Network Working Group
Internet-Draft
Intended status: Informational
Expires: February 12, 2011

A. Pfitzmann, Ed.
TU Dresden
M. Hansen, Ed.
ULD Kiel
H. Tschofenig
Nokia Siemens Networks

August 11, 2010

Terminology for Talking about Privacy by Data Minimization: Anonymity, Unlinkability, Undetectability, Unobservability, Pseudonymity, and Identity Management

draft-hansen-privacy-terminology-01.txt

Abstract

This document is an attempt to consolidate terminology in the field privacy by data minimization. It motivates and develops definitions for anonymity/identifiability, (un)linkability, (un)detectability, (un)observability, pseudonymity, identity, partial identity, digital identity and identity management. Starting the definitions from the anonymity and unlinkability perspective and not from a definition of identity (the latter is the obvious approach to some people) reveals some deeper structures in this field.

Note: In absence of a separate discussion list please post your comments to the IETF SAAG mailing list and/or to the authors. For information about that mailing list please take a look at https://www.ietf.org/mailman/listinfo/saag.

Status of This Memo

This Internet-Draft is submitted in full conformance with the provisions of BCP 78 and BCP 79.

Internet-Drafts are working documents of the Internet Engineering Task Force (IETF). Note that other groups may also distribute working documents as Internet-Drafts. The list of current Internet-Drafts is at http://datatracker.ietf.org/drafts/current/.

Internet-Drafts are draft documents valid for a maximum of six months and may be updated, replaced, or obsoleted by other documents at any time. It is inappropriate to use Internet-Drafts as reference material or to cite them other than as "work in progress."

This Internet-Draft will expire on February 12, 2011.

Copyright Notice

Table of Contents

1. Introduction

Early papers from the 1980ies about privacy by data minimization already deal with anonymity, unlinkability, unobservability, and pseudonymity and introduce these terms within the respective context of proposed measures.

Note:

> Data minimization means that first of all, the possibility to collect personal data about others should be minimized. Next within the remaining possibilities, collecting personal data should be minimized. Finally, the time how long collected personal data is stored should be minimized.
>
> Data minimization is the only generic strategy to enable anonymity, since all correct personal data help to identify if we exclude providing misinformation (inaccurate or erroneous information, provided usually without conscious effort at misleading, deceiving, or persuading one way or another [Wils93]) or disinformation (deliberately false or distorted information given out in order to mislead or deceive [Wils93]).
>
> Furthermore, data minimization is the only generic strategy to enable unlinkability, since all correct personal data provide some linkability if we exclude providing misinformation or disinformation.

We show relationships between these terms and thereby develop a consistent terminology. Then, we contrast these definitions with newer approaches, e.g., from ISO IS 15408. Finally, we extend this terminology to identity (as the the opposite of anonymity and unlinkability) and identity management. Identity management is a much younger and much less well-defined field - so a really consolidated terminology for this field does not exist.

The adoption of this terminology will help to achieve better progress in the field by avoiding that those working on standards and research invent their own language from scratch.

This document is organized as follows: First, the setting used is described. Then, definitions of anonymity, unlinkability, linkability, undetectability, and unobservability are given and the relationships between the respective terms are outlined. Afterwards, known mechanisms to achieve anonymity, undetectability and unobservability are listed. The next sections deal with pseudonymity, i.e., pseudonyms, their properties, and the corresponding mechanisms. Thereafter, this is applied to privacy-

enhancing identity management. To give an overview of the main terms defined and their opposites, a corresponding table follows. Finally, concluding remarks are given. In appendices, we (A1) depict the relationships between some terms used and (A2 and A3) briefly discuss the relationship between our approach (to defining anonymity and identifiability) and other approaches. To make the document readable to as large an audience as possible, we did put information which can be skipped in a first reading or which is only useful to part of our readership, e.g., those knowing information theory, in footnotes.

2. Terminology and Requirements Notation

Privacy: "Privacy is the claim of individuals, groups, or institutions to determine for themselves when, how, and to what extent information about them is communicated to others. Viewed in terms of the relation of the individual to social participation, privacy is the voluntary and temporary withdrawal of a person from the general society through physical or psychological means, either in a state of solitude or small-group intimacy or, when among larger groups, in a condition of anonymity or reserve.", see page 7 of [West67]

3. Setting

We develop this terminology in the usual setting of entities (subjects and objects) and actions, i.e., subjects execute actions on objects. In particular, subjects called that senders send objects called messages to subjects called recipients using a communication network, i.e., stations send and receive messages using communication

technology.

Note:

> To keep the setting as simple as possible, usually, we do not distinguish between human senders and the stations which are used to send messages. Putting it the other way round, usually, we assume that each station is controlled by exactly one human being, its owner. If a differentiation between human communication and computer communication is necessary or if the assumption that each station is controlled by exactly one human being is wrong, the setting has to be more complex. We then use sender and recipient for human beings and message for their communication. For computers and their communications, we use stations sending bit strings. If we have to look even deeper than bits which are "abstractions" of physical signals, we call the representation of bit strings signals.

For other settings, e.g., users querying a database, customers

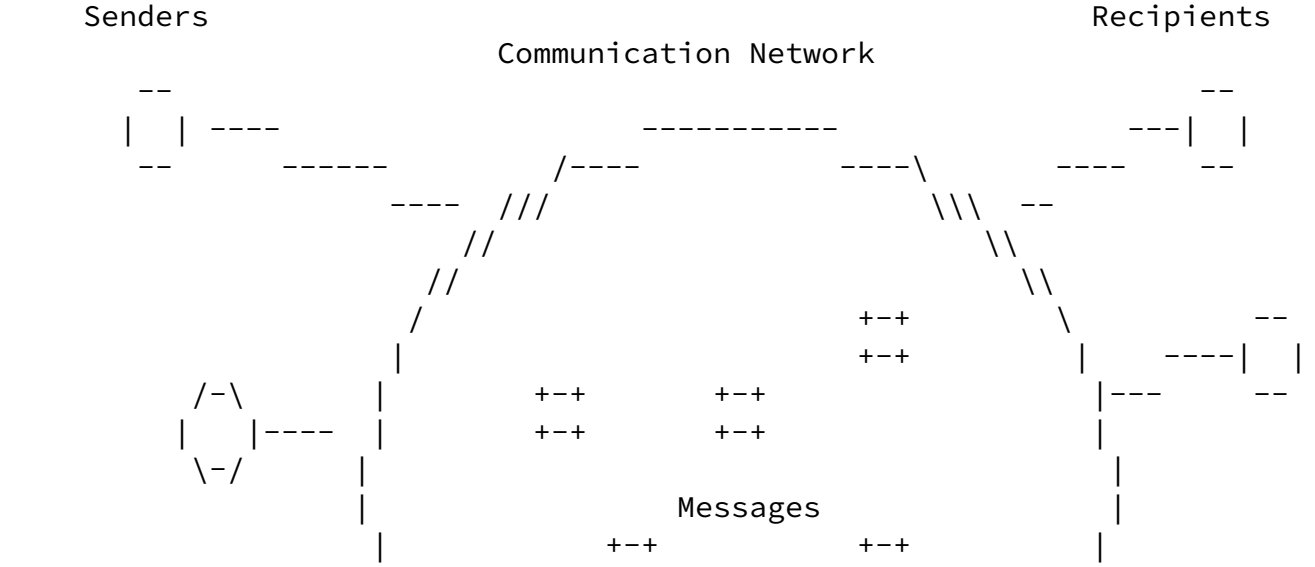
shopping in an e-commerce shop, the same terminology can be derived by instantiating the terms "sender", "recipient", and "message". But for ease of explanation, we use the specific setting here, see Figure 1. For a discussion in a broader context, we speak more generally about subjects, which might be actors (such as senders) or actees (such as recipients).

Irrespective whether we speak of senders and recipients or whether we generalize to actors and actees, we regard a subject as a human being (i.e., a natural person), a legal person, or a computer. An organization not acting as a legal person we neither see as a single subject nor as a single entity, but as (possibly structured) sets of subjects or entities. Otherwise, the distinction between "subjects" and "sets of subjects" would completely blur.

If we make our setting more concrete, we may l it a system. For our purposes, a system has the following relevant properties:

1. The system has a surrounding, i.e., parts of the world are "outside" the system. Together, the system and its surrounding form the universe.

2. The state of the system may change by actions within the system.



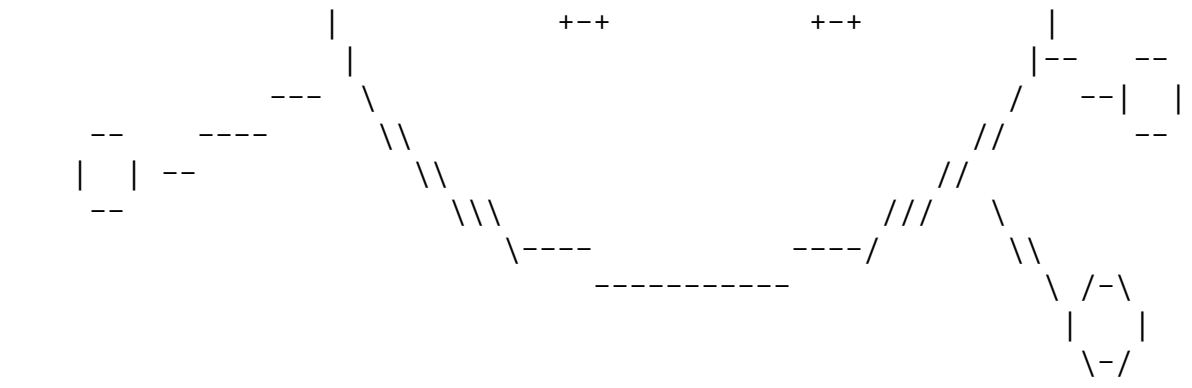


Figure 1: Setting

All statements are made from the perspective of an attacker , who may be interested in monitoring what communication is occurring, what patterns of communication exist, or even in manipulating the communication. The perspective describes the set of all possible observations. In the following, a property holds "from an attacker's perspective" iff it holds for all possible observations of that perspective. The attacker's perspective depends on the information the attacker has available. If we assume some limits on how much processing the attacker might be able to do, the information available to the attacker will not only depend on the attacker's perspective, but on the attacker's processing (abilities), too. The attacker may be an outsider tapping communication lines or an insider able to participate in normal communications and controlling at least some stations, cf. Figure 2. We assume that the attacker uses all information available to him to infer (probabilities of) his items of interest (IOIs), e.g., who did send or receive which messages. At this level of description, intentionally we do not care about particular types of IOIs. The given example would be an IOI which might be a 3-tupel of actor, action, and object. Later we consider

attribute values as IOIs. Attributes (and their values) are related to IOIs because they may be items of interest themselves or their observation may give information on IOIs: An attribute is a quality or characteristic of an entity or an action. Some attributes may take several values. Then it makes sense to make a distinction between more abstract attributes and more concrete attribute values. Mainly we are interested in attributes of subjects. Examples for attributes in this setting are "sending a message" or "receiving a

message".

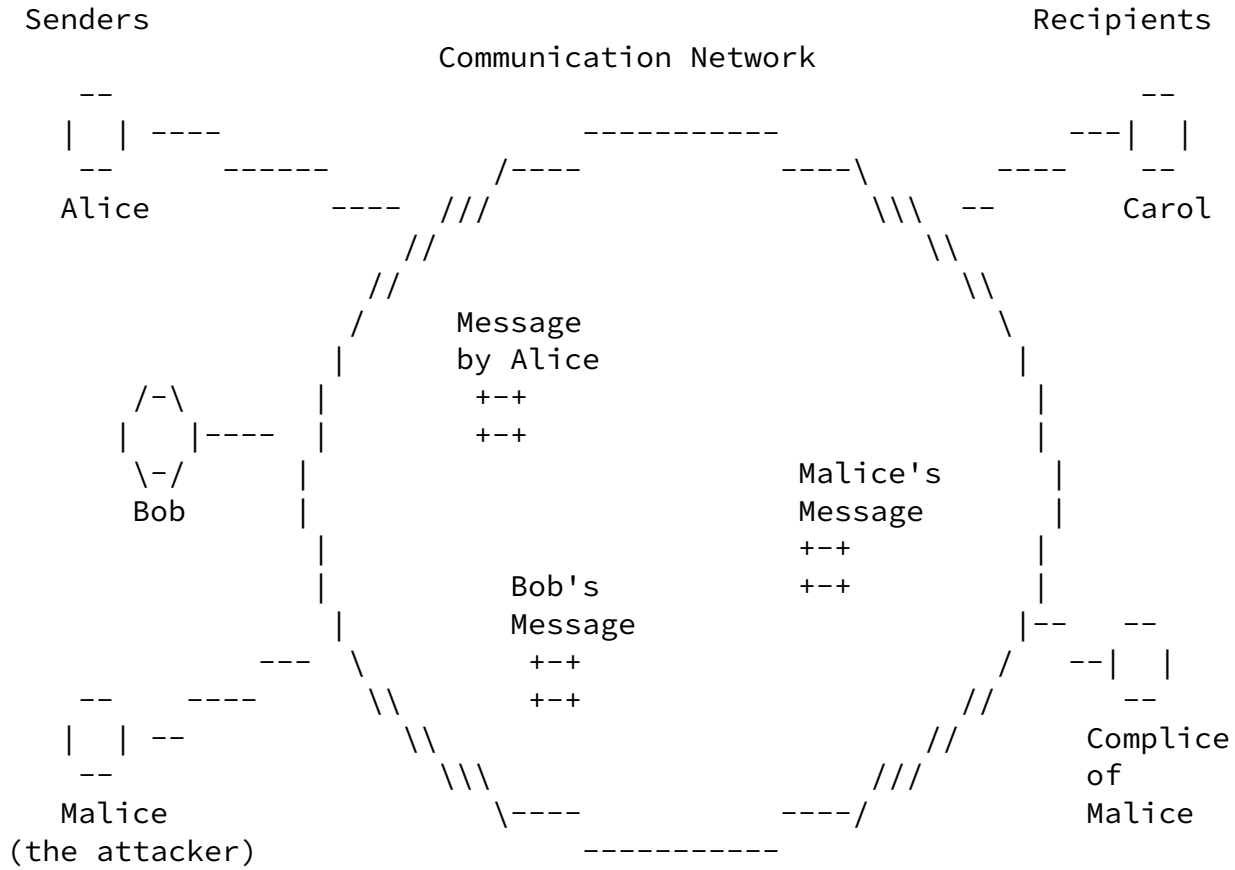


Figure 2: Example of an attacker's domain within the setting

Throughout the subsequent sections we assume that the attacker is not able to get information on the sender or recipient from the message content. Of course, encryption of messages provides protection of the content against attackers observing the communication lines and end-to-end encryption even provides protection of the content against all stations passed, e.g., for the purpose of forwarding and/or routing. But message content can neither be hidden from the sender nor from the recipient(s) of the message. Therefore, we do not mention the message content in these sections. For most applications it is unreasonable to assume that the attacker forgets something.

Thus, normally the knowledge of the attacker only increases.

"Knowledge" can be described by probabilities of IOIs. More knowledge then means more accurate probabilities, i.e., the probabilities the attacker assumes to be true are closer to the "true" probabilities.

4. Anonymity

To enable anonymity of a subject, there always has to be an appropriate set of subjects with potentially the same attributes . Since sending and receiving of particular messages are special cases of "attributes" of senders and recipients, this is slightly more general than the setting in Section 3. This generality is very fortunate to stay close to the everyday meaning of "anonymity" which is not only used w.r.t. subjects active in a particular context, e.g., senders and recipients of messages, but w.r.t. subjects passive in a particular context as well, e.g., subjects the records within a database relate to. This leads to the following definition:

Definition: Anonymity of a subject means that the subject is not identifiable within a set of subjects, the anonymity set.

Note:

"not identifiable within the anonymity set" means that only using the information the attacker has at his discretion, the subject is "not uniquely characterized within the anonymity set". In more precise language, only using the information the attacker has at his discretion, the subject is "not distinguishable from the other subjects within the anonymity set".

From [ISO99]: "Anonymity ensures that a user may use a resource or service without disclosing the user's identity. The requirements for anonymity provide protection of the user identity. Anonymity is not intended to protect the subject identity. [...] Anonymity requires that other users or subjects are unable to determine the identity of a user bound to a subject or operation." Compared with this explanation, our definition is more general as it is not restricted to identifying users, but any subjects.

The anonymity set is the set of all possible subjects. The set of possible subjects depends on the knowledge of the attacker. Thus, anonymity is relative with respect to the attacker. With respect to actors, the anonymity set consists of the subjects who might cause an action. With respect to actees, the anonymity set consists of the subjects who might be acted upon. Therefore, a sender may be anonymous (sender anonymity) only within a set of potential senders, his/her sender anonymity set, which itself may be a subset of all

subjects worldwide who may send a message from time to time. The same for the recipient means that a recipient may be anonymous (recipient anonymity) only within a set of potential recipients, his/her recipient anonymity set, cf. Figure 3. Both anonymity sets may be disjoint, be the same, or they may overlap. The anonymity sets may vary over time. Since we assume that the attacker does not forget anything he knows, the anonymity set cannot increase w.r.t. a particular IOI. Especially subjects joining the system in a later stage, do not belong to the anonymity set from the point of view of an attacker observing the system in an earlier stage. (Please note that if the attacker cannot decide whether the joining subjects were present earlier, the anonymity set does not increase either: It just stays the same.) Due to linkability, cf. below, the anonymity set normally can only decrease.

Anonymity of a set of subjects within an (potentially larger) anonymity set means that all these individual subjects are not identifiable within this anonymity set. In this definition, "set of subjects" is just taken to describe that the anonymity property holds for all elements of the set. Another possible definition would be to consider the anonymity property for the set as a whole. Then a semantically quite different definition could read: Anonymity of a set S of subjects within a larger anonymity set A means that it is not distinguishable whether the subject whose anonymity is at stake (and which clearly is within A) is within S or not.

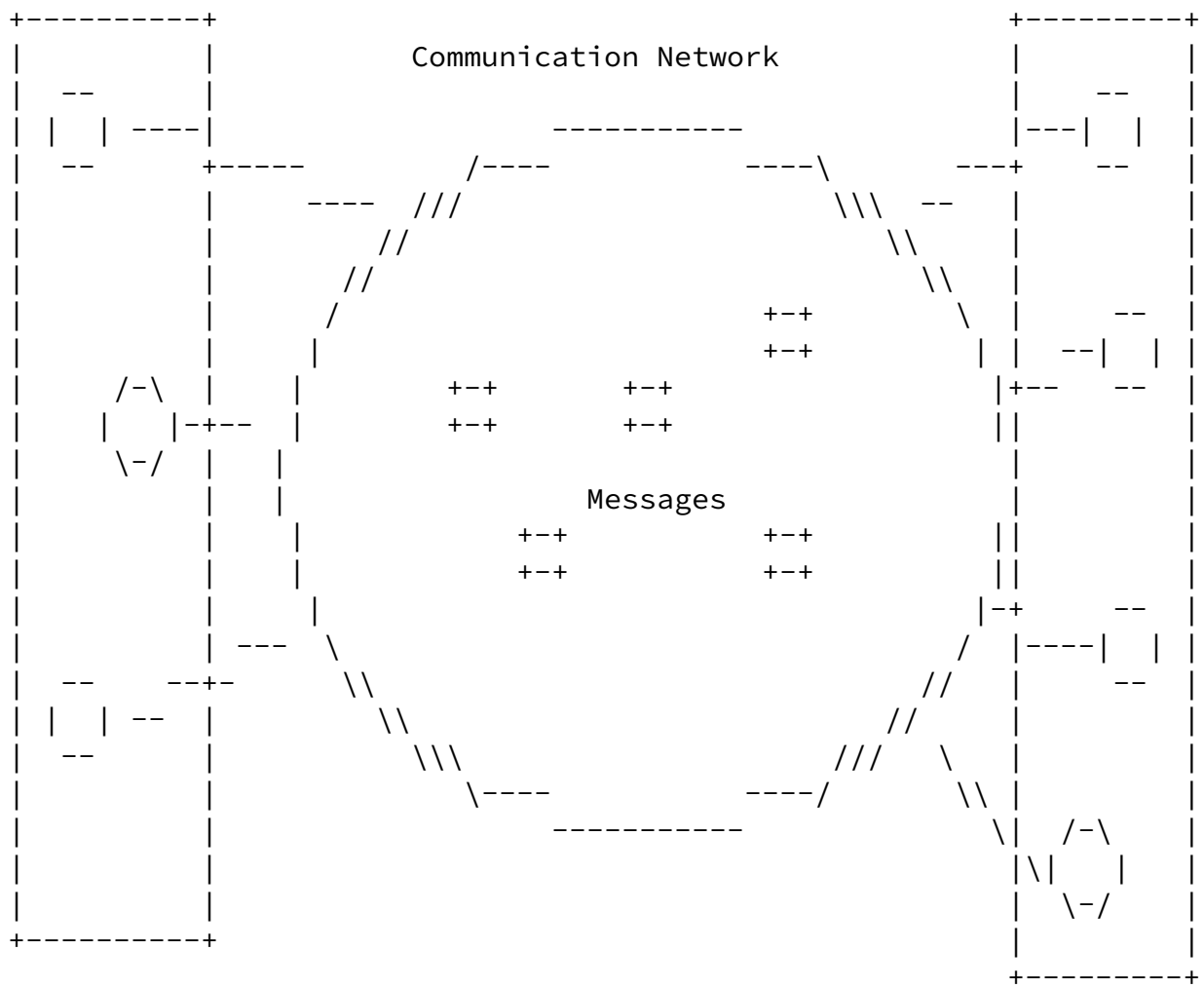


Figure 3: Anonymity sets within the setting

The definition given above for anonymity basically defines anonymity as a binary property: Either a subject is anonymous or not. To reflect the possibility to quantify anonymity in our definition and to underline that all statements are made from the perspective of an

attacker (cf. Figure 4), it is appropriate to work with a slightly more complicated definition in the following:

Definition: Anonymity of a subject from an attacker's perspective means that the attacker cannot sufficiently identify the subject within a set of subjects, the anonymity set.

In this revised definition, "sufficiently" underlines both that there is a possibility to quantify anonymity and that for some

applications, there might be a need to define a threshold where anonymity begins.

If we do not focus on the anonymity of one individual subject, called individual anonymity, but on the anonymity provided by a system to all of its users together, called global anonymity, we can state: All other things being equal, global anonymity is the stronger, the larger the respective anonymity set is and the more evenly distributed the sending or receiving, respectively, of the subjects within that set is.

Note:

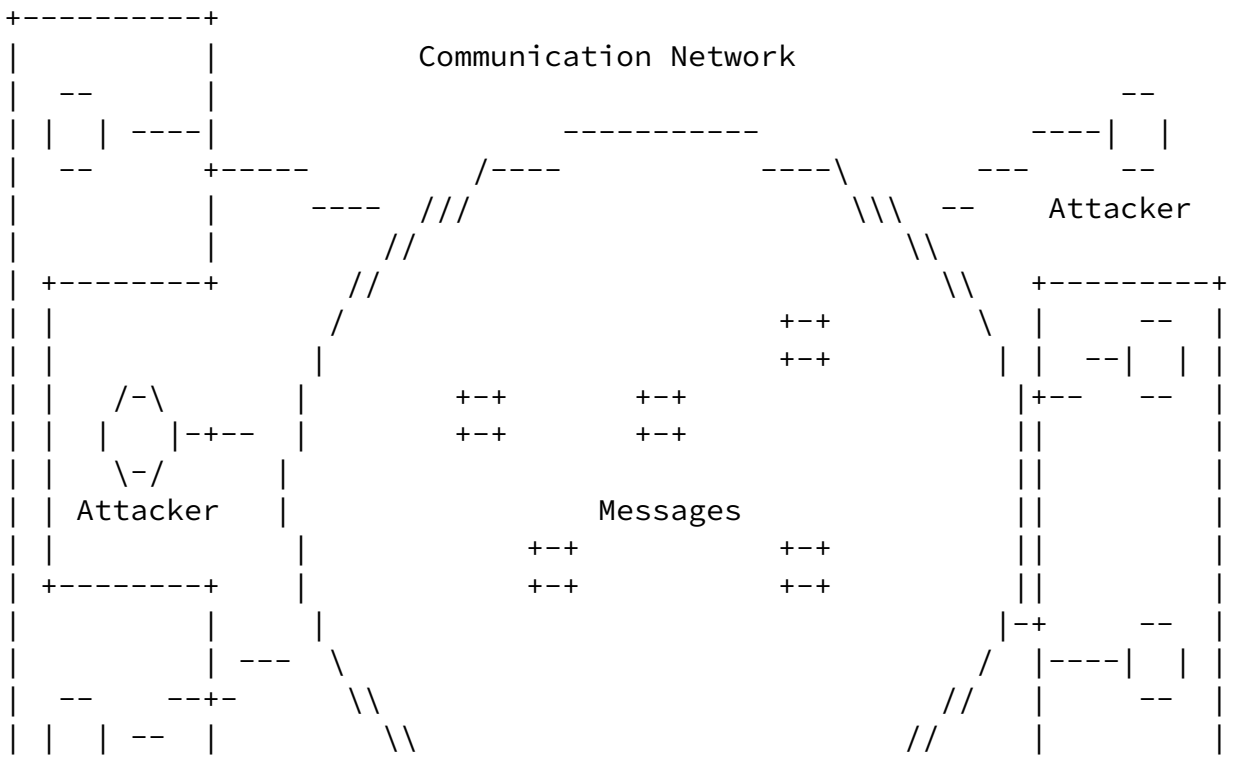
The entropy of a message source as defined by Claude E. Shannon [Shan48] might be an appropriate measure to quantify global anonymity - just take who is the sender/recipient as the "message" in Shannon's definition. For readers interested in formalizing what we informally say: "No change of probabilities" means "no change of knowledge" and vice versa. "No change of probabilities" (or what is equivalent: "no change of knowledge") implies "no change of entropy", whereas "no change of entropy" neither implies "no change of probabilities" nor "no change of knowledge". In an easy to remember notation: No change of probabilities = no change of knowledge => no change of entropy.

The definition of anonymity is an analog to the definition of "perfect secrecy" by Claude E. Shannon [Shan49], whose definition takes into account that no security mechanism whatsoever can take away knowledge from the attacker which he already has.

For a fixed anonymity set, global anonymity is maximal iff all subjects within the anonymity set are equally likely. Since subjects

may behave quite distinct from each other (and trying to persuade them to behave more equally may both fail and be not compatible with basic human rights), achieving maximal anonymity or even something close to it usually is impossible. Strong or even maximal global anonymity does not imply strong anonymity or even maximal anonymity of each particular subject. What maximal anonymity of one individual subject (maximal individual anonymity, for short) means is unclear. On the one hand, if her probability approaches zero, her Shannon entropy (as a measure for anonymity) gets larger and larger. On the other hand, if her probability gets zero, she is outside the anonymity set. Even if global anonymity is strong, one (or a few) individual subjects might be quite likely, so their anonymity is weak. W.r.t. these "likely suspects", nothing is changed if the anonymity set is made larger and sending and receiving of the other subjects are, e.g., distributed evenly. That way, arbitrarily strong global anonymity can be achieved without doing anything for the

"likely suspects" [ClSc06]. So there is need to define anonymity measures not only for the system as a whole, but for individual subjects (individual anonymity) or small sets of subjects.



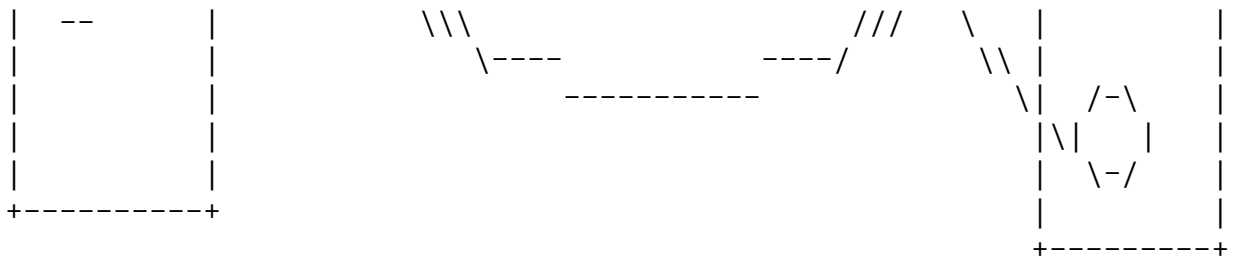


Figure 4: Anonymity sets w.r.t. attacker within the setting

From the above discussion follows that anonymity in general as well as the anonymity of each particular subject is a concept which is very much context dependent (on, e.g., subjects population, attributes, time frame, etc). In order to quantify anonymity within concrete situations, one would have to describe the system in sufficient detail, which is practically not (always) possible for large open systems (but maybe for some small data bases for instance). Besides the quantity of anonymity provided within a

particular setting, there is another aspect of anonymity: its robustness. Robustness of anonymity characterizes how stable the quantity of anonymity is against changes in the particular setting, e.g., a stronger attacker or different probability distributions. We might use quality of anonymity as a term comprising both quantity and robustness of anonymity. To keep this text as simple as possible, we will mainly discuss the quantity of anonymity in the following, using the wording "strength of anonymity".

The above definitions of anonymity and the mentioned measures of quantifying anonymity are fine to characterize the status of a subject in a world as it is. If we want to describe changes to the anonymity of a subject if the world is changed somewhat, e.g., the subject uses the communication network differently or uses a modified communication network, we need another definition of anonymity capturing the delta. The simplest way to express this delta is by the observations of "the" attacker.

Definition: An anonymity delta (regarding a subject's anonymity) from an attacker's perspective specifies the difference between the subject's anonymity taking into account the attacker's observations (i.e., the attacker's a-posteriori knowledge) and the subject's anonymity given the attacker's a-priori knowledge only.

Note:

In some publications, the a-priori knowledge of the attacker is called "background knowledge" and the a-posteriori knowledge of the attacker is called "new knowledge".

As we can quantify anonymity in concrete situations, so we can quantify the anonymity delta. This can be done by just defining: quantity(anonymity delta) := quantity(anonymity_a-posteriori) - quantity(anonymity_a-priori) If anonymity_a-posteriori and anonymity_a-priori are the same, their quantification is the same and therefore the difference of these quantifications is 0. If anonymity can only decrease (which usually is quite a reasonable assumption), the maximum of quantity(anonymity delta) is 0.

Since anonymity cannot increase, the anonymity delta can never be positive. Having an anonymity delta of zero means that anonymity stays the same. This means that if the attacker has no a-priori knowledge about the particular subject, having no anonymity delta implies anonymity. But if the attacker has an a-priori knowledge covering all actions of the particular subject, having no anonymity delta does not imply any anonymity at all. If there is no anonymity from the very beginning, even preserving it completely does not yield any anonymity. To be able to express this conveniently, we use

wordings like "perfect preservation of a subject's anonymity". It might be worthwhile to generalize "preservation of anonymity of single subjects" to "preservation of anonymity of sets of subjects", in the limiting case all subjects in an anonymity set. An important special case is that the "set of subjects" is the set of subjects having one or several attribute values A in common. Then the meaning of "preservation of anonymity of this set of subjects" is that knowing A does not decrease anonymity. Having a negative anonymity delta means that anonymity is decreased.

5. Unlinkability

Unlinkability only has a meaning after the system in which we want to describe anonymity properties has been defined and the attacker has been characterized. Then:

Definition: Unlinkability of two or more items of interest (IOIs, e.g., subjects, messages, actions, ...) from an attacker's perspective means that within the system (comprising these and possibly other items), the attacker cannot sufficiently distinguish whether these IOIs are related or not. ,

Note:

From [ISO99]: "Unlinkability ensures that a user may make multiple uses of resources or services without others being able to link these uses together. [...] Unlinkability requires that users and/or subjects are unable to determine whether the same user caused certain specific operations in the system." In contrast to this definition, the meaning of unlinkability in this text is less focused on the user, but deals with unlinkability of "items" and therefore takes a general approach.

As the entropy of a message source might be an appropriate measure to quantify (global) anonymity (and thereafter "anonymity" might be used as a quantity), we may use definitions to quantify unlinkability (and thereafter "unlinkability" might be used as a quantity as well). Quantifications of unlinkability can be either probabilities or entropies, or whatever is useful in a particular context.

Linkability is the negation of unlinkability:

Definition: Linkability of two or more items of interest (IOIs, e.g., subjects, messages, actions, ...) from an attacker's perspective means that within the system (comprising these and possibly other items), the attacker can sufficiently distinguish whether these IOIs are related or not.

For example, in a scenario with at least two senders, two messages sent by subjects within the same anonymity set are unlinkable for an attacker if for him, the probability that these two messages are sent by the same sender is sufficiently close to 1/(number of senders).

In case of unicast the same is true for recipients; in case of multicast it is slightly more complicated.

Definition: An unlinkability delta of two or more items of interest (IOIs, e.g., subjects, messages, actions, ...) from an attacker's perspective specifies the difference between the unlinkability of these IOIs taking into account the attacker's observations and the unlinkability of these IOIs given the attacker's a-priori knowledge only.

Since we assume that the attacker does not forget anything, unlinkability cannot increase. Normally, the attacker's knowledge cannot decrease (analogously to Shannon's definition of "perfect secrecy", see above). An exception of this rule is the scenario where the use of misinformation (inaccurate or erroneous information, provided usually without conscious effort at misleading, deceiving, or persuading one way or another [Wils93]) or disinformation (deliberately false or distorted information given out in order to mislead or deceive [Wils93]) leads to a growing uncertainty of the attacker which information is correct. A related, but different aspect is that information may become wrong (i.e., outdated) simply because the state of the world changes over time. Since privacy is not only about to protect the current state, but the past and history of a data subject as well, we will not make use of this different aspect in the rest of this document. Therefore, the unlinkability delta can never be positive. Having an unlinkability delta of zero means that the probability of those items being related from the attacker's perspective stays exactly the same before (a-priori knowledge) and after the attacker's observations (a-posteriori knowledge of the attacker). If the attacker has no a-priori knowledge about the particular IOIs, having an unlinkability delta of zero implies unlinkability. But if the attacker has a-priori knowledge covering the relationships of all IOIs, having an unlinkability delta of zero does not imply any unlinkability at all. If there is no unlinkability from the very beginning, even preserving it completely does not yield any unlinkability. To be able to express this conveniently, we use wordings like "perfect preservation of unlinkability w.r.t. specific items" to express that the unlinkability delta is zero. It might be worthwhile to generalize "preservation of unlinkability of two IOIs" to "preservation of unlinkability of sets of IOIs", in the limiting case all IOIs in the system.

For example, the unlinkability delta of two messages is sufficiently

small (zero) for an attacker if the probability describing his a-posteriori knowledge that these two messages are sent by the same sender and/or received by the same recipient is sufficiently (exactly) the same as the probability imposed by his a-priori knowledge. Please note that unlinkability of two (or more) messages of course may depend on whether their content is protected against the attacker considered. In particular, messages may be unlinkable if we assume that the attacker is not able to get information on the sender or recipient from the message content, cf. Section 3. Yet with access to their content even without deep semantical analysis the attacker can notice certain characteristics which link them together - e.g. similarities in structure, style, use of some words or phrases, consistent appearance of some grammatical errors, etc. In a sense, content of messages may play a role as "side channel" in a similar way as in cryptanalysis - i.e., content of messages may leak some information on their linkability.

Roughly speaking, no unlinkability delta of items means that the ability of the attacker to relate these items does not increase by observing the system or by possibly interacting with it.

The definitions of unlinkability, linkability and unlinkability delta do not mention any particular set of IOIs they are restricted to. Therefore, the definitions of unlinkability and unlinkability delta are very strong, since they cover the whole system. We could weaken the definitions by restricting them to part of the system: "Unlinkability of two or more IOIs from an attacker's perspective means that within an unlinkability set of IOIs (comprising these and possibly other items), the attacker cannot sufficiently distinguish whether these IOIs are related or not."

6. Anonymity in Terms of Unlinkability

To describe anonymity in terms of unlinkability, we have to augment the definitions of anonymity given in Section 4 by making explicit the attributes anonymity relates to. This is best explained by looking at an example in detail. In our setting, cf. Section 3, we choose the attribute "having sent a message" as the example. Then we have:

A sender s is anonymous w.r.t. sending, iff s is anonymous within the set of potential senders, i.e., within the sender anonymity set.

This mainly is a re-phrasing of the definition in Section 3. If we make the message under consideration explicit, the definition reads:

A sender s sends a message m anonymously, iff s is anonymous within

the set of potential senders of m, the sender anonymity set of m.

This can be generalized to sets of messages easily:

A sender s sends a set of messages M anonymously, iff s is anonymous within the set of potential senders of M, the sender anonymity set of M.

If the attacker's focus is not on the sender, but on the message, we can define:

A message m is sent anonymously, iff m can have been sent by each potential sender, i.e., by any subject within the sender anonymity set of m.

Again, this can be generalized to sets of messages easily:

A set of messages M is sent anonymously, iff M can have been sent by each set of potential senders, i.e., by any set of subjects within the cross product of the sender anonymity sets of each message m within M.

Of course, all 5 definitions would work for receiving of messages accordingly. For more complicated settings with more operations than these two, appropriate sets of definitions can be developed.

Now we are prepared to describe anonymity in terms of unlinkability.

We do this by using our setting, cf. Section 3. So we consider sending and receiving of messages as attributes; the items of interest (IOIs) are "who has sent or received which message". Then, anonymity of a subject w.r.t. an attribute may be defined as unlinkability of this subject and this attribute. In the wording of the definition of unlinkability: a subject s is related to the attribute value "has sent message m" if s has sent message m. s is not related to that attribute value if s has not sent message m. Same for receiving.Unlinkability is a sufficient condition of anonymity, but it is not a necessary condition. Thus, failing unlinkability w.r.t. some attribute value(s) does not necessarily eliminate anonymity as defined in Section 4; in specific cases (i.e., depending on the attribute value(s)) even the strength of anonymity may not be affected.

So we have: Sender anonymity of a subject means that to this potentially sending subject, each message is unlinkable.

Note:

The property unlinkability might be more "fine-grained" than anonymity, since there are many more relations where unlinkability

might be an issue than just the relation "anonymity" between subjects and IOIs. Therefore, the attacker might get to know information on linkability while not necessarily reducing anonymity of the particular subject - depending on the defined measures. An example might be that the attacker, in spite of being able to link, e.g., by timing, all encrypted messages of a transactions, does not learn who is doing this transaction.

Correspondingly, recipient anonymity of a subject means that to this potentially receiving subject, each message is unlinkable.

Relationship anonymity of a pair of subjects, the potentially sending subject and the potentially receiving subject, means that to this potentially communicating pair of subjects, each message is unlinkable. In other words, sender and recipient (or each recipient in case of multicast) are unlinkable. As sender anonymity of a message cannot hold against the sender of this message himself nor can recipient anonymity hold against any of the recipients w.r.t. himself, relationship anonymity is considered w.r.t. outsiders only, i.e., attackers being neither the sender nor one of the recipients of the messages under consideration.

Thus, relationship anonymity is a weaker property than each of sender anonymity and recipient anonymity: The attacker might know who sends which messages or he might know who receives which messages (and in some cases even who sends which messages and who receives which messages). But as long as for the attacker each message sent and each message received are unlinkable, he cannot link the respective senders to recipients and vice versa, i.e., relationship anonymity holds. The relationship anonymity set can be defined to be the cross product of two potentially distinct sets, the set of potential senders and the set of potential recipients or - if it is possible to exclude some of these pairs - a subset of this cross product. So the

relationship anonymity set is the set of all possible sender-recipient(s)-pairs. In case of multicast, the set of potential recipients is the power set of all potential recipients. If we take the perspective of a subject sending (or receiving) a particular message, the relationship anonymity set becomes the set of all potential recipients (senders) of that particular message. So fixing one factor of the cross product gives a recipient anonymity set or a sender anonymity set.

Note:

The following is an explanation of the statement made in the previous paragraph regarding relationship anonymity: For all attackers it holds that sender anonymity implies relationship anonymity, and recipient anonymity implies relationship anonymity.

This is true if anonymity is taken as a binary property: Either it holds or it does not hold. If we consider quantities of anonymity, the validity of the implication possibly depends on the particular definitions of how to quantify sender anonymity and recipient anonymity on the one hand, and how to quantify relationship anonymity on the other. There exists at least one attacker model, where relationship anonymity does neither imply sender anonymity nor recipient anonymity. Consider an attacker who neither controls any senders nor any recipients of messages, but all lines and - maybe - some other stations. If w.r.t. this attacker relationship anonymity holds, you can neither argue that against him sender anonymity holds nor that recipient anonymity holds. The classical MIX-net (cf. Section 9) without dummy traffic is one implementation with just this property: The attacker sees who sends messages when and who receives messages when, but cannot figure out who sends messages to whom.

7. Undetectability and Unobservability

In contrast to anonymity and unlinkability, where not the IOI, but only its relationship to subjects or other IOIs is protected, for undetectability, the IOIs are protected as such. Undetectability can be regarded as a possible and desirable property of steganographic systems (see Section 9). Therefore it matches the information hiding terminology [Pfit96], [ZFKP98]. In contrast, anonymity, dealing with the relationship of discernible IOIs to subjects, does not directly

fit into that terminology, but independently represents a different dimension of properties.

Definition: Undetectability of an item of interest (IOI) from an attacker's perspective means that the attacker cannot sufficiently distinguish whether it exists or not.

Note:

From [ISO99]: "Unobservability ensures that a user may use a resource or service without others, especially third parties, being able to observe that the resource or service is being used. [...] Unobservability requires that users and/or subjects cannot determine whether an operation is being performed." As seen before, our approach is less user-focused and insofar more general. With the communication setting and the attacker model chosen in this text, our definition of unobservability shows the method how to achieve it: preventing distinguishability of IOIs. Thus, the ISO definition might be applied to a different setting where attackers are prevented from observation by other means, e.g., by encapsulating the area of interest against third parties.

In some applications (e.g. steganography), it might be useful to quantify undetectability to have some measure how much uncertainty about an IOI remains after the attacker's observations. Again, we may use probabilities or entropy, or whatever is useful in a particular context.

If we consider messages as IOIs, this means that messages are not sufficiently discernible from, e.g., "random noise". A slightly more precise formulation might be that messages are not discernible from no message. A quantification of this property might measure the number of indistinguishable IOIs and/or the probabilities of distinguishing these IOIs.

Undetectability is maximal iff whether an IOI exists or not is completely indistinguishable. We call this perfect undetectability.

Definition: An undetectability delta of an item of interest (IOI) from an attacker's perspective specifies the difference between the undetectability of the IOI taking into account the attacker's

observations and the undetectability of the IOI given the attacker's a-priori knowledge only.

The undetectability delta is zero iff whether an IOI exists or not is indistinguishable to exactly the same degree whether the attacker takes his observations into account or not. We call this "perfect preservation of undetectability".

Undetectability of an IOI clearly is only possible w.r.t. subjects being not involved in the IOI (i.e., neither being the sender nor one of the recipients of a message). Therefore, if we just speak about undetectability without spelling out a set of IOIs, it goes without saying that this is a statement comprising only those IOIs the attacker is not involved in.

As the definition of undetectability stands, it has nothing to do with anonymity - it does not mention any relationship between IOIs and subjects. Even more, for subjects being involved in an IOI, undetectability of this IOI is clearly impossible. Therefore, early papers describing new mechanisms for undetectability designed the mechanisms in a way that if a subject necessarily could detect an IOI, the other subject(s) involved in that IOI enjoyed anonymity at least. The rational for this is to strive for data minimization: No subject should get to know any (potentially personal) data - except this is absolutely necessary. Given the setting described in Section 3, this means: 1. Subjects being not involved in the IOI get to know absolutely nothing. 2. Subjects being involved in the IOI only get to know the IOI, but not the other subjects involved - the other subjects may stay anonymous. Since in the setting described in

Section 3 the attributes "sending a message" or "receiving a message" are the only kinds of attributes considered, 1. and 2. together provide data minimization in this setting in an absolute sense. Undetectability by uninvolved subjects together with anonymity even if IOIs can necessarily be detected by the involved subjects has been called unobservability:

Definition: Unobservability of an item of interest (IOI) means

* undetectability of the IOI against all subjects uninvolved in it and

* anonymity of the subject(s) involved in the IOI even against the other subject(s) involved in that IOI.

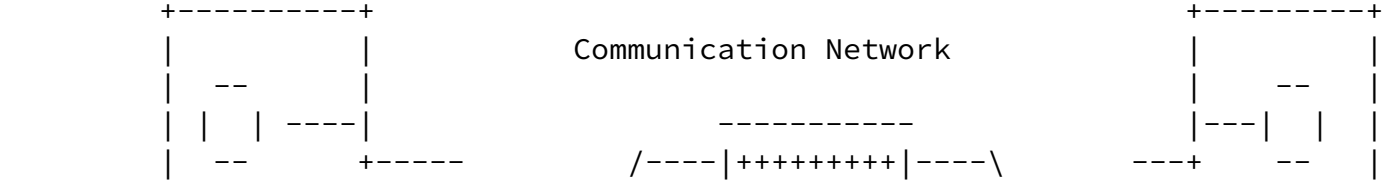
As we had anonymity sets of subjects with respect to anonymity, we have unobservability sets of subjects with respect to unobservability, see Figure 5. Mainly, unobservability deals with IOIs instead of subjects only. Though, like anonymity sets, unobservability sets consist of all subjects who might possibly cause these IOIs, i.e. send and/or receive messages.

Sender unobservability then means that it is sufficiently undetectable whether any sender within the unobservability set sends. Sender unobservability is perfect iff it is completely undetectable whether any sender within the unobservability set sends.

Recipient unobservability then means that it is sufficiently undetectable whether any recipient within the unobservability set receives. Recipient unobservability is perfect iff it is completely undetectable whether any recipient within the unobservability set receives.

Relationship unobservability then means that it is sufficiently undetectable whether anything is sent out of a set of could-be senders to a set of could-be recipients. In other words, it is sufficiently undetectable whether within the relationship unobservability set of all possible sender-recipient(s)-pairs, a message is sent in any relationship. Relationship unobservability is perfect iff it is completely undetectable whether anything is sent out of a set of could-be senders to a set of could-be recipients.

All other things being equal, unobservability is the stronger, the larger the respective unobservability set is, see Figure 6.



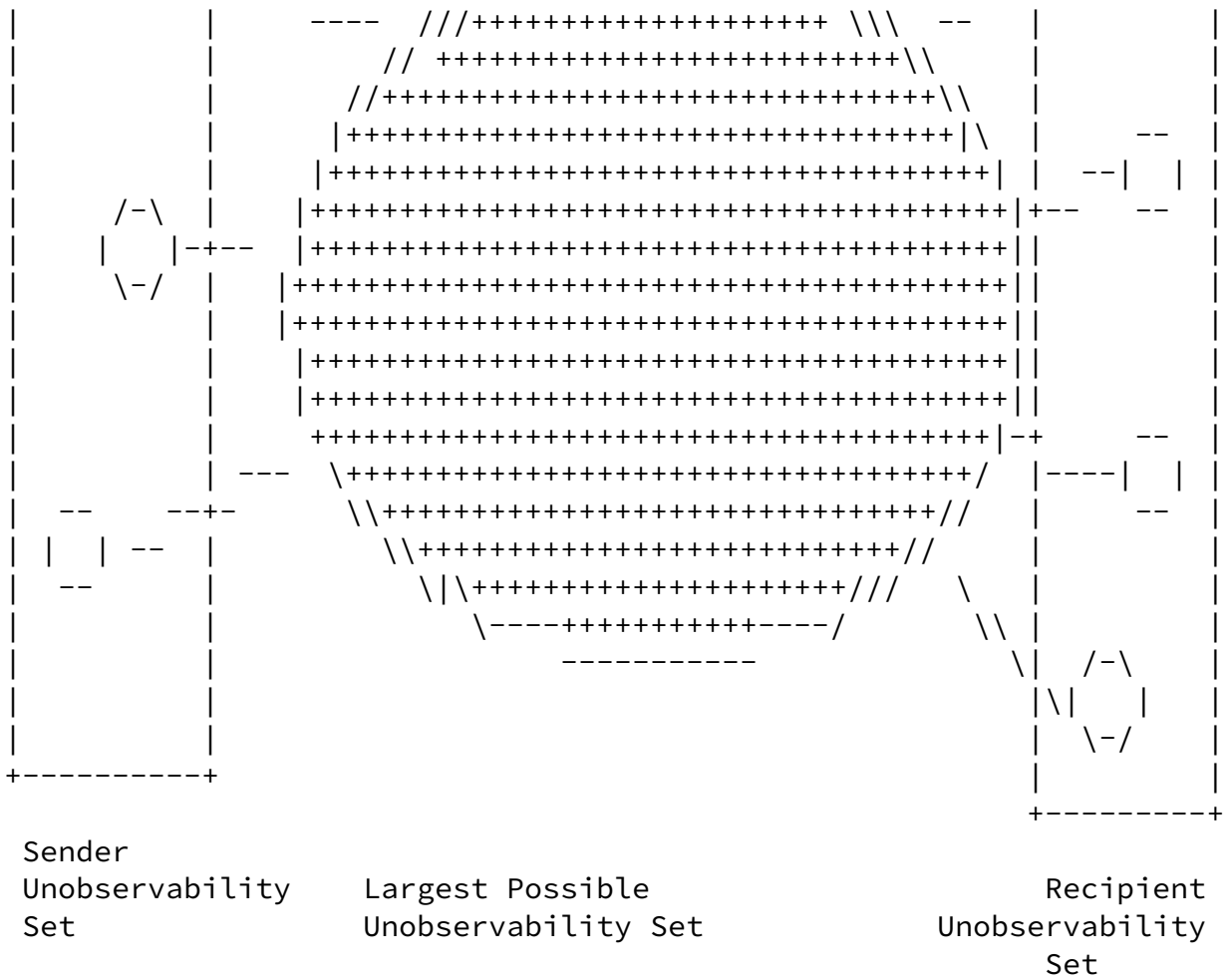


Figure 5: Unobservability sets within the setting

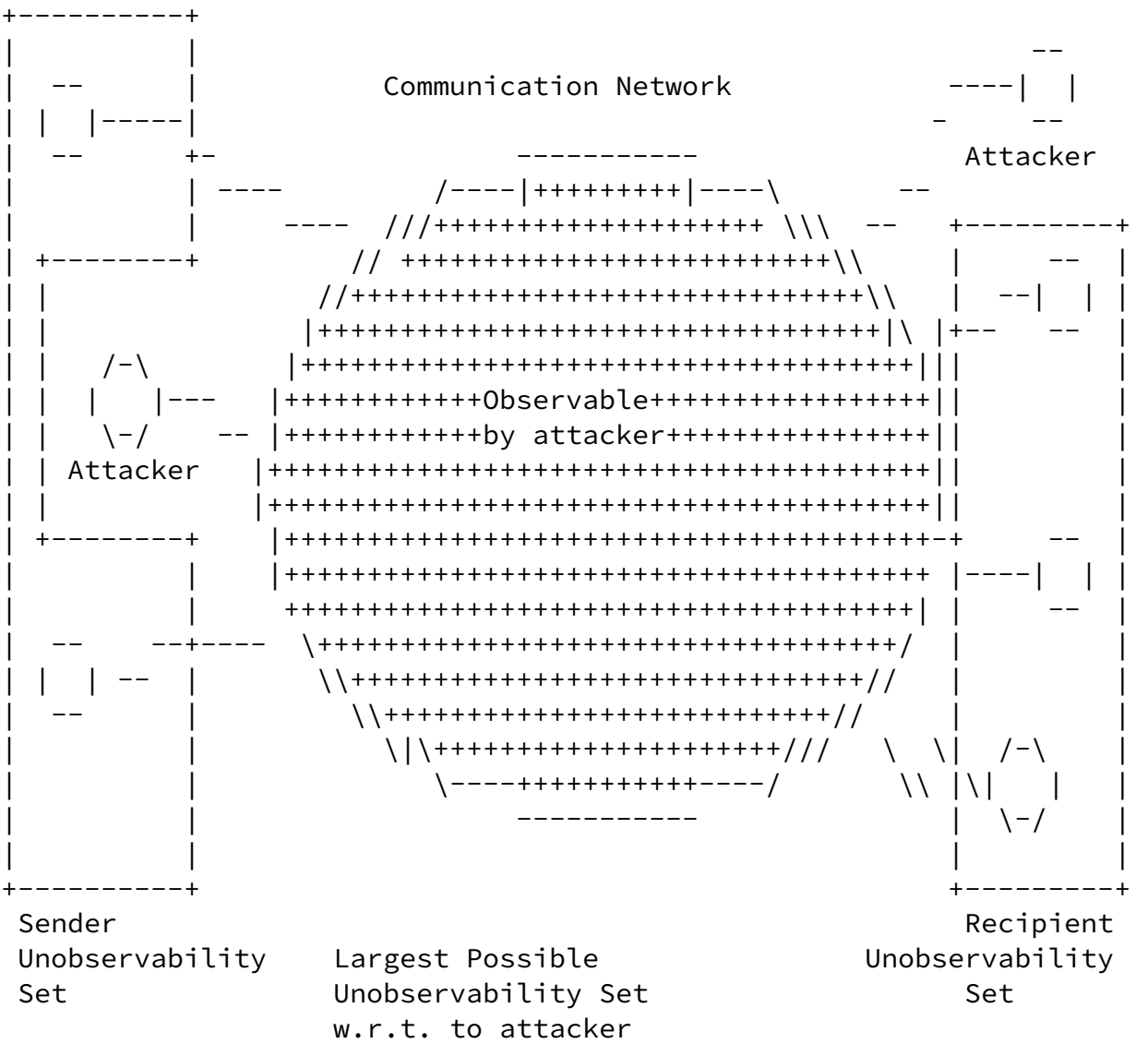


Figure 6: Unobservability sets w.r.t. attacker within the setting

Definition: An unobservability delta of an item of interest (IOI) means

* undetectability delta of the IOI against all subjects uninvolved in it and

* anonymity delta of the subject(s) involved in the IOI even against the other subject(s) involved in that IOI.

Since we assume that the attacker does not forget anything, unobservability cannot increase. Therefore, the unobservability delta can never be positive. Having an unobservability delta of zero w.r.t. an IOI means an undetectability delta of zero of the IOI against all subjects uninvolved in the IOI and an anonymity delta of

zero against those subjects involved in the IOI. To be able to express this conveniently, we use wordings like "perfect preservation of unobservability" to express that the unobservability delta is zero.

8. Relationships between Terms

With respect to the same attacker, unobservability reveals always only a subset of the information anonymity reveals. [ReRu98] propose a continuum for describing the strength of anonymity. They give names: "absolute privacy" (the attacker cannot perceive the presence of communication, i.e., unobservability) - "beyond suspicion" - "probable innocence" - "possible innocence" - "exposed" - "provably exposed" (the attacker can prove the sender, recipient, or their relationship to others). Although we think that the terms "privacy" and "innocence" are misleading, the spectrum is quite useful. We might use the shorthand notation

unobservability => anonymity

for that (=> reads "implies"). Using the same argument and notation, we have

sender unobservability => sender anonymity

recipient unobservability => recipient anonymity

relationship unobservability => relationship anonymity

As noted above, we have

sender anonymity => relationship anonymity

recipient anonymity => relationship anonymity

sender unobservability => relationship unobservability

recipient unobservability => relationship unobservability

With respect to the same attacker, unobservability reveals always only a subset of the information undetectability reveals

unobservability => undetectability

9. Known Mechanisms for Anonymity, Undetectability, and Unobservability

Before it makes sense to speak about any particular mechanisms for anonymity, undetectability, and unobservability in communications, let us first remark that all of them assume that stations of users do not emit signals the attacker considered is able to use for identification of stations or their behavior or even for identification of users or their behavior. So if you travel around taking with you a mobile phone sending more or less continuously signals to update its location information within a cellular radio network, don't be surprised if you are tracked using its signals. If you use a computer emitting lots of radiation due to a lack of shielding, don't be surprised if observers using high-tech equipment know quite a bit about what's happening within your machine. If you use a computer, PDA, or smartphone without sophisticated access control, don't be surprised if Trojan horses send your secrets to anybody interested whenever you are online - or via electromagnetic emanations even if you think you are completely offline.

DC-net [Chau85], [Chau88], and MIX-net [Chau81] are mechanisms to achieve sender anonymity and relationship anonymity, respectively, both against strong attackers. If we add dummy traffic, both provide for the corresponding unobservability [PfPW91]. If dummy traffic is used to pad sending and/or receiving on the sender's and/or recipient's line to a constant rate traffic, MIX-nets can even provide sender and/or recipient anonymity and unobservability.

Broadcast [Chau85], [PfWa86], [Waid90] and private information retrieval [CoBi95] are mechanisms to achieve recipient anonymity against strong attackers. If we add dummy traffic, both provide for recipient unobservability.

This may be summarized: A mechanism to achieve some kind of anonymity appropriately combined with dummy traffic yields the corresponding kind of unobservability.

Of course, dummy traffic alone can be used to make the number and/or length of sent messages undetectable by everybody except for the recipients; respectively, dummy traffic can be used to make the number and/or length of received messages undetectable by everybody except for the senders. (Note: Misinformation and disinformation may be regarded as semantic dummy traffic, i.e., communication from which an attacker cannot decide which are real requests with real data or which are fake ones. Assuming the authenticity of misinformation or disinformation may lead to privacy problems for (innocent) bystanders.)

As a side remark, we mention steganography and spread spectrum as two

other well-known undetectability mechanisms.

The usual concept to achieve undetectability of IOIs at some layer, e.g., sending meaningful messages, is to achieve statistical independence of all discernible phenomena at some lower implementation layer. An example is sending dummy messages at some lower layer to achieve, e.g., a constant rate flow of messages looking - by means of encryption - randomly for all parties except the sender and the recipient(s).

10. Pseudonymity

Having anonymity of human beings, unlinkability, and maybe unobservability is superb w.r.t. data minimization, but would prevent any useful two-way communication. For many applications, we need appropriate kinds of identifiers:

Definition: A pseudonym is an identifier of a subject other than one of the subject's real names.

Note:

The term 'pseudonym' comes from the Greek word "pseudonumon" and means "falsely named" (pseudo: false; onuma: name). Thus, it means a name other than the 'real name'. To avoid the connotation of "pseudo" = false, some authors call pseudonyms as defined in this paper simply nyms. This is nice and short, but we stick with the usual wording, i.e., pseudonym, pseudonymity, etc. However

the reader should not be surprised to read nym, nymity, etc. in other texts.

An identifier is a name or another bit string. Identifiers, which are generated using random data only, i.e., fully independent of the subject and related attribute values, do not contain side information on the subject they are attached to, whereas non-random identifiers may do. E.g., nicknames chosen by a user may contain information on heroes he admires; a sequence number may contain information on the time the pseudonym was issued; an e-mail address or phone number contains information how to reach the user.

In our setting 'subject' means sender or recipient.

The term 'real name' is the antonym to "pseudonym". There may be multiple real names over lifetime, in particular the legal names, i.e., for a human being the names which appear on the birth certificate or on other official identity documents issued by the State; for a legal person the name under which it operates and

which is registered in official registers (e.g., commercial register or register of associations). A human being's real name typically comprises their given name and a family name. In the realm of identifiers, it is tempting to define anonymity as "the attacker cannot sufficiently determine a real name of the subject". But despite the simplicity of this definition, it is severely restricted: It can only deal with subjects which have at least one real name. It presumes that it is clear who is authorized to attach real names to subjects. It fails to work if the relation to real names is irrelevant for the application at hand. Therefore, we stick to the definitions given in Section 4. Note that from a mere technological perspective it cannot always be determined whether an identifier of a subject is a pseudonym or a real name.

We can generalize pseudonyms to be identifiers of sets of subjects - see below -, but we do not need this in our setting.

Definition: The subject which the pseudonym refers to is the holder of the pseudonym.

Definition: A subject is pseudonymous if a pseudonym is used as identifier instead of one of its real names.

We prefer the term "holder" over "owner" of a pseudonym because it seems to make no sense to "own" identifiers, e.g., bit strings. Furthermore, the term "holder" sounds more neutral than the term "owner", which is associated with an assumed autonomy of the subject's will. The holder may be a natural person (in this case we have the usual meaning and all data protection regulations apply), a legal person, or even only a computer.

Fundamentally, pseudonyms are nothing else than another kind of attribute values. But whereas in building an IT system, its designer can strongly support the holders of pseudonyms to keep the pseudonyms under their control, this is not equally possible w.r.t. attributes and attribute values in general. Therefore, it is useful to give this kind of attribute a distinct name: pseudonym.

For pseudonyms chosen by the user (in contrast to pseudonyms assigned to the user by others), primarily, the holder of the pseudonym is using it. Secondarily, all others he communicated to using the pseudonym can utilize it for linking. Each of them can, of course, divulge the pseudonym and all data related to it to other entities. So finally, the attacker will utilize the pseudonym to link all data related to this pseudonym he gets to

know being related.

Defining the process of preparing for the use of pseudonyms, e.g., by establishing certain rules how and under which conditions civil identities of holders of pseudonyms will be disclosed by so-called identity brokers or how to prevent uncovered claims by so-called liability brokers (cf. Section 11), leads to the more general notion of pseudonymity, as defined below.

Note:

Identity brokers have for the pseudonyms they are the identity broker for the information who is their respective holder. Therefore, identity brokers can be implemented as a special kind

of certification authorities for pseudonyms. Since anonymity can be described as a particular kind of unlinkability, cf. Section 6, the concept of identity broker can be generalized to linkability broker. A linkability broker is a (trusted) third party that, adhering to agreed rules, enables linking IOIs for those entities being entitled to get to know the linking.

Concerning the natural use of the English language, one might use "pseudonymization" instead of "pseudonymity". But at least in Germany, the law makers gave "pseudonymization" the meaning that first personal data known by others comprise some identifiers for the civil identity and later these identifiers are replaced by pseudonyms. Therefore, we use a different term (coined by David Chaum: "pseudonymity") to describe that from the very beginning pseudonyms are used.

Definition: Pseudonymity is the use of pseudonyms as identifiers.

Note:

From [ISO99]: "Pseudonymity ensures that a user may use a resource or service without disclosing its user identity, but can still be accountable for that use. [...] Pseudonymity requires that a set of users and/or subjects are unable to determine the identity of a user bound to a subject or operation, but that this user is still accountable for its actions." This view on pseudonymity covers only the use of digital pseudonyms. Therefore, our definition of pseudonymity is much broader as it does not necessarily require disclosure of the user's identity and accountability. Pseudonymity alone - as it is used in the real world and in technological contexts - does not tell anything about the strengths of anonymity, authentication or accountability; these strengths depend on several properties, cf. below.

Quantifying pseudonymity would primarily mean quantifying the state of using a pseudonym according to its different dimensions (cf. Section 11 and Section 12), i.e., quantifying the authentication and accountability gained and quantifying the anonymity left over (e.g., using entropy as the measure). Roughly speaking, well-employed pseudonymity could mean in e-commerce appropriately fine-grained authentication and accountability to

counter identity theft or to prevent uncovered claims using, e.g., the techniques described in [BuPf90], combined with much anonymity retained. Poorly employed pseudonymity would mean giving away anonymity without preventing uncovered claims.

So sender pseudonymity is defined as the sender being pseudonymous, recipient pseudonymity is defined as the recipient being pseudonymous, see Figure 7. Providing sender pseudonymity and recipient pseudonymity is the basic interface communication networks have to provide to enhance privacy for two-way communications.

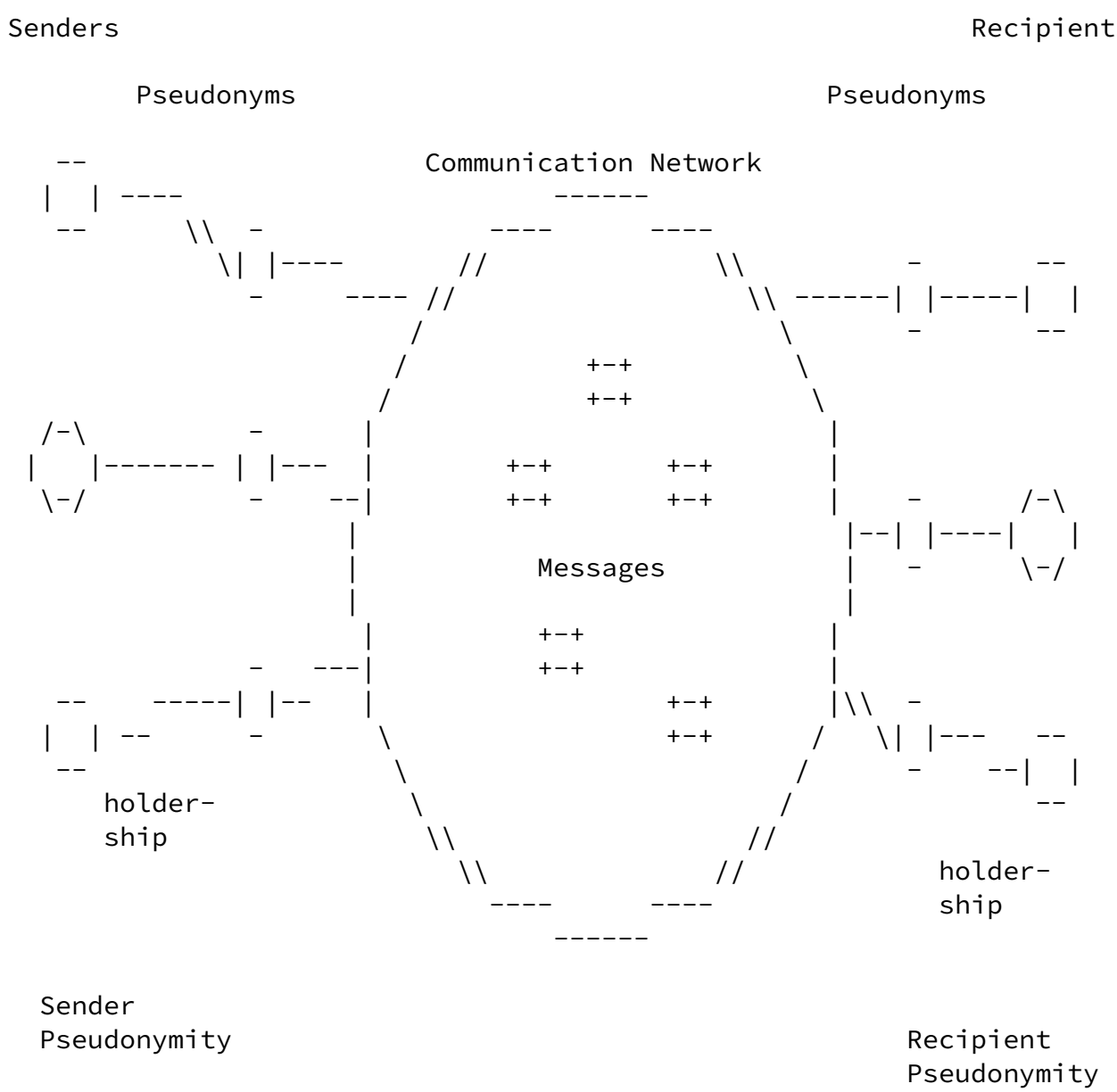


Figure 7: Pseudonymity

In our usual setting, we assume that each pseudonym refers to exactly one specific holder, invariant over time.

Specific kinds of pseudonyms may extend this setting: A group pseudonym refers to a set of holders, i.e., it may refer to multiple holders; a transferable pseudonym can be transferred from one holder to another subject becoming its holder.

Such a group pseudonym may induce an anonymity set: Using the information provided by the pseudonym only, an attacker cannot decide whether an action was performed by a specific subject within the set. Please note that the mere fact that a pseudonym has several holders

does not yield a group pseudonym: For instance, creating the same pseudonym may happen by chance and even without the holders being aware of this fact, particularly if they choose the pseudonyms and prefer pseudonyms which are easy to remember. But the context of each use of the pseudonym (e.g., used by which subject - usually denoted by another pseudonym - in which kind of transaction) then usually will denote a single holder of this pseudonym.

Transferable pseudonyms can, if the attacker cannot completely monitor all transfers of holdership, serve the same purpose, without decreasing accountability as seen by an authority monitoring all transfers of holdership.

An interesting combination might be transferable group pseudonyms - but this is left for further study.

11. Pseudonymity with respect to accountability and authorization

11.1. Digital pseudonyms to authenticate messages

A digital pseudonym is a bit string which, to be meaningful in a certain context, is

- unique as identifier (at least with very high probability) and
- suitable to be used to authenticate the holder's IOIs relatively to his/her digital pseudonym, e.g., to authenticate his/her messages sent.

Using digital pseudonyms, accountability can be realized with pseudonyms - or more precisely: with respect to pseudonyms.

11.2. Accountability for digital pseudonyms

To authenticate IOIs relative to pseudonyms usually is not enough to achieve accountability for IOIs.

Therefore, in many situations, it might make sense to either

- attach funds to digital pseudonyms to cover claims or to
- let identity brokers authenticate digital pseudonyms (i.e., check the civil identity of the holder of the pseudonym and then issue a

digitally signed statement that this particular identity broker has proof of the identity of the holder of this digital pseudonym and is willing to divulge that proof under well-defined circumstances) or

o both.

Note:

If the holder of the pseudonym is a natural person or a legal person, civil identity has the usual meaning, i.e. the identity attributed to that person by a State (e.g., a natural person being represented by the social security number or the combination of name, date of birth, and location of birth etc.). If the holder is, e.g., a computer, it remains to be defined what "civil identity" should mean. It could mean, for example, exact type and serial number of the computer (or essential components of it) or even include the natural person or legal person responsible for its operation.

If sufficient funds attached to a digital pseudonym are reserved and/or the digitally signed statement of a trusted identity broker is checked before entering into a transaction with the holder of that pseudonym, accountability can be realized in spite of anonymity.

11.3. Transferring authenticated attributes and authorizations between pseudonyms

To transfer attributes including their authentication by third parties (called "credentials" by David Chaum [Chau85]) - all kinds of authorizations are special cases - between digital pseudonyms of one and the same holder, it is always possible to prove that these pseudonyms have the same holder.

But as David Chaum pointed out, it is much more anonymity-preserving to maintain the unlinkability of the digital pseudonyms involved as much as possible by transferring the credential from one pseudonym to the other without proving the sameness of the holder. How this can be done is described in [Chau90] [CaLy04].

We will come back to the just described property "convertibility" of

digital pseudonyms in Section 13.

12. Pseudonymity with respect to linkability

Whereas anonymity and accountability are the extremes with respect to linkability to subjects, pseudonymity is the entire field between and including these extremes. Thus, pseudonymity comprises all degrees of linkability to a subject. Ongoing use of the same pseudonym allows the holder to establish or consolidate a reputation. Establishing and/or consolidating a reputation under a pseudonym is, of course, insecure if the pseudonym does not enable to authenticate messages, i.e., if the pseudonym is not a digital pseudonym, cf.

Section 11.1. Then, at any moment, another subject might use this pseudonym possibly invalidating the reputation, both for the holder of the pseudonym and all others having to do with this pseudonym. Some kinds of pseudonyms enable dealing with claims in case of abuse of unlinkability to holders: Firstly, third parties (identity brokers, cf.) may have the possibility to reveal the civil identity of the holder in order to provide means for investigation or prosecution. To improve the robustness of anonymity, chains of identity brokers may be used [Chau81]. Secondly, third parties may act as liability brokers of the holder to clear a debt or settle a claim. [BuPf90] presents the particular case of value brokers.

There are many properties of pseudonyms which may be of importance in specific application contexts. In order to describe the properties of pseudonyms with respect to anonymity, we limit our view to two aspects and give some typical examples:

12.1. Knowledge of the linking between the pseudonym and its holder

The knowledge of the linking may not be a constant, but change over time for some or even all people. Normally, for non-transferable pseudonyms the knowledge of the linking cannot decrease (with the exception of misinformation or disinformation, which may blur the attacker's knowledge.). Typical kinds of such pseudonyms are:

Public pseudonym: The linking between a public pseudonym and its holder may be publicly known even from the very beginning. E.g., the linking could be listed in public directories such as the entry of a phone number in combination with its owner.

Initially non-public pseudonym: The linking between an initially non-public pseudonym and its holder may be known by certain parties, but is not public at least initially. E.g., a bank account where the bank can look up the linking may serve as a non-public pseudonym. For some specific non-public pseudonyms, certification authorities acting as identity brokers could reveal the civil identity of the holder in case of abuse.

Initially unlinked pseudonym: The linking between an initially unlinked pseudonym and its holder is - at least initially - not known to anybody with the possible exception of the holder himself/herself. Examples for unlinked pseudonyms are (non-public) biometrics like DNA information unless stored in databases including the linking to the holders.

Public pseudonyms and initially unlinked pseudonyms can be seen as extremes of the described pseudonym aspect whereas initially non-public pseudonyms characterize the continuum in between.

Anonymity is the stronger, the less is known about the linking to a subject. The strength of anonymity decreases with increasing knowledge of the pseudonym linking. In particular, under the assumption that no gained knowledge on the linking of a pseudonym will be forgotten and that the pseudonym cannot be transferred to other subjects, a public pseudonym never can become an unlinked pseudonym. In each specific case, the strength of anonymity depends on the knowledge of certain parties about the linking relative to the chosen attacker model.

If the pseudonym is transferable, the linking to its holder can change. Considering an unobserved transfer of a pseudonym to another subject, a formerly public pseudonym can become non-public again.

12.2. Linkability due to the use of a pseudonym across different contexts

With respect to the degree of linkability, various kinds of pseudonyms may be distinguished according to the kind of context for their usage:

Person pseudonym: A person pseudonym is a substitute for the

holder's name which is regarded as representation for the holder's civil identity. It may be used in many different contexts, e.g., a number of an identity card, the social security number, DNA, a nickname, the pseudonym of an actor, or a mobile phone number.

Role pseudonym: The use of role pseudonyms is limited to specific roles, e.g., a customer pseudonym or an Internet account used for many instantiations of the same role "Internet user". See Section 14.3 for a more precise characterization of the term "role". The same role pseudonym may be used with different communication partners. Roles might be assigned by other parties, e.g., a company, but they might be chosen by the subject himself/herself as well.

Relationship pseudonym: For each communication partner, a different relationship pseudonym is used. The same relationship pseudonym may be used in different roles for communicating with the same partner. Examples are distinct nicknames for each communication partner. In case of group communication, the relationship pseudonyms may be used between more than two partners.

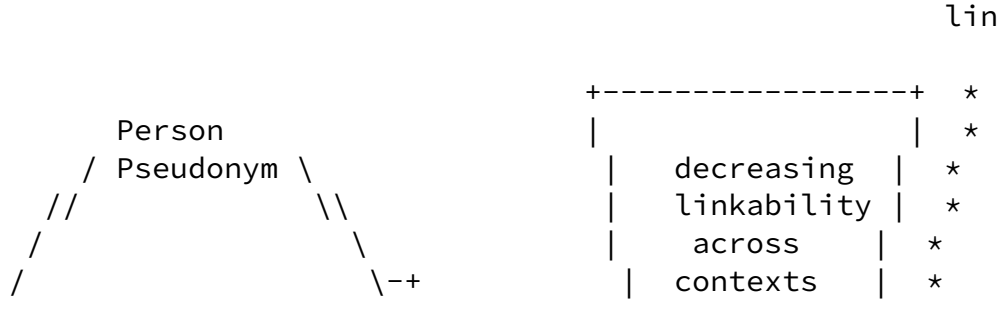
Role-relationship pseudonym: For each role and for each communication partner, a different role-relationship pseudonym is used. This means that the communication partner does not necessarily know, whether two pseudonyms used in different roles belong to the same holder. On the other hand, two different

communication partners who interact with a user in the same role, do not know from the pseudonym alone whether it is the same user. As with relationship pseudonyms, in case of group communication, the role-relationship pseudonyms may be used between more than two partners.

Transaction pseudonym: Apart from "transaction pseudonym" some employ the term "one-time-use pseudonym", taking the naming from "one-time pad". For each transaction, a transaction pseudonym unlinkable to any other transaction pseudonyms and at least initially unlinkable to any other IOI is used, e.g., randomly generated transaction numbers for online-banking. Therefore, transaction pseudonyms can be used to realize as strong anonymity as possible. In fact, the strongest anonymity is given when there is no identifying information at all, i.e., information that would

allow linking of anonymous entities, thus transforming the anonymous transaction into a pseudonymous one. If the transaction pseudonym is used exactly once, we have the same strength of anonymity as if no pseudonym is used at all. Another possibility to achieve strong anonymity is to prove the holdership of the pseudonym or specific attribute values (e.g., with zero-knowledge proofs) without revealing the information about the pseudonym or more detailed attribute values themselves. Then, no identifiable or linkable information is disclosed.

Linkability across different contexts due to the use of these pseudonyms can be represented as the lattice that is illustrated in the following diagram, see Figure 8. The arrows point in direction of increasing unlinkability, i.e., A -> B stands for "B enables stronger unlinkability than A". Note that "->" is not the same as "=>" of Section 8, which stands for the implication concerning anonymity and unobservability.



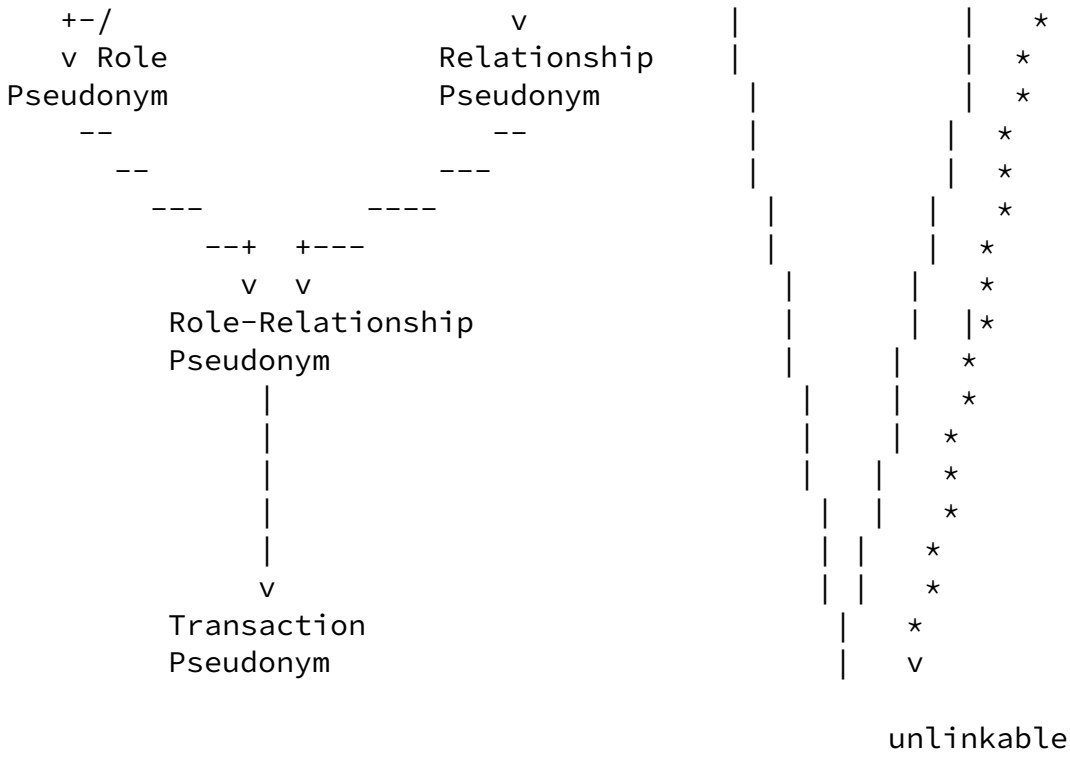


Figure 8: Lattice of pseudonyms according to their use across different contexts

In general, unlinkability of both role pseudonyms and relationship pseudonyms is stronger than unlinkability of person pseudonyms. The strength of unlinkability increases with the application of role-relationship pseudonyms, the use of which is restricted to both the same role and the same relationship. If a role-relationship pseudonym is used for roles comprising many kinds of activities, the danger arises that after a while, it becomes a person pseudonym in the sense of: "A person pseudonym is a substitute for the holder's name which is regarded as representation for the holder's civil identity." This is even more true both for role pseudonyms and relationship pseudonyms. Ultimate strength of unlinkability is obtained with transaction pseudonyms, provided that no other information, e.g., from the context or from the pseudonym itself, enabling linking is available.

Anonymity is the stronger, ...

- the less personal data of the pseudonym holder can be linked to the pseudonym;

- the less often and the less context-spanning pseudonyms are used and therefore the less data about the holder can be linked;

- the more often independently chosen, i.e., from an observer's perspective unlinkable, pseudonyms are used for new actions.

The amount of information of linked data can be reduced by different subjects using the same pseudonym (e.g., one after the other when pseudonyms are transferred or simultaneously with specifically created group pseudonyms) or by misinformation or disinformation. The group of pseudonym holders acts as an inner anonymity set within a, depending on context information, potentially even larger outer anonymity set.

13. Known mechanisms and other properties of pseudonyms

A digital pseudonym could be realized as a public key to test digital signatures where the holder of the pseudonym can prove holdership by forming a digital signature which is created using the corresponding private key [Chau81]. The most prominent example for digital pseudonyms are public keys generated by the user himself/herself, e.g., using PGP. In using PGP, each user may create an unlimited number of key pairs by himself/herself (at this moment, such a key pair is an initially unlinked pseudonym), bind each of them to an e-mail address, self-certify each public key by using his/her digital signature or asking another introducer to do so, and circulate it.

A public key certificate bears a digital signature of a so-called certification authority and provides some assurance to the binding of a public key to another pseudonym, usually held by the same subject. In case that pseudonym is the civil identity (the real name) of a subject, such a certificate is called an identity certificate. An attribute certificate is a digital certificate which contains further information (attribute values) and clearly refers to a specific public key certificate. Independent of certificates, attributes may be used as identifiers of sets of subjects as well. Normally, attributes refer to sets of subjects (i.e., the anonymity set), not to one specific subject.

There are several other properties of pseudonyms related to their use which shall only be briefly mentioned, but not discussed in detail in this text. They comprise different degrees of, e.g.,

- limitation to a fixed number of pseudonyms per subject [Chau81], [Chau85], [Chau90]. For pseudonyms issued by an agency that guarantees the limitation of at most one pseudonym per individual person, the term "is-a-person pseudonym" is used.

- guaranteed uniqueness [Chau81] [StSy00], e.g., "globally unique pseudonyms".

- transferability to other subjects.

- authenticity of the linking between a pseudonym and its holder (possibilities of verification/falsification or indication/ repudiation).

- provability that two or more pseudonyms have the same holder. For digital pseudonyms having only one holder each and assuming that no holders cooperate to provide wrong "proofs", this can be proved trivially by signing, e.g., the statement "<Pseudonym1> and <Pseudonym2> have the same holder." digitally with respect to both these pseudonyms. Putting it the other way round: Proving that pseudonyms have the same holder is all but trivial.

- convertibility, i.e., transferability of attributes of one pseudonym to another [Chau85], [Chau90]. This is a property of convertible credentials.

- possibility and frequency of pseudonym changeover.

- re-usability and, possibly, a limitation in number of uses.

- validity (e.g., guaranteed durability and/or expiry date, restriction to a specific application).

- possibility of revocation or blocking.

- participation of users or other parties in forming the pseudonyms.

- information content about attributes in the pseudonym itself.

In addition, there may be some properties for specific applications (e.g., an addressable pseudonym serves as a communication address which enables to contact its holder) or due to the participation of third parties (e.g., in order to circulate the pseudonyms, to reveal civil identities in case of abuse, or to cover claims).

Some of the properties can easily be realized by extending a digital

pseudonym by attributes of some kind, e.g., a communication address, and specifying the appropriate semantics. The binding of attributes

to a pseudonym can be documented in an attribute certificate produced either by the holder himself/herself or by a certification authority. The non-transferability of the attribute certificate can be somewhat enforced, e.g., by biometrical means, by combining it with individual hardware (e.g., chipcards), or by confronting the holder with legal consequences.

14. Identity management

14.1. Setting

To adequately address privacy-enhancing identity management, we have to extend our setting:

- It is not realistic to assume that an attacker might not get information on the sender or recipient of messages from the message content and/or the sending or receiving context (time, location information, etc.) of the message. We have to consider that the attacker is able to use these attributes for linking messages and, correspondingly, the pseudonyms used with them.

- In addition, it is not just human beings, legal persons, or simply computers sending messages and using pseudonyms at their discretion as they like at the moment, but they use (computer-based) applications, which strongly influence the sending and receiving of messages and may even strongly determine the usage of pseudonym.

14.2. Identity and identifiability

Identity can be explained as an exclusive perception of life, integration into a social group, and continuity, which is bound to a body and - at least to some degree - shaped by society. This concept of identity distinguishes between "I" and "Me" [Mead34] : "I" is the instance that is accessible only by the individual self, perceived as an instance of liberty and initiative. "Me" is supposed to stand for the social attributes, defining a human identity that is accessible by communications and that is an inner instance of control and consistency (see [ICPP03] for more information). In this

terminology, we are interested in identity as communicated to others and seen by them. Therefore, we concentrate on the "Me".

Note:

Here (and in Section 14 throughout), we have human beings in mind, which is the main motivation for privacy. From a structural point of view, identity can be attached to any subject, be it a human being, a legal person, or even a computer. This makes the

terminology more general, but may lose some motivation at first sight. Therefore, we start in our explanation with identity of human beings, but implicitly generalize to subjects thereafter. This means: In a second reading of this paper, you may replace "individual person" by "individual subject" throughout as it was used in the definitions of the Section 3 through Section 13. It may be discussed whether the definitions can be further generalized and apply for any "entity", regardless of subject or object.

According to Mireille Hildebrandt, the French philosopher Paul Ricoeur made a distinction between "idem and ipse. Idem (sameness) stands for the third person, objectified observer's perspective of identity as a set of attributes that allows comparison between different people, as well as unique identification, whereas ipse (self) stands for the first person perspective constituting a 'sense of self'.", see page 274 in [RaRD09]. So what George H. Mead called "I" is similar to what Paul Ricoeur called "ipse" (self). What George H. Mead called "Me" is similar to what Paul Ricoeur called "idem" (sameness).

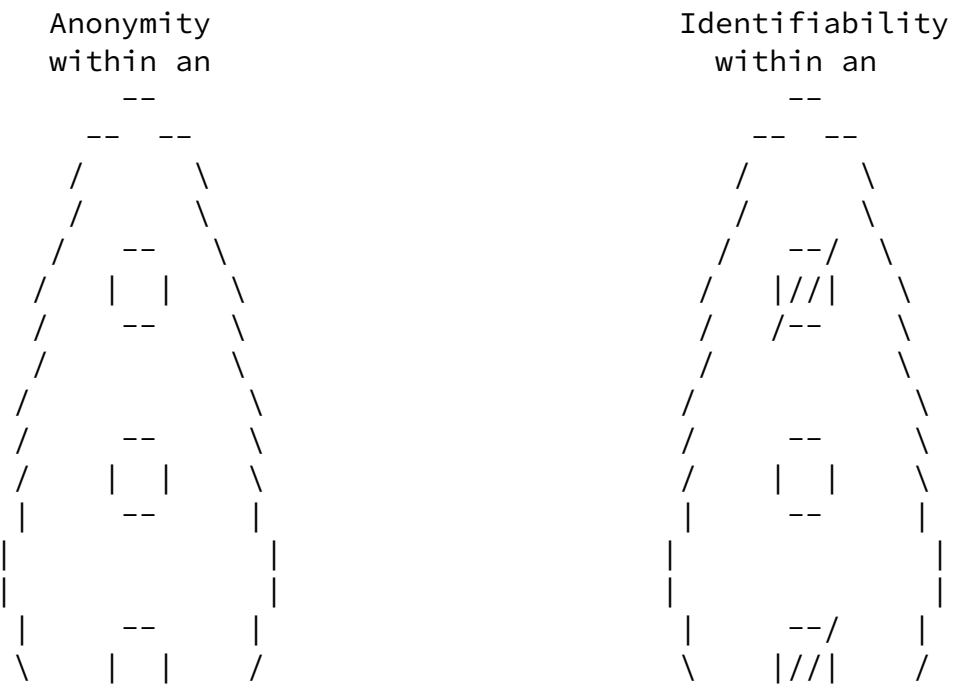
Motivated by identity as an exclusive perception of life, i.e., a psychological perspective, but using terms defined from a computer science, i.e., a mathematical perspective (as we did in the sections before), identity can be explained and defined as a property of an entity in terms of the opposite of anonymity and the opposite of unlinkability. In a positive wording, identity enables both to be identifiable as well as to link IOIs because of some continuity of life. Here we have the opposite of anonymity (identifiability) and the opposite of unlinkability (linkability) as positive properties. So the perspective changes: What is the aim of an attacker w.r.t. anonymity, now is the aim of the subject under consideration, so the

attacker's perspective becomes the perspective of the subject. And again, another attacker (attacker2) might be considered working against identifiability and/or linkability. I.e., attacker2 might try to mask different attributes of subjects to provide for some kind of anonymity or attacker2 might spoof some messages to interfere with the continuity of the subject's life.

Corresponding to the anonymity set introduced in the beginning of this text, we can work with an "identifiability set" [Hild03], which is the set is a set of possible subjects, to define "identifiability" and "identity". This definition is compatible with the definitions given in [HoWi03] and it is very close to that given by [Chi03]: "An identity is any subset of attributes of a person which uniquely characterizes this person within a community."

Definition: Identifiability of a subject from an attacker's perspective means that the attacker can sufficiently identify the subject within a set of subjects, the identifiability set.

Figure 9 contrasts anonymity set and identifiability set.



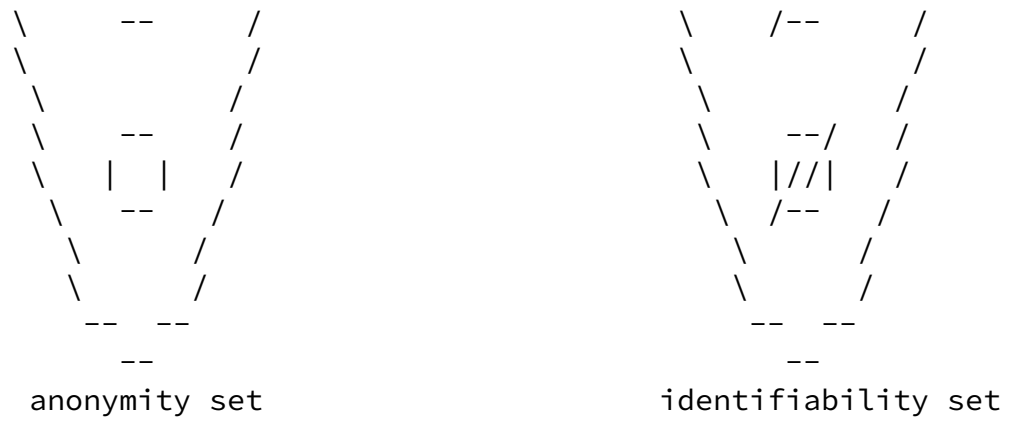


Figure 9: Anonymity set vs. identifiability set

All other things being equal, identifiability is the stronger, the larger the respective identifiability set is. Conversely, the remaining anonymity is the stronger, the smaller the respective identifiability set is.

Identity of an individual person should be defined independent of an attacker's perspective:

Definition: An identity is any subset of attribute values of an individual person which sufficiently identifies this individual person within any set of persons. So usually there is no such thing as "the identity", but several of them.

Note:

Whenever we speak about "attribute values" in this text, this shall comprise not only a measurement of the attribute value, but the attribute as well. E.g., if we talk about the attribute "color of one's hair" the attribute value "color of one's hair" is not just, e.g., "grey", but ("color of one's hair", "grey").

An equivalent, but slightly longer definition of identity would be: An identity is any subset of attribute values of an individual person which sufficiently distinguishes this individual person from all other persons within any set of persons.

Of course, attribute values or even attributes themselves may change over time. Therefore, if the attacker has no access to the change history of each particular attribute, the fact whether a particular subset of attribute values of an individual person is an identity or not may change over time as well. If the attacker has access to the change history of each particular attribute, any subset forming an identity will form an identity from his perspective irrespective how attribute values change. Any reasonable attacker will not just try to figure out attribute values per se, but the point in time (or even the time frame) they are valid (in), since this change history helps a lot in linking and thus inferring further attribute values. Therefore, it may clarify one's mind to define each "attribute" in a way that its value cannot get invalid. So instead of the attribute "location" of a particular individual person, take the set of attributes "location at time x". Depending on the inferences you are interested in, refining that set as a list ordered concerning "location" or "time" may be helpful.

Identities may of course comprise particular attribute values like names, identifiers, digital pseudonyms, and addresses - but they don't have to.

14.3. Identity-related terms

Role: In sociology, a "role" or "social role" is a set of connected actions, as conceptualized by actors in a social situation (i.e., situation-dependent identity attributes). It is mostly defined as an expected behavior (i.e., sequences of actions) in a given social context. So roles provide for some linkability of actions.

Partial identity: An identity of an individual person may comprise many partial identities of which each represents the person in a specific context or role. (Note: As an identity has to do with integration into a social group, on the one hand, partial identities have to do with, e.g., relationships to particular group members (or to be more general: relationships to particular subsets of group members). On the other hand, partial identities might be associated with relationships to organizations.) A partial identity is a subset of attribute values of a complete identity, where a complete identity is the union of all attribute values of all identities of this person. (Note: If attributes are

defined such that their values do not get invalid, "union" can have the usual meaning within set theory. We have to admit that usually nobody, including the person concerned, will know "all" attribute values or "all" identities. Nevertheless we hope that the notion "complete identity" will ease the understanding of "identity" and "partial identity".) On a technical level, these attribute values are data. Of course, attribute values or even attributes themselves of a partial identity may change over time. As identities, partial identities may comprise particular attribute values like names, identifiers, digital pseudonyms, and addresses - but they don't have to, either. A pseudonym might be an identifier for a partial identity. If it is possible to transfer attribute values of one pseudonym to another (as convertibility of credentials provides for, cf. Section 13), this means transferring a partial identity to this other pseudonym. Re-use of the partial identity with its identifier(s), e.g., a pseudonym, supports continuity in the specific context or role by enabling linkability with, e.g., former or future messages or actions. If the pseudonym is a digital pseudonym, it provides the possibility to authenticate w.r.t. the partial identity which is important to prevent others to take over the partial identity (discussed as "identity theft"). Linkability of partial identities arises by non-changing identifiers of a partial identity as well as other attribute values of that partial identity that are (sufficiently) static or easily determinable over time (e.g., bodily biometrics, the size or age of a person). All the data that can be used to link data sets such as partial identities belong to a category of "data providing linkability" (to which we must pay the same attention as to personal data w.r.t. privacy and data protection; "protection of individuals with regard to the processing of personal data" [DPD95]). Whereas we assume that an "identity" sufficiently identifies an individual person (without limitation to particular identifiability sets), a partial identity may not do, thereby enabling different quantities of anonymity. So we may have linkability by re-using a partial identity (which may be important to support continuity of life) without necessarily giving up anonymity (which may be important

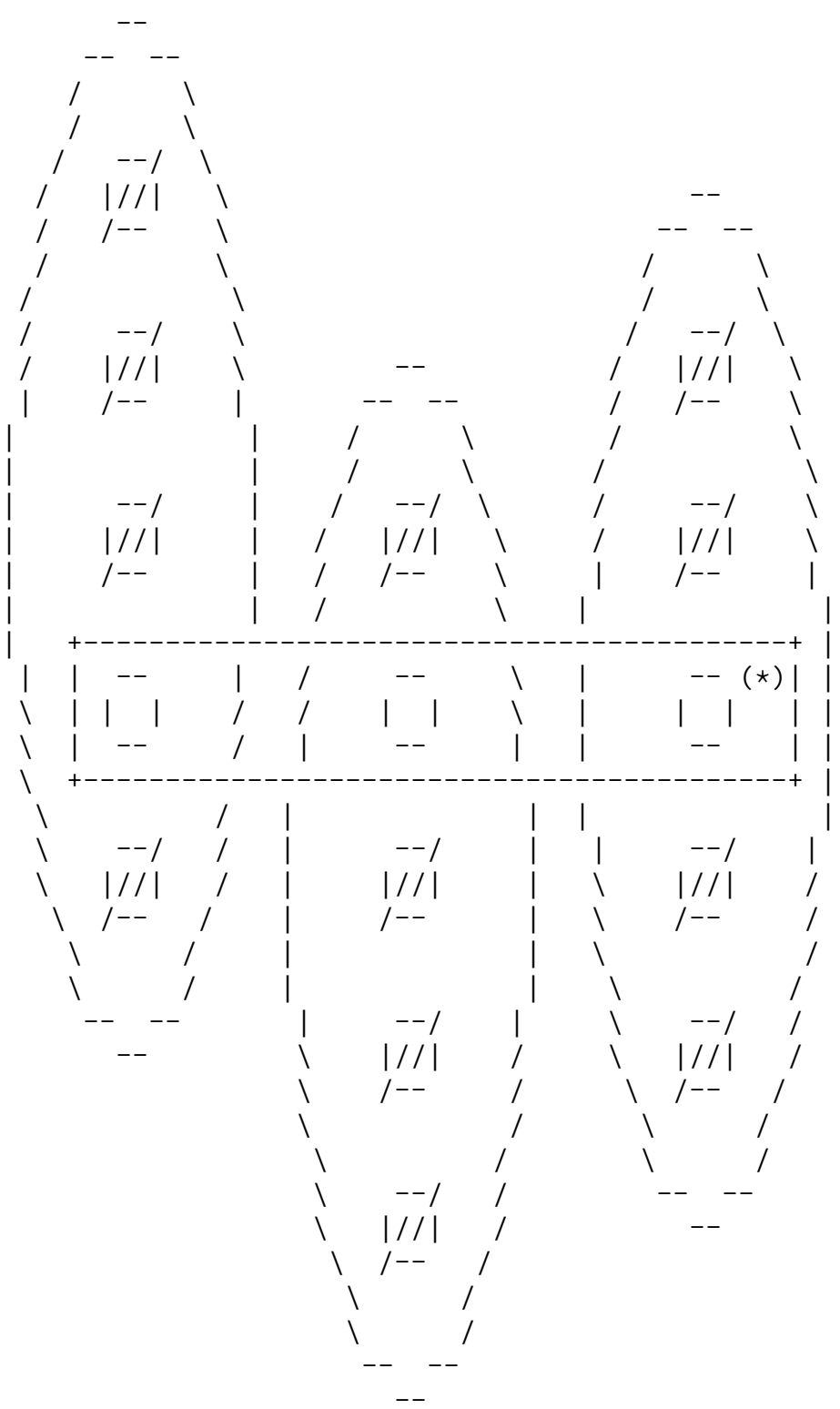
for privacy). But we may find for each partial identity appropriately small identifiability sets, where the partial identity sufficiently identifies an individual person, see Figure 10. For identifiability sets of cardinality 1, this is

trivial, but it may hold for "interesting" identifiability sets of larger cardinality as well. The relation between anonymity set and identifiability set can be seen in two ways:

1. Within an a-priori anonymity set, we can consider a-posteriori identifiability sets as subsets of the anonymity set. Then the largest identifiability sets allowing identification characterize the a-posteriori anonymity, which is zero iff the largest identifiability set allowing identification equals the a-priori anonymity set.

2. Within an a-priori identifiability set, its subsets which are the a-posteriori anonymity sets characterize the a-posteriori anonymity. It is zero iff all a-posteriori anonymity sets have cardinality 1.

As with identities, depending on whether the attacker has access to the change history of each particular attribute or not, the identifiability set of a partial identity may change over time if the values of its attributes change.



*: Anonymity set of a partial identity given
that the set of all possible subjects
(the a-priori anonymity set) can be partitioned
into the three disjoint identifiability sets
of the partial identity shown.

Figure 10: Relation between anonymity set and identifiability set

Digital identity Digital identity denotes attribution of attribute values to an individual person, which are immediately operationally accessible by technical means. More to the point, the identifier of a digital partial identity can be a simple e-mail address in a news group or a mailing list. A digital partial identity is the same as a partial digital identity. In the following, we skip "partial" if the meaning is clear from the context. Its owner will attain a certain reputation. More generally we might consider the whole identity as a combination from "I" and "Me" where the "Me" can be divided into an implicit and an explicit part: Digital identity is the digital part from the explicated "Me". Digital identity should denote all those personal data that can be stored and automatically interlinked by a computer-based application.

Virtual identity Virtual identity is sometimes used in the same meaning as digital identity or digital partial identity, but because of the connotation with "unreal, non-existent, seeming" the term is mainly applied to characters in a MUD (Multi User Dungeon), MMORPG (Massively Multiplayer Online Role Playing Game) or to avatars. For these reasons, we do not use the notions physical world vs. virtual world nor physical person vs. virtual person defined in [RaRD09] (pp. 80ff). Additionally, we feel that taking the distinction between physical vs. digital (=virtual) world as a primary means to build up a terminology is not helpful. First we have to define what a person and an identity is. The distinction between physical and digital is only of secondary importance and the structure of the terminology should reflect this fundamental fact. In other disciplines, of course, it may be very relevant whether a person is a human being with a physical body. Please remember Section 14.3, where the sociological definition of identity includes "is bound to a body", or law enforcement when a jail sentence has to be carried out. Generalizing from persons, laws should consider and spell out whether they are addressing physical entities, which cannot be duplicated easily, or digital entities, which can.

14.4. Identity management-related terms

Identity management Identity management means managing various partial identities (usually denoted by pseudonyms) of an individual person, i.e., administration of identity attributes including the development and choice of the partial identity and pseudonym to be (re-)used in a specific context or role. Establishment of reputation is possible when the individual person re-uses partial identities. A prerequisite to choose the

appropriate partial identity is to recognize the situation the person is acting in.

Privacy-enhancing identity management Given the restrictions of a set of applications, identity management is called privacy-enhancing if it sufficiently preserves unlinkability (as seen by an attacker) between the partial identities of an individual person required by the applications. Note that due to our setting, this definition focuses on the main property of Privacy-Enhancing Technologies (PETs), namely data minimization: This property means to limit as much as possible the release of personal data and for those released, preserve as much unlinkability as possible. We are aware of the limitation of this definition: In the real world it is not always desired to achieve utmost unlinkability. We believe that the user as the data subject should be empowered to decide on the release of data and on the degree of linkage of his or her personal data within the boundaries of legal regulations, i.e., in an advanced setting the privacy-enhancing application design should also take into account the support of "user-controlled release" as well as "user-controlled linkage". Identity management is called perfectly privacy-enhancing if it perfectly preserves unlinkability between the partial identities, i.e., by choosing the pseudonyms (and their authorizations, cf. Section 11.3) denoting the partial identities carefully, it maintains unlinkability between these partial identities towards an attacker to the same degree as giving the attacker the attribute values with all pseudonyms omitted. (Note: Given the terminology defined in Section 3 to Section 6, privacy-enhancing identity management is unlinkability-preserving identity management. So, maybe, the term "privacy-preserving identity management" would be more appropriate. But to be compatible to the earlier papers in this field, we stick to privacy-enhancing identity management.)

Privacy-enhancing identity management enabling application design An application is designed in a privacy-enhancing identity management enabling way if neither the pattern of sending/receiving messages nor the attribute values given to subjects (i.e., human beings, organizations, computers) reduce unlinkability more than is strictly necessary to achieve the purposes of the application.

User-controlled identity management Identity management is called user-controlled if the flow of this user's identity attribute values is explicit to the user and the user is in control of this flow.

Identity management system (IMS) An identity management system supports administration of identity attributes including the development and choice of the partial identity and pseudonym to be (re-)used in a specific context or role. Note that some publications use the abbreviations IdMS or IDMS instead. We can distinguish between identity management system and identity management application: The term "identity management system" is seen as an infrastructure, in which "identity management applications" as components, i.e., software installed on computers, are co-ordinated.

Privacy-enhancing identity management system (PE-IMS) A Privacy-Enhancing IMS is an IMS that, given the restrictions of a set of applications, sufficiently preserves unlinkability (as seen by an attacker) between the partial identities and corresponding pseudonyms of an individual person.

User-controlled identity management system A user-controlled identity management system is an IMS that makes the flow of this user's identity attribute values explicit to the user and gives its user control of this flow [CPHH02]. The guiding principle is "notice and choice".

Combining user-controlled IMS with PE-IMS means user-controlled linkability of personal data, i.e., achieving user-control based on thorough data minimization. According to respective situation

and context, such a system supports the user in making an informed choice of pseudonyms, representing his or her partial identities. A user-controlled PE-IMS supports the user in managing his or her partial identities, i.e., to use different pseudonyms with associated identity attribute values according to different contexts, different roles the user is acting in and according to different interaction partners. It acts as a central gateway for all interactions between different applications, like browsing the web, buying in Internet shops, or carrying out administrative tasks with governmental authorities [HBCC04].

15. Overview of main definitions and their opposites

o

o

Definition	Negation
Anonymity of a subject from an attacker's perspective means that the attacker cannot sufficiently identify the subject within a set of subjects, the anonymity set.	Identifiability of a subject from an attacker's perspective means that the attacker can sufficiently identify the subject within a set of subjects, the identifiability set.
Unlinkability of two or more items of interest (IOIs, e.g., subjects, messages, actions, ...) from an attacker's	Linkability of two or more items of interest (IOIs, e.g., subjects, messages, actions, ...) from an attacker's

perspective means that within the system (comprising these and possibly other items), the attacker cannot sufficiently distinguish whether these IOIs are related or not.	perspective means that within the system (comprising these and possibly other items), the attacker can sufficiently distinguish whether these IOIs are related or not.
Undetectability of an item of interest (IOI) from an attacker's perspective means that the attacker cannot sufficiently distinguish whether it exists or not.	Detectability of an item of interest (IOI) from an attacker's perspective means that the attacker can sufficiently distinguish whether it exists or not.
Unobservability of an item of interest (IOI) means undetectability of the IOI against all subjects uninvolved in it and anonymity of the subject(s) involved in the IOI even against the other subject(s) involved in that IOI.	Observability of an item of interest (IOI) means "many possibilities to define the semantics".

16. Acknowledgments

Before this document was submitted to the IETF it already had a long history starting at 2000 and a number of people helped to improve the quality of the document with their feedback. The original authors, Marit Hansen and Andreas Pfitzmann, would therefore like to thank Adam Shostack, David-Olivier Jaquet-Chiffelle, Claudia Diaz, Giles Hogben, Thomas Kriegelstein, Wim Schreurs, Sandra Steinbrecher, Mike Bergmann, Katrin Borcea, Simone Fischer-Huebner, Stefan Koepsell, Martin Rost, Marc Wilikens, Adolf Flueli, Jozef Vyskoc, Thomas Kriegelstein, Jan Camenisch, Vashek Matyas, Daniel Cvrcek, Wassim Haddad, Alf Zugenmair, Katrin Borcea-Pfitzmann, Thomas Kriegelstein,

Elke Franz, Sebastian Clauss, Neil Mitchison, Rolf Wendolsky, Stefan Schiffner, Maritta Heisel, Katja Liesebach, Stefanie Poetzsch, Thomas Santen, Maritta Heisel, Manuela Berg, Katrin Borcea-Pfitzmann, and Katie Tietze for their input.

The terminology has been translated to other languages and the result can be found here: http://dud.inf.tu-dresden.de/Anon_Terminology.shtml.

17. References

17.1. Normative References

[RFC2119] Bradner, S., "Key words for use in RFCs to Indicate Requirement Levels", BCP 14, RFC 2119, March 1997.

17.2. Informative References

[BuPf90] Buerk, H. and A. Pfitzmann, "Value Exchange Systems Enabling Security and Unobservability", Computers & Security , 9/8, 715-721, January 1990.

[CPHH02] Clauss, S., Pfitzmann, A., Hansen, M., and E. Herreweghen, "Privacy-Enhancing Identity Management", IEEE Symposium on Research in Security and Privacy , IPTS Report 67, 8-16, September 2002.

[CaLy04] Camenisch, J. and A. Lysyanskaya, "Signature Schemes and Anonymous Credentials from Bilinear Maps", Crypto , LNCS 3152, Springer, Berlin 2004, 56-72, 2004.

[Chau81] Chaum, D., "Untraceable Electronic Mail, Return Addresses, and Digital Pseudonyms", Communications of the ACM , 24/2, 84-88, 1981.

[Chau85] Chaum, D., "Security without Identification: Transaction Systems to make Big Brother Obsolete", Communications of the ACM , 28/10, 1030-1044, 1985.

[Chau88] Chaum, D., "The Dining Cryptographers Problem: Unconditional Sender and Recipient Untraceability",

Journal of Cryptology , 1/1, 65-75, 1988.

[Chau90] Chaum, D., "Showing credentials without identification: Transferring signatures between unconditionally unlinkable pseudonyms", Auscrypt , LNCS 453, Springer, Berlin 1990, 246-264, 1990.

[Chi03] Jaquet-Chiffelle, D., "Towards the Identity", Presentation at the the Future of IDentity in the Information Society (FIDIS) workshop , http://www.calt.insead.edu/fidis/workshop/workshop-wp2-december2003/, December 2003.

[ClSc06] Clauss, S. and S. Schiffner, "Structuring Anonymity Metrics", in A. Goto (Ed.), DIM '06, Proceedings of the 2006 ACM Workshop on Digital Identity Management, Fairfax, USA, Nov. 2006, 55-62, 2006.

[CoBi95] Cooper, D. and K. Birm, "Preserving Privacy in a Network of Mobile Computers", IEEE Symposium on Research in Security and Privacy , IEEE Computer Society Press, Los Alamitos 1995, 26-38, 1995.

[DPD95] European Commission, "Directive 95/46/EC of the European Parliament and of the Council of 24 October 1995 on the protection of individuals with regard to the processing of personal data and on the free movement of such data", Official Journal L 281 , 23/11/1995 P. 0031 - 0050, November 2005.

[HBCC04] Hansen, M., Berlich, P., Camenisch, J., Clauss, S., Pfitzmann, A., and M. Waidner, "Privacy-Enhancing Identity Management", Information Security Technical Report (ISTR) , Volume 9, Issue 1, 67, 8-16, Elsevier, UK, 35-44, 2004.

[Hild03] Hildebrandt, M., "Same selves? Identification of identity: a social perspective from a legal-philosophical point of view", Presentation at the the Future of IDentity in the Information Society (FIDIS) workshop , http://www.calt.insead.edu/fidis/workshop/workshop-wp2-december2003/, December 2003.

[HoWi03] Hogben, G., Wilikens, M., and I. Vakalis, "On the Ontology

of Digital Identification", , in: Robert Meersman, Zahir Tari (Eds.): On the Move to Meaningful Internet Systems 2003: OTM 2003 Workshops, LNCS 2889, Springer, Berlin 2003, 579-593, 2003.

[ICPP03] Independent Centre for Privacy Protection & Studio Notarile Genghini, "Identity Management Systems (IMS): Identification and Comparison Study", Study commissioned by the Joint Research Centre Seville, Spain , http://www.datenschutzzentrum.de/projekte/idmanage/study.htm, September 2003.

[ISO99] ISO, "Common Criteria for Information Technology Security Evaluation", ISO/IEC 15408 , 1999.

[Mart99] Martin, D., "Local Anonymity in the Internet", PhD dissertation , Boston University, Graduate School of Arts and Sciences, http://www.cs.uml.edu/~dm/pubs/thesis.pdf, December 2003.

[Mead34] Mead, G., "Mind, Self and Society", Chicago Press , 1934.

[PfPW91] Pfitzmann, A., Pfitzmann, B., and M. Michael Waidner, "ISDN-MIXes -- Untraceable Communication with Very Small Bandwidth Overhead", 7th IFIP International Conference on Information Security (IFIP/Sec '91) , Elsevier, Amsterdam 1991, 245-258, 1991.

[PfWa86] Pfitzmann, A. and M. Michael Waidner, "Networks without user observability -- design options", Eurocrypt '85 , LNCS 219, Springer, Berlin 1986, 245-253; revised and extended version in: Computers & Security 6/2 (1987) 158-166, 1986.

[Pfit96] Pfitzmann, B., "Information Hiding Terminology -- Results of an informal plenary meeting and additional proposals", Information Hiding , NCS 1174, Springer, Berlin 1996, 347-350, 1996.

[RaRD09] Rannenberg, K., Royer, D., and A. Deuker, "The Future of Identity in the Information Society - Challenges and Opportunities", Springer, Berlin 2009. , 2009.

[ReRu98] Reiter, M. and A. Rubin, "Crowds: Anonymity for Web Transactions", ACM Transactions on Information and System Security , 1(1), 66-92, November 1998.

[Shan48] Shannon, C., "A Mathematical Theory of Communication", The

Bell System Technical Journal , 27, 379-423, 623-656, 1948.

[Shan49] Shannon, C., "Communication Theory of Secrecy Systems", The Bell System Technical Journal , 28/4, 656-715, 1949.

[StSy00] Stubblebine, S. and P. Syverson, "Authentic Attributes with Fine-Grained Anonymity Protection", Financial Cryptography , LNCS Series, Springer, Berlin 2000, 2000.

[Waid90] Waidner, M., "Unconditional Sender and Recipient Untraceability in spite of Active Attacks", Eurocrypt '89 , LNCS 434, Springer, Berlin 1990, 302-319, 1990.

[West67] Westin, A., "Privacy and Freedom", Atheneum, New York , 1967.

[Wils93] Wilson, K., "The Columbia Guide to Standard American English", Columbia University Press, New York , 1993.

[ZFKP98] Zoellner, J., Federrath, H., Klimant, H., Pfitzmann, A., Piotraschke, R., Westfeld, A., Wicke, G., and G. Wolf, "Modeling the security of steganographic systems", 2nd Workshop on Information Hiding , LNCS 1525, Springer, Berlin 1998, 345-355, 1998.

Authors' Addresses

Andreas Pfitzmann (editor)
TU Dresden

EMail: pfitza@inf.tu-dresden.de

Marit Hansen (editor)
ULD Kiel

EMail: marit.hansen@datenschutzzentrum.de

Hannes Tschofenig
Nokia Siemens Networks
Linnoitustie 6
Espoo 02600
Finland

Phone: +358 (50) 4871445
EMail: Hannes.Tschofenig@gmx.net
URI: http://www.tschofenig.priv.at