 0,\ u_l^\alpha = 0,\ r_{\mathrm{u},l}^\alpha = 0) \quad \forall l \in L \ . \end{aligned} \tag{3}$$

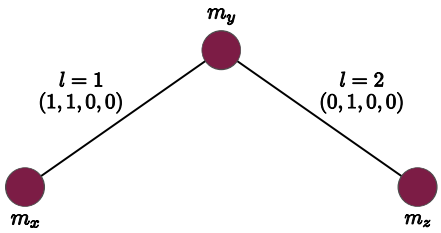
This means that the attacker is unable to decide whether links exist ($b_l^\alpha = 0.5$) and that this result is completely unreliable ($r_{\mathrm{b},l}^\alpha = 0$). Furthermore, he ignores no observation ($u_l^\alpha = 0$) and this result is completely unreliable ($r_{\mathrm{u},l}^\alpha = 0$), because no observation is made at all.

Although we define one representation for the absence of a priori knowledge, the linkability measures according to Section II-C accept any representation for which $b_l^\alpha \cdot r_{\mathrm{b},l}^\alpha = 0$ and $u_l^\alpha \cdot r_{\mathrm{u},l}^\alpha = 0$ hold true.

[2]also: certainty or confidence in correctness of a value




(a) elaborate version




(b) concise version

Fig. 2. Example linkability graph for the attacker's a posteriori state

Figure 2 depicts two equivalent graphical representations of the linkability graph for a simple example. We assume $\widehat{r}_{\mathrm{b}}^\alpha = 2$. Figure 2(a) shows two subjects distinguished by vertex fillings. The first subject has sent the three messages $m_x, m_y$ and $m_z$. The second has sent $m_w$ only. In the example, the attacker believes with $b_1^\alpha = 1$ and reliance $r_1^\alpha = 1$ that $m_x$ and $m_y$ have the same owner. For $m_y$ and $m_z$, he believes with $b_2^\alpha = 0$ and reliance $r_1^\alpha = 1$, i. e. that they are unlinked. For example, the common characteristic of $m_y$ and $m_z$ differs in its value and the attacker interprets it as evidence for different owners. All other links are unobserved. For example, $m_w$ may have no common characteristic with the other messages and $m_x$, $m_z$ may be only observed during different periods of time.

The elaborate version Fig. 2(a) explicates the a posteriori states even for unobserved links ($l = 3, 4, 5, 6$). This can impact clarity in larger settings. Hence, the concise version Fig. 2(b) omits unobserved IOIs and links. It implicitly refers to the initial values according to (3). We henceforward prefer the concise version.

### C. Linkability Measures

The linkability graph from Section II-B captures the attacker's state for each IOI-pair. This is our basis to quantify the breach of unlinkability per subject. However, the unprocessed linkability graph suffers from the following drawbacks.

1) It is hard to grasp the breach of unlinkability per subject from a large linkability graph. Thus, it is hard to compare the breach of unlinkability for different subjects.
2) Comparison of unprocessed linkability graphs can become costly. Thus, comparison of results for different settings can become costly.

3) The linkability graph captures beliefs in links but not any resulting impact. In other words, all links are considered identical in breach of privacy, irrespective of the characteristics of linked IOIs. To obtain privacy risk estimates, linkability and impact must be considered. We do not examine this further.
4) The linkability graph does not help in defining any threshold as to when unlinkability has to be regarded as broken. However, the linkability graph is not special in this respect. Other models suffer from this issue as well. This is left for further study.

To address the first two drawbacks, we need few, concise measures that represent the linkability graph and the resulting breach of unlinkability per subject. In the following, we derive such linkability measures.

We now take a global outsider perspective with knowledge about true ownerships and the particular attack $\mathcal{A}$. However, we do not assume that this is known to any subject. This perspective enables to assess the true breach of unlinkability per subject caused by attack $\mathcal{A}$. We start by formalisation of the true ownerships and proceed to the evaluation of the linkability graph.

Let $\sim_r$ be an equivalence relation on $I$ that formalises the relation 'owned by same subject'. This equivalence relation partitions $I$ and yields the equivalence classes

$$I_{\sim_r} = I \,/\, \sim_r = \{I_1, I_2, ..., I_k\} \quad .$$

There are $|S|$ subjects and also $k = |S|$ equivalence classes. The equivalence class $I_j$ contains those IOIs which are truly owned by subject $j \in S$. Further, for two IOIs $p$ and $q$ we denote $p \equiv q$ if $p, q \in I_j$, i. e. they are truly owned by the same subject $j$. Otherwise we denote $p \not\equiv q$.

It is known that an equivalence relation $\sim_r$ on $I$ is also an equivalence relation on any subset of $I$. Hence, the equivalence relation $\sim_r$ also partitions $I^\alpha$ with respect to true ownerships. The attacker can at best draw correct conclusions about $|I^\alpha / \sim_r|$ subjects. However, as we postulate that the attacker has no a priori knowledge about $S$, wrong conclusions can yield an arbitrary number of supposed subjects up to $|I^\alpha|$.

We now classify links into two subject specific classes. First, cohesive links relate IOIs of one subject. They reinforce correct clustering of IOIs. Second, adhesive links relate IOIs of different subjects. They increase similarity of subjects, and thus reduce certainty about ownerships. In the following, we formalise both classes.

For subject $j \in S$, the set of observed cohesive links is

$$L^\alpha_{\text{co},j} = \{l \in L^\alpha | l \leftrightarrow (p, q) : p \equiv q \ \wedge \ p \in I_j\}$$

and the set of observed adhesive links is

$$L^\alpha_{\text{ad},j} = \{l \in L^\alpha | l \leftrightarrow (p, q) : p \not\equiv q \ \wedge \ (p \in I_j \vee q \in I_j)\} \quad .$$

These sets, in combination with the mapping $\alpha$, describe what the attacker obtained about subject $j$. Hence, linkability measures can focus on these sets.

We exploit the distinction of links and define the linkability measures

$$F^\alpha_{\text{co}}(j) = \frac{2}{|I_j| \cdot (|I_j| + 1)} \cdot \left( |I^\alpha_j| + \sum_{l \in L^\alpha_{\text{co},j}} b^\alpha_l \cdot \frac{r^\alpha_{\text{b},l}}{\widehat{r}^\alpha_{\text{b}}} \right)$$

and

$$F^\alpha_{\text{ad}}(j) = \frac{1}{|I_j| \cdot (|I| - |I_j|)} \cdot \sum_{l \in L^\alpha_{\text{ad},j}} b^\alpha_l \cdot \frac{r^\alpha_{\text{b},l}}{\widehat{r}^\alpha_{\text{b}}}$$

with codomain $F^\alpha_{\text{co}}(j), F^\alpha_{\text{ad}}(j) \in [0, 1]$. Further, we define their difference

$$F^\alpha_\Delta(j) = F^\alpha_{\text{co}}(j) - F^\alpha_{\text{ad}}(j) \tag{4}$$

with codomain $F^\alpha_\Delta(j) \in [-1, 1]$.

The linkability measure $F^\alpha_{\text{co}}(j)$ is the proportion of fully and reliably believed cohesive links, i. e. each belief is weighted with its normalised reliance. It captures the amount of correct conclusions. The maximum is reached if all possible cohesive links are actually observed and reliably believed. We note that observability of subject $j$'s IOIs is itself a threat to unlinkability. This means that observability is a prerequisite for linkability. Therefore, the number of observed IOIs $|I^\alpha_j|$ is explicitly considered by $F^\alpha_{\text{co}}(j)$.

The linkability measure $F^\alpha_{\text{ad}}(j)$ is the proportion of fully and reliably believed adhesive links. It captures the amount of wrong conclusions. The maximum is reached if all possible adhesive links are actually observed and reliably believed.

Additionally, we define the measures for imposed ignorance

$$U^\alpha_{\text{co}}(j) = \frac{1}{|L^\alpha_{\text{co},j}|} \cdot \sum_{l \in L^\alpha_{\text{co},j}} u^\alpha_l \cdot \frac{r^\alpha_{\text{u},l}}{\widehat{r}^\alpha_{\text{u}}}$$

and

$$U^\alpha_{\text{ad}}(j) = \frac{1}{|L^\alpha_{\text{ad},j}|} \cdot \sum_{l \in L^\alpha_{\text{ad},j}} u^\alpha_l \cdot \frac{r^\alpha_{\text{u},l}}{\widehat{r}^\alpha_{\text{u}}}$$

with codomain $U^\alpha_{\text{co}}(j), U^\alpha_{\text{ad}}(j) \in [0, 1]$. These are the proportions of ignored observations for cohesive and adhesive links respectively. They reach their maximum if all observations are ignored (for causes refer to Section II-B).

By these measures we exploit the fact that any unobserved link $l \in L \setminus L^\alpha$ has initial values according to (3), which yield $b^\alpha_l \cdot r^\alpha_{\text{b},l} = 0$ and $u^\alpha_l \cdot r^\alpha_{\text{u},l} = 0$. Thus, an unobserved link neither requires explicit representation in the linkability graph nor consideration in any measure.

### D. Discussion

We first interpret the linkability measures from Section II-C and discuss their properties. Then we address prospects opened by our model in terms of safeguard design.

*1) Equivalence of Linkability Graphs:* Our linkability measures imply the following equivalence, which is irrespective of how the attacker states are reached. In other words, whether different attackers reach equivalent states by mounting the same attack $\mathcal{A}$ or by different attacks does not matter. Instead, the equivalence compares the completeness and correctness of probabilistically concluded links.

TABLE I
EXAMPLE A POSTERIORI STATES OF LINKS FOR ATTACKER $\alpha$ AND $\beta$

| link number | attacker $\alpha$ | attacker $\beta$ |
|---|---|---|
| 1 | $(1,\ 1,\ \ldots)$ | $(0.5,\ 2,\ \ldots)$ |
| 2 | $(0,\ 1,\ \ldots)$ | $(0.5,\ 2,\ \ldots)$ |

We assume that the a posteriori state of attacker $\alpha$ is known. The state of another attacker $\beta$ is equivalent with respect to the imposed ignorance for subject $j$ if

$$U^{\alpha}_{\text{co}}(j) = U^{\beta}_{\text{co}}(j) \qquad \text{and} \qquad U^{\alpha}_{\text{ad}}(j) = U^{\beta}_{\text{ad}}(j) \ .$$

Analogously, the state of attacker $\beta$ is equivalent to attacker $\alpha$ with respect to breach of unlinkability for subject $j$ if

$$F^{\alpha}_{\text{co}}(j) = F^{\beta}_{\text{co}}(j) \qquad \text{and} \qquad F^{\alpha}_{\text{ad}}(j) = F^{\beta}_{\text{ad}}(j) \ .$$

We further analyse this equivalence. For simplicity we assume that both attackers observe identical amounts of IOIs, i. e. $|I^{\alpha}_j| = |I^{\beta}_j|$. Any attacker $\beta$ is equivalent to attacker $\alpha$ with respect to breach of unlinkability for subject $j$ if

$$\frac{1}{\widehat{r}^{\alpha}_{\text{b}}} \cdot \sum_{l \in L^{\alpha}_j} b^{\alpha}_l \cdot r^{\alpha}_{\text{b},l} \stackrel{!}{=} \frac{1}{\widehat{r}^{\beta}_{\text{b}}} \cdot \sum_{l \in L^{\beta}_j} b^{\beta}_l \cdot r^{\beta}_{\text{b},l} \ . \tag{5}$$

We now revisit the a posteriori attacker state of the example from Fig. 2(b), where $\widehat{r}^{\alpha}_{\text{b}} = 2$. We focus on beliefs and their reliance and omit ignorance variables. We assume, for sake of clarity, that the sets of observed IOIs and links are identical, i. e. $I^{\alpha}_j = I^{\beta}_j$ and $L^{\alpha}_j = L^{\beta}_j$. For attacker $\beta$ we assume $\widehat{r}^{\beta}_{\text{b}} = 4$ and the a posteriori states according to Table I.
Obviously, for these links (5) holds:

$$\frac{1}{2} \cdot (1 \cdot 1 + 0 \cdot 1) \stackrel{!}{=} \frac{1}{4} \cdot (0.5 \cdot 2 + 0.5 \cdot 2) \ .$$

This allows for the following interpretation: full belief in the first link paired with full denial of the second link is equivalent to two links for which the attacker is unable to decide about their existence. This equivalence is intrinsic to our linkability measures and should be kept in mind.

*2) Lower and Upper Bounds of Linkability Measures:* In addition to the definition of equivalent attacker states, the linkability measures use well-defined extreme cases as references. On the one hand, the minimum of $F^{\alpha}_{\text{co}}(j)$ and $F^{\alpha}_{\text{ad}}(j)$ is reached if and only if the attacker knows nothing about subject $j$. Formally this means $I^{\alpha}_j = \emptyset$ and $L^{\alpha}_j = \emptyset$. On the other hand, the maximum of $F^{\alpha}_{\text{co}}(j)$ is reached if and only if all possible cohesive links $L_{\text{co},j}$ are concluded with

$$b^{\alpha}_l = 1 \ , \ r^{\alpha}_{\text{b},l} = \widehat{r}^{\alpha}_{\text{b}} \quad \forall l \in L_{\text{co},j} \ .$$

This is, there are no cohesive links left unobserved and all cohesive links are fully and reliably believed. Analogously, the maximum of $F^{\alpha}_{\text{ad}}(j)$ is reached if and only if all possible adhesive links $L_{\text{ad},j}$ are concluded with

$$b^{\alpha}_l = 1 \ , \ r^{\alpha}_{\text{b},l} = \widehat{r}^{\alpha}_{\text{b}} \quad \forall l \in L_{\text{ad},j} \ .$$

*3) Practicability:* Although linkability quantification is our focus, we should not forget about practicability. The linkability graph from Section II-B is based on dynamic characteristics of IOIs. This implies that for any simulative analysis of, for example, already deployed systems, models or prototypes of future systems, the linkability graph must be manageable with today's technology. Hence, practicability considerations are important. We consider space complexity of the linkability graph and both space and time complexity of the evaluation to our linkability measures.

According to (2), attacker $\alpha$ associates four variables to any link. The set of all possible links is defined by (1). The maximum cardinality is reached for a complete linkability graph. Links are undirected and there is only one link for any IOI-pair. Thus, the space complexity of a linkability graph is quadratic in the number of IOIs.

The evaluation according to Section II-C operates on the four variables associated to any observed link. Thus, the number of additions and multiplications scales with the number of links. In the worst-case, $L^{\alpha}$ is identical to $L$. As a result, the time complexity of the evaluation is quadratic in the number of IOIs. We evaluate the linkability graph for all subjects $S$ concurrently and thereby substitute time by space complexity. Thus, the space complexity is linear in the number of subjects.

Given today's technology, these complexities are manageable. We therefore regard our linkability graph and linkability measures as practical for simulative studies, especially considering the application to partial identities (see Section III).

*4) Prospects for Safeguard Design:* The introduced classification into cohesive and adhesive links in combination with the attacker model presents four unlinkability safeguards. First, a subject can target cohesive links and either reduces their pure number or influences the attacker's belief in their existence. Second, a subject can target adhesive links and either influence their pure number or the attacker's belief in their existence. Third and fourth, a subject can target the ignorance imposed on the attacker with respect to cohesive and adhesive links respectively. This means, a subject avoids disclosure of IOIs which allow any kind of conclusion by attacker $\alpha$. However, four reflections must be noted.

First, adhesive links are not under control of any single subject. Instead, they are influenced by many subjects, which are typically unknown to each other. For example, messages sent by different subjects, all containing identical date of birth, can cause adhesive links. A subject will hardly prevent others from sending such messages. This argument applies to manipulation of both belief and ignorance. Hence, we assume that influence of a subject is negligible unless subjects extensively cooperate in this respect. Such defensive alliances necessitate exchange of characteristics of IOIs, which is contrary to privacy protection aims.

Second, it is illusionary to assume that subjects will correctly guess properties of attack $\mathcal{A}$, as required for effective avoidance of conclusions. For example, a subject needs to guess which characteristics are exploited by the attack and how. Here, a worst-case assumption would help. This means,

finding the best attack on unlinkability that could be mounted. We do not further investigate in this direction.

Third, cohesive links are solely influenced by the considered subject itself and provide a starting point for safeguard design. Thus, influencing the number of cohesive links, the belief in their existence or the imposed ignorance are options. For the last, the unknown properties of attack $\mathcal{A}$ limit the effectiveness. Yet, a subject may exhaustively enumerate the characteristics of its IOIs and act according to worst-case assumptions.

Fourth, assume an extreme case where one subject, aware of linkability issues, succeeds in hiding all of its cohesive links. All other subjects are unaware. Due to unawareness, we assume that the other subjects' cohesive links exist with large beliefs, while adhesive links are generally low in their beliefs. It becomes clear that the linkability aware subject did not gain anything. The attacker may, by inversion, conclude that IOIs with only low-believed adhesive links are owned by one subject, which is an optimistic assumption for clarity of the issue. Thus, the relation between cohesive and adhesive links matters, which is expressed by (4). Consequently, IOIs of a subject are protected if either of the following cases applies. First, the attacker failed to conclude anything, i. e. both linkability measures are zero. Second, the attacker concluded about links but still fails in correct distinction of ownerships. In both cases $F^{\alpha}_{\Delta}(j)$ is at best zero. As rule of thumb: The smaller $|F^{\alpha}_{\Delta}(j)|$, the less distinguishable are IOIs of subject $j$ from others, and the better protected is subject $j$.

These safeguards are substantiated by our attacker state model and linkability measures. However, we leave actual design and realisation for further study.

## III. Linkability of Partial Identities

We illustrate the linkability graph by application to a service platform implementing an Identity Management System, which offers multiple partial identities as a privacy safeguard. The consumers of the service platform are privacy-aware users. Consumers correspond to subjects and their partial identities to IOIs. Each partial identity includes a set of data. Available services are offered by service providers. A service is described by the mandatory data for its function. Consumers decide for any service consumption which partial identity to use. Any chosen partial identity must at least provision the mandatory data of the addressed service.

We aim to evaluate the influence of chosen partial identities on linkability. We assume that service providers may be attackers. This conforms to Section II-A and [15]. Partial identities have unique identifiers that allow for charging. Thus, consumptions using the same partial identity are linkable at the provider. However, they would be initially unlinkable if different partial identities were used.

We now provide an example attack $\mathcal{A}$ similar to [15], which yields a concrete linkability graph. We consider a scenario where service providers limit their efforts for attacks. Precisely, no retrieved data is stored after session end, although our attacker model from Section II-A would allow. Only concurrently active partial identities $I^{\alpha}_{\text{active}}$ are subject to linkability attacks by means of regular retrieved data, i. e. that no access control is bypassed.

To model the attacker $\alpha$, we must compute the four continuous variables according to Section II-B. We interpret them as empirical probabilities and derive them from three directly managed integers. For each link we count the total number of observations $v_l$, conclusions $c_l$ and positive conclusions $e^+_l$. The initial values are $v_l = c_l = e^+_l = 0$. We define the variables to

$$b^{\alpha}_{l} = \begin{cases} \frac{e^+_l}{c_l} & \text{if } c_l \neq 0 \\ 0 & \text{otherwise} \end{cases}$$

$$r^{\alpha}_{\mathrm{b},l} = \widehat{r}^{\alpha}_{\mathrm{b}} \cdot \frac{f^{\alpha}_{\mathrm{b}} \cdot c_l}{1 + f^{\alpha}_{\mathrm{b}} \cdot c_l}$$

$$u^{\alpha}_{l} = \begin{cases} \frac{v_l - c_l}{v_l} & \text{if } v_l \neq 0 \\ 0 & \text{otherwise} \end{cases}$$

$$r^{\alpha}_{\mathrm{u},l} = \widehat{r}^{\alpha}_{\mathrm{u}} \cdot \frac{f^{\alpha}_{\mathrm{u}} \cdot v_l}{1 + f^{\alpha}_{\mathrm{u}} \cdot v_l}$$

with $0 < f^{\alpha}_{\mathrm{b}}, f^{\alpha}_{\mathrm{u}} \in \mathbf{R}$, which influence convergence to the state 'fully reliable'. For our setting we assume $f^{\alpha}_{\mathrm{b}} = f^{\alpha}_{\mathrm{u}} = 1$ and $\widehat{r}^{\alpha}_{\mathrm{b}} = \widehat{r}^{\alpha}_{\mathrm{u}} = 1$.

We assume that any data is constant over time. Further, by the end of any session, all data accessible via a partial identity is actually retrieved. Given these conditions, the linkability graph requires update only in case a session is either set up or released. We focus on session setup for an arbitrary partial identity $p$ and collapse access to all data to the very start of the session. The update procedure is depicted in Fig. 3.

The time complexity of our example linkability graph update is linear in the number of active sessions. This is due to our assumption that any data is constant. In presence of time-variant data, time complexity is at most quadratic in the number of active sessions.

We have validated our proposal by an implementation in Java. It uses probability distributions to construct partial identities and services. It applies event-driven simulation to mimic service consumption behaviour, which can cause linkability of partial identities.

We have executed this simulation on state-of-the-art hardware for 1,000 consumers, a total of 10,000 partial identities and 10 service providers, all attacking concurrently. Thus, each linkability graph has a maximum size of approx. 50 million links. In our implementation each link requires three 32 bit integers and thus 12 bytes of memory. Hence, in the worst-case the simulation requires approx. 575 MB of memory per linkability graph and a total of approx. 6 GB. Due to statistical properties of the configuration, total memory usage is below 1 GB and thus much better than in the worst-case.

A detailed explanation of our simulator is out of scope of this paper. Furthermore, actual evaluations based on our linkability graph and measures are for further study. Nonetheless, the above example attack and memory requirements indicate the practicability of our approach for scenarios with a meaningful number of partial identities.

**Input**: $p \in I$
**Data**: $I^\alpha, L^\alpha, \{(v_l, c_l, e_l^+) | l \in L^\alpha\}, I^\alpha_{\text{active}}$

$P = I^\alpha_{\text{active}}$ ;
**while** $P \neq \emptyset$ **do**
  select $q \in P$ ;
  $P = P \setminus \{q\}$ ;
  **if** $q = p$ **then** continue ; /* skip loops */
  $l \leftrightarrow (p, q) = (q, p)$ ; /* get link number */
  $v_l = v_l + 1$ ;
  $A = q \cap p$ ;
  **if** $A = \emptyset$ **then** continue ; /* incomparable */
  $c_l = c_l + 1$ ; /* inc. conclusion count */
  $r = \bot$ ; /* assume unlinked */
  **while** $r = \bot \wedge A \neq \emptyset$ **do**
    select $a \in A$ ;
    $A = A \setminus \{a\}$ ;
    /* check for identical values */
    **if** $v(p, a) = v(q, a)$ **then** $r = \top$
  **end**
  /* cond. inc. positive evidence */
  **if** $r = \top$ **then** $e_l^+ = e_l^+ + 1$
**end**
$I^\alpha_{\text{active}} = I^\alpha_{\text{active}} \cup \{p\}$ ;
$I^\alpha = I^\alpha \cup \{p\}$ ;

Fig. 3. Linkability graph update for example attack

## IV. Related Work

Privacy studies [16], [17], [18], [19] started with focus on subject identification, i. e. identifying a subject in a known set of subjects. These studies model the attacker's a posteriori state by a probability density function (PDF) on the known subject set and use information theory [20] to quantify anonymity. They establish the 'degree of anonymity' as the quotient of the a posteriori entropy and the maximum entropy. Despite their success in capturing identifiability, quantifying (un)linkability is left open. The successive studies [21], [15], [22], [23], [24] take up information theory and apply it to models for unlinkability. They use an appropriately defined equivalence relation as a representation for an attack. Any equivalence relation partitions the considered universe of IOIs. In case of a successful attack, the resulting partition corresponds to the truth. However, the attacker may conclude a wrong partition due to lack of correct information. In essence, each partition is concluded with a certain probability, which yields a PDF on all possible partitions. The PDF is evaluated to the 'degree of unlinkability' analogous to the 'degree of anonymity'. These analytical models imply the following limitations.

First, all IOIs must have an identical set of characteristics, and characteristics differ in their values only. In the example from [21], the characteristic named 'filling' is common to all IOIs. If this is not case, the precondition for evaluation of the equivalence relation is not met and the model is not applicable. Our linkability model differs in this assumption and covers heterogeneous, time-variant and disjoint sets of characteristics.

Second, an attacker is forced to conclude about an IOI-pair under all circumstances. This is the closed world assumption. Consequently, there is no explicit notion of ignorance. At best, ignorance is implicitly represented by a uniform probability density function over all possible partitions. In realistic settings, ignorance can be caused by two cases. First, IOIs are heterogeneous in their characteristics, as opposed to the assumptions, and an observed IOI-pair has none in common. Second, partial knowledge contradicts the closed world assumption. This means that distinct values of a shared characteristic do not necessarily imply that the IOIs are truly unlinked and the attacker decides to not conclude at all.

Further, modelling the attacker state is achieved by use of equivalence relations, which is inappropriate for probabilistic linkability. For example, for three IOIs $a, b, c$ the attacker may well probabilistically conclude that $a \sim b$ and $b \sim c$ but $a \not\sim c$. This clearly contradicts transitivity of linkability and equivalence relations, which is also detectable by an attacker. Such conflicts arise from partial knowledge and probabilistic linkability, i. e. from uncertainty in general and not from our model. We do not investigate how an attacker may resolve such conflicts by utilising transitivity as a requirement and/or obtained reliance in beliefs.

Our model captures the attacker's state for probabilistic linkability, including wrong and conflicting conclusions. We adopt the equivalence relation approach to model the true ownerships, which are the reference for our linkability measures.

The space and time complexities of the approaches are important for simulative studies of dynamic systems. Any of the above partition-based models establishes a PDF on all possible partitions. The total number of possible partitions is the $n$-th Bell number $B_n$, where $n$ is the number of IOIs. The asymptotic behaviour of the $n$-th Bell number is derived on pages 102–108 in [25] to

$$\begin{aligned} \frac{\log B_n}{n} &= \log n - \log\log n - 1 + \frac{\log\log n}{\log n} + \frac{1}{\log n} + \\ &\quad + \frac{1}{2} \cdot \left(\frac{\log\log n}{\log n}\right)^2 + O\left(\frac{\log\log n}{(\log n)^2}\right) \qquad (n \to \infty) \end{aligned}$$

and shows exponential growth in $n$. Further, computing the 'degree of unlinkability' from the PDF is linear in $n$. Thus, the time complexity is also exponential in $n$. Given today's technology, this is still impractical for any meaningful simulative analysis. For example, for $n = 11$ and one 32 bit integer per partition, the PDF requires in the worst-case approx. 373 GB of memory.

The authors of [26] provide an analytical framework for validation of security and privacy properties based on the function view concept under the open world assumption. A function view represents the attacker's partial knowledge. The focus is on non-probabilistic attacks and privacy as secrecy of relationships. The authors indicate extensibility to probabilistic attacks. Our model addresses probabilistic linkability and differs in the assumed attacker model. In particular, we assume that correspondent nodes are attackers.

## V. Conclusion

We presented an attacker state model named linkability graph that captures the results of probabilistic linkability attacks. From this model we derived linkability measures. We exemplified how a linkability graph can be obtained in a simulative privacy analysis. Although our model is inspired by previous work, it offers distinct properties. To name few: first, the lower and upper bounds of measures correspond to well-understood extreme cases of the linkability graph. Second, the linkability graph is more practical in terms of space and time complexity and enables simulative analysis of larger settings. Third, it distinguishes between linked items owned by one subject and by different subjects and thus separates quantification of distinct aspects.

In future, we will validate our model in two regards. First, we will use our implementation in simulation studies to quantify unlinkability in as realistic as possible settings. Second, we want to apply our model to other attacks than the presented example. Last but not least, we aim for privacy risk assessment, i. e. considering non-uniform impact associated with linked partial identities.

## Acknowledgment

The author would like to thank Andreas Reifert, Detlef Saß and Jochen Kögel for spending valuable time on discussions and their beneficial comments, as well as Sandra Steinbrecher and Sebastian Clauß for their unreserved correspondence.

This work is partly funded by the German Research Foundation (DFG) through the Collaborative Research Centre (SFB) "Nexus – Spatial World Models for Mobile Context-Aware Applications".

## References

[1] "Lidl im Visier der Datenschützer," *Zeit Online*, March, 27th 2008, http://www.zeit.de/online/2008/14/lidl-entschuldigung (in German).

[2] "Datenaffäre – Mehdorn gesteht noch mehr Spitzel-Aktionen," *Zeit Online*, February, 10th 2009, http://www.zeit.de/online/2009/07/bahn-zwischenbericht (in German).

[3] "Brisante Steuerdaten: Sünder-Daten sind erstklassig," *Zeit Online*, February, 22nd 2008, http://www.zeit.de/news/artikel/2008/02/22/2480191.xml (in German).

[4] R. Mullins, F. Mahon, C. Kuhmünch, M. Crotty, J. Mitic, and T. Pfeifer, "Daidalos: A platform for facilitating pervasive services," in *Advances in Pervasive Computing 2006: Adjunct Proceedings of the 4th International Conference on Pervasive Computing*, vol. 207, May 2006.

[5] R. Lange, N. Cipriani, L. Geiger, M. Großmann, H. Weinschrott, A. Brodt, M. Wieland, S. Rizou, and K. Rothermel, "Making the world wide space happen: New challenges for the nexus context platform," in *Proceedings of the 7th Annual IEEE International Conference on Pervasive Computing and Communications (PerCom)*, March 2009.

[6] M. Weiser, "The computer for the twenty-first century," *Scientific American*, vol. 265, no. 3, pp. 94–104, September 1991.

[7] K. Lyytinen and Y. Yoo, "Issues and challenges in ubiquitous computing," in *Communications of the ACM*, vol. 45, no. 12, December 2002, pp. 62–65.

[8] A. Pfitzmann and M. Hansen, "Anonymity, unlinkability, undetectability, unobservability, pseudonymity, and identity management – a consolidated proposal for terminology, version 0.31," February 15th 2008.

[9] D. L. Chaum, "Untraceable electronic mail, return addresses, and digital pseudonyms," *Communications of the ACM*, vol. 24, no. 2, pp. 84–90, February 1981.

[10] D. Kesdogan, J. Egner, and R. Büschkes, "Stop-and-go-MIXes providing probabilistic anonymity in an open system," in *Proceedings of the 2nd International Workshop on Information Hiding (IH)*. London, UK: Springer, 1998, pp. 83–98.

[11] O. Berthold, H. Federrath, and S. Köpsell, "Web MIXes: A system for anonymous and unobservable Internet access," in *Proceedings of Designing Privacy Enhancing Technologies: Workshop on Design Issues in Anonymity and Unobservability*. New York, USA: Springer, 2001, pp. 115–129, july 25–26, 2000.

[12] M. G. Reed, P. F. Syverson, and D. M. Goldschlag, "Anonymous connections and onion routing," *IEEE Journal on Selected Areas in Communications*, vol. 16, no. 4, pp. 482–494, May 1998.

[13] A. Nanda. (2006, December) A technical reference for the information card profile v1.0. Microsoft Corporation.

[14] G. J. Klir and T. A. Folger, *Fuzzy Sets, Uncertainty, and Information*, E. Cliffs, Ed. Prentice Hall, New Jersey, 1988.

[15] S. Clauß, "A framework for quantification of linkability within a privacy-enhancing identity management system," in *Proceedings of the International Conference on Emerging Trends in Information and Communication Security (ETRICS)*, June 2006.

[16] C. Díaz, S. Seys, J. Claessens, and B. Preneel, "Towards measuring anonymity," in *Proceedings of Privacy Enhancing Technologies (PET)*, ser. Lecture Notes in Computer Science (LNCS), vol. 2482/2003. Springer, April 14–15th 2002, pp. 54–68.

[17] A. Serjantov and G. Danezis, "Towards an information theoretic metric for anonymity," in *Proceedings of Privacy Enhancing Technologies (PET)*, ser. Lecture Notes in Computer Science (LNCS), vol. 2482/2003. Springer, April 14–15th 2002, pp. 41–53.

[18] C. Díaz, S. Seys, J. Claessens, and B. Preneel, "Information theory and anonymity," in *Proceedings of the 23rd Symposium on Information Theory in the Benelux*, ser. Werkgemeenschap voor Informatie- en Communicatietheorie, B. Macq and J. Quisquater, Eds., May 2002, pp. 179–186.

[19] S. Clauß and S. Schiffner, "Structuring anonymity metrics," in *Proceedings of the second ACM workshop on Digital identity management (DIM)*. New York, USA: ACM, November 2006, pp. 55–62.

[20] C. E. Shannon, "A mathematical theory of communication," in *The Bell System Technical Journal*, vol. 27, July and October 1948, pp. 379–423 and 623–656.

[21] S. Steinbrecher and S. Köpsell, "Modelling unlinkability," in *Privacy Enhancing Technologies (PET)*, ser. Lecture Notes in Computer Science (LNCS), R. Dingledine, Ed., vol. 2760. Springer, March 2003, pp. 32–47, proceedings of the Workshop on Privacy Enhancing Technologies 2003.

[22] M. Franz, B. Meyer, and A. Pashalidis, "Attacking unlinkability: The importance of context," in *Privacy Enhancing Technologies (PET)*, ser. Lecture Notes in Computer Science (LNCS), vol. 4776/2007. Springer, September 2007, pp. 1–16.

[23] S. Steinbrecher, "Mehrseitige Sicherheit in Reputationssystemen — Anforderungsanalyse und Umsetzungsmöglichkeiten," Ph.D. dissertation, Technischen Universität Dresden — Fakultät Informatik, July 2008.

[24] S. Clauß, "Towards quantification of privacy within a privacy-enhancing identity management system," Ph.D. dissertation, Technische Universität Dresden, May 2008.

[25] N. G. De Bruijn, *Asymptotic Methods in Analysis*, 2nd ed., ser. Bibliotheca Mathematica: A Series of Monographs on Pure and Applied Mathematics, N. G. De Bruijn, J. De Groot, and A. C. Zaanen, Eds. North-Holland Publishing Co. – Amsterdam, 1961, vol. IV.

[26] D. Hughes and V. Shmatikov, "Information hiding, anonymity and privacy: A modular approach," *Journal of Computer Security*, vol. 12, no. 1, pp. 3–36, 2004.