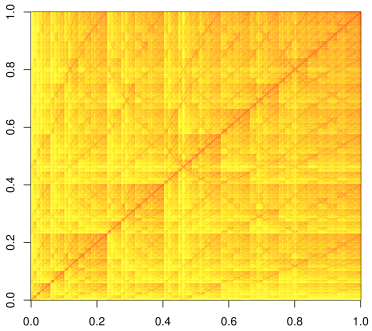
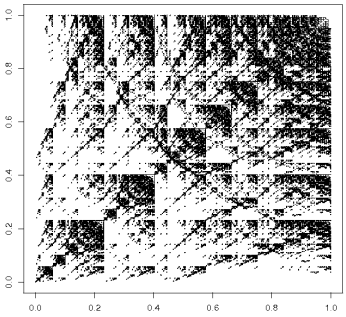
Measuring Unlinkability for Privacy Enhancing Technologies





Vom Fachbereich Informatik der Technischen Universität Darmstadt
zur Erlangung eines Doktors der Naturwissenschaften (Dr. rer. nat.)
genehmigte Dissertation
vorgelegt von Dipl. Inf. Lars Fischer,
aus Marburg an der Lahn

Gutachter: Prof. Dr. Claudia Eckert
Prof. Dr. Stefan Katzenbeisser
Prof. Dr. Dogan Kesdogan

Tag der Einreichung: 21.09.2010
Tag der Prüfung: 29.11.2010

Darmstadt 2010
D 17

Abstract

This work contributes to the field of unlinkability. *Unlinkability* describes a situation where attackers are unable to correctly cluster items of interest. The focus of this work is on modelling and measuring unlinkability. Additionally a protocol for provision of unlinkable certificates, tailored to the requirements of vehicular communication is introduced. The vehicular scenario is used throughout this work as a motivating example.

In this work, a linkability graph is developed and used to illustrate the tight relation between unlinkability and anonymity in certification schemes. The main distinction between anonymity and unlinkability problems is captured by the notion of *inner structure*, which is developed in this work. Inner structure describes the inherent difference between hypotheses of unlinkability problems. This difference is expressed by metrics that reflect the attackers' objectives and world-view. Inner structure motivates consistency of an attacker's world-view as criterion for assessment of attacker quality.

Previous entropy-based global unlinkability measures do not reflect the consistency of attacker probability mass assignments. In this work, these shortcomings are analysed, and criteria for unlinkability measures are developed. These criteria are then discussed for a new unlinkability measure introduced herein. This *expected distance unlinkability measure* is defined using the inner structure of a space of hypotheses. It thus reflects the consistency of an attacker's assignment and not only the attacker's certainty as done by entropy-based measures.

This work also contributes a certification protocol that provides revocable-anonymous protocols that are not linkable to each other or the authenticated requester by the issuing certification authority or others. Revocation of anonymity is possible only by a quorum of independent revocation authorities. The objective is to reduce the trust that has to be put into single authorities to reduce the risk of misuse. The protocol is based on a previously known cut-and-choose blind signature protocol, but made more flexible to suit vehicular communication scenarios. An extensive security analysis of newly introduced protocol-parameters is provided.

Concluding, this work provides a contribution to the ongoing discussion on unlinkability, unlinkability measures and analysis. It draws motivation from a communication scenario that is intensively researched at the moment and contributes to the protocol development there.

Zusammenfassung

In dieser Arbeit wird die quantitative Analyse von Unverkettbarkeit[1] behandelt. Unter Unverkettbarkeit versteht man die Unmöglichkeit, dass ein Angreifer den Zusammenhang zwischen Datenobjekten herausfinden kann. Das bedeutet insbesondere, dass kein Angreifer in der Lage ist anonyme Nachrichten miteinander zu verketten.

Wir konkretisieren zunächst die Unterscheidung zwischen Anonymität und Unverkettbarkeit indem wir formal Anonymitäts- und Verkettbarkeitsprobleme definieren. Der Zusammenhang zwischen beiden Problemklassen, insbesondere, dass Anonymitätsprobleme spezielle Verkettbarkeitsprobleme darstellen wird im Anschluss gezeigt. Der Zusammenhang wird dann in einem Analysemodell weiter konkretisiert und auf ein konkretes Szenario, die Zertifikatsstruktur in Fahrzeugkommunikation, angewendet. Das in dieser Arbeit verwendete Beispiel der Fahrzeugkommunikation ist besonders geeignet weil dabei zum einen hohe Anforderungen an die Privatheit gestellt werden, andererseits aber sehr genaue und umfangreiche Positionsdaten verwendet werden und zusätzlich besondere Dringlichkeit der Informationszuverlässigkeit besteht.

In unserer und anderen Arbeiten zu Unverkettbarkeitsmaßen wird angenommen, dass der Angreifer eine Wahrscheinlichkeitsverteilung über den gesamten Hypothesenraum eines Verkettbarkeitsproblems liefert. Der Hypothesenraum umfasst dabei alle Partitionierungen der Menge der betrachteten Objekte (*Items of Interest (IOI)*). In dieser Arbeit entwickeln wir die Unterscheidung zwischen dieser *outer structure* und der *inner structure* des Hypothesenraums. Die äußere Struktur beschreibt die durch den Angreifer induzierte Bewertung des Hypothesenraums. Die innere Struktur wird durch Definition einer Metrik[2] der ähnlichkeit von Hypothesen aufgebaut.

Bisher bekannte Verkettbarkeits-Maße verwenden ausschließlich die äußere Struktur und ignorieren die innere Struktur des Hypothesenraums. Als Resultat werden unterschiedlich gute Angreifer von bisherigen Maßen gleich bewertet. Wir entwickeln den Begriff der Konsistenz eines Angreifers (*consistency*) aus dieser Kritik heraus, welcher sich auf die innere Struktur stützt.

In dieser Arbeit werden Kriterien für Verkettbarkeitsmaße entwickelt und diskutiert und ein neues Maß (*Expected Distance Unlinkability Measure*) vorgestellt, welches den zu erwartenden Fehler des Angreifers zur Bewertung verwendet. Der erwartete Fehler ist die, durch die Angreifer-bestimmten Wahrscheinlichkeiten (externe Struktur) gewichtete, Summe der Distanzen (innere Struktur) aller Hypothesen von einer gegebenen Referenzhypothese.

[1] Übers. *Unlinkability*

[2] oder Quasi-Metrik

Wir unterscheiden die zwei Analysearten *Black Box* und *White Box*. In der *Black Box* Analyse ist die korrekte Hypothese, welche mit der Realität korreliert, unbekannt, weshalb sinnvollerweise die vom Angreifer höchstbewertete Hypothese als Referenz genommen. In der *White Box* Analyse wird die korrekte Hypothese als Referenz verwendet. In beiden Fällen korreliert der darauf berechnete erwartete Fehler des Angreifers mit dem von uns eingeführten Konsistenzbegriff.

Die Verwendung von Verkettbarkeitsmaßen an realen Beispielen ist im Allgemeinen schwierig aufgrund der Kardinalität des Hypothesenraums und der Summe die über diesem Raum gebildet werden muß. Wir stellen in dieser Arbeit erste Ansätze vor mit der die Komplexität der Berechnung für einfache Angreifer deutlich reduziert werden kann. Diese soweit zu entwickeln, daß Verkettbarkeitsmaße in realistischen Szenarien mit beliebigen Angreifern durchgeführt werden können ist Teil zukünftiger Forschung.

Unverkettbarkeit setzt Anonymität voraus. In früheren Arbeiten haben wir Protokolle entwickelt die kontrolliert widerrufbar-anonyme Zertifizierung unter der besonderen Prämisse der Trennung von Zertifizierungs-[3] und Widerrufsstelle[4] vorgestellt. Eine Zertifizierungsstelle hat die Aufgabe die Authentizität und Authorisierung eines Antragsstellers zu prüfen und ist authorisiert anonyme Zertifikate auszustellen ohne diese allerdings später dem Antragssteller zuordnen zu können. Die Widerrufsstelle ist authorisiert diese Verbindung herzustellen und damit die Anonymität des Antragsstellers aufzuheben. Das Ziel ist unverkettbare Kommunikation auch gegenüber der Zertifizierungsstelle.

In dieser Arbeit erweitern wir ein bestehende Protokoll Zertifizierungs-Protokoll. In einer vollständigen Analyse wird anschließend gezeigt wie die Sicherheit der Revozierung gegen Berechnungsaufwand in Verschiedenen Protokollabschnitten abgewogen werden kann. Am Beispiel der Fahrzeugkommunikation wird aufgezeigt, dass dadurch eine ausreichend kurze Verifikationszeit während des Kommunikationsvorgangs bei ausreichender Sicherheit erreicht werden kann.

Erst durch Protokolle wie das hier vorgestellte ist es möglich Unverkettbarkeit zu implementieren, die starke Anonymität und häufige Zertifikatswechsel voraussetzt. Nicht nur im behandelten Szenario ist eine Vielzahl von auswertbaren Informationen enthalten die letztendlich Objekte verkettbar machen.

Protokollentwickler benötigen präzise Maße zur Entwicklung von privatheitserhaltenden Systemen. Der hier verfolgte Ansatz, ein Maß anhand von festen Kriterien zu entwickeln trägt dem Rechnung. Die aufgezeigten

[3] *Certification Authority*
[4] *Revocation Authority*

Probleme bestehender Maße werden diskutiert und durch die Entdeckung der inneren Struktur des Hypothesenraums erklärt. Die von uns vorgestellte *Expected Distance Unlinkability Measure* vermeidet die Probleme bestehender Maße. Zudem liefert unser Ansatz die erste Diskussion der Umsetzbarkeit der bisherigen Maße, die in Zukunft weiter geführt werden muss um quantitative Abschätzungen von Verkettbarkeit realistisch zu implementieren.

Contents

List of Figures

List of Tables

Acknowledgements

This work would never have happened, if my advisor Claudia Eckert had not given me the opportunity to become a member of her research group at the Technische Universität Darmstadt. I have enjoyed her support and have been nurtured with critique where necessary. I would never have finished this work without her support.

I am very much indebted to Stefan Katzenbeisser, who provided guidance and help during the very important phase of putting it all together.

This particular text has in part been reviewed by Aron who improved my outlandish english and Andreas who went through some of the proofs. Any mistakes I managed to slip by them are mine and mine alone.

Finally, I have been supported by family and friends who give me plenty of reason to get up in the morning to live in the same world as them.

Always try the problem that matters most to you.

Andrew Wiles

1 Introduction

With the raising use of computer and network technology in daily life, one security goal is becoming more and more important, while being increasingly difficult to achieve: *privacy*. While technology progressively supports our daily life, the individual loses more and more control over its personal data. Some security technologies, e. g., surveillance measures, potentially compromise privacy as a side-effect. In a computer-aided life, every action leaves a mark, denoted as *sample*, in the data-sphere. Linked samples, denoted as *trace*, provide an increasingly complete image of the individual. This work is motivated by the wish to empower individuals to know and control the amount of personal information they necessarily emit.

There are many informal — and nowadays even judicial — definitions of privacy. The first example is "the right to be left alone." [70]. Another one is the "informationelle Selbstbestimmung"[1] that became law in Germany in 1983. The common factor of these definitions is that the individual person shall be able to determine which of his personal information is revealed to whom. The definitions of *personal information* vary in the details but in general every information on a person or information that can be related to a person has to be considered as personal information.

Obviously anybody should also be free to disclose his personal information. But in order to choose disclosure freely it must be possible to keep the information secret. While in the past it has been complicated to distribute information it has nowadays become increasingly difficult to participate in

[1] informational self-determination

the modern communication world and control the amount or distribution of personal information. Whereas gathering of information has been a complex task and secrecy was the default state in the past, the situation is almost reversed in the present.

This work contributes to privacy research, especially to the field of *unlinkability*. Unlinkability describes the inability of an attacker to correctly derive how items of interest (IOI) are related with respect with respect to equal sender or other equalities. For an example take anonymous emails as items of interest. Emails are anonymous if the sender of an mail cannot be derived, even if a pseudonym is used, emails from the same sender are still *linkable*. Emails are unlinkable if the attacker cannot even derive which emails were sent by the same sender.

The problem of unlinkability is related to anonymity. While a sender might be anonymous with respect to a message's content, by relating messages of the same sender, an attacker gains knowledge from multiple messages which can lead to an anonymity compromise. At least, a (seemingly anonymous) sender is identified, with any of his messages serving as pseudonym. The attacker can then derive behavioural patterns from linked messages and thus uncover the identity of the sender step by step. To have perfect anonymity, messages have to be unlinkable.

Our main contribution is the introduction and discussion of a new unlinkability measure, which respects the *inner structure* of the unlinkability hypotheses-space. This inner structure allows us to assess the consistency of an attack either with respect to the real relation between items of interest, or measure the consistency of an attacker's world view in itself. Consistency of an attacker is the most important out of three criteria for unlinkability measures we define in this work. Previous unlinkability measures do not consider this consistency, but solely rely on the attacker's opinion which resembles what we denote the *outer structure*.

This work also contributes to secure revocable anonymous certification protocols to provide unlinkable certificates. We use a mobile communications scenario — safety messages in vehicular networks — as motivating example. It is believed that technology for vehicular communication will improve safety, efficiency, and convenience of traffic in the future. By exchanging safety messages that carry vehicle status information and traffic safety warnings, future vehicles should be enabled to better support the driver and prevent accidents. Of course, vehicular communication provides new vectors of attack for malicious actors. Thus strong security is required. Also time constraints are very tight which proves to be very challenging for protocol developers.

In mobile communication the importance of unlinkability is of great im-

portance because mobile devices are usually directly connected to a user. Movements of a device directly translate to movements of the user. Thus, if the device can be located and traced, the user can be traced as well.

To provide unlinkability between messages, frequent pseudonym changes are considered necessary [22]. Certificates provide pseudonyms for vehicles. A prerequisite for unlinkability thus is an anonymous, efficient and secure certification protocol. Communication has to be anonymously authenticated and communication partners have to be authorised to send messages. Further requirements in the vehicular scenario are the ability of traffic authorities to revoke anonymity of vehicles and isolate them from further access to the network. In Chapter 6 we propose a protocol for certification with scalable security parameters and a minimum of necessary authority-trust.

This work is structured as follows. In Chapter 2, the unlinkability problem is defined and introduced. In Chapter 3, the vehicular communications scenario is introduced. A security analysis in Chapter 4 leads to a formal model of unlinkability and anonymity problems in this scenario. In Chapter 5 unlinkability measuring is discussed and our expected distance unlinkability measure is introduced. Chapter 6 introduces a secure revocable anonymous certification protocol. Discussion of related work is included in the individual chapters. This work is closed by a conclusion in Chapter 7.

All life is problem solving.
Karl Popper

2 Fundamental Definitions and Models

Contents

This work deals with unlinkability of messages in mobile environments. In this chapter, basic definitions for the whole work are given, fundamental terms are introduced, and basic properties of unlinkability and anonymity are discussed.

Unlinkability is a sub-problem of the body of problems on anonymity and privacy. Unlinkability describes the inability of an attacker to relate certain *items of interest* with respect to a common attribute. In the following sections anonymity and unlinkability are introduced. Formalisations of the introduced concepts are provided where needed throughout this work. Where appropriate related terms and concepts are mentioned. The relationship between anonymity and unlinkability is discussed briefly.

2.1 Privacy

The distinction between public and private life has already been made by Aristotle [17]. A currently urgent need to protect this "right to be left alone" is created by modern technology [70]. *Privacy* is the notion used to generally describe this separation. Simplified it can be said that an individual that is able to control the border between its public and private life enjoys *privacy*.

In reality neither the separation in public and private, nor the understanding of privacy are simple descisions, but are highly subjective. The need for — and legal amount of — privacy to which an induvidual is entitled are subject of highly controversial political and social debates and beyond the scope of this work. But to enable the debate on privacy, sensible definitions for *amount of privacy* have to be found. With respect to privacy, the objective of our work is to propose means to assess unlinkability, a subproblem of privacy, in given scenarios.

In the following privacy-related terminology, based on [52], is introduced.

2.2 Items of Interest and Identification Anchors

In this work the terms *action*, *role*, *message*, and *sample* are used in many places. Considering anonymity and unlinkability, these terms describe *items of interest* (IOI). Items of interest are objects that an attacker may observe and which he tries to cluster or link to subjects as *senders* or *receivers*. Please refer to [52] for a detailed definition.

Depending on the context items of interest often model messages, events, or actions. If items of interest are the objects acted upon or related to actions, subjects are actors, e. g., senders, creators, and recipients. The term *identity* commonly describes either the abstract concept of uniqueness of a subject or some data which identifies subjects in a given context.

Because of this ambiguity between abstract concept and concrete data, this work introduces the term identification anchor. An *Identification an-*

chor (IA) denotes a set of distinguishable data objects which can be mapped onto subjects. We say that an identification anchor *identifies* a subject. Identification anchors may either be identities used explicitly by a subject or any data which is used to represent a, potentially unknown, subject. For example an attacker who observes anonymous emails, may nonetheless relate emails of the same sender and use any email as an identification anchor for this unknown sender. Identification anchors allow to use concrete objects for identification although they are not intended to function as identities.

IA are related to IOI with respect to a role which expresses the semantic of the relationship of individual anchors and items, e. g., an IA is the sender of an IOI. IOI are related to other IOI if they share a given mapping to the same IA, e. g., have the same sender. Many possible maps can be defined on IOI and IA, but one map has has to be considered special, which is herein called true relation. The *true relation* is the relation that describes how IOI are related to IOI or IA.

The distinction between IOI and IA is used in the following to distinguish anonymity and unlinkability problems.

2.3 Anonymity

Anonymity, in general, describes that the actions of an individual cannot be traced back to that individual. In [51], a terminology paper on the subject, anonymity is defined as follows. "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*."

An *anonymity set* is comprised of a set of identification anchors. The term has been introduced by Chaum in [15]. The cardinality of an anonymity set provides a basic anonymity metric.

In [59] Serjantov and Danezis defined a model to measure anonymity, similar to [37], as following. Let Ψ be a set of identity anchors which represent subjects. Let $R = \{\text{sender}, \text{receiver}, \text{none}\}$ be a set of roles a subject can have with regard to a message (an item of interest) $m \in M$. Let further be $r \in R$ be the role of a user $u \in \Psi$ with respect to a message m. The objective of an anonymity-attacker is to determine for a given role r, which $u \in \Psi$ has this role with respect to m.

We follow this model and define *anonymity problem* to later distinguish between anonymity and unlinkability problems. An *anonymity problem* is a problem that comprises a set of IA and a set of IOI. The problem for an attacker is to find the match between IOI and IA that represents reality with respect to a given role. An attacker searches a map $f_{\text{sender}} : M \to \Psi$ that relates every message to the sender of this message.

Definition 2.1 (Anonymity Problem)
An anonymity problem *denotes the problem of finding a function $f : IOI \rightarrow IA$ that maps every item of interest (IOI) onto the one identification anchor (IA) and is isomorphic to the true relation.*

Anonymity problems provide different levels of complexity to solve with respect to the number of IOI and IA as well as the attackers context. As *anonymity* is the inability of an attacker to distinguish which of all possible maps between IOI and IA is the *true relation*, an estimation of this inability of attackers provides a measure for anonymity.

The separation of items in two classes is the main feature that distinguishes anonymity from unlinkability (as defined below). Other models for anonymity also emphasise this distinction of IA and IOI, e. g., the PROB-Channel for modelling time delays in Mix Networks [65].

The definition of the sets IOI and IA depends on the modelled scenario. The terminology is based on the "classical" scenario containing a set of messages, a set of potential senders, and an attacker interested to infer which message has been sent by which sender. However, other scenarios can be modelled as well, such as relationships between messages and receivers, senders and receivers, anonymous certificates and certificate holders, entries in two distinct databases, and others. In other works (e. g., [45]) this latter scenario has been denoted as an unlinkability problem. We refrain from this terminology here for the sake of clarity.

In the following sections we distinguish further attributes of anonymity and anonymity problems. These attributes have been used in different works on anonymity, and as they can be transferred to unlinkability, it is useful to discuss them in this work.

2.3.1 Perfect/Conditional Anonymity

The term *perfect anonymity* has been used in [62], but has not been sufficiently defined. In the context of anonymity, the notion *perfect* describes a situation where an attacker has no means whatsoever to derive any knowledge on the subject(s) related to actions (IOI).

In terms of anonymity sets, perfect anonymity means that the anonymity set of each IOI is the whole set of identification anchors, without the attacker being able to distinguish between elements of the anonymity set. Considering results from [45], perfect anonymity implies that the anonymity set for all IOI is equal.

By *conditional anonymity* we refer to a form of anonymity which is unveiled — by design — within a defined context. Unless defined conditions are met anonymity is perfect, otherwise anonymity is revoked. The term

revocable anonymity is used synonymously for anonymity schemes that contain a revocation operation. The objective of revocable anonymity is to achieve perfect anonymity with the exception of authorised entities which, under defined circumstances, are able to break anonymity. An example for conditioned anonymity revocation is provided in Chapter 6.

2.3.2 Local/Dispersed Anonymity

A concept that has not been discussed in the context of unlinkability, but is already used ambiguously in the context of anonymity, is the notion of *local anonymity*. In [46] the term *local anonymity network* is used for the opposite of a *dispersed anonymity network*. The distinction is one of control. Local anonymity defines anonymity networks that exclusively rely on the "local administrative domain", meaning that they reside within a domain controlled by the anonymity subject itself. Dispersed anonymity describes networks that rely also on "external administrative domains". Administrative domains denote spheres of influence. The local administrative domain describes the domain connected to the current point of view.

We are not using this concept in this work. The reason for that is, that most anonymity-preserving protocols explicitly are designed to spread anonymity measures over as many administrative domains, especially as possible to disguise any relation of IOI to any administrative domain. This obviously increases the anonymity set related to given IOI, which otherwise would only consist of those subject local to a given domain. There might be scenarios where it is sufficient to only provide anonymity within a local domain, but this would also need for the local administrative domain to be completely trustworthy with respect to anonymity. This would lead to the adminstrators of that domain becoming a single point of failure, which we explicitly try to avoid in the methods proposed in this work.

2.3.3 Individual/Global Anonymity

In [66] the notion of *locality* is used to distinguish different anonymity measures. The term *global anonymity* is used for models that give a global (average) metric on the anonymity of a scenario. The term *local anonymity* is used to describe models that assess the anonymity for an individual subject in a given scenario.

Because of the latent ambiguity in the usage of *local anonymity* the notion of *individual anonymity* has been introduced in [51]. In this work the latter understanding of local anonymity is used.

2.4 Pseudonym

According to Pfitzmann and Hansen, "A *pseudonym* is an identifier of a subject other than one of the subject's real names" [51]. Every pseudonym is an identification anchor as well as "real names" are.

Pseudonyms are subject-identifiers that are often used within a distinct context. The process of generating pseudonyms is crucial for anonymity preservation, because at the point of pseudonym generation multiple identification anchors of a subject may be put in relation to each other. It becomes even more crucial if a subject has to authenticate to be entitled to receive pseudonyms. For example if a subject has to prove control of an email-account before being granted an pseudonym in a web-forum.

The process of generating secure, i. e., authentic pseudonyms, is herein referred to as *certification*. The relation between pseudonyms and subjects is a surjective map. Pseudonyms are a classical solution to provide anonymity, if neither the relation between pseudonym and subject nor the relation between different pseudonyms is revealed to an attacker.

By changing pseudonyms between contexts, the subject can make it impossible to link between identities in different context spheres. A simple example is the usage of different email addresses as pseudonyms to provide unlinkability of business and private communication. Other context classes could be temporally, spatially or content related [27][22].

2.5 Anonymity Revocation

Anonymity has its drawbacks. An example is given by the well known blackmail-case in [69] or the vehicular communication scenario as described in Section 3. In some scenarios it should be possible to revoke anonymity under certain circumstances. Revocation means, that the relation between a pseudonym and a subject is revealed. In terms of anonymity this is similar to *conditional anonymity* as described in Section 2.3.1. Using conditional anonymity means, of course, that privacy is endangered if the entities, that are authorized for anonymity revocation, turn into attackers.

Revocation is useful in scenarios where misuse may lead to grave danger and where the authorised entities are trusted not to misuse the power of revocation. The vehicular communication scenario described in Section 3 provides our motivation for this type of anonymity. But other applications like e-cash [69] motivate revocable anonymity as well. In Chapter 6, protocols for generation of certificates with revocable anonymity are discussed.

We informally define revocation of anonymity, as follows.

> *Revocation of anonymity* is the authorised recovery of the relation between a subject and the pseudonyms used by this subject.

Obviously revocation should only be possible for authorized entities under determined circumstances. The term *authority view* denotes the knowledge, e. g., data learned during a certification process, of entities authorised to revoke anonymity. In the following, we will abbreviate *revocation of anonymity* as *revocation* if the meaning is unambiguous within its context.

In [62], a useful distinction of based on the direction of anonymity revocation is made. We use this terminology in the remainder of this work.

We translate Type-I anonymity revocation denotes the derivation of pseudonyms from subject identities.

Definition 2.2 (Type-I Anonymity Revocation)
The mapping of a subject onto corresponding identification anchors (pseudonyms).

Type-II Anonymity Revocation denotes the other direction of revocation, where the subjects identification anchor is derived from a pseudonym.

Definition 2.3 (Type-II Anonymity Revocation)
The mapping of a subject onto corresponding pseudonyms.

Using the terminology of IA and IOI, both revocation is always mappings between identification anchors, i. e., between an IA directly related to a subject and pseudonymous IA. It is useful to distinguish those types because often one type is required while the other is prohibited. For an example see Chapter 3 on a vehicular communications scenario.)

2.6 Isolation

After anonymity revocation, the next escalation of misuse protection measures is the *isolation* of the subject from the system. In computer networks this would mean to prevent further access to the network by malicious devices. *Isolation* is the action of preventing a node from affecting other nodes or the network. As the examples in this work do not have secrecy as security objectivesl, isolation has only to prevent manipulation of the internal state of nodes through network communication, but no eavesdropping. Whether isolation is used as some kind of punishment or purely for protection is up to the system.

The damages a malicious nodes may inflict, while it is not isolated, are the costs of slow isolation. We cannot estimate these costs outside a specific scenario, but may assume that these cost are related to the time while a malicious node is not isolated. We denote by *isolation time* the time between recognition of malicious behaviour and isolation becoming effective. We use

this to estimate the efficiency of isolation and the most interesting parameter of an isolation scheme is the upper bound of its isolation time, which we denote *maximum isolation time*:

Definition 2.4 (Maximum Isolation Time [23])
The upper bound of the length of the time interval between recognition of an attacker and its effective isolation from the network is denoted as maximum isolation time.

The isolation time starts with the actual detection of an attack. It ends at the point where the subject is effectively isolated. Recognising an attack means recognising its effects, which means that beforehand no attack has been recognised by the system. Thus the point of recognition is the moment when an attack becomes imminent to the system, which makes it a natural measure point.

2.7 Unlinkability

In the extremal case every subject may enjoy perfect anonymity, e. g., by using a different pseudonym for every action or better, no identification at all. From the anonymity point of view this subject enjoys perfect privacy. However an attacker might still be able to relate different actions of a subject. The attacker may even do so without using identification anchors (IA). Some data related to individual actions might link actions of a subject with each other. Furthermore, behavioural statistics or usage patterns may reveal further information on the relation between individual IOIs.

The notion used to describe that an attacker can not (correctly) relate IOI is *unlinkability* [51, 63, 25].

> "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 [...] distinguish whether these IOIs are related or not." [51]

Taking the viewpoint of an attacker, unlinkability is the inability to find a clustering of a set of IOI that corresponds to the true relation between IOI. *Relation* in this sense normally denotes *related with respect to sender equivalence*, but in general other relations can be considered as well. *Sender-equivalence* denotes a relation where IOI are in the same equivalence class if they originated from the same subject[1].

[1] Or if they originate from a subject using the same identification anchor.

The notion of the *unlinkability problem* is used herein to explicitly denote the problem of finding an equivalence relation over the set of IOI that correlates to reality. Equivalence relations are represented by a partition π from the set of all partitions Π_M of the set M of IOI.

Definition 2.5 (Unlinkability Problem)
An unlinkability problem *denotes the problem of finding a partition π from the set of all partitions Π_M of a set M that is isomorph to the true relation.*

The opposite of unlinkability is *linkability*, which is also sometimes mentioned in this work. Linkability describes the ability of an attacker to correctly relate IOI to each other. Similar to anonymity, unlinkability can be distinguished as *perfect/conditional* or *global/individual*.

The main topic of this work is measuring unlinkability. In Chapter 5 an unlinkability measure is proposed that provides a quantification of unlinkability.

2.8 Anonymity vs. Unlinkability

In the following we shown that anonymity problems are sub-problems of unlinkability. This means that every anonymity problem can be modelled as an unlinkability problem with additional context information.

By definition of unlinkability problems, an unlinkability-attacker is not interested in, nor wants to, identify subjects. Its objective is to discover the equivalence relation on messages, which correctly relates messages according to the sender. An anonymity-attacker wants to know *who* is the sender/receiver of a particular message. The question of an unlinkability-attacker is *which* messages are in the same equivalence class. In most examples, as in the vehicular scenario used herein, the equivalence class will be a same-sender relation.

Unlinkability is stronger in terms of privacy than anonymity. Every anonymity problem can be mapped onto an unlinkability problem but not the other way round. In Appendix A we show how such a mapping can be constructed using hint-classes for unlinkability problems.

2.9 Conclusion

In this chapter we have explained basic concepts and terminology of unlinkability and anonymity. While anonymity conceptually prevents an attacker from relating items of interest to identification anchors, unlinkability directly considers relations between of items of interest.

Anonymity and unlinkability are terms that are often used ambiguously. Unlinkability is the broader notion — encapsulating anonymity. The definition of these two concepts as *problems* allows us to clearly distinguish anonymity and unlinkability.

3

Vehicular Communication Scenario

Contents

This chapter introduces the vehicular communication scenario which serves as a motivating example in this work. Vehicular communication provides an interesting field for privacy research, because of the strong relation between users and devices. Another interesting point is the ubiquity of vehicle usage which provides a large footprint for any developments.

Vehicular communication is currently a thriving field for research, applications from traffic safety over toll collection to amenity services are being developed. Aside from critical privacy protection, the scenario's tight constraints require new and advanced solutions ranging from networking to cryptographic protocols. Furthermore, vehicular communication is still an emerging field, where research can directly influence the development of technology without being blocked by existing standards.

Vehicular communication describes wireless communication between vehicles (Car-2-Car or C2C communication). C2C communication often also is called *inter-vehicle communication* (IVC). Communication between vehicles

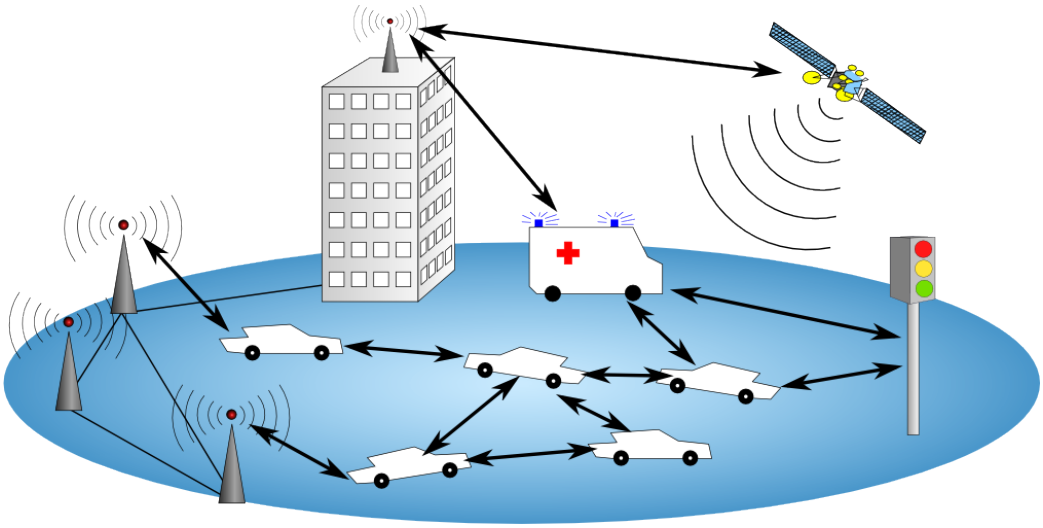


Figure 3.1: Vehicular Communication (derived from [49])

and the road infrastructure, denoted as Car-2-Infrastructure or C2I communication. In general, communication between vehicles and either infrastructure or vehicles is abbreviated as C2X.

For the remainder of this chapter, possible applications, security objectives and assumptions for the vehicular scenario are described.

3.1 Applications

In this section, example applications for C2X technologies are discussed. The purpose of this section is to motivate security objectives and assumptions.

Many applications have been envisioned for future vehicular networks. Comprehensive information can be found, e. g., in [14], [48], and [33]. Figure 3.1 shows some possible communication channels in vehicular communication. The image includes a central authority that could advice emergency vehicles or broadcast information through satelites. Vehicles may communicate with other vehicles, active traffic signs or via static antennae and backbone networks with central authorities and information sources.

Applications are generally classified into safety and non-safety applications. The class *safety application* contains applications that "reduce traffic accidents and to improve general public safety" [14], such as electronic brake lights, collision avoidance, congestion information, priority vehicles, and smart traffic lights. A *non-safety application* is an application that is not inherently safety-related, i. e., failures of these applications do not directly threaten life or health. Another common distinction is based on the

applications' field, e. g., traffic related communication, vehicle maintenance, travel information, or entertainment. The solutions in this work are focussed on traffic-safety applications.

Safety applications often demand very efficient and very secure communication For example, latency requirements of 0.1 seconds, non-interactive one-way communication combined with an absolute need for trustworthiness of received information are constraints of the *electronic brake light.* The term *trustworthiness* here describes that the contents of a message must correctly describe reality.

Many applications explicitly require information on the sender's position and movement. This information has to be considered private because it is related to the driver's movement, even outside of the context of the traffic scenario. Applications that do not explicitly reveal positional information do, nonetheless, emit radio messages that can be related to the emitters position.

Another problem that directly hinders the introduction of vehicular communication into the real world is the penetration rate. Many, especially C2C-applications, can only be effective if a sufficient rate of vehicles has these applications installed. In [39] the required penetration rates for the success of certain applications are provided.

3.2 Traffic Scenarios

In this section, variants of traffic scenarios are introduced. In the microscopic traffic simulation tool *Simulation of Urban Mobility* (SUMO)[1], six parameters are used to define vehicles: acceleration, deceleration, length of vehicle, maximum speed and the driver's imperfection. We have used only one type of vehicle in our preliminary simulations.

A scenario is partially defined by *streams of vehicles* comprised by trajectory pattern, vehicle emission frequency, number of emitting vehicles and stream timing. Vehicle streams are timed in a way, that the middle of streams meet in the middle of the paths. Varied frequency and number of vehicles should be used independently in multiple experiments. The *trajectory patterns* described below, based on a grid like road-map, are considered in this work. These basic trajectory patterns represent all possible crossings of two traffic streams.

A crossing of two trajectories is, similar to a mix proxy, a point where traces can be confused. Single crossings of two trajectories are the minimum building blocks of which larger scenarios can be constructed. (See Figure 3.2.) Theses basic traffic scenarios are standard situations where unlinkability of different emission protocols can be analysed. We distinguish

[1] `http://sumo.sourceforge.net/`

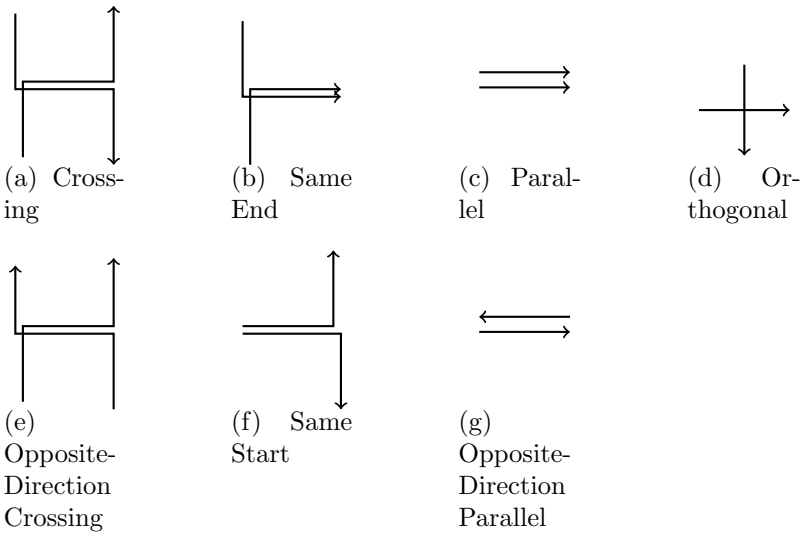


Figure 3.2: Types of paths in the Simulation.

different scenarios by their trajectory pattern: crossing paths, parallel paths, opposite-direction paths, and orthogonal paths. These Trajectory patterns for the general case have been discussed in greater detail in [53].

Crossing-Paths (see Figure 3.2(a)) describe a scenario of two paths sharing a certain segment of the road network and separating into opposite directions. The idea is to mingle two streams for a short part of the itinerary. The three attributes that describe scenarios are length of the start-, shared-, and end-segment as well as direction. Length may be expressed as absolute or relative length.

The *same-end* and *same-start* trajectories, in Figures 3.2(b) and 3.2(f), describe trajectory pairs that start or end in the same point. They can be seen as crossing-paths scenarios with zero-length end or start segments. *Parallel-Paths* (see Figure 3.2(c) and Figure 3.2(g)) share the whole segment. *Orthogonal-Paths* (see Figure 3.2(d)) are paths that meet in only one point, i. e., at a crossing. *Opposite-Direction* (see Figure 3.2(e) and Figure 3.2(g)) scenarios describe that paths are defined in opposite directions at the shared segment.

More complex shapes — multiple streams, looping streams, and others — can be modelled by combining these basic types. The objective here is to reduce the complexity of an analysis by reducing the complexity of scenarios.

3.3 Communication Primitives

In this section, basic primitives of vehicular communication are introduced as a foundation for a common nomenclature and understanding. In the following the individual players and devices are introduced, the basic communication structure is explained, and the security infrastructure is described.

The central device in vehicular communication is called the *on-board unit* (OBU), which is connected to the vehicles' internal communication system and processes all inter-vehicle communication. An OBU fulfills all operations related to vehicular communication. Inside the OBU, all cryptographic keys, state information and protocol logic is stored. OBU offer only low computing power for obvious economical reasons. It is assumed that an OBU is tamper-proof and inseparable from the vehicle. Thus the terms *vehicle* and *OBU* can be used interchangeably in the vehicular communication scenario.

OBU can be distinguished from *road side units* (RSU) which denote any communication device with static position that is part of the infrastructure. RSU may be connected to a high-bandwidth communication backbone.

The anchor for authentication of an OBU is the OBU-certificate. In this work it is assumed that any OBU can be authenticated by way of an *OBU-certificate*, which is bound to the vehicle and granted and signed by a *traffic authority* (TA) on registration of the vehicle. The TA is assumed to be the single authorised entity that commands and controls all traffic related operations. An OBU-certificate fulfils the function to an electronic license plate.

3.3.1 Authorities

In this work we assume that there exist at least the following functional entities. The vehicles appear in the functions of *emitters* of messages in vehicular communication, as well as *requesters* of *Inter Vehicle Communication certificates*. IVC certificates are used to authenticate vehicles during communication and to prove that the vehicle is authorised by the TA. IVC certificates may only be provided to OBU that have authenticated themselves against an authority and are authorised to take part in IVC. IVC certificates are assumed to be short lived and are granted by a *certification authority* (CA). The CA grants IVC certificates on behalf of the TA.

In case a vehicle misuses its privileges, its authorisation has to be revoked and the identity of the vehicle and its owner have to be derived for prosecution; thus the anonymity of the vehicle is revoked (see Section 2.5) and the vehicle has to be isolated (see Section 2.6). It is assumed that a *Certificate Revocation List* (CRL) of OBU certificates that have been revoked is

managed by the CA.

In this scenario two levels of revocation can be distinguished. The first level revokes the authorisation of requesters/vehicles to acquire new IVC certificates. The second level directly revokes IVC certificates and thus has to be broadcasted to all concerned vehicles. The second level of revocation is only used if a current attack has to be stopped immediately, meaning that all vehicles must be prevented from accepting message of a *rogue vehicle* (see below).

The usage of CRL for IVC certificates is otherwise discouraged, because this considerably increases the time and memory complexity for all vehicular communications. Optimisation by regional and time-limited certificate revocation lists should be applied.

Long term isolation of vehicles is done by the first type of CRL. This type of CRL does not critically influence the efficiency of vehicular communication because it is only distributed to the certification authorities. It is assumed that CA have sufficient computing power and bandwidth.

3.4 Security Objectives

The main objective of vehicular communication is to increase traffic safety. Vehicular communication is envisioned to directly influence the actions of drivers or even vehicles. Further, vehicles are related to persons. The driver generally is held responsible for any actions of the vehicle and all information on the vehicle is directly related to the driver or the owner of the vehicle. Because erroneous data may have grave consequences in the vehicular scenario, and because of the personal information involved, security is paramount in vehicular communication.

Security objectives have been discussed in almost every paper on vehicular communication. Exemplary security objectives can be found in [33],[49], and [48]. In 2007 the *Car 2 Car Communication* Consortium (C2C-CC) mentioned correctness, trustworthiness, privacy of participants, and robustness of the whole system as major (security) objectives [1]. The adversary considered normally is passive and listening globally to inter-vehicle communication.[20] The structure that is used in the following has first been used by us in [23].

The following objectives describe the focus of this work towards secure vehicular communication.

Authenticity and Isolation. The risk of misuse of a safety message system can be assessed as high due to the serious damage that might be inflicted by attacks. Therefore the first directive for any C2C-communication scheme is that every safety message has to be authenticated.

Trust is understood as a transitive relation that between natural persons that is only relayed onto devices [32]. The correctness of messages is rooted in the trust in the manufacturer's ability to produce a working car. Anyone relying on the correctness of a message may want to verify that the vehicle is indeed a vehicle, produced by a trusted manufacturer. Furthermore it is to be ensured that the vehicle is not tampered. The usual way of implementing this is by certificates and authentication.

As vehicles and OBU may break and malfunction nonetheless, misbehaving OBU must be isolated from the communication network.

We assume that under normal circumstances vehicles are emitting only information that is correctly describing reality. But obviously vehicles have to expect messages that carry information which is not correct. One can distinguish two causes for messages that are not correct: malfunction or malignancy. As the cause cannot be deducted from the message itself we summarise vehicles that emit such messages under one term. A vehicle that emits messages with illicit content or content that does not reflect the reality correctly is a called *rogue vehicle*.

Within the scope of this work it is not important whether a vehicle acts maliciously or is simply malfunctioning. To prevent damage to vehicles and humans, faulty information has to be prevented from interfering with vehicles operation.

Maximum Isolation Time. As mentioned above, if a vehicle has been identified as rogue vehicle it has to be isolated from the network. Isolation is the *ultima ratio* of the network to prevent accidents resulting from wrong information. Isolation of a vehicle normally should be initiated by traffic authorities, or even through voting processes among vehicles. In most existing societies it will probably be mandatory by law to inform the owner of an isolated vehicle.

There should also be a defined maximum isolation time, as introduced in Section 2.5. Two systems can be compared based on their maximum isolation time. A system with smaller isolation time will usually be considered more secure than the one with a larger time.

Non-Interactivity and Low Overhead. Because of the high velocity of vehicles and thus short duration within which communication between two vehicles is possible, the communication has to be non-interactive and the message overhead has to be very low. Non-interactivity means that data is transferred without handshake protocols or other challenge-response schemes. The urgency of safety messages implies that authentication must be instant without additional communication. The size of messages has been estimated

to be around 200 bytes [67].

Anonymity and Unlinkability. Safety messages include data that erodes privacy of vehicle owners or drivers. Most relevant is the danger of tracking a vehicle through positional information and linkable pseudonyms. Thus, while preserving the security goals mentioned above, a scheme has to provide privacy for vehicles and their drivers. More precisely, it must be provided that drivers and vehicles can neither be traced by an observer nor, that unauthorised attackers may acquire the relation between messages and vehicle. Moreover, to prevent misuse, a driver should not be linkable to a vehicle and the messages it sends by a single authority.

Revocable Anonymity. Prerequisite to isolation is the revocation of anonymity. In case of emergencies and investigations of malicious or criminal behaviour of single subjects, legal authorities must be able to link messages to an OBU and the OBU to the vehicle owner.

3.5 Conclusion

Security in vehicular communication is a field of conflicting objectives. While traffic safety recommends high level authentication and the means to find and prosecute misbehaviour, strict time-constraints are set for communication processes. For privacy reasons, on the other hand, it is recommended to emit as little information as possible to the public and to restrict the potential misuse by other participants or even authorities.

Vehicular communication provides an excellent example for *privacy technologies*. Vehicular scenarios contain personalised objects (OBUs) whose behaviour can directly be related to the driver. Furthermore many applications in vehicular communications need explicit positional information, which generally is very sensitive information if it can be related to a person. Thus, providing privacy is critical for vehicular communication scenarios.

Vehicular communication scenarios are particularly well suited as an example for *unlinkability research* because they provide a field with messages that are related to spatial and temporal positions. The messages thus are samples of vehicles' trajectories. Because of the constrained movement of vehicles and available context information, a broad spectrum of different attackers and movement scenarios can be used for analysis. For a start, vehicular movement is restricted to some physical constraints, e. g., spatial restrictions (roads), or maximum velocity. One may very well assume that vehicles stay on the road, keep roughly to speed regulations or show more

specific behaviour. The examples used in Chapter 5 make use of velocity statistics to relate vehicular messages.

t

Security is mostly a superstition. It does not exist in nature, nor do the children of men as a whole experience it. Avoiding danger is no safer in the long run than outright exposure. Life is either a daring adventure,
or nothing.

Helen Keller

4 Unlinkability Graph Model

Contents

This chapter provides a graph model for anonymity and unlinkability problems in certification structure. The modes is used to represent the two types of relationships that comprise anonymity and unlinkability problems, i. e., from IA to IOI between IOI. Based on the model an attacker classification is derived to distinguish attackers based on objectives and resources. The model is based on a certification structure as it could be used in a vehicular scenario similar to the one described in the previous chapter. The model is realised linkability graphs that also visualise the relation between anonymity and unlinkability problems.

4.1 Linkability Model

In this section we provide a model of linkability for the communication scenario as described in Chapter 3. This section extends the linkability model as introduced by the author in [24]. The main characteristic of the vehicular scenario is that pseudonymously authenticated messages are broadcast on

a public medium. Pseudonyms are provided through a PKI-like structure, manifested through *certification authorities* (CA).

This work distinguishes identity authorities from certification authorities (CA). An *identity authority* manages the relations between identities and devices. In the vehicular example the identity authority is synonymous with the traffic authority (TA), which controls the relations between vehicle owners and their vehicles. A CA grants certificates to vehicles, represented by their on-board units (OBU), and thus manages the relation between vehicle and certificates.

The authority knowledge in a vehicular scenario can be modelled as k-partite graph of disjunct sets. Vertexes represent sets of *items of interest (IOI)*: *subjects* $s \in S$, *devices* $d \in D$, *certificates* $c \in C$, and *messages* $m \in M$. Anonymity problems are setup by a pair of sets, where one set represents identification anchors and the other items of interest. Unlinkability problems are comprised by a single sets. The set of subjects can be taken as identification anchors (IA), while the set of messages naturally provides items of interest. Vehicles and certificates are either used as items of interest or identification anchors, depending on the attacker's objective.

Edges in the graph represent relations between elements in a given knowledge sphere. A *knowledge sphere* contains every information on relations between IOI known to some given entity. In terms of this model a knowledge sphere is described by sets of edges on which the entity has information, e. g., the identity authority observes some event which provides information on a certain subject being related to some device. Edges are distinguished into horizontal and vertical relations. By *horizontal relation* we denote relations between different sets, e. g., between certificates and messages. Relations within a set are denoted *vertical relation*, i. e., relations between messages. According to common notations of related items, we will denote an edge between two items a and b as $a \sim b$ (see [63]).

The graph representation allows formalisation of the knowledge of an onlooker, e. g., an attacker, about the relations between items of interest. Edges in the graph may be weighted to reflect the belief in the correctness of that relation.

4.1.1 Horizontal Linkability Graph

A *horizontal linkability graph* [24], describing possible relations between different sets of items of interest is constructed as:

$$\begin{aligned} \mathcal{G}_h &= (\mathcal{V}_h, \mathcal{E}_h, \mathcal{P}_h) \\ \mathcal{V}_h &= S \cup D \cup C \cup M \\ \mathcal{E}_h &\subseteq (S \times D) \cup (D \times C) \cup (C \times M) \\ \mathcal{P}_h &: \mathcal{E}_h \to [0, 1], \end{aligned} \tag{4.1}$$

where $\mathcal{G}_h$ is a graph consisting of vertexes $\mathcal{V}_h$, edges $\mathcal{E}_h$, and a weighting function $\mathcal{P}_h$. $\mathcal{G}_h$ is, in this scenario, a reduced 4-partite graph. A horizontal graph only has edges between vertex sets. The subset $S \times D$ of $\mathcal{E}_h$ describes the relations between owners and vehicles (devices) D. Similarly, the edge subsets $D \times C$ and $C \times M$ denote the relations between vehicles and certificates, as well as between certificates and vehicles. Other relations may be modelled by introducing dummy element in the sets in between, e. g., to model information on a relation $S \times M$ on introduces a temporary dummy element each in D and C.

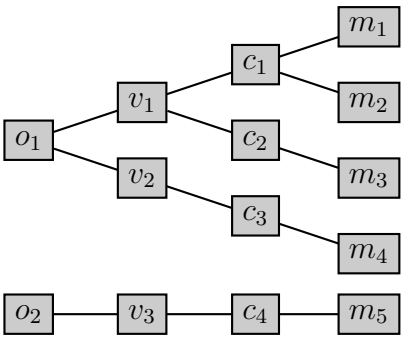


Figure 4.1: Horizontal Linkability Graph

Edge weights are used to express an attacker's belief in two items being related. An edge-weight $\mathcal{P}_h$ of zero denotes that the edge's endpoints are not related. We will use the notions of *edge has weight zero* and *edge does not exist* synonymously with *edge is not in graph*. A horizontal linkability graph consist of one or more trees whose roots are the elements of S. In Figure 4.1 an example graph is shown.

One may observe that $\mathcal{G}_h$ represents distinct layers of anonymity problems, e. g., anonymity with IA subjects and IOI vehicles, between IA vehicles and IOI certificates, as well as between IA vehicles and IOI messages. Every choice of two sets of vertexes, one representing IA and the other IOI, poses a distinct anonymity problem.

4.1.2 Vertical Linkability Graph

A *vertical linkability graph* [24] represents relations between IOI of the same set. A vertical linkability graph is a collection of cliques, because it represents an equivalence relation, which is transitive. A vertical graph is constructed as follows:

$$\begin{aligned} &\mathcal{G}_v = (\mathcal{V}_v, \mathcal{E}_v, \mathcal{P}_v) \\ &\mathcal{V}_v = S \cup D \cup C \cup M \\ &\mathcal{E}_v \subseteq (S \times S) \cup (D \times D) \cup (C \times C) \cup (M \times M) \\ &\mathcal{P}_v : \mathcal{E}_v \to [0, 1]. \end{aligned} \tag{4.2}$$

Each set of nodes in $\mathcal{G}_v$ provides a separate unlinkability problem, e. g., (un-)linkability of vehicles with respect to equal subjects. The most interesting unlinkability problem in a vehicular scenario probably is the unlinkability of messages M, as messages represent spatial points from which trajectories of vehicles can be derived.

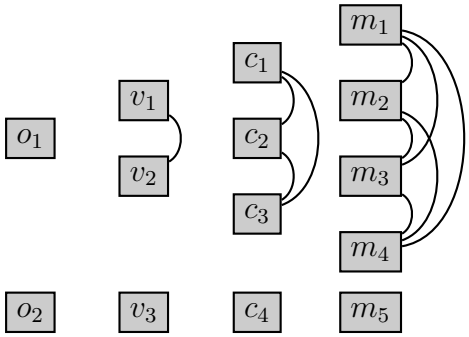


Figure 4.2: Vertical Linkability Graph

An example vertical linkability graph, using the same vertexes as in the previous example, is shown in Figure 4.2. Each clique represents an equivalence class of IOI. In this example, equivalence denotes that IOI are related to the same owner.

4.1.3 Linkability Graph

Horizontal and vertical graphs provide a representation that shows the relation between anonymity and unlinkability. The connection between $\mathcal{G}_h$ and $\mathcal{G}_v$ is the node set, which contains the same elements, i. e., subjects, vehicles, certificates, and messages. One may observe the duality between both graph types: Given $\mathcal{E}_h$ and $\mathcal{P}_h$ the corresponding $\mathcal{E}_v$ and $\mathcal{P}_v$ can be inferred. Obviously the other direction needs additional information on the relation between vertex sets.

As $\mathcal{V}_h = \mathcal{V}_v$ the notation can be simplified by using $\mathcal{V} := \mathcal{V}_h$ as vertex set. As the domains of $\mathcal{P}_h$ and $\mathcal{P}_v$ are disjoint, we can combine them to simplify the notation using $\mathcal{P} : \mathcal{E}_h \cup \mathcal{E}_v \to [0, 1]$. These observations allow for a combined graph that includes vertical and horizontal relations:

$$\begin{aligned} \mathcal{G} &= (\mathcal{V}, \mathcal{E}, \mathcal{P}) \\ \mathcal{V} &= S \cup D \cup C \cup M \\ \mathcal{E} &\subseteq \mathcal{E}_h \cup \mathcal{E}_v \\ \mathcal{P} &: \mathcal{E} \to [0, 1]. \end{aligned} \tag{4.3}$$

The combined graph G allows representation of the complete knowledge of an attacker using knowledge from the domain of horizontal relations $\mathcal{E}_h$ and from the domain of vertical relations $\mathcal{E}_v$.

In graph $\mathcal{G}$, we divide *horizontal relations* into *authority knowledge* consisting of $(S \times D) \cup (D \times C)$, and *public knowledge* consisting of $(C \times M)$. *Authority knowledge* denotes information on relations between subjects and vehicles, or vehicles and certificates. Authority knowledge can be further subdivided into knowledge of the identity authority domain $S \times D$ and knowledge from the certification authority domain $D \times C$. *Public knowledge* denotes relations that are obvious to anybody seeing the items of interest. In the scenario described in Section 3 and used in Chapter 6, relationship between certificates and messages are publicly readable in the messages.

Information on *vertical relations* is mostly based on *context information*. By *context information* we denote any knowledge on the relation between IOI derived from external evidence. External evidence is all information not directly derived from the registration, certification, or communication process. This includes probabilistic information on vehicular behaviour, which can be exploited by statistical inference.

4.1.4 Using Linkability Graphs

A linkability graph can be used to represent the true relation as well as the world view of different attackers. We denote by reality graph (G, V, E, P) the linkability graph corresponding to the true relation. This means that *iff* two items of interest a and b are related in the real world, there exists an edge (a, b) in the reality graph with $\mathcal{P}(a, b) = 1$. If a and b are not related then $\mathcal{P}(a, b) = 0$ in the reality graph, sometimes denoted as *not existent*. An example reality graph for two vehicle owners is depicted in Figure 4.3. It can be noted that the reality graph is a forest of trees with the owners being the roots of trees.

This model provides a notation to express a scenario as it is seen by an observer. We can express the knowledge of an attacker A as the weighted

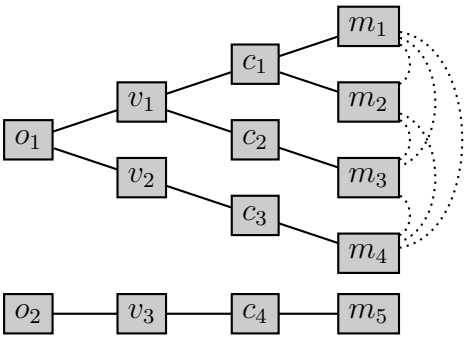


Figure 4.3: Example reality graph: o_1, o_2 denote vehicle owner identities, $v_1, \ldots, v_3$ denote vehicle identities, $c_1, \ldots, c_4$ vehicle communication certificates (IVC), and $m_1, \ldots, m_5$ denote positional messages. Continuous lines denote edges in the reality graph. Dotted lines connect related messages.

graph $\mathcal{G}_A = (\mathcal{V}, \mathcal{E}, \mathcal{P}_A)$. The belief of attacker A on whether or not two items of interest are related is expressed by the weights $\mathcal{P}_A : \mathcal{E} \to [0, 1]$. A similar notion of probability is also used in [63, 25] to define the *degree of unlinkability*. This is discussed further in Chapter 5.

An alternative way to use this model is to create attack trees for defined scenarios: a security analyst is able to derive an attack tree if he uses complexity measures as edge weights. These weights then are used to find a minimum spanning tree. This threat model provides an estimation of the (complexity) costs for the modelled attacker in relation to the attacker's success.

4.2 Attacker Classification

Protection mechanisms for secrecy and authenticity depend on secrecy of single keys. Anonymity and unlinkability depends on secrecy of relationships. While the secrecy of keys is controlled by the holder of the keys, secrecy of relationships depends on the inability of attackers to observe communication channels, e. g., read address information in messages. Therefore a worst case attacker in the vehicular scenario with complete knowledge makes little sense. In Chapter 6 protocols are introduced that prevent single authorities from having complete knowledge, reducing the strength of this former worst case attacker. No security analysis is complete without at least noticing the existence of different types of attackers, which is provided in this section.

Attackers can be distinguished by their objectives and resources at their disposal. In unlinkability attacks, the resources of an attacker are the items of interest and any contextual information usable to relate items of interest.

Authority	from	$2^{\{IA,CA\}}$
Activity	from	$\{active, passive\}$
Space	is	r-global, $r \in [0,1]$
Objective	from	$\{M \to M,$
		$C \to M,$
		$V \to M,$
		$O \to M,$
		$M \to \Pi_M\}$

Table 4.1: Attacker Classification (excerpt)

Obviously this includes the quality of information as well as the amount of information. The objective of an attacker is to correctly link items of interest.

Similar to the unlinkability model in Section 4.1, we distinguish resources of an attacker into *information known by authorities*, *attacker activity*, and *observation space*. Literature knows different classes of ressources that might be applied in combination, e. g., Aijaz et al. [4] distinguish attackers by quality of access to vehicles' on-board units. The objective of an attack is to discover the trajectory of a vehicle in terms of the sequence of positional messages. A trajectory T is a sequence of positions in space-time corresponding to positional messages from M. For simplicity we denote $T \subseteq M$.

Attacker objectives can be distinguished by the vertex-set that is used for identification, i. e., the set used as identification anchors in a given scenario. This might be either a message certificate, vehicle identity, or owner identity. Messages can be used as identity anchors if we agree on using e. g., the temporal first message in an trajectory as makeshift identity. Certificates, vehicles, or owner ID obviously are IA only for specific (sub-)trajectories, e. g., an IA representing a certain vehicle is related to the trajectory of that vehicle which is a sub-trajectory of the owner/driver.

This classification of attackers is used here as a working hypotheses. For different purposes it might be useful to adapt the classification. This example classification is summarized in Table 4.1.

The *authority class* describes whether an attacker has knowledge that could only be known to a identity and/or certification authority — or not. We represent this by $2^{\{IA,CA\}}$, which denotes the power set of both types of authority knowledge. This classification becomes interesting if privacy protection in the presence of untrustworthy authorities or malicious employees is an objective (see Chapter 6).

The *activity class* describes whether the attacker is limited to passive sniffing or whether he can actively intercept, initiate, or modify communications.

An active attacker tries to change the behaviour of communication partners; for example, he may be able to provoke a higher frequency of heartbeat messages, which increases the number of samples for traffic analysis. This in turn might increase the possibility that vehicles can be tracked. (See [57] for an analysis of unlinkability related to emission frequency.)

The *spatial class* describes the relative amount of items of interest that can be observed by an attacker. In this example the r-global ratio describes the relative size of the area in which an attacker is able to observe messages. This example is directly related to the analysis undertaken in [12]. Therein success of an attacker has been examined in relation to strength of an attacker in terms of the number of observed zones. Please note that our notion of "observed" differs essentially from the notion "observed area" as used in [12].

The *objective class* describes the goal of an attacker. All attackers try to recover the trajectory of an vehicle identified either by a message, a communications certificate, a vehicle identification or an owner's identity. For simplicity reasons, we omitted unlikely objectives in Table 4.1, e. g., partitioning of certificates or vehicles.

An attacker class can now be defined by choosing an element of each class. Disregarding the continuous ratio of the space-class, this attacker model defines 40 distinct attacker classes. This classification is used as a reference in the following unlinkability analysis.

4.3 Unlinkability Analysis

The model and attacker classification described above allows description of unlinkability problems. Using this model we are able to express uncertain attackers knowledge and calculate the plausibility of unknown relations.

The following example is focused on *pseudonym changes*, a technique that has been discussed widely in vehicular communication (e. g., [22]).

Imagine an attacker who has access to a database containing relations between certificates and vehicles, i. e., an attacker in the CA-class. Compare this attacker to an attacker without this or any authority knowledge. The first attacker is able to relate messages m_1 and m_3 using this knowledge from the database. The second attacker may only gather information on relations between messages from repeated use of single certificates[1]. However, he has no information on relations between vehicles and certificates. In Figure 4.4, the additional authority knowledge of the first attacker is depicted by dotted lines. The second attacker is hindered by certificate changes, while the first

[1]The attacker may obviously also derive information from the messages' content.

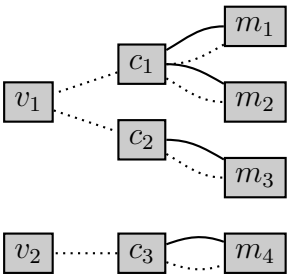


Figure 4.4: Example linkage graph: Adversary with CA-knowledge (dotted lines) vs. adversary without CA-knowledge, owner identities omitted.

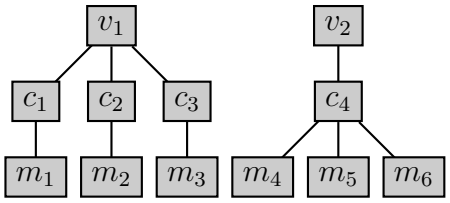


Figure 4.5: Example of pseudonym change interval. Vehicle v_1 changes the pseudonym with every message. Vehicle v_2 does not change the pseudonym. v_2 can be tracked without knowledge of $v_2 \sim c_4$.

attacker is not. If seen in linkability graph representation, it becomes obvious that there exists no path between messages related to different certificates in the second attacker's graph.

In the following, scenarios with different pseudonym-change intervals are examined. Herein the interval is expressed as the maximum number of messages n_C sent using a single certificate. Other definitions for pseudonym change intervals are discussed in the literature, see [57], [27].

Assume, vehicle v has sent the messages $M_v \subseteq M$. Thus the maximum value for n_C to be considered in analysis is $|M_v|$. In the example depicted in Figure 4.5 vehicle v_1 changes its pseudonym with every message ($n_C = 1$) while v_2 uses its pseudonym for (at least) three messages ($n_C \geq 3$).

Let vehicle v be the requester of certificates $c_1, \ldots, c_n$. Then v has used each certificate c_i, $1 \leq i \leq n$ to sign messages $m_{i,1}, \ldots, m_{i,k_i}$, $k_i \leq n_C$ and $1 \leq i \leq n$. If an attacker wants to know whether two different messages $m_{i,k}$ and $m_{j,l}$ are sent by the same vehicle he has to see whether there is a path from $m_{i,k}$ to $m_{j,k}$ or use other evidence from context information. The relation between $m_{i,k}$ and c_i is known to any receiver of the message. The events "observation of $c_i \sim v$" and "observation of $v \sim c_j$" are independent, therefore we can interpret $\mathcal{P}$ as probability distribution and multiply along the path:

$$P_A(m_i \sim m_j) = \mathcal{P}_A(V \sim C_i) \cdot \mathcal{P}_A(V \sim C_j), \tag{4.4}$$

where P_A denotes the attacker's probability of messages m_i and m_j being related. Observe that this is not the same as the context information mentioned above. Context information would provide further evidence towards or against this relation.

From the attackers belief, the certainty of an attacker that a set of messages M was sent by vehicle v can be calculated. The set of possible relations between messages $\{(m, m') \in M \times M | m \neq m'\}$ is not a set of independent events, thus we cannot apply the product rule here. In [30] the minimum of all evidence supporting a hypotheses (e. g., "$m_i \sim m_j$ are in G") is used for the basic weighting function $P_A(T)$ of equivalence cluster[2],

$$\begin{aligned} P_A(T) = &\min\{P_A(m_i \sim m_{i+1}) | 1 \leq i < |T|\}, \\ &T =< m_1, \ldots, m_{|T|} > . \end{aligned} \tag{4.5}$$

Which states that the probability of the sequence T is the minimum the probabilities of adjacent messages in that sequence.

Further combination of equivalence classes and following normalization yields a probability density function for equivalence relations. This probability density can be used for calculation of the *degree of anonymity* as described in [63].

4.4 Conclusion

This linkability model provides a structured approach to anonymity and unlinkability problems in a credential-based scenario. By using this graph notation attackers can be described with respect to their resources and objectives. Effects of information disclosure become obvious in the weighted graph.

Furthermore this model graphically shows the strong dependencies between anonymity and unlinkability. Generally we can assume that if one can be broken by an attacker, the other one is broken too.

[2] cluster, i. e., subsets of M

Errors using inadequate data are much less than those using no data at all.

Charles Babbage

5 Unlinkability Measure

Contents

This chapter introduces a global unlinkability measure, and contributes to the struggle for a formal definition of unlinkability.

When mobile devices travel and communicate they leave a trace of emitted messages in their wake. The number and distribution of messages is commanded by device movement and communication protocols. An attacker who is able to correctly link messages can trace the movement of devices. An attacker may derive information on the relations between messages from many sources. Road-maps, traffic patterns, constraints on vehicular movement, information leaked from the certification process, radio-frequency fingerprinting, CCTV observation — any information that restricts the hypothesis space may be useful to an attacker. These sources of information are summarised under the term *context information*.

Unlinkability analysis may have different objectives. Protocol designers who develop Privacy Enhancing Technologies need to analyse the quality of their methods. Users may want to know the amount of privacy they enjoy in a certain situation, some of which are unlinkability-related. Customers may need to weight quality of service against privacy level against price.

Designers of privacy preserving communication protocols need tools to assess the quality of their protocols. The topic of this chapter is the assessment of the amount or quality of unlinkability. Unlinkability can be understood as the inability of an attacker to track moving devices from the messages emitted by them.

An *unlinkability measure* is a quantitative assessment of unlinkability for a given scenario. A scenario is defined by message emissions and contents, attacker models and the context information available to attackers. In this chapter a measure is introduced that assesses unlinkability depending on a specific attacker and a given message set in a defined scenario. We use the scenarios described in Section 3.

Input variables which influence the unlinkability measure are the attributes and distribution of items of interest and the type and amount of context information that is available to an attacker. The attributes of items of interest and their distribution depend on the used communication protocols, e. g., message emission protocols or certification protocols (see Chapter 6). The context information depends on the resources available to an attacker.

Considering the different situations under which an analysis can be undertaken, different levels of knowledge on the side of the analyst must be considered. In this work white-box and black-box analysis are considered. In an *white-box analysis* details of the attacker and the vehicular scenario (especially the true sender-relation) are known to an analyst (but still not to the attacker). The analyst uses the measure to compare the attacker with the real situation. In a *black-box analysis* only information known to the attacker is available to estimate the quality of an attack. White-box analysis describes, for example, the situation of a protocol designer who simulates and analyses communication protocols. Black-box analysis may be undertaken by an attacker to assess the quality of his attack.

This chapter is structured as follows. In Section 5.1, fundamental terms such as unlinkability measures, items of interest, attackers, and unlinkability-analysis are introduced. The notion of consistency motivates most of the concepts discussed in this thesis and is introduced in Section 5.2. In Section 5.3, the model of inner and outer structure is introduced, which is used to criticise related measures in Section 5.4. A new expected distance unlinkability measure is proposed in Section 5.5. The practical measurement is

finally discussed in Section 5.6 and Section 5.7. The chapter is closed by a conclusion in Section 5.8.

5.1 Basics and Notation

In this section, basic notations are introduced. It can be used as a reference for symbols and notation in the sections that follow afterwards.

5.1.1 Abstraction of Vehicular Messages

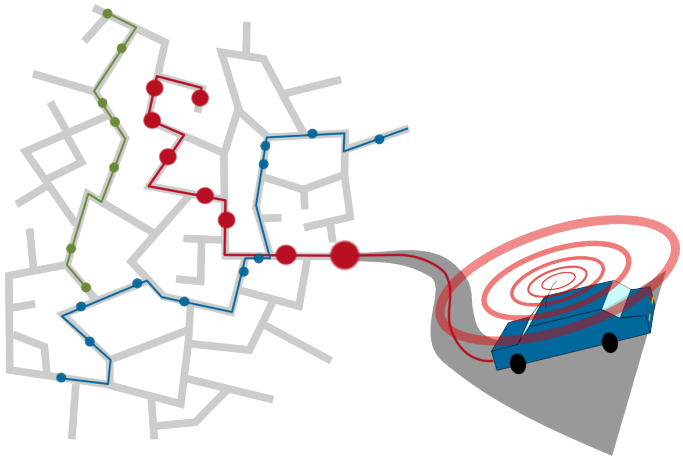
An *item of interest* is an element m from a set of IOI denoted M. M does not necessarily contain every existing IOI, but only the IOI known to an attacker. In this work, a global passive adversary (see Section 4.2) who observes all *emitted messages* in a vehicular scenario is taken as example adversary. Thus the set IOI will often be called set of *messages* and an individual IOI is often denoted *message*. If the need arises to distinguish between emitted and observed messages, this will be noted explicitly.

Each message/IOI represents a sample point in spatial, temporal and other spaces often represented as attributes of the message. Without loss of generality, only two attributes, *temporal position* (or emission time) and *spatial position*, are used as an example in this chapter.

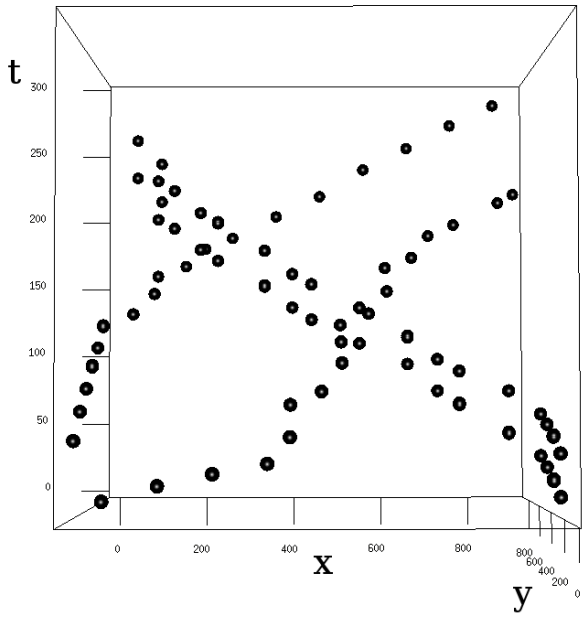
In Figure 5.1, two different illustrations of the vehicular message scenario are provided. Figure 5.1(a) shows traces of vehicles with the spatial points where messages have been emitted marked by coloured circles. The image shows the vehicles' itineraries on a road network which restricts the movement of the emitting vehicles. The temporal information obviously cannot be shown here. In Figure 5.1(b), the temporal and spatial information of position samples is represented as position in a 3-dimensional plot. The trajectories, and thus the relation between sample points, have been omitted as well as the road network. While the attacker probably has knowledge of the road network, the relation between the itineraries is not previously known. The latter image depicts the situation of an attacker where a subset of position samples within a spatial and temporal interval is received and has to be clustered to reveal the trajectories of individual vehicles.

5.1.2 Sample Points

Temporal positions are denoted as $t \in \mathcal{T}$, where $\mathcal{T}$ denotes the space of time, e. g., $\mathcal{T} = \mathbb{R}_{+0}$. The mapping $t : M \to \mathcal{T}$ is used for explicitly denoting the point in time $t(m)$ related to a given message/sample point $m \in M$. Likewise, spatial positions are denoted $a \in \mathcal{A}$, where $\mathcal{A}$ denotes an arbitrary



(a) Communicating vehicles leave traces of position messages while navigating through the traffic. Attackers are interested in linking position samples to discover the paths of single vehicles.



(b) An example set of anonymous, position samples in spatial space (x and y) and temporal space (t) from multiple (here 4) sources (vehicles).

Figure 5.1: Vehicular communication generates anonymous samples in time and space that attackers can use to trace vehicles.

spatial space. Again, $a : M \to \mathcal{A}$ denotes the position $a(m)$ in relation to a message m. In practise spatial and temporal information is either observed by the attacker or explcitly encoded in the message.

In this work directed graphs are used to model road-map networks. This approach does not restrict generality, as other spaces, such as spherical coordinates commonly used in GPS-navigation, work similarly. The road-network is modelled as a graph $G = (V, E)$, where V denotes junctions, and $E \subseteq V \times V$ denotes roads, i. e., directed connections between junctions. Additionally, the length of roads may be given by additional edge-weight-functions.

An attacker tries to correctly link sample points with respect to some unknown attribute, e. g., the real sender. Linking sample points means to determine a subset of elements of M. The positions encoded in the messages of a subset can be interpreted as samples on a *trajectory*, e. g., the trajectory of an vehicle in space and time. For simplicity τ is used for two different notations of "trajectory". First as a mapping from time to space $\tau : \mathcal{T} \to \mathcal{A}$ and second for a subset of sample points $\tau \subseteq M$. This can be considered valid in the example scenario, as trajectories of vehicles are derived from subsets of the message set M.

5.1.3 Partitions and Hypotheses

Determining all trajectories within M amounts to determine an equivalence relation on M. An *equivalence relation* on a set M is isomorphic to a set partition π on M, consisting of *clusters* τ, which represent trajectories. The relation symbol $m \sim_\pi m'$ denotes that two messages m, m' are equivalent with respect to the equivalence relation defined by π.

We denote the set of all partitions on M by Π_M. The objective of an *global attacker* is to find the partition of M that is isomorph to the true relation, i. e., corresponds to the reality. Thus every partition on M is a *hypothesis*, and Π_M is sometimes called *hypotheses space*.

As usual, all clusters of a partition are pairwise disjoint and the union of all clusters of a partition is M. The *number of clusters* in a partition π, i. e., the cardinality of π, is denoted as $|\pi|$, where $\pi = \{M_1, \ldots, M_n\}$. The notion $m \sim_\pi m'$ is equivalent to stating that m, m' "are related" under hypothesis π. Usually this notion means *sender-equivalence*, i. e., m and m' "have the same sender". Relation between IOI is denoted as

$$\begin{aligned} m \sim_\pi m' &: \Longleftrightarrow \exists i : m \in M_i \wedge m' \in M_i \\ m \nsim_\pi m' &: \Longleftrightarrow \forall i : m \in M_i \Rightarrow m' \notin M_i, \end{aligned} \tag{5.1}$$

where $M_i \in \pi$.

The process of partitioning a set into disjoint subsets, i. e., clusters, is known as *clustering* in literature. The quality of *clustering algorithms* can be measured. Those measures provides no complete estimation of the unlinkability situation, because it evaluates only one candidate partition.

Besides the common set notation for partitions, we use an alternative "cluster notation" to define clusterings of M. In *set notation* one set partition is given as a disjoint set of subsets of M. For example $\{\{m_1, m_2, m_3\}, \{m_4, m_5\}, \{m_6\}\}$ provides a set partition of $M = \{m_1, \ldots, m_6\}$. In *cluster notation* a sequence of set partition identifiers is used. The i-th element in a cluster notation sequence identifies the cluster of the i-th element of the strictly ordered set M. For example, the above set partition can be written as $\langle 1, 1, 1, 2, 2, 3\rangle$ in cluster notation. This example defines an equivalence relation where $m_1 \sim m_3$, but $m_2 \nsim m_5$.

Without loss of generality, we assume that a total order can be imposed on M, e. g., by an order on temporal and/or spatial spaces. The ordered set M is denoted $\{m_1, \ldots, m_n\}$, where the index denotes the ordinal number.

By the total order on M we can inflict a total order on Π_M, e. g., by using a lexicographic order on partitions in cluster-notation. Single partitions are denoted by π, often with an index subscript that denotes its ordinal number $ord(\pi_i) = i$ in $\Pi_M = \{\pi_1, \ldots, \pi_k\}$. Note that cluster notation and the definition of $\sim_\pi$ imply that π is invariant to the order of messages within clusters, otherwise π would not induce an equivalence relation, i. e., an reflexive, symmetric, and transitive relation.

In the following discussions in Sections 5.3.4 and 5.6, a toy example, consisting of the set M of cardinality 3 and set partitions Π_M, is used:

$$\begin{aligned} M &:= \{m_1, m_3, m_3\} \\ \Pi_M &:= \{\pi_1, \pi_2, \pi_3, \pi_4, \pi_5\} \\ &= \{\langle 1,1,1\rangle, \langle 1,1,2\rangle, \langle 1,2,1\rangle, \langle 1,2,2\rangle, \langle 1,2,3\rangle\}, \end{aligned} \tag{5.2}$$

where Π_M is ordered lexicographically in cluster-notation. The single set partitions are denoted $\pi_1, \ldots, \pi_5$ according to the order given above.

Set partitions directly correspond to a trace of devices, e. g., vehicles. Set partition π_1 would be interpreted as "only one vehicle has been sending messages, it successively passed the positions given in m_1, m_2, m_3". Set partition π_4 denotes "one vehicle has passed point m_1 and a different vehicle successively passed the positions m_2 and m_3".

Note that only a very limited example can be used in this work, as the number of set partitions $|\Pi_M|$ grows exponentially in $|M|$. $|\Pi_M|$ is given by the *Bell Number*[1] $B_{|M|}$ of $|M|$. For better readability, we use $B(n)$ to denote

[1] Weisstein, Eric W. "Bell Number." From MathWorld–A Wolfram Web Resource. `http://mathworld.wolfram.com/BellNumber.html`

B_n. The Bell Number of n can be calculated much faster[2] than generating B_n elements.

One partition represents a very special hypothesis. The *true hypothesis* π^* is the one partition that corresponds to the reality, i. e., all messages m, m' with $m \sim_\pi m'$ are sent by the same sender. The objective of an attacker is to determine the true relation.

5.1.4 Attacker's Assignment and Choice

In this work we assume that an attacker is generating a probability mass assignment $P : \Pi_M \to [0, 1]$ with $\sum_{\pi \in \Pi_M} P(\pi) = 1$. An *attacker's assignment* represents the attacker's global knowledge/belief about which equivalence relation represents the true relation best.

An attacker's probability mass assignment P induces a partial order on Π_M. It is reasonable to assume that an attacker bases its decision on the assigned probabilities, i. e., chooses the set partition with the highest probability mass as his hypothesis of the true partition. To make the choice deterministic the partition's ordinal number $ord : \Pi_M \to \mathbb{N}$ is used as a second order attribute. We define the strict *attacker's order* $<_P$ as:

$$\begin{aligned} \pi_i <_P \pi_j :\Longleftrightarrow\ & P(\pi_i) > P(\pi_j) \vee \\ & [P(\pi_i) = P(\pi_j) \wedge ord(\pi_i) < ord(\pi_j)] \,. \end{aligned} \tag{5.3}$$

Please not that the attacker's order reverts the order of the probability mass assignment P.
An attacker A chooses the first set partition, i. e., the "smallest" π_A from Π_M. We denote by π_A the *attackers choice* of attacker A. π_A is chosen such that

$$\forall \pi \in \Pi_M : \pi_A <_P \pi \tag{5.4}$$

An attacker's choice π_A represents the solution of attacker A for the (unlinkability) problem of clustering M accordingly to the unknown equivalence.

5.2 Consistency

Consistency is the main concept that distinguishes the unlinkability measure introduced in this work. An assignment's consistency is an important property that reflects the quality of an attacker's choice for a given scenario.

[2] e. g., Dobiński's formula Weisstein, Eric W. "Dobiński's Formula." From MathWorld–A Wolfram Web Resource. `http://mathworld.wolfram.com/DobinskisFormula.html` 1.4.2008

Consistency is the motivation behind the distinction of inner and outer structure as introduced in Section 5.3. *Consistency* of an attacker denotes that the attacker's assignment does not, or does to a low degree, support contradictory hypotheses.

Consistency may be explained by the following example. Let there be three hypotheses π_1, π_2, and π_3, where π_1 and π_2 are very similar, while π_3 contradicts both π_1 and π_2. By *contradiction* we describe that hypotheses represent very different relations. Let A and B be two attackers with probability mass assignments P_A and P_B, where

$$P_A(\pi) = \begin{cases} 0 & \pi = \pi_3 \\ 0.5 & otherwise, \end{cases} \qquad P_B(\pi) = \begin{cases} 0 & \pi = \pi_1 \\ 0.5 & otherwise. \end{cases}$$

As π_1 and π_2 are very similar, it makes not a big difference whether an attacker chooses one or the other. Now observe, as π_3 contradicts π_2, that P_B reflects a very ambiguous attacker's choice. The attacker is not able to find a consistent view from context information and scenario but is indecisive between two contradicting hypotheses π_2 and π_3. In contrast, P_A reflects a consistent attacker's view as non-zero weights are assigned to all similar hypotheses but not to others. Obviously the quality of both assignments is very different. The term we use to describe this difference is *consistency*.

The attacker's assignment depends on probabilistic *context information* that is (always) tainted with uncertainty due to coarse granularity and inaccuracy. This uncertainty is transferred to the assignment. Assume, for example, an attacker whose context information consists of a distribution of vehicles' velocities derived from previous traffic observation. Context information is granular, i. e., it is a summary of information of one specific spatial or temporal interval. This context information also is inaccurate as it is learned from an earlier observation. Furthermore, statistical context information carries some uncertainty in itself.

Thus, an attacker has to take into account that his probability mass assignment maybe flawed. Slightly different messages M or the inaccuracy of the context information may lead to changes of the assigned probabilities. As a result, the attacker's order of partitions (5.3) is tainted with uncertainty too. Thus also the attacker's choice (5.4) is tainted with uncertainty.

An attacker's assignment, herein, is said to be *consistent*, if the attacker's choice is robust[3] against small changes in the assignment, i. e., a choice is robust if small changes in the probability mass assignment produce only small differences in the choice. This implies that, if the attacker's choice changes due to a small change in a consistent attacker's assignment, the new choice

[3]Another term describing a similar property in statistics is *robustness*.

is very similar to the old choice.

5.3 Inner and Outer Structure

In this section, the notions of inner and outer structure are introduced. While outer structure describes measures imposed onto hypotheses space Π_M by the attacker, inner structure denotes an intrinsic distance on Π_M. Inner structure is a novel concept in unlinkability measures that has not been considered in previous works.

The quality of clustering algorithms in a *white-box analysis* is often measured by notions of difference between the cluster algorithms result, herein called the attacker's choice π_A, and the true relation π^*. This distance between π_A and π^* captures only punctual information about the quality of the clustering algorithm, i. e., considers only two elements of the hypotheses space. It thus does not consider the complete attacker's view as it is represented by an attacker's assignment. As this work is focused on situations where an attacker's world-view can be expressed as probability mass assignment, ignorance of this information must reduce the expressiveness of the metric.

In the remainder of this section, structures within the measured object, namely an attacker's probability mass assignment P_A, are uncovered. The notions of inner and outer structure are explained.

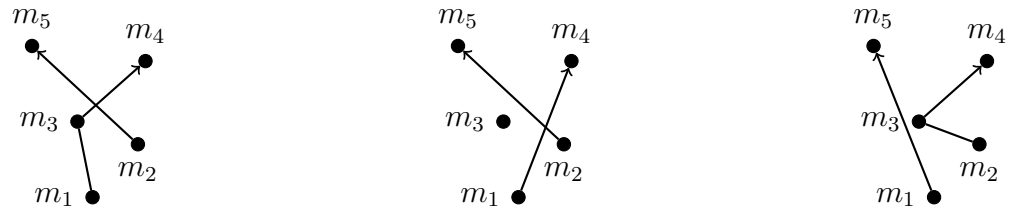
5.3.1 Outer Structure

The term *outer structure* denotes a measure that is externally imposed on Π_M, e. g., by an attacker. Outer structure is, contrary to inner structure, dependent on context information and is not an intrinsic attribute of the hypotheses space.

Outer structure in this work is defined by P_A. It provides the attacker's order and thus defines the attacker's choice. P_A defines a probability distribution on Π_M. Outer structure generally denotes a quantification of elements of Π_M that induces an order.

5.3.2 Inner Structure

The term *inner structure* captures an intrinsic notion of *distance* between set partitions in Π_M. The notion of distance between two set partitions reflects dis-similarity of two hypotheses for a given scenario. The notion of distance should be based on the objectives of an attacker.



(a) π_{17} (b) π_{37} (c) π_{29}

Figure 5.2: Example set partitions $\pi_{17}, \pi_{37}, \pi_{29}$ from Π_{M_5} with $M_5 = \{m_1, \ldots, m_5\}$ depicted as possible traces. ($\pi_{17} = \langle 1, 2, 1, 1, 2\rangle$, $\pi_{37} = \langle 1, 2, 3, 1, 2\rangle$, $\pi_{29} = \langle 1, 2, 2, 2, 1\rangle$)

Take for example an attacker who is not interested in trajectories of devices, but only in endpoints of an vehicle's journey[4]. This attacker would sensibly interpret trajectories with similar endpoints as equal. The definition of distance between set partition should reflect this interpretation of hypotheses.

Distance between hypotheses is a (metric) function $\delta : \Pi_M \times \Pi_M \to \mathbb{R}_{+0}$. Two hypotheses are considered equal if and only if the distance is zero. A larger distance denotes that the hypotheses are less similar.

Motivation

The distance δ between two set partitions is large if they represent very different trajectories. The distance between two set partitions is small if they represent similar traces. In Figure 5.2 three set partitions of the same set $M = \{m_1, \ldots, m_5\}$ are shown as spatial traces of vehicles.

In Figure 5.2(a) an example partition $\pi_{17} = \langle 1, 2, 1, 1, 2\rangle$ is depicted. Two trajectories, one consisting of m_1, m_3, m_4 and the other consisting of m_2, m_5 are depicted. In Figure 5.2(b) the same sample points are partitioned into $\pi_{37} = \langle 1, 2, 3, 1, 2\rangle$, which defines three trajectories. The first trajectory consists of points m_1, m_4, the second of the single point m_3 and the third of m_2, m_5. Figure 5.2(c) again shows two trajectories that comprise partition $\pi_{29} = \langle 1, 2, 2, 2, 1\rangle$ with clusters m_1, m_5 and m_2, m_3, m_4.

Assume that an attacker is interested in tracking the trajectories of vehicles. Intuitively, trajectories in Figures 5.2(a) and 5.2(b) are more similar than 5.2(a) and 5.2(c), because at least the endpoints of the single traces are the same in 5.2(a) and 5.2(b).

In different scenarios an attacker might only be interested in correctly

[4]Assume that there exists a way to distinguish "endpoints" from "normal" items of interest.

	$m \sim_{\pi_j} m'$	$m \nsim_{\pi_j} m'$
$m \sim_{\pi_i} m'$	a	b
$m \nsim_{\pi_i} m'$	c	d

Table 5.1: Binary Contingency Table. The four fields contain the number of tuples $(m, m') \in M \times M, m \neq m'$, for which the row-condition AND the column condition holds.

$\delta_H(\pi_i, \pi_j)$	π_1	π_2	π_3	π_4	π_5
π_1	0	1	1	1	2
π_2	1	0	1	1	1
π_3	1	1	0	1	1
π_4	1	1	1	0	1
π_5	2	1	1	1	0

Table 5.2: Partition Distance δ_H of example set partitions $\Pi_M = \{\pi_1, \ldots, \pi_5\}$ defined as sum $\delta_h(\pi_i, \pi_j) := b + c$. (See Table 5.1)

guessing the number of vehicles. To this attacker π_{17} and π_{29} with trajectories of two vehicles would have an equal interpretation while π_{37} with three vehicles yields a different scenario.

A metric on a hypotheses-space has to reflect the criteria that are important to an attacker, independent of any context information on the relation between elements. This distance provides information on the semantic difference which can be used to estimate errors, as will be done in Section 5.5. The notion of inner structure is fundamental for expressing *structural consistence*, which is introduced in Section 5.2 and should be reflected by any useful unlinkability measure.

5.3.3 Partition Distance

A simple example for a hypotheses distance is the partition distance. The *partition distance* between two partitions π and π' of M is the minimum number of elements that have to be deleted from M so that both partitions, restricted to the remaining elements, are equal. This partition distance may be thought of as a "natural" inner structure with respect to similarity of partitions.

Definition 5.1 (Partition Distance)
The partition distance $\delta_H(\pi, \pi')$ *of two partitions* π, π' *of a given set* M *is the minimum number of elements that have to be deleted from* M *that* π *and* π'*, restricted to the remaining elements, are equal.*

Partition distance can be calculated by using a binary contingency table

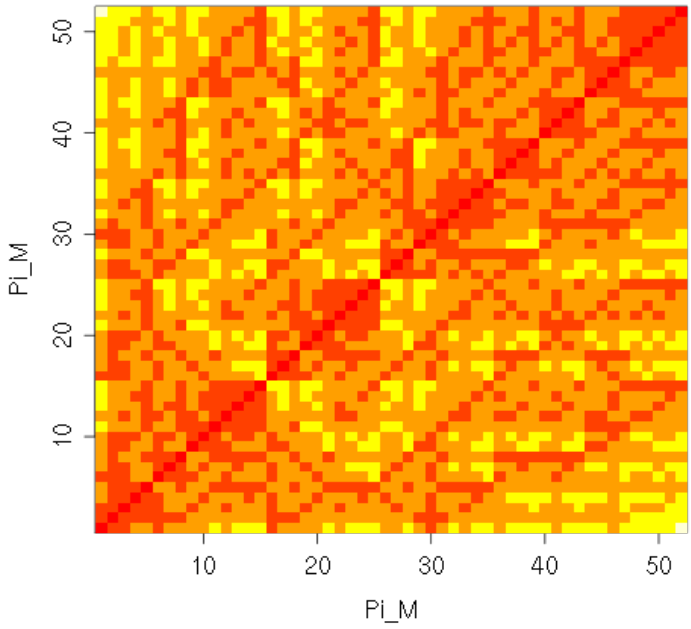


Figure 5.3: Heat-plot of set partition distance for partitions of M with cardinality 5. The axes show ordinal numbers of partitions, the colour encodes the distance between the x and y-partition. Darker colours denote smaller distance between partitions.

as given in Table 5.1, or by a polynomial time algorithm using minimum assignments as described in [29]. In Table 5.2, partition distances for the example of a 3-elemental set M is given. In Figure 5.3 a heat-plot of partition distances for a set of cardinality 5 is given. The pattern of the plot shows that even the simple partition distance provides a complex inner structure.

The partition distance is a metric that induces an inner structure for the most general case where messages contain no attributes, e. g., an underlying spatial or temporal structure. This is, of course, not the case in the vehicular scenario where we assume that messages contain information on temporal and spatial positions as well as possibly a movement vector.

5.3.4 Trajectory-Based Partition Distance

To incorporate information from the underlying temporal and spatial structure of the sample space, partition distances are insufficient. In this section, a partition distance is introduced that inherits the metric attributes of an underlying distance metric defined on subsets of M. A subset τ of the mes-

sage set M is interpreted as a trajectory or path on a road map. The road map herein is represented as directed graph.

The *Euclidean distance* gives the "length of the direct way" between to points. This direct route is not necessarily a possible way in the road-map which defines many restrictions on movements, e. g., one-way-streets, no-right-turn-junctions and the general restriction to certain subspaces of the plane (i. e., roads). Because of this unwanted simplification we decided on a graph-based road model and distance measure.

In the following, a trajectory distance is introduced. Afterwards it is shown how trajectory distance can be used to define a partition distance.

Trajectory Distance in Directed Graphs

In this section, a trajectory distance is defined for a directed-graph road-map and message set M. Formally, a *trajectory* is a function $\tau : \mathcal{T} \to \mathcal{A}$ that maps temporal values $\mathcal{T}$ onto spatial positions $\mathcal{A}$, i. e., trajectories are formally considered functions that map time onto position.

The *trajectory distance* is defined as average distance between points of two continuous trajectories (taken from [68]):

$$D(\tau_1, \tau_2) = \frac{\int_{t_s}^{t_s+|T|} d(\tau_1(t), \tau_2(t))dt}{|T|}, \tag{5.5}$$

where τ_1, τ_2 represent trajectories. The function $d : \mathcal{A}^2 \to \mathbb{R}_{+0}$ represents the spatial distance between two points as defined below. T is the length of the considered time-interval, and t_s is its start.

Trajectory distance D is based on a point distance d. As mentioned in Section 5.1 items of interest resemble points in space and time. In this section, spatial positions are represented in a directed road-map graph $G = (V, E, o)$ with V being the vertices of the graph, $E \subseteq V^2$ being the directed edges of that graph and $o : E \to \mathbb{R}_{+0}$ defining the length of edges. The spatial point $a \in E \times \mathbb{R}_{+0}$, related to message m is given by the tuple (e, o), where e denotes the edge where the point is situated and o denotes the offset from the beginning of the edge. The distance d is defined as the length of the shortest path between two points a, a', situated on edges e, e' accordingly.

The functions *to* and *from* map edges onto their target- or source-node. The shortest path between a and a' is the shortest path between the vertexes $to(e)$ and $from(e')$ plus the remaining offsets on e and e'. The shortest path sp is defined as the sequence of edges which has the smallest sum of edge weights. The distance between a and a' is thus the remaining length of e, plus the length of $sp(to(e), from(e'))$, plus the offset o' of a' on e', denoted

$$d(a, a') := o' - o + \begin{cases} 0 & e = e' \wedge o \leq o' \\ o(e) + ||sp(to(e), from(e'))|| & otherwise, \end{cases} \tag{5.6}$$

where $e, e' \in E$ denote the edges where $a, a' \in M$ are situated, $o, o' \in \mathbb{R}_{+0}$ denote the corresponding offset of a from the start of edge e, e', and $o(e)$ the length of edge e. The function d distinguishes the cases where a' from a are on the same directed edge and a' lies in the edge's direction from a (first case) and thus a' can be reached from a without crossing any vertices. The second case describes any other situation where at least one vertex has to be crossed.

It is assumed that all edge lengths $o(e)$ are non-negative and that the triangle-inequality holds for the length of the shortest path *sp* in the directed graph G. Obviously this distance is a *quasi-metric* (i. e., it is not symmetric) because of the underlying non-symmetric paths in directed graphs.

Metric Attributes of Spatial Distance

In this section, we show that d defined in (5.6) is a quasi-metric. This is done by successively showing that d is non-negative, that identical samples are indiscernible by d, and that the triangle-inequality holds for d. In the following we use sample positions $a = (e, o)$, $a' = (e', o')$, and $a'' = (e'', o'')$.

Non-Negativity. *Non-negativity* of d obviously follows non-negativity of edge weights $o(e)$. Every offset o of every $a = (o, e)$ is non-negative and $o < o(e)$. Thus $d(a, a') \geq 0$ for every a and a' on the road map.

Identity of Indiscernibles. We have $d(a, a') = 0 \iff a = a'$, if the road-map graph has no negative edge length and all point offsets o are well formed, i. e., not larger than the corresponding edge length.

Identity $m = m'$ of two points is defined as equivalence of edges $e = e'$ and offset $o = o'$.

Lemma 5.1 (Identity of Indiscernibles of d)
For non-negative edge-weights $o(e)$ the point distance $d(a, a')$ is zero if and only if $a = a'$.

Proof 5.1 ($d(a, a') = 0 \iff a = a'$)
The contra-positive of the implication $d(a, a') = 0 \Rightarrow a = a'$ is given by $d(a, a') \neq 0 \Leftarrow o \neq o' \vee e \neq e'$. First the implication is shown for $e \neq e'$ and then for $o \neq o'$.

For $e \neq e'$ we can conclude that

$$\begin{aligned} e \neq e' \quad &\Rightarrow \quad d(a,a') = o' - o + o(e) + ||sp(to(e), from(e'))|| \\ &\overset{o<o(e)}{\Rightarrow} \quad d(a,a') \neq 0, \end{aligned}$$

since $||sp||$ is always greater than or equal to zero.

For the second condition $o \neq o'$ we have to consider the two cases of d separately. The first case $o < o'$ considers the situation that the shortest path from a to a' never traverses any vertex. We thus have

$$\begin{aligned} o < o' \wedge e = e' \quad &\Rightarrow \quad o' - o > 0 \\ &\Rightarrow \quad d(a,a') \neq 0. \end{aligned}$$

The second case $o > o'$ considers the case where at least one vertex has to be traversed on the shortest path. Therefore we have

$$\begin{aligned} o > o' \wedge e = e' \quad &\overset{o<o(e)}{\Rightarrow} \quad o' - o + o(e) > 0 \\ &\Rightarrow \quad d(a,a') \neq 0. \end{aligned}$$

The other direction $d(a,a') = 0 \Leftarrow a = a'$ is a metric property which must hold for $d(a,a')$ by definition. This concludes the proof. □

Triangle-Inequality Triangle-inequality states that a metric measures the shortest path between two points. If an intermediary point is inserted in a path, then the sum of the distances between start and intermediary point plus intermediary point to end can be not smaller than the original distance.

Lemma 5.2 (Triangle-Inequality for $d(a,a')$)
For all points a, a', and a'' we have $d(a,a'') \leq d(a,a') + d(a',a'')$ for a non-negative sample point distance d as defined in (5.6).

Proof 5.2
To show that the *triangle-inequality* $d(a,a'') \leq d(a,a') + d(a',a'')$ holds for d, one can distinguish four cases of relations between the edges e, e', e''. In general, this proof runs along the same lines as standard proofs for triangle-inequality of shortest-path distance in graphs. If the intermediary point a' is on the shortest path from a to a'' then the right and left hand side of the inequality are equal. If the intermediary is not on the shortest path, then the right hand side is larger.

Let $a, a'' \in \mathcal{A}$ be sample points called start- and end-point. Let $a' \in \mathcal{A}$ be another point called intermediary point.

The proof distinguishes first whether the intermediary point a' is on the same edge as any or both of the endpoints a, a'' or not. Then four cases of edge equalities have to be considered:

(a) $e = e' \neq e''$

(b) $e \neq e' = e''$

(c) $e \neq e' \neq e''$

(d) $e = e' = e''$

First, consider the case that the intermediary point a' is on the shortest path from a to a''. Then, in case (d) $o \leq o' \leq o''$, case (c) $e' \in sp(to(e), from(e''))$, case (b) $o' \leq o''$, and case (a) $o \leq o'$.

As a' is on the shortest path, then it cuts the path on an edge. For $e = e'$ the length $o' - o$ before a' is part of the path, otherwise the whole length o' on the edge e' is on the path. For $e' = e''$ the length $o'' - o'$ is on the path after a', otherwise it is $o(e') - o'$.

For case (a) and a' on the shortest path from a to a'' the triangle inequality holds as follows

$$\begin{aligned} d(a, a'') &\overset{def.}{=} o'' - o + o(e) + ||sp(to(e), from(e''))|| \\ &= o' - o + o'' - o' + o(e) + ||sp(to(e), from(e''))|| \\ &\overset{e=e'}{=} d(a, a') + d(a', a''). \end{aligned}$$

The proof for case (b) runs along the same lines and case (d) is a trivial combination of both.

Case (c) takes advantage of the definition of the length of the shortest path which is the sum of the length-attribute of the edges in the path. Equality then is derived from the definition of d as

$$\begin{aligned} d(a, a'') &\overset{def.}{=} o'' - o + o(e) + ||sp(to(e), from(e''))|| \\ &= o'' - o + o(e) + ||sp(to(e), from(e'))|| \\ &\quad + o(e') + ||sp(to(e'), from(e''))|| \\ &= o' - o + o(e) + ||sp(to(e), from(e'))|| \\ &\quad + o'' - o' + o(e') + ||sp(to(e'), from(e''))|| \\ &\overset{def.}{=} d(a, a') + d(a', a''). \end{aligned}$$

Thus, if a' is on the shortest path from a to a'', the triangle-inequality holds for d.

Now consider cases (a), (b), (c), and (d) with a' not being on the shortest path from a to a''. The triangle inequality again follows from the definition of d.

(a) $e = e'$, if $o' \geq o$ then a' is on the shortest path from a to a''. If $o' \leq o$ then, due to the direction of the edge e, it is not on the shortest path. The combined path, using a', is longer than the direct path because the direct path is a part of the combined path.

$$\begin{aligned} d(a,a') + d(a',a'') &\overset{def.}{=} o' - a + o(e) + ||sp(to(e), from(e'))|| \\ &\quad + o'' - o' + o(e') + ||sp(to(e'), from(e''))|| \\ &\overset{e=e'}{=} o' - o' + o(e) + |sp(to(e), from(e))|| \\ &\quad + o'' - a + o(e) + ||sp(to(e), from(e''))|| \\ &\overset{def.}{=} o(e) + ||sp(to(e), from(e))|| + d(a,a'') \\ &\geq d(a,a'') \end{aligned}$$

From non-negativity of $||sp||$ and edge length $o(e)$, it follows that the direct distance $d(a,a'')$ is at most as long as the combined distance $d(a,a') + d(a',a'')$.

(b) The proof for case (b) runs along parallel lines to (a) but uses the condition $e' = e''$. It turns out that

$$\begin{aligned} d(a,a') + d(a',a'') = &\quad o(e'') + |sp(to(e''), from(e'')| + d(a,a'') \\ \geq &\quad d(a,a''). \end{aligned}$$

(c) To show that the triangle inequality holds in the case that a' is not on the shortest path from a to a'' we only have to consider the path $sp(to(e), from(e''))$ as every itinerary has to pass $to(e)$ as the first vertex and $from(e'')$ as the last vertex of the road map. Thus the problem is reduced to the question of whether the triangle-inequality holds for sp in directed graphs, which was postulated at the beginning.

Thus in case (c), if a' is not on the shortest path, the triangle-inequality holds for d.

(d) To show that the triangle inequality holds in case that $e = e' = e''$ and a' not on the shortest path from a to a'', five possible permutations of a, a', a'' have to be considered with respect to their respective offsets o, o', o'' on edge e. The permutations to be considered here are (a, a'', a'), (a', a, a''), (a', a'', a), (a'', a, a'), and (a'', a', a).

The argument is, again, that every direct path is included in the combined path, and that the combination is at least as long as the direct

path. The proof for the first of the above permutations of a, a', a'' is

$$\begin{aligned} d(a,a') + d(a',a'') &\overset{def.}{=} o' - o \\ &\quad + o'' - o' + o(e') + ||sp(to(e'), from(e''))|| \\ &\overset{def.}{=} d(a,a'') + o(e) + ||sp(to(e), from(e))|| \\ &\geq d(a,a''). \end{aligned}$$

The other permutations follow along the same reasoning and are omitted here. Thus, given that $||sp||$ and $o(e)$ are non-negative, the triangle inequality holds for case (d).

As the triangle-inequality holds for every case it is proven that it holds for $d(a,a')$. □

Symmetric Spatial Distance

Spatial distance in directed graphs is a very natural distance in the vehicular scenario but provides no symmetry. As symmetry is an important feature needed for the unlinkability measure provided in Section 5.5, the directed graph model has to be relaxed.

There are plenty of ways to create symmetry but considering that it is sensible to respect road features like one-way roads, it is useful to keep as many attributes of the road map as possible. We define the point distance of two points a, a' as minimum of the distance from a to a' and the reverse direction a' to a:

$$d_{sym}(a,a') := min(d(a,a'), d(a',a)). \tag{5.7}$$

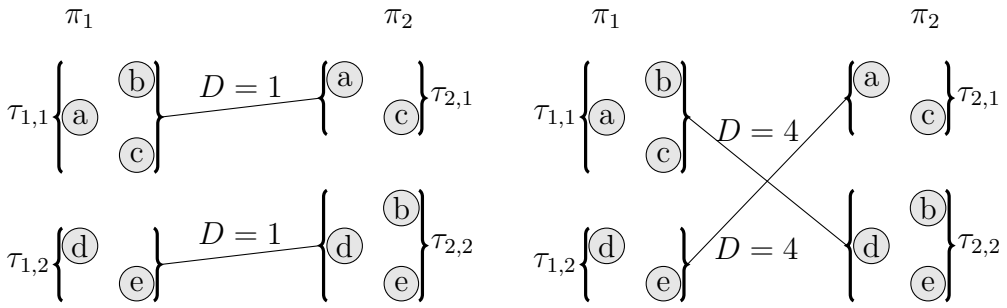
The minimum of $d(a,a')$ and $d(a',a)$ provides a symmetric distance d_{sym} but keeps directions of edges. In the following, we use this distance if symmetry is required.

Trajectory-Based Partition Distance

Trajectory or path distances, as defined in Section 5.3.4, can be used to define a distance on partitions. The notion of distance caries the notion of the minimum effort that is necessary to transform one state into another state. Similarly we herein define trajectory-based distance between partitions relative to the minimum amount of changes necessary to transform one partition into another.

A partition consists of clusters (subsets of M), thus a transformation from one partition to another means transforming every cluster of the first partition into a cluster of the second partition. This can be modelled by the

well known assignment problem. In this section, we define partition distance by way of a minimum assignment of trajectories between two partitions.



(a) $D_{f_1}(< 1, 1, 1, 2, 2 >) = 2$,
where $f_1(\{a, b, c\}) \mapsto \{a, c\}$
and $f_1(\{d, e\}) \mapsto \{b, d, e\}$

(b) $D_{f_2}(< 1, 1, 1, 2, 2 >) = 8$,
where $f_2(\{a, b, c\}) \mapsto \{b, d, e\}$
and $f_1(\{d, e\}) \mapsto \{a, c\}$

Figure 5.4: Example for minimum assignment for distance between partition π_1 and π_2.

Observe the example in Figure 5.4 which shows the two assignments $f_1, f_2 : \pi_1 \to \pi_2$ that map all elements (cluster) in π_1 onto elements in π_2. Further be D a distance function[5] of subsets of $M = \{a, b, c, d, e\}$. The distances in Figure 5.4(a) are smaller than those in Figure 5.4(b), because e. g., partition $\tau_{1,1}$ contains two elements a and$c$ which are also contained in $\tau_{2,1}$, but not in $\tau_{2,2}$. Assignment f_1 is a more direct way to transform π_1 into π_2 and should be chosen if the minimum effort is to be expressed. Assignment f_1 is said to be the minimum assignment in this example.

A *minimum assignment* is a one-on-one mapping of elements of one set onto elements of another set that is minimum with respect to some cost function. A minimum assignment can be found in polynomial time. [40] In general it is the problem of finding a bijective map between two sets which is minimal with respect to costs defined for each pair. For simplicity reasons assume that $min.assign_d : \pi \to \pi'$ is a solution[6] to the assignment problem of assigning π to π' weighted by d.

Assume there is a distance function $D : \mathfrak{T} \times \mathfrak{T} \to \mathbb{R}_{+0}$ for trajectories (e. g., Equation (5.5)) and a solution $min.assign_D$ to the assignment problem, as described above. This function can be used to construct a distance function $\delta_{\mathfrak{T}} : \Pi_M \times \Pi_M \to \mathbb{R}_{+0}$ between set partitions,

$$\delta_{\mathfrak{T}}(\pi, \pi') := D_f(\pi) \text{ where } f = min.assign_D. \tag{5.8}$$

[5]We chose the edit distance in this example for simplicity.

[6]For the examples used in this work we used an implementation of the Hungarian Algorithm, see Appendix B.

The distance D_f is defined as edit distance to reflect the amount of work to change partition π into partition π',

$$D_f(\pi) := \sum_{\tau \in \pi} D(\tau, f(\tau)). \tag{5.9}$$

The map $f : \pi \to \pi'$ has to be bijective. In the case that $|\pi| \neq |\pi'|$, the smaller partition is "filled up" with empty sets. Assume without loss of generality that $|\pi| < |\pi'|$. Then π is substituted by a new partition $\pi'' = \pi \cup \pi_\emptyset$ with $|\pi_\emptyset| = |\pi'| - |\pi|$ and $\forall \tau \in \pi_\emptyset : \tau = \emptyset$.

Metric Attributes of the Trajectory-Based Partition Distance

In the following, it is shown that the metric attributes, (*non-negativity*, *identity of indiscernibles*, *symmetry.* and *triangle inequality*) are inherited from the trajectory distance D as defined by Equation (5.9).

Non-Negativity. Trivially, every sum of non-negative values D is itself non-negative. Thus $\delta_{\mathfrak{T}}$ is non-negative if D is non-negative.

Identity of Indiscernibles. Given that D is non-negative, identity of indiscernibles is inherited by $\delta_{\mathfrak{T}}$.

Lemma 5.3 (Identity of Indiscernibles in $\delta_{\mathfrak{T}}$)
If D is non-negative and $D(\tau, \tau') = 0 \iff \tau = \tau'$,
then we have $\delta_{\mathfrak{T}}(\pi, \pi') = 0 \iff \pi = \pi'$.

Proof 5.3
The identity of indiscernibles of $\delta_{\mathfrak{T}}$ follows if every (non-negative) trajectory distance $D(\tau, f(\tau))$ for every trajectory-pair in a minimal assignment f is zero.

$$\begin{aligned}
\delta_{\mathfrak{T}}(\pi, \pi') = 0 \quad &\overset{def.}{\iff} \quad \exists f. \sum_{\tau \in \pi} D(\tau, f(\tau)) = 0 \\
&\overset{D \geq 0}{\iff} \quad \exists f. \forall \tau \in \pi : \; D(\tau, f(\tau)) = 0 \\
&\overset{D \text{ identity}}{\iff} \quad \exists f. \forall \tau \in \pi : \; \tau = f(\tau) \\
&\iff \quad \pi = \pi'.
\end{aligned}$$

If $\delta_{\mathfrak{T}}$ is zero, then, because of non-negativity of D, there must be a bijection f for which $D(\tau, f(\tau))$ is also zero for every $\tau \in \pi$. Because identity of indiscernibles holds for D, abbreviated by D identity above, $\tau = f(\tau)$ is valid for every $\tau \in \pi$. As every cluster in π is equal to its image in π' both partitions are equal. □

Symmetry Given that D is symmetric, then $\delta_{\mathfrak{T}}$ is also symmetric:

Lemma 5.4 (Symmetry of $\delta_{\mathfrak{T}}$)
If for all $\tau, \tau' \in \mathfrak{T} : D(\tau, \tau') = D(\tau', \tau)$,
then for all $\pi, \pi' \in \Pi_M : \delta_{\mathfrak{T}}(\pi, \pi') = \delta_{\mathfrak{T}}(\pi', \pi)$.

Proof 5.4
To show that $\delta_{\mathfrak{T}}$ inherits symmetry, the inverse map f^- to f is used, which is a minimum assignment if f is a minimum assignment.

$$\begin{aligned}
\delta_{\mathfrak{T}}(\pi, \pi') &\overset{def.}{=} \sum_{\tau \in \pi} D(\tau, f(\tau)) \\
&\overset{D \text{ symmetric}}{=} \sum_{\tau \in \pi} D(f(\tau), \tau) \\
&= \sum_{\tau \in \pi} D(f(\tau), f^-(f(\tau))) \\
&\overset{f(\tau)=\tau'}{=} \sum_{\tau' \in \pi'} D(\tau', f^-(\tau')) \\
&= \delta_{\mathfrak{T}}(\pi', \pi).
\end{aligned}$$

As D is symmetric, denoted $Dsymmetric$, $D(\tau, f(\tau))$ equals $D(f(\tau), \tau)$. The inverse images exists and is bijective because f is bijective and $\tau = f^-(f(\tau))$. Thus for every $\tau' \in \pi'$ there exists $\tau \in \pi$ with $f^-(\tau') = \tau$. As f is a minimum assignment f^- also is a minimum assignment and $\sum_{\tau' \in \pi'} D(\tau', f^-(\tau'))$ is equal to $\delta_{\mathfrak{T}}(\pi', \pi)$. □

Thus, symmetry of $\delta_{\mathfrak{T}}$ follows from symmetry of the trajectory distance D.

Triangle-Inequality The inheritance of the triangle-inequality from D to $\delta_{\mathfrak{T}}$, again, needs non-negativity of D.

Lemma 5.5 (Triangle-Inequality of $\delta_{\mathfrak{T}}$)
If the triangle-inequality of D holds for all $\tau, \tau', \tau'' \in \mathfrak{T}$,
then for all $\pi, \pi'\pi'' \in \Pi_M : \delta_{\mathfrak{T}}(\pi, \pi'') \leq \delta_{\mathfrak{T}}(\pi, \pi') + \delta_{\mathfrak{T}}(\pi', \pi'')$.

Proof 5.5
We show that $\delta_{\mathfrak{T}}$ fulfils the triangle-inequality, because it is valid for every addend in $\delta_{\mathfrak{T}}$. If D is non-negative, then, if the triangle-inequality holds for every part of the sum, it holds for the whole sum.

$$\begin{aligned}
\delta_{\mathfrak{T}}(\pi,\pi') + \delta_{\mathfrak{T}}(\pi',\pi'') &\overset{def.}{=} \sum_{\tau\in\pi} D(\tau, f(\tau)) + \sum_{\tau'\in\pi'} D(\tau', f'(\tau')) \\
&= \sum_{\tau\in\pi} [\underbrace{D(\tau, f(\tau)) + D(f(\tau), f'(f(\tau)))}_{\geq D(\tau, f'(f(\tau)))}] \\
&\geq \delta_{\mathfrak{T}}(\pi,\pi'').
\end{aligned}$$

□

It has been shown that any metric attribute that holds for the underlying trajectory distance measure D holds for $\delta_{\mathfrak{T}}$, as long as D is non-negative.

5.4 Related Work

This work is part of a growing number of initiatives to quantify privacy. The main focus in the community is on anonymity metrics; only relatively few works concentrate on unlinkability of either databases and messages. As we have shown in Chapter 2, anonymity and unlinkability are strongly related. In this section we provide an overview on existing metrics for anonymity and unlinkability.

The main motivation for this work on unlinkability measures is based on the criticism that current measures ignore the inner structure of hypotheses space. This is discussed in Section 5.4.3.

5.4.1 Anonymity Measures

Most privacy research has focused on the notion of *anonymity*. As anonymity and unlinkability are tightly related, one can not develop unlinkability measures without relating to the work done in anonymity beforehand.

Anonymity, as defined in Section 2.3, is concerned with the relations between two distinct classes of elements, e. g., the class of subjects or subject identifiers and a class of items of interest. Proxy-based anonymity mixes are the technique most widely discussed for anonymity preservation. Therefore, most works on anonymity measures focus on analysis of mix networks.

The discussion on anonymity metrics is mostly attack-driven. A measure is used for a period of time, until an attack is found that works although the known measure deduced high anonymity. A comprehensive collection of anonymity measures and attacks is provided in [34].

The development of anonymity measures probably was started by the introduction of *anonymity sets* by Chaum [15]. An anonymity set consists

of subjects that are possibly related to an action. The cardinality of this anonymity set is a basic and descriptive measure for individual anonymity.

In 2002, Serjantov and Danezis [59] showed that a uniform probability distribution, as implied by anonymity sets[7], is not sufficient to describe anonymity scenarios in general. They introduced the *anonymity probability distribution* and measured the effective size of an anonymity set using Shannon's entropy [60]. This work was shortly followed by the introduction of the *degree of anonymity* by Díaz et al. . [19].

In a criticism of these global anonymity measures Tóth, Hornák, and Vajda [66] introduced a *local anonymity measure*[8]. The sender unlinkability is given by an upper bound of all probabilities P_{β_k,S_I} of senders S_I being related to IOI β_k.

Other anonymity measures are based on success probabilities of an attacker or the amount of ressources necessary for an attacker to succeed. As an example, in [36] success probabilities for disclosure attacks are provided from given message emission rate distributions. In [3] the expected number of observations needed by an attacker to successfully relate communication endpoints is measured in relation to the number of participants in the scenario.

The anonymity measures discussed above focused on anonymity from the perspective of single subjects or single actions. Edman, Sivrikaya, and Yener [21] introduced a combinatorial approach to measure a global system's anonymity level. They defined a *global degree of anonymity* derived from the number of plausible perfect matchings between subjects and items of interest. Plausible edges between subjects and IOI are modelled by an $n \times n$ $(0,1)$-adjacency matrix A. The global degree of anonymity is given by $log(per(A))/log(n!)$, where $per(A)$ is the permanent of A.

Edman et al. [21] used their measure on graphs consisting of vertices that represent messages entering and leaving a mix network. The problem observed is exactly the anonymity problem as in Definition 2.1. Gierlichs, Troncoso, Diaz, and Preneel [28] introduce equivalence classes of messages. Instead of observing message entry or exit points they observe only the anonymity set of senders or receivers. The main idea is that an attacker is most often more interested in the question of which subject communicates with which other subject, instead of the usual question on mix networks on the relation between entry and exit events (messages). This combinatorial trick reduces the number of hypotheses that have to be considered.

[7] An anonymity set has no probability assignment which, if one element has to be chosen, for practical reasons means that each element has to be considered as equal likely.

[8] In [51] this type of anonymity measure is denoted as *individual anonymity* to distinguish it from the definition of *local anonymity* as defined in [46]

5.4.2 Database Privacy

The term *unlinkability* is used in the context of databases to describe the inability of an attacker to correctly derive relations between entries in different, but related databases. For an example, consider a database of medical records that is used in different contexts. In every context some values are blacked out or generalised. The objective is to provide needed information in the context, but provide anonymity of the subjects. Furthermore, we want to prevent that entries from one context can be related to entries from another context, which would reveal more information and brings the danger of individuals being identifiable. Considering our formal distinction between anonymity and unlinkability problems, *database unlinkability* is an anonymity problem, not an unlinkability problem.

Although named *k-unlinkability*, the concept introduced by Malin in [45] corresponds better to the notion of anonymity as used in this work. Given a bipartite graph representing two partially cloaked tables of data "trails" associated by being derived from the same source, k-unlinkability describes that each table can be mapped to at least k entries in the other table.

The notion of the *anonymity set* is related to k-anonymity, where k describes the minimum number of indistinguishable records in a database. Machanavajjhala et al. [44] describe a homogeneity attack and a context information attack that break k-anonymity. They introduce *l-diversity*, which denotes that a database contains records in a way that every selection of records "contains at least ℓ *well represented* values". This term *well represented* implies that an attacker with access to the database cannot reduce the selection in a way that it contains less than ℓ different values, e. g., identities.

Again, ℓ-diversity is criticised by Li et al. [41] because it is susceptible to a *skewness* or *similarity attack* condition. Therein *t-closeness* is defined that describes the similarity of prior and posterior knowledge of an attacker who gets hold of the (generalised) database.

5.4.3 Unlinkability Measures

In the following, we will discuss example unlinkability measures and explain shortcomings of entropy-based measures.

Holczer and Buttyan [12] use a link-wise attacker's success ratio to measure unlinkability. Their objective is to analyse the number of pseudonym changes needed for unlinkability of messages in vehicular communication. Their work is based on the Mix Zones Model by Beresford and Stajano [6][8] (see Section 5.4.5).

An *attacker's success ratio* in correctly guessing single links is taken as measure for unlinkability. The attacker uses context knowledge in the form

of a correlation matrix to estimate probabilities of single links, and makes a greedy guess. The correlation matrix is generated from the real vehicle movements in a mix zone. The ratio of correct guesses is taken as global unlinkability measure for a given scenario.

Success ratios condense repeated experiments to a ratio of success and failure. Such measures obviously do not reflect the consistency of an assignment but only the similarity of the attacker's choice to the true hypothesis. Furthermore, measures of this class can only be used for a white-box analysis where the analyst knows the real relation.

Sampigethaya et al. [57] introduce another unlinkability measure based on the success of an attacker. Therein unlinkability is measured as the expected time a vehicle can be followed by way of its messages. The expectation is calculated from traffic density and a vehicle movement model. Their measure is used to support the usage of a random silence interval to keep the time of the next message emission by an vehicle unpredictable.

Steinbrecher and Köpsell [63] defined the *degree of unlinkability* as the quotient of the Shannon-entropy [60] of the posterior and prior knowledge of an attacker. There, posterior knowledge describes the knowledge that an attacker gains by adding context information to his prior knowledge. Prior knowledge denotes the knowledge of an attacker without context information, i. e., prior knowledge is a uniform distribution. This measure is derived from the degree of anonymity [19] and was defined first in [63]. However, for its superior clarity we will use the definition in [25].

Shannon's *entropy* H and the *degree of unlinkability* $\mathcal{D}$ are defined as:

$$H(P) := -\sum_{\pi \in \Pi_M} P(\pi) \log_2(P(\pi)), \tag{5.10}$$

$$\mathcal{D}(P) := \frac{H(P)}{H_{max}}, \tag{5.11}$$

where P denotes a probability distribution, or posterior knowledge, of the finite random variable which is the hypotheses space Π_M. $H_{max} = \log_2(|\Pi_M|)$ denotes the maximum entropy, which is, provided the attacker has no prior knowledge, achieved by the uniform distribution on the domain. $|\Pi_M|$ is the cardinality of Π_M, which is the Bell Number of the cardinality of M.

In [50] Pashalidis uses a subset of all possible binary relations on M to provide the elements of an unlinkability hypotheses space. This definition is generally broader than considering only equivalence relations. It stands to debate whether this level of abstraction is useful here as this model includes hypotheses that are not solutions to linkability problems, e. g., if the implied relation is not an equivalence relation. Similarly an attacker's probability mass assignment on the hypotheses space is assessed by calculating entropy

$P : \Pi_M \to [0,1]$	π_1	π_2	π_3	π_4	π_5	H	$\mathcal{D}$
P_A	0.60	0.05	0.15	0.05	0.15	1.70	0.73
P_B	0.05	0.15	0.05	0.15	0.60	1.70	0.73
P_C	1.00	0.00	0.00	0.00	0.00	0.00	0.00
P_D	0.20	0.20	0.20	0.20	0.20	2.32	1.00
P_E	0.40	0.05	0.11	0.05	0.39	1.84	0.79
P_F	0.05	0.39	0.11	0.05	0.40	1.84	0.79

Table 5.3: Example probability mass assignments of attackers A, B, C, D, E, F together with Shannon entropy H and degree of unlinkability $\mathcal{D}$.

or normalised entropy[9].

In the following section it will be shown by example situations that the degree of unlinkability — or basically any measure ignoring the inner structure of the hypothesis space — fails to discern fundamentally different situations.

Structural Consistence and the Degree of Unlinkability

In this section, we show that the known measure *degree of unlinkability* fails to distinguish attackers' assignments that are nonetheless different with respect to unlinkability. We argue that this disqualifies the degree of unlinkability as an unlinkability measure.

In the following, the toy example (5.2) from Section 5.1 is used with the partition distance given in Table 5.2 from Section 5.3.3. Although this example is very simple it already shows the problems of unlinkability measures based solely on outer structure. We use the partition distance in our discussion for simplicity.

In Table 5.3, six example probability mass assignments P_A, P_B, P_C, P_D, P_E, and P_F are shown. For later recognition of the distributions, small barplots depicting scaled probability values are drawn. On the right-hand side of the table, Shannon-entropy H and degree of unlinkability $\mathcal{D}$ for the assignments are given. Note that degrees of unlinkability for P_A and P_B, as well as the values for P_E and P_F, are equal while the degree for P_C is minimum and for P_D is maximum. Equal degree values mean that both attacker assignments provide similar quality, i. e., both assignments leave equal unlinkability. The maximum/minimum degree of unlinkability is interpreted as maximum/minimum unlinkable scenario.

Comparing probability assignments P_E and P_F in Table 5.3, one may recognise that they are merely assign the same probabilities to different par-

[9]The terminology used is *opaqueness* and *degree of opaqueness*.

titions. Entropy, and thus the degree of unlinkability, is invariant to permutations. The reason why this does not reflect the quality of this assignment is that inner structure of the hypotheses space is not considered by H or $\mathcal{D}$.

The most likely hypotheses of attacker E are π_1 and π_5 whose partition distance is, according to Table 5.2, two, which is the maximum distance possible in this hypotheses space, thus π_1 and π_5 are very dissimilar. The most likely hypotheses for attacker F are π_2 and π_5, whose partition distance is only one, the smallest partition distance for non-equal partitions. In both cases the difference between the first and second partition in attacker's order is small. This means that a small shift of probability mass from the first to the second partition results in a different attacker's choice. Considering consistency, it can be said that P_E is less consistent than P_F because a similar small change to both assignments has, in the worst case, a stronger impact on the attacker's choice of E than on F's choice, but still the degree of unlinkability is equal.

A meaningful measure of unlinkability should reflect consistency within the assignment of an attacker. The degree of unlinkability fails here because it does not use the inner structure of the support of the probability mass function, i.e. the distances between partitions.

Attacker assignments P_A and P_B from Table 5.3 are, again, permutations of equal probability masses. Both have one singular peak, assigned to hypothesis π_1, respectively π_5. In terms of partition distance, π_5 is generally closer to all other partitions than π_1. This is a slight advantage of P_B because two set partitions with small distance are supportive hypotheses for each other. Set partitions with large distance, on the other hand, contradict each other. Thus, π_1 is relatively unsupported by other hypotheses — compared to π_5.

Depending on their assigned probability mass, hypotheses in support or contradiction of the attacker's choice determine the consistency of the whole assignment. It should be obvious that entropy-based measures do not consider consistency.

What is reflected well by entropy-based measures is the *uncertainty* within an attacker's assignment. The two extreme assignments, the degenerate distribution P_C and the uniform distribution P_D, can clearly be distinguished by the degree of unlinkability.

P_C describes the situation where the attacker has maximum belief in one particular partition being the true relation. This distribution is the best case for an attacker: it denotes the situation where he has no uncertainty left, although the attacker's choice might still be wrong. P_D, on the contrary, models the situation of maximum uncertainty, where the attacker has no context information at all.

But if attacker D considers the inner structure example in Table 5.2, he could recognise one distinguished distance value. The distance between π_1 and π_5 is two, while all other distances have value one. The interpretation of this scenario is that π_1 and π_5 contradict each other more than any other pair of partitions.

Assume an attacker chooses any of π_2, π_3, and π_4, and compare this scenario to a second attacker choosing π_1 or π_5. The first attacker would have chosen a hypothesis that is very similar to every other hypotheses. If his choice where wrong he would have made an error of one in the worst case. The maximum error of the second attacker would be two. Given, that from a point of view of the attacker, every hypothesis is equally likely, the second choice has the possibility of a higher error and thus has to be considered inferior to the first choice.

This implies that the choice of an attacker should not rely solely on outer structure. We introduce such an attacker in Section 5.7.3. Furthermore we can state that the uniform distribution, from external structure alone, is not the case where an attacker is completely indifferent of which partition to chose but still can make an improved choice.

The entropy-based measure of unlinkability — as the entropy itself — has not been developed to reflect the inner structure. Entropy and derived measures can only reflect the outer structure. It stands to debate how exactly to weight consistency against certainty. An optimum unlinkability measure probably should reflect both attributes. In Section 5.5 we introduce a measure that uses both the internal structure of the partitions and the probability assignment of the attacker in a single measure of unlinkability.

Quality Measures for Clustering Techniques

An unlinkability attacker basically tries to find the optimum clustering of a set. Clustering techniques have been studied for a long time and are still a main topic of scientific research. Nonetheless, the connection between the field of privacy and the field of clustering has seldom been directly mentioned.

The quality of a clustering is often measured in two distinct ways that can be directly related to the white-box analysis and black-box analysis from Section 5.7. Cluster quality evaluation by *internal criteria* uses no information available to the clustering algorithm. *External criteria* use additional, external information, not available to the clustering algorithm to assess the quality of clustering results. One external criterion normally is similarity of the attacker's choice to the true hypothesis. An introduction to clustering and cluster quality metrics can be found in [38] and [31].

In the context of inner and outer structure as defined in Section 5.3, above cluster quality measures are based on inner structure, but consider

only the distance between two hypotheses space as the result of a clustering algorithm is a single hypothesis. This hypothesis is then compared to results from previous experiments or, if known, the true hypothesis. Considering the development of attacker algorithms, this method suffices to develop clustering algorithms. But given the scenario with knowledge of the attacker's assignment, valuable information is discarded.

5.4.4 Notes on Entropy-based Anonymity Metrics

In the previous section, examples have been shown where unlinkability measures that solely rely on outer structure fail to discern different attackers. As the degree of unlinkability is directly derived from the degree of anonymity, it has to be discussed why the degree of anonymity does not fail in the same way to measure anonymity.

The difference between the *degree of unlinkability* [63, 25] and the *degree of anonymity* [19] is the domain of the attacker's assignment. The assignment's domain in unlinkability is the set of all set partitions. Set partitions are not atomic but are themselves collections of subsets of items of interest. The domain of attacker probability assignments in anonymity problems is a set of identification anchors. Identities normally have no inner structure and generally are atomic. The "natural" distance for identities is equality of identity but there is normally no useful[10] metric besides this.

Thus, our argument from the previous Section 5.4.3 cannot be applied to the degree of anonymity. Without inner structure, consistency of an attacker's assignment, as discussed in Section 5.2, becomes unimportant. Without inner structure, entropy based measures do not neglect important information.

5.4.5 Location Privacy

A different approach for analysing privacy in mobile communication scenarios is expressed by the notion *location privacy*. Location privacy is achieved if the location of a subject cannot be derived by an attacker. In mobility scenarios the objective in location privacy is very closely related to unlinkability. Location privacy can be achieved by preventing any location information from being known to an attacker, or by achieving anonymity and unlinkability.

In their work on location privacy, Beresford and Stajano [8, 6, 7] developed the Mix Zones Model. A Mix Zone is an area that is not observed by

[10] When semantic interpretation is concerned there is always space for debate, but at this point we may safely assume that there is no interpretation that defines a notion for "almost the same person". Two IA trivially either reference the same individual or two distinct subjects.

the attacker. The attacker is only able to observe entry- and exit-events at a zone's border. Zone border and time are modelled as discrete spaces. An attacker has to find the optimum *perfect matching* between entries and exits given a correlation matrix[11]. The correlations are derived from observation of movements through the zone some time before the actual attack. A correlation matrix contains values that describe probabilities for times needed for traveling through the zone and probable exit-events given the entry-event. An attacker's probability distribution over the set of perfect matchings is calculated and Shannon's entropy is proposed as a global anonymity measure.

Furthermore, in [8], a notion of individual anonymity is sketched as the uncertainty of an attacker to find a distinct exit-event to a given entry-event. In other words: to track a single movement through the zone. Again, uncertainty is measured as entropy over the probability distribution over the possible exits.

5.5 Measuring Unlinkability

In this section, we describe an alternative unlinkability measure that assesses the quality of an attacker's probability mass assignment, called *expected distance unlinkability measure*. Unlike the degree of unlinkability, this measure considers *consistency*. Basically, the expected distance unlinkability measure is an expectation of the error an attacker makes.

In the following, criteria for unlinkability metrics are defined, the expected distance unlinkability measure is introduced, and its properties are discussed.

5.5.1 Criteria for Unlinkability Measures

In this section we define criteria which should be satisfied by measures[12] of unlinkability. Consistency has repeatedly been mentioned as an important quality criterion for an attacker's assignment. From the lessons learned in Section 5.2 and Section 5.4.3, we derive three semantic criteria for unlinkability measures.

Criterion 1: The measure quantifies the consistency of an assignment.

Criterion 2: The measure quantifies the certainty of an assignment.

[11] a matrix that reflects the correlation between pairs of elements, e. g., between entry- and exit-events

[12] Formally, it has not been shown that the space, measured by unlinkability measures, i. e., Π_M and P_A, constitutes a σ-algebra. We are using the term *measure* nonetheless, because unlinkability is similar to a volume measure, e. g., like probability.

Criterion 3: The measure can be used to quantify the correctness of the assignment with respect to a correct hypothesis.

Criterion 1 relates to the consistency of an attacker's assignment with respect to an inner structure as discussed in Section 5.2. Criterion 2 relates to certainty of an attacker[13]. Criterion 3 again is useful only with respect to some inner structure that provides a (similarity-)distance.

The need for assessing consistency is, to the best of our knowledge, a criterion that has been missed by previous work. Thus, we deem it the most important criterion in our work. Certainty, as an attribute of the outer structure, has been captured by the notion of entropy and other measures. As both criteria use different structures of the hypotheses space, a measure of unlinkability has to combine both properties and therefore must weight one criterion against the other. On the other hand it will always be possible to us measure each property individually.

The third criterion is prerequisite to any white-box analysis. An unlinkability measure should be able to consider the correctness of the attacker's assignment if the real partition is known to the analyst. A simple measure for correctness is expressed by the distance between attacker's choice and real partition.

The following properties should additionally be fulfilled by an unlinkability measure.

Continuity. An unlinkability measure should be continuous with respect to context information, items of interest, and inner and outer structure. Small changes to the attacker's assignment, for example, should have only a small effect on the unlinkability measure. Likewise for message positions, context information and the used hypotheses distance.

Symmetry. An unlinkability measure should be symmetric with respect to the order of hypotheses which includes the order of messages and clusters.

Bounds. An unlinkability measure should map into the positive real numbers $\mathbb{R}_{+0}$ and be bound by an infimum and supremum, where the infimum should be 0. Bounds are required to represent the two states of *perfect unlinkability* and *no unlinkability*. Obviously both states exist, and it must be defined for every measurement function which of the inputs to the function will determine these these bounds.

In the following section, the expected distance unlinkability measure is introduced. Afterwards, our measure is analysed with respect to the criteria listed above.

[13] which often is represented by entropy

5.5.2 Expected Distance Unlinkability Measure

In this section, we define unlinkability as the expected distance from a reference hypothesis. While entropy measures the uncertainty of an attacker, the *expected distance unlinkability measure* provides a notion of error of an attacker's assignment. The larger the expected distance, the worse is the attacker's assignment.

The idea is to interpret the attacker's assignment $P_A : \Pi_A \to [0,1]$ as probabilities in a stochastic experiment. Distance $\delta_{\pi^*} : \Pi_M \to \mathbb{R}_{+0}$ to a reference partition π^* is used as random variable. This allows calculation of an expectation over all hypotheses to a given reference partition π^*.

Definition 5.2 (Expected Distance Unlinkability Measure)
The Expected Distance Unlinkability Measure *is defined as*

$$Ed_{P_A,\delta}(\pi^*) := \sum_{\pi \in \Pi_M} P_A(\pi) \cdot \delta_{\pi^*}(\pi),$$

where $\delta_{\pi^*}(\pi) := \delta(\pi, \pi^*)$ *and* $\delta : \Pi_M \times \Pi_M \to \mathbb{R}_{+0}$ *is a symmetric distance function on* Π_M.

By way of choosing either the real partition or the attacker's choice as the reference partition, white-box analysis and black-box analysis are distinguished. In *white-box analysis* the set partition related to the true sender relation π^* is known to the analyst. Thus the analyst is able to estimate the amount of error that an attacker makes with respect to attackers assignment. In *black-box analysis* the reference partition is the attacker's choice π_A, e. g., the maximum likely partition as defined in Equation (5.4).

5.5.3 Analysis of Measure

In this section, the expected distance unlinkability measure is discussed with respect to the criteria introduced in Section 5.5.1. The focus is on the measure's ability to grasp the semantic notion of consistency of the attacker's assignment. Furthermore, the certainty criterion is discussed and it is shown that the formal criteria, as defined above, are fulfilled. We will not discuss the third criterion because the proposed measure is effectively an estimation of error from a assumed solution, thus this criterion is trivially fulfilled.

Consistency

To show that the measure fulfils the consistency criterion, changes to the unlinkability measure are observed under minimum changes of the attacker's

probability mass assignment. The idea is to exchange the attacker's assignment of two hypotheses and observe that the expected unlinkability measure behaves in a sensible manner.

Let π_i, π_j be two hypotheses, P_A be the attacker probability mass assignment and δ be a distance measure on the hypotheses space. Let π^* be the reference hypothesis. Values P_A or δ_{π^*} for both hypotheses can be either in the relation *less*, *equal*, or *greater*. Thus we have to consider three times three cases of relations between $P_A(\pi_i), P_A(\pi_j)$ and $\delta_{\pi^*}(\pi_i), \delta_{\pi^*}(\pi_j)$.

		$\delta_{\pi^*}(\pi_i)\ ?\ \delta_{\pi^*}(\pi_j)$		
		$<$	$=$	$>$
$P_A(\pi_i)\ ?\ P_A(\pi_j)$	$<$	fall	stay	raise
	$=$	stay	stay	stay
	$>$	raise	stay	fall

Table 5.4: Expected behaviour of Ed on exchange of attacker assignments P_A for hypotheses π_i, π_j for a given relation between values of P_A and δ_{π^*}.

In the following, we observe changes to Ed once the assignments P_A for π_i and π_j are exchanged. Exchanging the assignments means to define a changed assignment P'_A with

$$\begin{aligned} P'_A(\pi_i) &:= P_A(\pi_j) \\ P'_A(\pi_j) &:= P_A(\pi_i) \end{aligned}$$

Exchanging two values may either reduce, increase, or maintain the consistency of the assignment. This behaviour depends both on the assignments P_A and the distance between hypotheses δ. In the following, three different cases of expected behaviour (fall, raise, or stay) of Ed are distinguished, in Table 5.4 nine cases of different results are summarised. The case *stay* denotes situations where either one or both of the P_A or the δ_{π^*} values for π_i and π_j are equal. In the case of *fall* both values are in the same inequality relation, and in the case of *raise* both values are of opposite inequality. In the following, we will argue that Ed behaves as expected in all cases.

Equal Values If one or both relations are equal, i. e., $P_A(\pi_i) = P_A(\pi_j)$ or $\delta_{\pi^*}(\pi_i) = \delta_{\pi^*}(\pi_j)$ it is expected that the consistency of the attackers assignment is not changed if the assignments are exchanged. Thus, if the assignments for π_i, π_j are exchanged, an unlinkability measure should not change.

Obviously, for any case where $P_A(\pi_i) = P_A(\pi_j)$, the expected distance unlinkability Ed is unchanged. Furthermore, the same can be said in the case of $\delta_{\pi^*}(\pi_i) = \delta_{\pi^*}(\pi_j)$.

Both Less or Both Greater Than. Considering the case where both relations show the same inequality, take for example the case of $P_A(\pi_i) < P_A(\pi_j)$ and $\delta(\pi_i, \pi^*) < \delta(\pi_j, \pi^*)$. Exchanging the P_A assignment for π_i and π_j means to change the attacker's order in a way that hypothesis π_i is now higher up in that order. As π_i is closer with respect to δ, to the reference partition than π_j, the dissimilarity of the higher ordered hypotheses is reduced. This means the consistency of the assignment is improved.

In the expected distance unlinkability measure, two terms of the sum are changed, namely the values for P_A in two of the summands. Given the relation between the values of δ and P_A the new sum is larger for $\pi_j <_P \pi_i$ as

$$P_A(\pi_i) \cdot \delta(\pi_i, \pi^*) + P_A(\pi_j) \cdot \delta(\pi_j, \pi^*) > \\ P_A(\pi_j) \cdot \delta(\pi_i, \pi^*) + P_A(\pi_i) \cdot \delta(\pi_j, \pi^*).$$

In the case of $P_A(\pi_i) > P_A(\pi_j)$ and $\delta(\pi_i, \pi^*) > \delta(\pi_j, \pi^*)$ the same effect is observed. A smaller expected distance is interpreted as lower unlinkability which correctly reflects the higher consistency.

One Greater, The Other Lesser Than. We now consider the case where the two relations are different inequalities, i. e., $P_A(\pi_i) < P_A(\pi_j)$ and $\delta(\pi_i, \pi^*) > \delta(\pi_j, \pi^*)$. Again, exchanging the P_A assignment for π_i and π_j means to change the attacker's order (5.3) in a way that hypothesis π_i is now higher up in that order. As π_i is further away, with respect to δ, from the reference partition than π_j this increases the dissimilarity of higher order hypotheses globally. This means the consistency of the assignment is reduced.

The effect on the expected distance unlinkability measure is that the measure is larger after exchanging the assignments of π_i and π_j as with the original attacker's assignment. The calculation is similar as in the previous case. For the case of $P_A(\pi_i) > P_A(\pi_j)$ and $\delta(\pi_i, \pi^*) < \delta(\pi_j, \pi^*)$, again, the same argument applies and similarly this behaviour is expected.

It can be concluded that the expected distance unlinkability measure fulfils the first criterion of reflecting the consistency of an attacker's probability mass assignment.

Certainty

As mentioned before, Shannon's entropy can be considered a good measure for *certainty* of an attacker's assignment. As entropy is a property of only the outer structure, *Ed* cannot reflect certainty in the same way as entropy.

In the expected distance, the probability assignment is combined with relative distance values δ_{π^*}. This means that the effect to Ed by any modification to probability assignments is biased by δ_{π^*}.

Consider the case that probability mass is shifted towards, with respect to δ, the reference hypothesis. Then the unlinkability measure changes toward "less unlinkability", i. e., Ed grows. On the other hand, the behaviour of entropy is not clear, meaning, that entropy may change in both directions or even stay unchanged. Thus entropy and Ed are not correlated, which means that Ed is not reflecting consistency in the same sense of entropy.

Certainty, if reflected by Ed, cannot be seen independent from inner structure distance metric δ. Which is not related to entropy. Any notion of certainty — expressed by Ed — is correlated with above notion of consistency. Thus, if certainty in terms of entropy is needed, Ed is insufficient and has to be supplemented.

Formal Criteria

The criteria consistency, certainty and correctness define semantics of unlinkability measures. The criteria continuity, symmetry and boundedness describe more formal properties of unlinkability measures. In the following, it is shown that Ed has these properties.

Continuity Concerning *continuity* of Ed with respect to small changes of P_A and δ, addition and multiplication are continuous operations. Distance δ and P_A are non-negative, thus, small changes to P_A only have little effect on Ed. Small changes to δ also have only a small effect on ed. Thus Ed must be continuous.

Symmetry *Symmetry* of Ed, with respect to the order of hypotheses, is trivially inherited from commutativity of the addition operation.

Bounds Because of the different[14] definitions for δ, values for δ_{π^*} generally only have a lower bound. But because of the finite cardinality of Π_M, there is an upper bound to Ed governed by this inner structure. P_A is defined with range $[0, 1]$, $1 = \sum_{\pi \in \Pi_M} P_A(\pi)$ and thus trivially has upper and lower bound.

Given a distance function δ, probability mass assignments P_{min}, P_{max}, for minimum, respectively maximum values of Ed can be found.

[14]The definition of δ is chosen with respect to the objectives of the attacker.

Minimum Trivially, an attacker assigning maximum probability to the reference hypothesis provides a minimum for Ed. The minimum assignment is denoted as

$$P_{min}(\pi) = \begin{cases} 1, & \text{if } \pi = \pi^* \\ 0, & \text{if } \pi \neq \pi^*, \end{cases}$$

As $\delta_{\pi^*}(\pi^*) = 0$, the expected distance $Ed_{P_{min},\delta}(\pi^*) = 0$. Furthermore, P_{min} is the only minimum because we have $\delta(\pi, \pi') = 0 \iff \pi = \pi'$. Therefore the minimum expected distance unlinkability is zero, as long as the underlying inner structure provides identity of indiscernibles.

Maximum For a maximum assignment, we distinguish between black-box analysis, where the reference partition is given by the attacker's choice and white-box analysis, where the reference partition is independent of the attacker's choice.

In black-box analysis, the reference partition is the attacker's choice. Given that the attacker's choice is defined in Equation (5.4), the maximum expected distance is achieved when two hypotheses π, π' are chosen in a way that $\delta(\pi, \pi')$ is maximum. P_A then is constructed in a way that $P_A(\pi)$ equals $P_A(\pi')$ and $P_A(\pi) = 0.5$. Assume that, without loss of generality, $ord(\pi) < ord(\pi')$ and thus π becomes the attacker's choice because of the second criterion of the attacker's order (5.3). The *upper bound* for Ed in black-box analysis is given by

$$Ed_{P_{max},\delta} = 0.5 * \delta(\pi, \pi')$$

where $\pi, \pi' \in \Pi_M$ produce maximum distance δ, i. e., $\delta(\pi, \pi') \geq (\delta(\pi'', \pi'''))$ for all (π'', π''') in $\Pi_M \times \Pi_M$.

Constructing an *upper bound* for Ed in white-box analysis runs along the same lines as before, we choose the element with maximum distance to the reference partition. An attacker's assignment P_{max} that produces a maximum $Ed_{P_{max},\delta}$ for a given δ in white-box analysis is

$$P_A(\pi) = \begin{cases} 1, & \text{if } \pi = \hat{\pi} \\ 0, & \text{if } \pi \neq \hat{\pi}, \end{cases}$$

where $\hat{\pi} \in \{\pi \in \Pi_M : \forall \pi' \in \Pi_M : \delta(\pi, \pi^*) \geq \delta(\pi', \pi^*)\}$ is a hypothesis with maximum distance to π^*.

Knowing the bounds of Ed allows calculation of a normalised expected distance unlinkability measure. We leave this to the discussion and further work.

5.6 Approaching Efficiency

In this section, efficient approaches for analysis of unlinkability are discussed using the expected distance unlinkability measure Ed. Efficient computation of Ed in realistic environments is still work in progress. The expected distance unlinkability comprises a sum over Π_M that is not efficiently computable because of the large cardinality of Π_M. Attacker probability mass assignments also often need normalisation by summation over Π_M. This makes a direct approach for simulating the measure infeasible. In the following only preliminary works toward a simulative analysis can be presented.

The main contribution of this section is a complexity reduction of Ed for an example attacker assignment. In the following, the information flow of a simulation is introduced. Afterwards, we explain a layer model of context information that leads to the example attacker assignment.

5.6.1 Information Flow

The difference between white-box and black-box is in the knowledge of the real partition. Aside of this knowledge, the flow of information is the same for both types of analysis.

Information flow in white-box analysis is depicted in Figure 5.5. *Devices* emit messages M. The real partition π^* is determined by an device's movement. An algorithm representing the *attacker* then generates P_A which defines the attacker's choice π_A. The *analyst* derives the unlinkability measure $Ed_{P_A,\delta}$ from the inner structure distance δ, π_A, P_A, and π^*.

The three boxes in Figure 5.5 denote points of computational operations in a simulation. Simulating movements of devices is out of the scope of this work, leaving the operations of *attacker* and *analyst* for simulation. Both operations turn out to be computationally complex for large $|M|$.

5.6.2 Layers of Context Information

The notion of *context information* has repeatedly been used in an very abstract way. In this section the sources of context information are discussed with focus on simulative analysis. Sources are classified by denoting the building blocks of hypotheses: set partitions Π_M, sets (trajectories) 2^M, and tuples of elements (messages) M^2. These classes, denoted *layers*, provide a model for the definition of attacker probability mass assignments that combine information from all sources.

Context information, in general, can be modelled by measure-like plausibility functions $pl : \mathfrak{D} \to \mathbb{R}_{+0}$, where the domain $\mathfrak{D}$ consists of elements relevant to the specific layer. Herein $\mathfrak{D}$ on each layer is one of $\Pi_M, 2^M, M^2$.

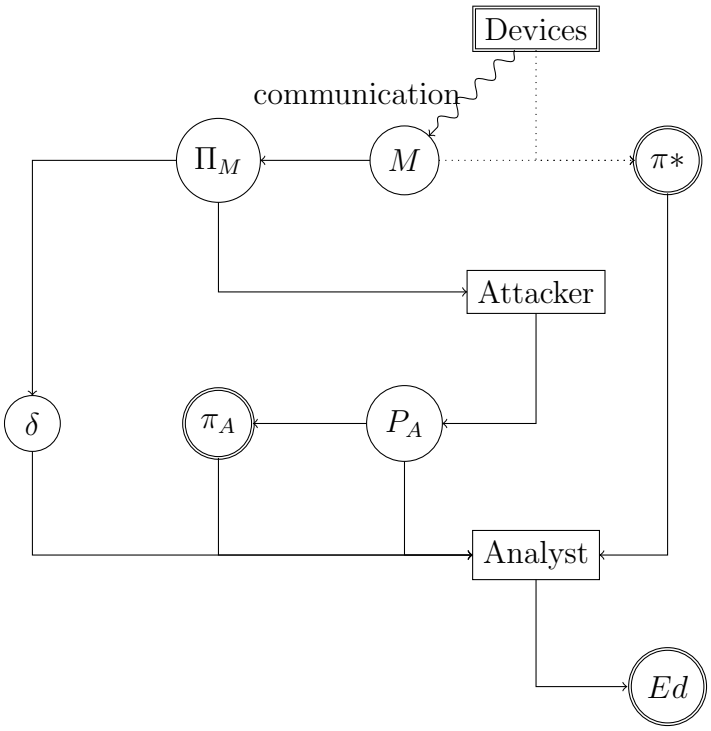


Figure 5.5: White-box Analysis Information Flow.

Interpretation of *pl* is only guided by the directive that a higher value means a higher plausibility and zero comprises the minimum value.

The *partition layer* is described by a partition plausibility function $pl : \Pi_M \to \mathbb{R}_{+0}$. Information used on this layer concerns statements or statistics on the partition, e. g., an expectation of the number of devices, respectively trajectories. This layer comprises the top layer. Partition plausibilities can be transformed into probability mass assignments, e. g., by normalisation.

On the *trajectory layer*, context information on the plausibility of individual trajectories is gathered and combined. A set plausibility function $pl : 2^M \to \mathbb{R}_{+0}$ defines the plausibility of trajectories. Examples for information on the set layer are plausibilities of trajectory endpoints and plausibility of turns at junctions.

Context information on the *link layer* describes the plausibility of pairs of sample points to be contained in the same trajectory. A link plausibility function $pl : M^2 \to \mathbb{R}_{+0}$ considers independent information on the relation between two distinct pairs. Dependencies to links between other messages are not included on this layer.

The three layers are naturally ordered, with lower layers providing information to higher layers. Being the direct prerequisite for probability mass

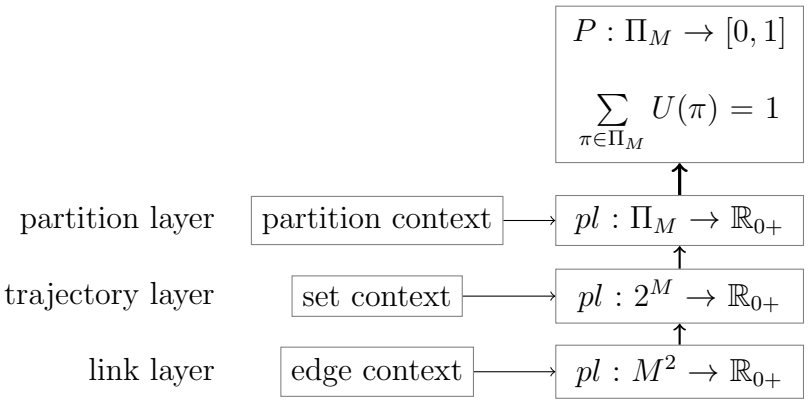


Figure 5.6: Layer Model

assignments, the partition layer is on top. Plausibility values from the trajectory layer can be combined to provide information on the plausibility of partitions. The link layer is on the bottom of the stack as a pair of messages comprises the smallest set for a link. The link layer provides plausibility values that can be used on the set layer. From the bottom to the top, each layer uses plausibility values from the lower layer, refines them and weights them by layer specific context information. Information flows from the bottom to the top.

Sources of context information depend only on the restrictions imposed on an attacker. In a simulation, an attacker can be provided with movements patterns of defined granularity and precision. In a real-life situation, an attacker may have collected prior knowledge in the form of traffic statistics, or educated guesses. These statistical distributions describe how devices move through the sample space. In preliminary simulations, the distribution of velocity has been measured and provided to the attacker.

The whole stack can be used to define an attacker's probability mass assignment. Implementing an assignment already proves to be a hard problem, because of the large number of set partitions that have to be considered at least in the normalisation-step from partition plausibility to probability mass assignment.

This classification of context information can be seen as an addition to the hint given by Franz et al. [25]. Whereas they provide individual hints, our model is more concerned with the combination of hints for implementation in a simulation.

5.6.3 Sum-Sum-Norm Attacker Assignment

In this section an example attacker's probability assignment is introduced. The approach is based on the layer model from Section 5.6.2. In this approach, we assume that the attacker has some context information on the relation between single messages, e. g., a velocity distribution.

The *sum-sum-norm* attacker assignment is a very naïve method of defining P_A. Assuming that the attacker is able to quantify the plausibility of two sample points being related, we model the attacker's probability of a partition as sum of pairwise plausibilities of samples in the same cluster of this partition. This sum is normalised by the sum of all partitions to gain a probability distribution.

This assignment obviously favours large trajectories, resulting in the highest probability mass being assigned invariably to the partition containing only one trajectory. The Mean-Mean-Norm Attacker Assignment in Section 5.6.4 provides a more useful measure. Herein the sum-sum-norm assignment is used to show an exemplary solution for complexity reduction.

Assume that the attacker is able to construct a link plausibility function $pl : M^2 \to \mathbb{R}_{+0}$. Sample points M then form a complete undirected graph $G = (M, E, pl : E \to \mathbb{R}_{+0})$, where M is the set of samples, $E = \{e \in 2^M : |e| = 2\}$, and pl defining edge weights. The edge weights herein define a link plausibility.

The sum of link plausibilities in a set of edges can be defined as the plausibility of a set of edges.

Definition 5.3 (Plausibility Value of Edge Set)
The plausibility of a set E of edges is the sum of plausibility values of the edges in this set,

$$pl(E) := \sum_{e \in E} pl(e),$$

where $pl : E \to \mathbb{R}_{+0}$ is a link plausibility function.

This implicitly defines the trajectory plausibility $pl : 2^M \to \mathbb{R}_{+0}$ as sum of edge plausibilities of all edges contained in a subgraph of G which contains exactly only those sample points contained in the trajectory.

On the partition layer, the plausibility of an partition is defined as the sum of trajectory plausibilities. For simplicity, the trajectory plausibility is skipped and the partition plausibility defined directly from the link plausibility.

The notion that a *partition π contains an edge* denotes that both sample points m, m' forming an edge $e = (m, m')$ are related with respect to π, denoted $m \sim_\pi m' \iff e \in \pi$. We further re-define $E : \Pi_M \to 2^E$ as a function that maps each partition to the subset of edges contained in that

partition.

Definition 5.4 (Edges in Partition)
The set of edges of the graph $G = (M, E, pl : E \to \mathbb{R}_{+0})$ contained *in the partition* π *of* M *is defined as*

$$E(\pi) := \{e \in E : \forall M_i \in \pi,\ m \in M_i \iff m' \in M_i\},$$

where $\pi = \{M_1, \ldots, M_n\}$ *with* $M_i \cap M_j = \emptyset \iff i \neq j$, $M = \bigcup M_i$, *and* $e = (m, m')$.

With the set of edges in a partition defined, the plausibility value of edge sets can be applied to the set of edges in a partition. Furthermore, we can now summarise the plausibility values of all partitions needed for normalisation.

Definition 5.5 (Plausibility Value of Partitions)
The plausibility value $pl(\pi)$ *of a single partition* π *is the plausibility value of the contained set of edges,*

$$pl(\pi) := pl(E(\pi)).$$

The plausibility value $pl(\Pi_M)$ *of the set of partitions over* M *is the sum of plausibility values of all single partitions in* Π_M,

$$pl(\Pi_M) := \sum_{\pi \in \Pi_M} pl(\pi).$$

Now a probability mass assignment can be defined for an attacker based on plausibility values for relations between samples as

$$P_A(\pi) := \frac{pl(\pi)}{pl(\Pi_M)}, \tag{5.12}$$

with $P_A : \Pi_M \to [0, 1]$ and $\sum_{\pi \in \Pi_M} P_A(\pi) = 1$. This defines a world view of the sum-sum-norm attacker.

Complexity Reduction

As has been mentioned repeatedly, due to the exponentially growing cardinality of Π_M, even for small $|M|$, $pl(\Pi_M)$ cannot be computed directly in an efficient manner. In the following, Theorem 5.1 is introduced which reduces the computational complexity of $pl(\Pi_M)$.

Theorem (5.1) allows computation of $pl(\pi)$ as $B(|M| - 1)pl(E)$. The following Lemma 5.6 explains the relation between $B(|M| - 1)$ and the number of partitions containing a specific edge[15].

[15]It has been brought to our attention that Lemma 5.6 is a special case of Equation (5) in [25] which has not been proven there. It should be trivial to induce this Equation (5) from our lemma

Lemma 5.6 (Number of Partitions Containing an Edge)
For all edges $e \in E$, the number of partitions in partition set Π_M that contain the edge e is equivalent to the Bell-Number of the cardinality of M reduced by one element,

$$|\{\pi \in \Pi_M : e \in E(\pi)\}| = B(|M| - 1), \quad \forall e \in E \tag{5.13}$$

The following proof of Lemma 5.6 is based on construction of $\Pi_{M\setminus\{m\}}$. For an example observe the partitions of a 3-elemental set: $\langle 1,1,1\rangle$, $\langle 1,1,2\rangle$, $\langle 1,2,1\rangle$, $\langle 1,2,2\rangle$, $\langle 1,2,3\rangle$. Choose, without loss of generality, the edge between the first and the second element. The partitions containing this edge are $\langle 1,1,1\rangle$ and $\langle 1,1,2\rangle$. The number of partitions containing this edge is the number of partitions of a two-elemental set. The number of all partitions containing any edge in this example edge is equal to $B(2)$, or $B(|M|-1)$ in general. The reader is encouraged to try the same with a four-elemental (or any larger) set to get the intuition right.

Proof 5.6
It is to be shown that

$$|\{\pi \in \Pi_M : e \in E(\pi)\}| = B(|M| - 1)$$

for every $e \in E$.

Given, without loss of generality $e = (m, m')$, $m, m' \in M$ and choose m so that

$$B(|M| - 1) = |\Pi_{M\setminus\{m\}}|,$$

and define $\Pi' := \{\pi \in \Pi_M : e \in E(\pi)\}$.

Now define a map $f : \Pi' \to \Pi_{M\setminus\{m\}}$ that maps every partition in Π' onto a partition without element m, $f : \pi \mapsto \pi \setminus \{m\}$. Where the set-substraction $\setminus$ denotes that m is removed from the subset $M_i \in \pi$. As $\pi \in \Pi_M$ obviously $f(\pi) \in \Pi_{M\setminus\{m\}}$

To show the lemma it has to be shown that f is a bijection.

a) f is surjective, because

$$\forall \pi \in \Pi_{M\setminus\{m\}} \exists M_i \in \pi : m' \in M_i$$

let $M_i' = M_i \cup \{m\}$ and let $\pi' = \{M_1, \ldots, M_{i-1}, M_i', M_{i+1}, \ldots, M_n\}$ where $M_1, \ldots, M_n \in \pi$. Then obviously $f(\pi') = \pi$ and surjectivity of f is shown.

b) f is injective because negation of implication leads to contradiction. Assume there exists $\pi, \pi' \in \Pi'$ with $f(\pi) = f(\pi')$ and $\pi \neq \pi'$. This means that there exists a bijective map $b : f(\pi) \to f(\pi')$ with $M_i = b(M)$ for all $M_i \in f(\pi)$.

Especially here must be $m' \in M_j \in f(\pi), M_j = b(M_j)$. Because all partitions in Π' have m, m' in the same cluster, all clusters are disjoint and f removes only m, m, m' must be in the same cluster in pi and pi'. As no other clusters are changed $pi = \{M_1, \ldots, M_j \cup \{m\}, \ldots, M_{|f(\pi)|}\} = \pi'$ which contradicts the assumption.

As f is bijective and $|\Pi'| = |\Pi_{M\setminus\{m\}}|$ the lemma is proven. □

Lemma 5.6 provides the cornerstone to reduce the complexity of calculating the *plausibility of all set partitions* (Def. 5.5) from exponential complexity to polynomial complexity.

The following theorem on the *Sum of Partition Plausibility* exploits the relation between the number of partitions containing a specific edge and the number of occurrences of each edge in the sum of $pl(\Pi_M)$. This theorem states that each edge is contained exactly $B(|M| - 1)$ times in the sum of plausibilities of all partitions.

Theorem 5.1 (Sum of Partition Plausibility)
If the plausibility of a partition $pl(\pi)$ is defined as the sum of plausibilities of all edges contained in that partition (Def. 5.5) and the plausibility of a set of edges $pl(E)$ is defined as the sum of all edge plausibilities as well (Def. 5.3), then the sum of plausibilities of all partitions can be calculated as

$$\sum_{\pi \in \Pi_M} \quad pl(\pi) \quad = \quad B(|M| - 1) \quad pl(E). \tag{5.14}$$

The correctness of Theorem 5.1 follows from the definitions of pl and Lemma 5.6.

Proof 5.7

$$
\begin{aligned}
B(|M|-1)pl(E) &\overset{def}{=} B(|M|-1)\sum_{e\in E} pl(e) \\
&= \sum_{e\in E}\left[|\Pi_{M\setminus\{m\}}| \quad pl(e)\right] \\
&\overset{Lemma\ 5.6}{=} \sum_{e\in E}[|\{\pi\in\Pi_M : e\in E(\pi)\}| \quad pl(e)] \\
&= \sum_{e\in E}\left[\sum_{\pi\in\Pi_M : e\in E(\pi)} pl(e)\right] \\
&= \sum_{\pi\in\Pi_M}\left[\sum_{e\in E : e\in E(\pi)} pl(e)\right] \\
&\overset{def}{=} \sum_{\pi\in\Pi_M} pl(\pi)
\end{aligned}
$$

□

The only new step is the reordering of sums between the fourth and fifth line which is justified because addition is both commutative and associative. What is done there is counting all occurrences of all edges in the plausibility value pl for all partitions. From Lemma 5.6, we know that the number of occurrences is equal for all edges and depends only on the cardinality of the base set M. As $\sum_{\pi\in\Pi_M} pl(E\pi)$ can be efficiently calculated, $P_A(\pi)$ can be efficiently calculated.

5.6.4 Mean-Mean-Norm Attacker Assignment

The Sum-Sum-Norm Assignment from Section 5.6.3 can be easily improved in terms of computational complexity, but provides a very biased attacker. A different approach that does not favour large trajectories is to use mean values instead where summation is used in the sum-sum-norm. In the following, some hints towards a generalised complexity reduction are given.

Plausibilities pl for links, trajectories, and partitions and attacker's assignment P_A are then defined along the lines of the previous section as means

$\lvert M\rvert$	$B(\lvert M\rvert)$	$P_A/\sum_{e\in E} pl(e)$
2	2	1
3	5	$\frac{1}{2}+\frac{1}{3}$
4	15	$\frac{3}{6}+\frac{1}{2}+\frac{1}{3}$
5	52	$\frac{3}{12}+\frac{1}{10}+\frac{3}{9}+\frac{3}{6}+\frac{1}{4}+\frac{3}{3}+\frac{1}{2}$
6	203	$\frac{4}{20}+\frac{6}{18}+\frac{1}{15}+\frac{10}{12}+\frac{12}{9}+\frac{4}{6}+\frac{1}{5}+\frac{6}{4}+\frac{7}{3}+\frac{1}{2}$
7	877	$\frac{1}{2}+\frac{15}{3}+\frac{25}{4}+\frac{10}{5}+\frac{6}{6}+\frac{35}{9}+\frac{40}{12}+\frac{5}{15}+\frac{30}{18}+\frac{10}{20}+\frac{1}{21}+\frac{10}{40}+\frac{15}{30}$

Table 5.5: Mean-Mean-Norm-Quotient of P_A

$$
\begin{aligned}
&pl(e) \text{ given}\\
pl(E) &:= \frac{\sum\limits_{e\in E} pl(e)}{|E|}\\
pl(\pi) &:= \frac{\sum\limits_{E\in \pi} pl(E)}{|\pi|}\\
P_A(\pi) &:= \frac{pl(\pi)}{\sum\limits_{\pi'\in \Pi_M} pl(\pi')}
\end{aligned}
\tag{5.15}
$$

The complexity of direct calculation of P_A again is exponential because of the exponential summation over all elements of Π_M. We have not been able to find a complexity reduction as for the Sum-Sum-Norm from the previous section. In the following, preliminary work that hints toward a possible complexity reduction is given.

Where the Sum-Sum-Norm (5.12) describes a sum of plausibility values, Mean-Mean-Norm describes a quotient of a sum of plausibility values of edges. Evidently, it turns out that this seems, again, to be a factor to the sum of all link-plausibilities $\sum_{e\in E} pl(e)$. One can compute this factor $P_A/\sum_{e\in E} pl(e)$ by counting the number of times an edge occurs within a certain trajectory and partition size combination.

In Table 5.5 the factor $P_A/\sum_{e\in E} pl(e)$ is given. The numbers have been computed by calculating the factor to each link-plausibility $pl(e)$. For cardinalities of M up to 8 this factor has been equal for all link-plausibilities. Given the symmetry of occurrence of edges in Π_M this is no surprise.

The sum of numerators in each row of Table 5.5 is $B(|M|-1)$, obviously because the number of occurrences of a single edge in Π_M is the same as in the Sum-Sum-Norm assignment. The order and occurrences of denominators is not explicitly known, neither is the order of numerators to denominators.

The next step would be to determine the denominators to find the relation

$P : \Pi_M \to \mathbb{R}_{+0}$	H	$\mathcal{D}$	π^*	$Ed_{P_x,\delta}(\pi^*)$
P_A 	1.70	0.73	π_1	0.55
P_B 	1.70	0.73	π_5	0.45
P_C 	0.00	0.00	π_1	0.00
P_D 	2.32	1.00	π_1	1.00
P_E 	1.84	0.79	π_1	0.99
P_F 	1.84	0.79	π_5	0.65

Table 5.6: Entropy H, Degree of Unlinkability $\mathcal{D}$, and Expected Distance (Ed) in Black-box Analysis.

$P_A / \sum_{e \in E} pl(e)$. This general structure may help to find a general formular for reduction of complexity of computation of P_A.

5.7 Discussion

In Section 5.5, the expected distance unlinkability measure Ed has been defined. In this section we discuss Ed with respect to the example used in the criticism of the degree of unlinkability in Section 5.4.3.

Within this section the example Π_M and example attackers A, B, C, D, E, and F from Table 5.3 are used. A and B provide distributions with a single peak, C, D are the degenerate/uniform distributions, and E, F are distributions consisting of two peaks.

5.7.1 Black-Box Analysis

Table 5.6 provides entropy based measures (H and $\mathcal{D}$) and the corresponding expected distance measure Ed in black-box analysis. Black-box analysis means that the reference partition is chosen to be the attacker's choice (5.4), which are $\pi_1, \pi_5, \pi_1, \pi_1, \pi_1, \pi_2$ for attackers A, B, C, D, E, F. Recall that a higher expected distance Ed is interpreted as higher unlinkability. In this example we use the partition distance from Definition 5.1 for the hypotheses distance δ.

The main difference between the expected distance Ed and the degree of unlinkability $\mathcal{D}$ in Table 5.6 is that Ed distinguishes assignments that that are not distinguished by $\mathcal{D}$.

Recalling the discussion from Section 5.4.3, the difference between attacker A and B is that the attacker's choice of B is better supported by other hypotheses than the attacker's choice of A. P_B therefore is more consistent than P_A, which is reflected by Ed. The same argument is valid for

attackers E and F, where P_F is considered to be more consistent than P_E, which again is reflected by Ed.

Analysing the representation of certainty by Ed is complicated — as has been discussed in Section 5.5.3 — because every comparison is biased by the inner structure distance metric. However, the examples presented above indicate that Ed is not indifferent to certainty.

Informally speaking: the wider P is distributed over the hypotheses space, the more probability mass is given to hypotheses with large distance δ_{π^*}. This obviously increments the value of Ed. Compare for example the single peak distributions P_A, P_B with the double-peak distributions P_E and P_F. The latter distributions can be interpreted as *less certain* assignments, because the probability mass is spread wider. This is reflected already by $\mathcal{D}$. As mentioned in Section 5.5.1, a single measure probably cannot perfectly reflect certainty and consistency at the same time.

Up to now, mostly assignments with equal entropy and different Ed have been compared. It is however not difficult to construct examples where Ed is invariant and $\mathcal{D}$ varies. An example of the latter is an assignment P'_A as given in Table 5.7 which differs from P_A only in that the assignments from π_3, π_4 are shifted onto the assignment of π_2 as given in Table 5.7. It can be observed that after the manipulation $\mathcal{D}$ is lower while $Ed_{P_A,\delta} = Ed_{P'_A,\delta}$.

	π_1	π_2	π_3	π_4	π_5	$\mathcal{D}$	Ed
P'_A	0.60	0.25	0.00	0.00	0.15	0.58	0.55

Table 5.7: Example assignment P'_A with $Ed_{P_A} = Ed_{P'_A}$ but $\mathcal{D}(P_A) \neq \mathcal{D}(P'_A)$

Logically, the change of certainty is not reflected because, considering this inner structure δ, π_2, π_3, and π_4 are equidistant from the reference partition, which is π_1 in both P_A and P'_A. Thus, concerning inner structure $\pi 2$, π_3, and π_4 are indiscernible, and thus identical, with respect to δ_{π_1}.

This shows, again, that every change to a distribution is measured by Ed only with respect to δ_{π^*}. Meaning, only if probability mass is shifted closer to or farther away from π^*, in terms of inner structure, does Ed reflect a change. Shifting mass equidistant does not change Ed, independent whether mass is concentrated on a single hypothesis or spread over hypotheses with equal δ_{π^*}.

For entropy based measures the highest unlinkability is achieved by uniform distributions. In Section 5.4.3 we have already discussed, that uniform distributions are not providing the highest unlinkability if the attacker considers the inner structure. The same is not true for Ed.

Using bounds from Section 5.5.3 we can calculate the maximum unlinkability value for attacker assignments, which is one in this example. Here

already more than one attacker's assignment provides $Ed = 1$. Due to the ordering of Π_M assignment P_D already provides maximum unlinkability. But the *canonical maximum unlinkability assignment* in this scenario for black-box analysis would be P'_D with the pairwise farthest hypotheses, here π_1, π_5 are assigned 0.5 probability mass each as given in Table 5.8.

	π_1	π_2	π_3	π_4	π_5	$\mathcal{D}$	Ed
P'_D ■|||■	0.50	0.00	0.00	0.00	0.50	0.43	1

Table 5.8: Maximum unlinkability assignment in black-box analysis P'_D.

One may observe that, again, $\mathcal{D}$ provides a higher value for P_D than for P'_D. This is, because by focusing probability mass on only two hypotheses in P'_D, the attacker's certainty, in terms of entropy, is higher than the certainty of a uniform distribution.

From these examples we make out one situation where different probability mass assignments are indistinguishable to Ed while being distinguished by entropy. Assuming that inner structure is defining a sensible distance metric the distribution of probability mass between hypotheses with equal distance from the current attacker's choice is unimportant to unlinkability. Further research has to show whether this is a weakness of Ed or if this is a genuine property of unlinkability itself.

5.7.2 White-Box Analysis

Consider now the white-box analysis, where the analyst has external knowledge on which set partition corresponds to reality. Assume that a white-box analyst will be testing privacy-enhancing measures against a variety of attackers. A white-box analyst is not only interested in improving the consistency of attack algorithms, but also in the correctness of attacks. Very different expected errors would hint toward flaws in the privacy-enhancing measures that are tested.

The expected distance unlinkability Ed in white-box analysis has slightly different attributes compared to Ed in black-box analysis. Most remarkable is the different upper bound which equals the largest distance between any two hypotheses (see Section 5.5.3).

In Table 5.9, a complete compilation of expected distances Ed is given for the considered examples. As the expected distance unlinkability measure depends on π^*, a white-box analyst can use the measure to determine the expectation error of an attacker in relation to the true relation, or any assumed attacker's choice. Thus the analyst is not only able to quantify the absolute quality of an attack, but is able to quantify the absolute success of countermeasures against a given attacker.

$P : \Pi_M \to \mathbb{R}_{+0}$	$\pi^* =$ π_1	π_2	π_3	π_4	π_5
P_A 	0.55	0.95	0.85	0.95	1.45
P_B 	1.55	0.85	0.95	0.85	0.45
P_C 	0	1	1	1	2
P_D 	1.0	0.8	0.8	0.8	1.0
P_E 	0.99	0.95	0.89	0.95	1.01
P_F 	1.35	0.61	0.89	0.95	0.65

Table 5.9: Expected Distance $Ed_{P_x,\delta}(\pi^*)$ in White-box analysis, calculated for different reference hypothesis and attacker probability distributions.

The data from Table 5.9 provides two types of information for either analysis or attack. For the analyst it provides knowledge which scenarios provide high unlinkability. For the attacker it provides an improved attacker's choice as described in Section 5.7.3.

A white-box analyst naturally calculates only one expected distance for a given attacker's assignment and takes measures from different attackers to estimate the unlinkability of a scenario. In the following we describe how Table 5.9 can be utilised in throughout analysis.

Consider for example Attacker A, whose attacker's choice is π_1. Assume, for example, that the real hypothesis is π_3, the attacker's choice of A would be have an expected error of 0.85. Attacker B would be only slightly worse, with Ed of 0.95, although its attacker's choice π_5 is very much the opposite to A.

If different attacker's provide equally wrong results, than the scenario is not vulnerable to specific attacks. Both attacker's choices would be wrong, of course, but considering that their context information hinted at opposite directions, this scenario can be considered as not being vulnerable.

A different situation would exist if the real partition would be π_1 or π_5. Then, one of the attackers provides a sufficiently better attack, which could hint at flaws in any privacy enhancing techniques use in this scenario. This, in return, might help to find better techniques to provide privacy.

Thus, white-box analysis yields two types of information. First absolute information on the unlinkability of privacy-enhancing techniques in the face of selected attackers. Second, information on potential flaws in privacy-enhancing techniques.

Furthermore, Table 5.9 provides information on how to improve the attacker's choice by considering the expected error. Attacker D and F provide example cases where a different attacker's choice would reduce the expected distance with respect to the context information condensed in the attacker's

assignment.

The choice of attacker D is π_1. Making this choice, the expected error, assuming the attacker makes a black-box estimation, is 1. But, if attacker D chooses any of π_2, π_3, π_4 he would reduce Ed to 0.8. Attacker F could improve his choice, in terms of Ed, from 0.65 to 0.61 by choosing π_2 instead of π_5. An attacker can improve his choice, with respect to error estimation, by choosing based on Ed.

In the following section, a revised attacker's choice is introduced that uses a modified attacker's order.

5.7.3 Attackers Choice Revisited

Taking an attacker's point of view, the inner structure from Section 5.3.2 introduces additional context information that can be used to improve the attacker's choice. The idea is that an attacker may define its *attacker's order* with respect to the expected distance Ed. By using Ed, the expected error is minimised based on the attacker's assignment. The attacker calculates Ed with different set partitions for π^* and chooses the partition that leads to the smallest error.

Essentially this means to define a new attacker's order $<_P^*$ that supersedes the partial order induced by Ed before the attacker's order $<_P$ as defined by (5.3). The revisited attacker's order is defined as follows

$$\begin{aligned} \pi_i <_P^* \pi_j : \Longleftrightarrow\ & Ed_{P_A,\delta}(\pi_i) < Ed_{P_A,\delta}(\pi_j) \\ & \vee\ (Ed_{P_A,\delta}(\pi_i) = Ed_{P_A,\delta}(\pi_j) \wedge \pi_i <_P \pi_j). \end{aligned} \tag{5.16}$$

This second probability mass assignment is then used as order in the attacker's choice.

Effectively, the attacker has now included the inner structure in his assignment. The attacker's choice, which is basically a guess based on context information, is now modified to reduce the actual error. In terms of Ed this automatically improves the attacker's choice.

5.8 Conclusion

In this chapter, the expected distance unlinkability measure has been introduced. This measure is motivated by the shortcomings of existing entropy based measures (see Section 5.4.3).

The main contribution of this work is the distinction between *inner* and *outer* structure in global unlinkability metrics. Hypotheses in unlinkability problems are not atomic like in hypotheses in anonymity problems. A

hypotheses-space in unlinkability problems generally has a complicated structure that represents the objectives of the analysed attacker. Ignoring the inner structure, produces metrics that do not sufficiently represent the quality of an attacker and thus are not fit to be used as unlinkability metric.

The *degree of unlinkability*, discussed in Section 5.4.4, is an example of an unlinkability measure ignoring the *inner structure* of the hypothesis-space. The *expected distance unlinkability measure* introduced in Section 5.5 distinguishes situations that are not distinguished by the previous measure. Both unlinkability measures have the disadvantage that they are not efficiently computable in a direct way. Concerning efficiency, the expected distance unlinkability is at least not worse than the degree of unlinkability.

Unlinkability measures, with hypotheses space Π_M, have been considered in only few previous works. Thus it is to be expected that this chapter leaves plenty of open questions.

The simulative analysis still provides future work. The application of Ed to a realistic scenario is yet missing. We provided two simple attacker algorithms based on a layer model of context information. We could reduce the complexity of the Sum-Sum-Norm Assignments (Section 5.6.3) to polynomial complexity. Direct implementation of Ed is expensive, but we provided a first step towards substantial complexity reduction.

Pending computational feasibility, the expected distance can be useful as a defining measure for the amount of unlinkability. For the development of privacy enhancing emission protocols standard measures for unlinkability are useful. The development of measures needs a coherent and uniform definition of the amount of unlinkability.

Furthermore, standard scenarios and standard attackers have to be found to provide test-cases for protocol developers. The perfect unlinkability measure would measure scenarios independent of specific attackers, similar to measures like ℓ-diversity. Starting from the notion of distance, standards for inner structure should be agreed upon by the privacy community. An urgent next step in this area would be to combine the lessons learned from the database-privacy measures and unlinkability measures towards this goal.

Principle 11
The protocol designer should know which trust relations his protocol depends on, and why the dependence is necessary. The reasons for particular trust relations being acceptable should be explicit though they will be founded on judgement and policy rather than logic.

Martin Abadi and Roger Needham in [2]

6 Anonymous Certification

Contents

In this chapter, we introduce an anonymous certification protocol with secure anonymity revocation. We emphasize *secure* as attribute to hint at our focus on protection against single points of failure in the revocation process. This especially targets authorities which usually have unlimited power over the relation between user and certificate, e. g., certification authorities.

We use the vehicular communications scenario described in Chapter 3 as backdrop and general motivation. The main feature that distinguishes this protocol from other works in vehicular communication is its focus on the principle of *separation of duty*. Another focus of this work is on the tradeoff between privacy, security, and communication overhead.

Unlinkability for certificates provides multiple problems for communication protocols. These problems range from the authenticity of data to the question of which data is included in a certificate and which data is revealed

to certification authorities. The need for revocation adds questions about which data is stored at the revocation authorities, of whom and how this data is generated and by what methods a vehicle is prevented from gaining non-revocable certificates. Some of these topics are addressed here.

The scenario considered here is determined on one side by strict time and security constraints and on the other side by high privacy risks. [55] Safety-related messages in the vehicular context have to be processed — from sending to reception till reaction — within few hundred milliseconds. Errors, e. g., injected messages by an attacker, might easily result in humans being harmed. On the other hand there are privacy concerns. Positional and identifying information is included in these messages and possibly open to the public. In the overall debate, there is always a compromise between security (e. g., by authentication) and privacy.

Specific choices for security parameters of an implementation of vehicular communication infrastructure are subject to political decisions. The objective of our work is to provide the means to support a wide range of parameters to allow for an optimum trade-off between privacy, security, efficiency and costs. The addressed security objectives are discussed in Section 6.1.

One technique to provide unlinkability is to authenticate vehicles using pseudonymous certificates that are changed frequently to prevent tracking of vehicles citeeichler2007:strategiespseudonymchanges[13][48]. Certificates are provided by certificate authorities which generally have complete control over data and thus have to be trusted by all involved parties. We herein refer to a notion of trust as given by Audun Jøsang:

> "*Trust* is the extent to which one party is willing to depend on something or somebody in a given situation with a feeling of relative security, even though negative consequences are possible." [32]

While *trusted third parties* often have been used as a panacea for all kinds of security problems, they are actually not a good solution but a kludge to solve situations where real security cannot be provided[1]. While *trust* is a concept very common in human interaction, it cannot be considered as being equal to security.

Certification authorities concentrate a huge amount of power: The power to grant services (communication certificates here) and knowledge (here relations between certificates and identities). Practice, however, shows that such parties are not necessarily trustworthy. Certification authorities can fail in many ways: technical and administrative errors as well as non-trustworthy

[1] Or where secure solutions are to costly, unknown or no one cared to research for a solution.

employees with access to crucial information. We approach this problem by enforcing the many-eye-principle on crucial operations that have been controlled previously by single authorities. We present a solution that combines fair blind signatures with shared secret computation.

Similar to our previous work [64] single authorities are herein considered as potential adversaries. Privacy incidents, where authorities misused their privileges[2], have shown that this scenario is justified in practise. The objective of this is to increase trust in authorities by enforcing privacy protection through many-eye-protected anonymity revocation.

The general objective of this part of our work is to reduce the amount of needed trust within the whole system to the unavoidable minimum.

This chapter is structured as follows. In Section 6.1, security objectives for certification in the vehicular scenario are discussed. The certification protocol described in Section 6.3 is based on protocol concepts that are introduced in Section 6.2. The protocol is then analysed with respect to security in Section 6.4 and complexity in Section 6.5. The chapter is concluded in Section 6.7.

6.1 Prerequisites

In this section, security objectives for revocable anonymous certification in vehicular communication are discussed. These objectives refine the security objectives from Chapter 3. While many problems that occur in vehicular communication scenarios have been sufficiently tackled in the past, the problem of non-trustworthy authorities has mostly been avoided. Certification authorities have been declared as *trusted* and thus failure has been declared impossible.

In the following, our notion of *authority failure* is introduced. Afterwards, we describe assumption that have to be fulfilled for the protocol to work. This is followed by specific objectives for certification and revocation. Finally complexity constraints for the protocol are discussed.

6.1.1 Authority Failures

We distinguish two different types of authority failures, based on certain responsibilities of certification authorities, that shall be used in the analysis of the protocol later. An *authority failure* is a situation where a (trusted) authority fails to fulfil its duty.

[2]For example a big german telecommunications corporation used their superior resources to spy on journalists and the directorate.

We distinguish two different tasks for authorities. *Certification authorities* (CA) are entrusted with the power to sign or deny certificates. *Revocation authorities* (RA) are entrusted with the power to link certificates to identities. In standard *Public-Key Infrastructure* (PKI)-scenarios CA and RA are implemented as one entity.

The first type of failure considered here is the *CA-failure*. A CA can fail in basically two ways. Either it fails to deny a certificate to an requester that has no right to receive a certificate. or the CA falsely denies a certificate to an authorised requester.

An *RA-failure* occurs if that the information between a certificate and the identity of the requester is disclosed without legal reason to do so. The failure of falsely refusing revocation is not discussed here because such a failure is based on faulty interpretation of revocation policies, e. g., privacy law, and is out of the scope of this work.

The main problem of RA-failures is, that the effects of revocation are not reversible. In this aspect privacy is very similar to the security objective of *secrecy*. A secret can only be considered a secret until it is disclosed for the first time. *Disclosure* in this context denotes that information leaves the *sphere of influence* of the bearer of the secret. Outside of this sphere of influence, further distribution generally cannot be controlled. This attribute provides the motivation for prevention of RA-failures.

Requesters of certificates only win from privacy protection. Given that the requester may, at any point, choose to disclose the relation between certificates and identity, malicious behaviour, with respect to privacy, is not to be expected. Nonetheless, it is sensible to recognise possible erroneous behaviour of a requester. Please refer to the security discussion in Section 6.4 where protocol specific requester-failures are discussed.

6.1.2 Security Assumptions

The protocol proposed in this chapter provides only a part of the whole protocol stack. In this section we introduce assumptions on the security properties of the used scenario and authorities.

We assume that all communication between requester and CA, as well as communication between CA and RA, is encrypted and two-way authenticated. Authentication between CA and RA is prerequisite to prevent that non-authorised third parties are able to inject information into the RA's database.

A requester has to be authenticated and authorised before a CA grants an Inter-Vehicle Communication certificate (IVC). Authorisation has to be checked against an up-to-date revocation lists by the CA. The certificate request of a vehicle that fails to correctly authenticate itself or whose autho-

risation is revoked must be rejected. The case of a CA that fails to reject an illegitimate certificate request is described as CA failure in Section 6.1.1.

Misuse of information by a communication-network provider, e. g., UMTS or GSM provider, or other adversaries which overlook all network traffic is out of scope of this work. This includes any kind of traffic analysis, high-frequency fingerprinting, triangulation or similar techniques.

We assume further that no owner's or vehicle's identities or identifying data is hidden in certificates and messages by the vehicles themselves.

The protocol introduced in Section 6.3 makes use of cryptographic functions. In this chapter, it is assumed that the cryptographic algorithms used are strong. Used hash functions are assumed to be secure in the random oracle model. The reader is referred to the original works on these cryptographic methods for further information[3].

Furthermore, misbehaviour by the CA in form of denial of service is out of the scope of this work. It is assumed that a sufficient number of CA is available in every region at any given time.

6.1.3 Protocol Objectives

In this section objectives for certification in vehicular communication are described. These objectives motivate the changes we introduced to the original revocation protocol given in [62]. In the following we distinguish objectives for certification, revocation and complexity.

Certification Objectives

Only an authorised CA must be able to sign valid certificates, and integrity of certificates must be protected. A signing CA must be able to verify the correctness of certain information that is included in the certificate without seeing the whole certificate. One possibility for such data is the *validity time* of a certificate. A CA must be able to verify that this data is indeed embedded in the certificate and complies with given rules. This data must not be removable or overruled by other data in the certificate.

The protocol must ensure that it is possible for any vehicle to verify that no hidden identification is inserted by the CA. During certification, no information that can be used by a CA to identify the certificate must be disclosed to the CA. As a vehicle has to authenticate towards the CA (Section 6.1.2) this would give the CA sufficient information to link a certificate to the vehicle's identity and thus counter our main privacy objective. More general, it must be prevented that information sufficient to identify a certificate's owner,

[3] An excellent overview on applied cryptography is given in [47].

aside from the information within the revocation process, is disclosed to any entity.

Explicitly, certification authorities must never be able to relate identification of certificates to requester identification. Certificates must not contain information which links them to other certificates of the same requester.

Revocation Objectives

The notion of *revocation* covers the whole process starting from the detection of malevolent messages to the revocation of an vehicle's authorisation. The process, as sketched in Chapter 3, includes Type-II anonymity revocation and revocation of the vehicle's — respectively the vehicle's owner's — authorisation to emit vehicular messages.

Certificate revocation often is a necessity in anonymous certification to prevent perfect crimes [69]. This is especially true in vehicular communications [67]. Revocation in vehicular networks has been tackled in various ways in different works [56, 23, 64]. If certification authorities are completely trusted to not misuse their revocation powers, standard public key infrastructures, e. g., as described in [55], are a valid solution.

RA-Failure Prevention An RA-failures as described in Section 6.1.1, implies the disclosure of information. Information disclosure, as opposed CA-failures cannot be made undone. Therefore, it is critical to prevent them. For this work we set the following objectives.

No single authority must be able to revoke a certificate. A certification/revocation protocol should provide a choice of a minimum number t of authorities that have to agree on the revocation of anonymity of a requester/certificate. For any number of authorities less than t it must then be impossible to do either Type-I or Type-II revocation (see Section 2.5).

Obviously, no onlooker must be able to revoke anonymity, meaning no certificate or message must include sufficient information to identify the vehicle.

Revocation Enforcement Revocation is prerequisite to isolation of rogue vehicles. The main security attribute of an isolation protocol, the maximum isolation time (see Section 2.6), depends on the complexity of the revocation process. Thus, any revocation process should be guaranteed to succeed not only in finite, but also short time.

The first method to revoke the authorisation of a vehicle is to limit the validity time of certificates. Certificates should not be valid for an infinite time interval. The main reason for this is the limited security of cryptographic

key material. It has been prudent engineering practise to regularly change keys, which are a crucial part of certificates.

A second reason, more specific to the vehicular scenario is to avoid the use of certificate revocation lists (CRL). Revocation lists increase the complexity of certificate validation. Verification is part of the strictly time-constrained part of vehicular communication as has been described in Section 3. In [56] it has also been discussed that timely distribution of revocation lists in vehicular networks requires distribution through road-side infrastructure, which may not always be available.

A certification protocol should allow for multiple means of revocation. Revocation of certificates must be possible without distribution of certificate revocation lists, which may be reserved for urgent cases. Revocation of authorisation can be implemented at the certification authority and thus is not as time critical, and revocation lists are reasonable here.

To ensure revocability of anonymity, each IVC certificate needs a revocation anchor. An *revocation anchor* is information encoded in the certificate that allows for an authorised entity to derive the corresponding identity. As certificates are only granted to authorised and authentic vehicles, information gained during the certification process, denoted *protocol trace*, can be linked to the identity of the vehicle and thus to the vehicle's owner. It must be prevented that revocation anchors can be removed from certificates.

A quorum of RA should be able to derive the vehicle's identity from this anchor, or conversely to recognise IVC certificates by vehicle identity. For secure privacy protection only a sufficient quorum of RA must be able link identity with certificates.

6.1.4 Complexity Objectives

It has been mentioned previously that the verification of IVC certificates is very time constrained. The allowable latency for traffic-safety messages ranges from 100 milliseconds to 1 second [14], with the most critical applications being those most strictly constrained.

Pseudonym changes are widely seen as the first and basic privacy enhancing method, which implies, that the certification process is executed frequently. In the worst case, one certification process has to be undertaken for every message sent. As certification can be undertaken independently before current events make it necessary to send a message, it is not as time-constrained as immediate IVC communication process. Further, certification protocols should not use the low bandwidth and high noise channel that probably will be used for inter-vehicle communication, but use dedicated communication channels.

Revocation processes are initiated if rogue vehicles are detected. Since revocation ensures that attackers can be identified (and prosecuted if necessary) the risk of an attacker can be assumed to be sufficient to reduce malicious attacks to a minimum. Thus the major part of revocation is probably due to vehicles with faulty sensors or OBU. Obviously there are no statistics on systems not yet implemented. We assume that the number of vehicles whose authorisation has to be revoked is not very large. To have some numbers to estimate size of revocation lists and number of revocation processes, we postulate, that no more than ten percent of all vehicles' anonymity has to be revoked within each year. The only limiting objective is the maximum isolation time, given that an attacker may try a denial of service attack on the revocation authorities, sufficient computer power has to be available at the revocation authorities.

Considering these points, revocation is the least restricted protocol part of the three mentioned parts. The focus of efficiency in IVC protocol designs thus must be on the verification process. Considering that neither the roadside infrastructure, nor the hardware inside the vehicle is currently decided upon, any proposed protocol should be configurable with respect to complexity of calculations. In Section 6.3 we propose a protocol modification that allows to balance the complexity of different protocol parts.

6.1.5 Summary Prerequisites

Concluding this section, three classes of failures have been described based on the entity that fails. Revocation authority (RA), certification authority (CA) or requester. Each class of failures can be divided further. Figure 6.1 provides an overview of the items considered herein.

In this work only *unauthorised disclosure* (Category 1.1) is considered under *RA-failures* (Category 1). Category 1.1 is subdivided into the two types of revocation defined in Section 2.5. Scenarios where a RA fails by other means are not a responsibility of the protocols discussed herein and thus not considered in the hierarchy.

CA-failures (Category 2) are subdivided into four different categories. Categories 2.1 and 2.2 denote failures where a CA makes an erroneous decision on the authorisation of the requester. Categories 2.3 and 2.4 denote failures where the CA fails to detect misbehaviour of the requester.

A *requester failure*, with respect to privacy only, means that the requester disclosed too much information. In the certification process, two failures can be distinguished. Failing to blind information (Category 3.1) leads to, at least, disclosure of information to the CA, i. e., the CA gains knowledge which is sufficient to revoke the anonymity of the certificate. Including too much information in the certificate (Category 3.2) leads to certificates that

1 RA-failures

- 1.1 unauthorised disclosure
  - 1.1.1 Type-I revocation
  - 1.1.2 Type-II revocation

2 CA-failures

- 2.1 false deny certificate
- 2.2 false grant certificate
- 2.3 accept non-revocable certificate where...
  - 2.3.1 Requester inserts fake $S\sharp' : S\sharp' \neq S\sharp$ in certificate
  - 2.3.2 Requester uses false revocation encryption scheme (D_x, E_y)
- 2.4 certificate with false validity interval

3 Requester-failure

- 3.1 unblinded certification
- 3.2 identifying information in certificate
  - 3.2.1 unencrypted revocation anchor $S\sharp$

Figure 6.1: Failure Tree Certification in Vehicular Networks.

can be related to the vehicle (anonymity problem) or to other certificates of the vehicle (unlinkability problem).

6.2 Protocol Concepts

During our work on SRAAC and T-SRAAC [23, 64] we introduced methods to meet objectives similar to those discussed above. In this section, protocol concepts are introduced that are intended to enhance privacy. All concepts are driven by the idea to reduce the knowledge and privileges of individual authorities.

6.2.1 Unpredictable Behaviour

Literature describes some methods to reduce linkability of position samples from statistical inference. In [57] *random silence periods* and *grouping of vehicles* are described and quantified using the notion of expected tracking time. In [12] the effects of *pseudonym changes* are examined in terms of attackers' success ratio. Other concepts that have not been discussed yet are *unpredictable behavior*, e. g., driving in circles or changing velocity. By behaving contrary to the expected attackers' behaviour, the attacker's statistical information becomes useless to track this individual vehicle. Adding *noise* to the positional data, e. g., randomising time information, can have a similar effect. Communication protocol privacy techniques can be classified into methods that increase the overall entropy of the probabilities of the (M^2) edges and methods that hide an individual vehicle.

6.2.2 Separation of Privilege

In Section 6.1.1 we described the distinction between revocation and certification. This distinction already indicates a concept that has also been described as *separation of privilege* by Papadimitratos et al. in [49]. Alternative names are *separation of duty* and *separation of knowledge*, which both are correct in this context.

To implement the need-to-know principle in certification we have to separate the processes of revocation and certification. In practise, this means that authorities that grant certificates, and authorities that are empowered to link certificates to identities are institutionally separated.

The objective of this method is to reduce the "amount of trust" that has to be put into a single institution. Without implementation of this principle all power and knowledge is given to a single institution. Using separation of duty, the public only has to entrust certification authorities (CA) with the

verification of whether a vehicle is granted an pseudonymous certificate or not. The CA is not able to relate that certificate to the vehicles identity.

In the example of anonymity revocation this concept can be implemented even further. As a revocation authority (RA) is empowered to revoke anonymity, and given the high misuse potential of this process, special checks might be put in place to ensure that an RA is indeed trustworthy. Again separation of privilege may be used to force multiple, separate entities to collaborate during revocation processes. We refine this general concept into the concept of *Quorum-RA* in Section 6.2.3.

6.2.3 Privacy Enhancing CA-Structures

In the following we describe an ideal communication structure for inter-vehicle communication. Herein, vehicles that play the role of *requesters of certificates* and *senders of inter-vehicle communication messages* (IVC messages) are distinguished from vehicles that are *receiver of IVC messages*. *Inter-vehicle communication certificates* (IVC certificates) contain credentials to authenticate IVC messages and are normally sent along with the message. IVC certificates are signed by *certification authorities* (CA). IVC certificates are pseudonymous and must not be linked to each other or to the identity of the holder of the certificate. *revocation authorities* (RA) hold the keys to identify vehicles given a valid IVC certificate.

Herein certification authorities are assumed to be organised hierarchically by region as proposed in [48]. This *spatial separation* reduces the knowledge (see Section 4.1) of single authorities. Observe Figure 6.2, only the CA at the bottom are aware of the relation between different IVC certificates. The higher layers of CA only authorise the bottom CA and do not grant IVC certificates. Thus the CA in contact with privacy-critical information observe only a defined spatial subspace. This concept of *separation of knowledge* is one of the protocol concepts proposed in Section 6.2.2. This structure is actually common in public key infrastructures.

Revocation authorities build a quorum that shares responsibility for revocation in a defined set of certification regions. A defined subset of RA within a quorum has to agree on revocation of a given certificate-user's anonymity for revocation to become possible. This is depicted in the example of Figure 6.2 as a connected set of RA, responsible for revocation within a whole area, e. g., Germany.

There are different parameters that control the security and overhead of the whole system. The partition of regions, the depth of a CA-hierachy, the number of RA, and the number and validity time of IVC certificates held by a single vehicle are some examples. Further parameters control attributes of the used protocols. Security-relevant parameters will be introduced in the

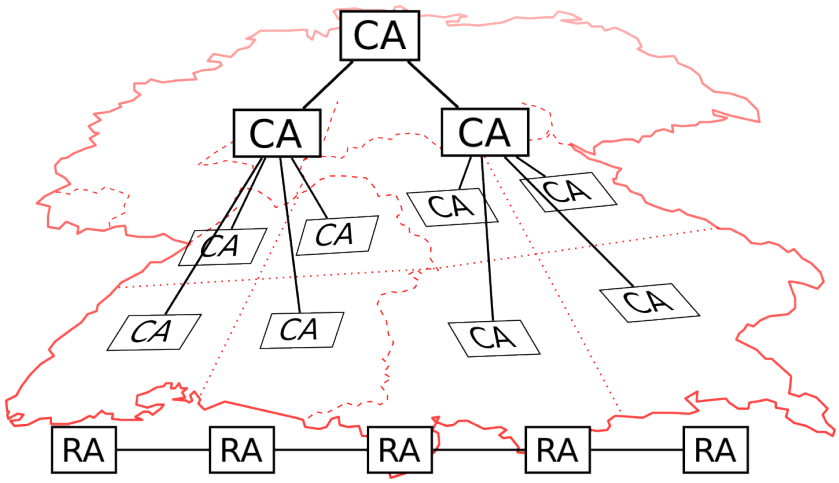


Figure 6.2: CA and RA spatial separation and separation of duty

following sections. Effects of these parameters will be discussed in Section 6.4.

In the following sections, different privacy-preserving certification structures and their effects on the unlinkability-model from Chapter 4 are discussed. This section extends our presentation in [24] and introduces additional privacy enhancing structures.

Quorum RA

In [64] we proposed a protocol that protects revocation from single points of failure. This protocol forces certification authorities to collaborate to relate a given certificate to a vehicle. This concept is denoted *quorum-RA*. Certificates and certification process must be constructed in a way that makes revocation possible if a previously defined number of RA collaboratively revokes the anonymity. This concept has been proposed as *federated databases* in [27].

The general concept is visualised in Fig. 6.3. Each member of a *quorum RA* is involved in the certification but only gets to know a share of the information needed to identify the vehicle from a certificate. By acting together (broken lines) a quorum of $t = 3$ of $n = 5$ RA can link certificate c and vehicle v.

In terms of linkability graphs $\mathcal{G}$ from Section 4.1, introducing a quorum of RA removes all $(V \times C)$-edges unless the attacker compromises a sufficient number of certification authorities. Other relations in the graph would not be concerned.

In a threat analysis, briefly, the costs for discovery of $V \sim C$ would increase by the costs needed to compromise the necessary number of author-

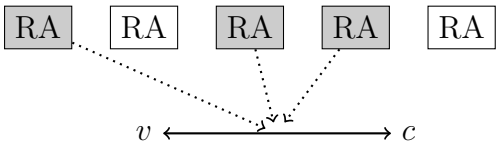


Figure 6.3: Quorum-RA Principle: The certification request is blinded. Certificate is signed blindly using magic-ink blind signatures. Revocation need collaboration of t of n RA.

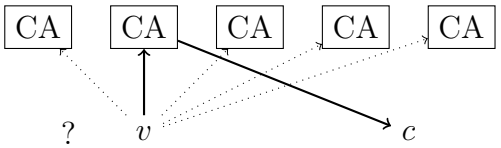


Figure 6.4: Vehicle-Selected CA: A vehicle changes not only the used certificates but as well the certification authority.

ities. These costs have to be calculated against the costs of implementing a system with corresponding number of RA.

Vehicle-selected CA

Vehicle-selected CA is a certification structure aiming at reducing the number of certificates that a single CA can link to the related certificate requester. Vehicle-selected CA is the concept of letting the vehicle, as agent of its owner, decide which CA is chosen for each independent certification. The principle is depicted in Figure 6.4. The question mark beside the vehicle denotes the place where the choice is made of which certification authority to use.

This principle has an effect similar to pseudonym changes, but affects certification protocols instead of message authentication protocols. Additional to the free choice of the pseudonym to use, a vehicle gets to choose which CA it uses. Provided a sufficient number of independent certification authorities are available at the point of certification, a vehicle is able to control which certificates each single CA is able to link to its identity.

Using graph terminology from Section 4.1, a passive attacker with the knowledge of a (single) CA is able to relate only a small part of all messages to a vehicle's identity. In the linkability graph this means that the attacker knows only a subset of relations $V \sim C$. Thus, only part of a vehicle's trajectory is disclosed to any, potentially compromised, CA.

Requested certificates have to be used sensibly by vehicles, as otherwise large parts of a trajectory would be disclosed to a single CA. Take the extremal cases as an example: first, a vehicle uses certificates granted from

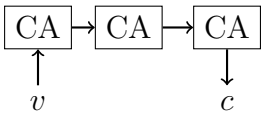


Figure 6.5: Chained CA with three intermediates

different CA for each subsequent message. Secondly, a vehicle for some time uses certificates granted by one authority and afterwards never again. The view of a malicious CA on the first vehicle would be rather complete, with respect to the general directions of that vehicle. It would though be rather coarse and lack precision. For the second vehicle the malicious CA would have a part of the vehicle's trajectory with high precision, but no information on the time before or after that period.

Chained CA

Chained CA is a concept that shall provide enhanced privacy by implementing each certification process as chained communication over intermediate authorities. Similar to mix networks (see [16]) a vehicle v selects a certification path. A certificate is requested either through a series of communication processes with the certification authorities or by an onion protocol without interaction at every CA. This concept is depicted in Figure 6.5.

The general idea is that the vehicle authenticates itself only at the first CA in the chain. At the last CA the certificate is completed and returned to the requester. Each step is protected and authenticated by credentials from the previous communication step. The objective is to prevent any CA from observing more than two neighbours during a certification process. Therefore, an attacker has to compromise all involved CA to map one certificate to the corresponding vehicle identity.

In terms of an linkability graph, an attacker with authority knowledge of a single CA knows only a single step of the certification chain. Provided the protocol is secure this leaves the attacker with no hint as to whether a certificate c is related to vehicle v.

The main problem of this concept lies in the design of the protocol, which involves creation of multiple cryptographic keys for temporary authentication, anonymous network communication to prevent out-of-band identification and various issues concerning the data included in the certificate itself. This is the main reason why we have settled for the much simpler concepts of hierarchical CA, separation of knowledge and quorum-CA in our solution. This concept therefore is not used in this work.

Hierachical CA

A *hierarchical CA* is related to and can be combined with the vehicle-selected CA. Actually, the hierarchy of identity authorities and certification authorities used in the scenario in Chapter 3 already is an instance of this very principle.

The idea is depicted clearly in Figure 6.2. Various layers of CA are introduced. Only the lowest layer actually grants IVC certificates to vehicles, the upper layers are authorisation layers that authorise the signature keys of the lower CA. This is actually the very same as a public-key infrastructure.

In terms of linkability graphs, this scheme works like the vehicle-selected CA from Section 6.2.3. Each CA knows only a subset of vehicle-certificate links. In the hierarchical CA structure, the knowledge of each CA is restricted to a defined spatial area. If no further measures are taken, each CA would have a complete view of the movements within its area, given it can get hold of the vehicular communication.

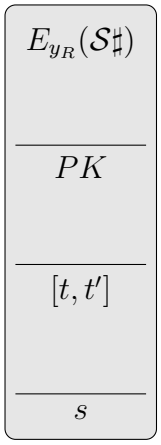
6.3 Certification Protocol

In this section we describe, a certification protocol using partially blinded signatures with secure anonymity revocation, which is suitable for use in vehicular communication scenarios. The protocol allows for implementation of the privacy enhancing principles *separation of duty*, *hierarchical CA*, and *quorum-RA*. Besides the implementation of these principles the design focusses on a large number of parameters that control complexity and security of different parts of the protocol. The idea is to provide a protocol that can be adapted to a wide range of privacy levels.

We combine blind signatures with "cut-and-choose" by Stadler, Piveteau and Camenisch [62] with Shared RSA proposed by Boneh and Franklin [9]. Fair blind signatures provide an approach that can be easily extended to allow for partially blinded signatures and shared key cryptography. The partially blinding technique provides the means for verification of certain attributes of a certificate by the CA. Shared key cryptography implements a Quorum-RA.

Revocation is ensured by forcing the requester to include data that identifies a distinguished certification process into the certificate. Unlinkability of the certificate is ensured by blinding identifying parts during the whole certification process. Authenticity is ensured by encapsulation of the whole communication in an authenticated tunnel, e. g., two-way-authenticated Transport Layer Secutity (TLS).

The authenticated tunnel is not described here. We assume that every communication described in the following is authenticated and encrypted. A



encrypted revocation anchor (readable by RA only)

pseudonym/public key (blinded to CA)

validity time (verified by CA)

signature (created by CA)

Figure 6.6: IVC Certificate Structure

requester has authenticated itself to the CA prior to the protocol below. The whole communication is authentic, integrity-protected and confidential.

An IVC certificate, as depicted in Figure 6.6 consists of an encrypted revocation anchor $\mathcal{S}\sharp$, a public key PK and further attributes $[t, t']$ which can be verified by the CA, as well as a cryptographic signature s over the whole certificate. The revocation anchor is an identity anchor usable by revocation authorities that are able to decrypt it. The public key is used as a pseudonym for IVC. Within the remaining attributes, information like "validity time" or distinct usage information could be encoded. Obviously the signature can be verified using the corresponding public key of the CA which is assumed to be known to all vehicles concerned.

6.3.1 Partially-Blinded Certification Protocol

The following certification protocol is a modification of fair blind signatures with "cut-and-choose" as proposed in [62]. Two modifications are introduced: First, blind signatures are turned into partially blind signatures. Second, the protocol is extended by two security parameters that control the cut-and-choose algorithm and the signature selection. This gives us flexibility to fine-tune the tradeoff between security and complexity. The original protocol does not offer this flexibility which leads to certificate verification that is too costly for use in vehicular communication.

The idea behind partially blinded signatures is that the CA is able to verify some defined certificate parameters, e. g., the validity time interval or the purpose of the certificate, while the vehicle still controls these parameters. The CA has to ensure that the vehicle is not provided with an unlimited certificate and that the revocation anchor is included in the blinded certificate.

The vehicle wants to be assured that the CA is not inserting some identifying data into the certificate. (See Section 4.3 for a discussion of "identifying data".)

The complete protocol is shown in Figure 6.7 with the differences to the Stadler-protocol marked by boxes. The protocol uses parameters K, k, and l for controlling the number of certificate parts to be created, inspected and signed. The requester generates K certificate parts by encrypting and randomising the blinded part of the certificate. He then creates randomised hashes of the randomised revocation anchor. Here *randomising* denotes the concatenation/addition of random numbers to otherwise low-entropy or known data.

The CA then chooses a subset of k certificate parts for inspection. The random values corresponding to these parts are revealed and the CA verifies that the parts contain the revocation anchor and that the unblinded certificate part c_o is valid and included in these parts. This verification technique is called *cut-and-choose*[4].

After successful inspection, the CA signs l of the non-inspected certificate parts. These parts are revealed to the requester which can then calculate and build the signed certificate from this signature. The parameter l linearly influences the complexity of signature verification. The this parameter allows to chose between security of the certification and the complexity of certificate verification.

After a protocol run, the requester holds a valid certificate

$$c = (c_b||c_o, s, \mathcal{T} = \{\alpha_i, v_i : i \in \mathcal{S}\}), \tag{6.1}$$

signed with the CA's private signing key d. Where c_b denotes data blinded to the CA, c_o denoted data disclosed to the CA, s denotes the signature and $\mathcal{T}$ a set of tuples of blinding factors α_i, β_i and the encrypted revocation anchor $\mathcal{S}\sharp$.

The CA itself has gained no knowledge from the process except a *protocol trace* represented by the tuple

$$(R_{ID}, S\sharp, c_o, \{\tilde{m}_i, \mathcal{S}, L, \{(r_i, u_i, \beta_i) : i \notin \mathcal{S}\}, \tilde{s}).$$

Which means the CA knows the vehicles identification R_{ID} from the encapsulating authentication, the revocation anchor $\mathcal{S}\sharp$, the content of the revealed parts $(m_i, i \notin \mathcal{S})$ and the blinded parts $(m_i, i \in \mathcal{S})$ themselves.

The elements $u_i : i \in L$ and the revocation anchor $S\sharp$ are sent securely to the revocation authorities. Values u_i are cipher-texts encrypted with the

[4]Literature analysing cut-and-choose is interestingly sparse, although, it might be considered the standard technique in blind signatures.

Requester		**Cert. Authority**
prior knowledge:		
	K, k security parameters R_{ID} requester ID	
c_b blind data c_o open data		$\mathcal{S}\sharp$ revocation anchor
	$\xleftarrow{S\sharp}$	
for i in $1, \ldots, K$ do:		
$r_i \in_R \mathbb{Z}$		
$\alpha_i, \beta_i \in_R \mathcal{R}$		
$u_i := E_{y_R}(c_b \,\|\|\, \alpha_i)$ $\boxed{\|\|\mathbf{c_o}}$		
$v_i := E_{y_R}(S\sharp \,\|\|\, \beta_i)$		
$\tilde{m}_i := r_i^e \mathcal{H}(u_i \,\|\|\, v_i)\,(mod n)$		
	$\xrightarrow{\{\tilde{m}_i : 1 \le i \le K\}}$	
		select $\mathcal{S} \subseteq \{1, \ldots, K\}$
	$\xleftarrow{\mathcal{S}}$	
check if $\|S\| = K - k$		
	$\xrightarrow{\{(r_i, u_i, \beta_i) : i \notin \mathcal{S}\}}$	
		for $i \notin \mathcal{S}$ verify $\tilde{m}_i \equiv r_i^e \mathcal{H}(u_i \|\| v_i)(mod\ n)$ if all are true: select $\boxed{\mathbf{L} \subseteq \mathcal{S}}, \|L\| = l$ $\tilde{s} := (\prod_{i \in \boxed{\mathbf{L}}} \tilde{m}_i)^d (mod n)$
	$\xleftarrow{\tilde{s}}$	
$s := \tilde{s} / \prod_{i \in \boxed{\mathbf{L}}} r_i$		
signature: $(c_b, c_o, s, \mathcal{T})$ $\mathcal{T} = \{(\alpha_i, v_i) : i \in \boxed{\mathbf{L}}\}$		send to RA: $S\sharp, \{u_i : i \boxed{\in \mathbf{L}}\}, \boxed{\mathbf{R_{ID}}}$

Figure 6.7: "Cut-and-Choose" Fair Partially-Blind Signatures. Adapted from [61]. Changes to the protocol marked by boxes.

public revocation key y_R. Thus, no information on c_b can be gained from the values as long as the used encryption is secure.

A receiver of an inter-vehicle message signed with the private key c^{-1}

corresponding to the certificate c has to verify that the certificate is valid. He does so by verifying that

$$s^e \equiv \prod_{(\alpha,v)\in\mathcal{T}} \mathcal{H}(E^{\alpha}_{y_R}(c_b)||c_o) \ (modn) \tag{6.2}$$

and that the validity date in c_o is good.

The complexity of different protocol parts depends on protocol parameters K, k, and l. Furthermore, security of the protocol depends on these parameters as well. The effects on security and complexity of the described protocol parameters are discussed in detail in Section 5.7.

6.3.2 Quorum-Protected Revocation Protocol

It might be necessary to revoke the certificates associated with a given vehicle. Consider that vehicles might malfunction or could be used for communication-based traffic manipulations (see Chapter 3). In such circumstances it must be possible to revoke the anonymity of a certificate holder to *a)* prevent issuance of new certificates and *b)* to revoke all IVC certificates of that vehicle.

Step *a)* Revocation of anonymity is possible by way of the revocation anchor $S\sharp$ encrypted in v_i in each certificate. This operation is denoted *Type-II Revocation* in [62] and is based solely on knowledge of the private revocation key x_R. For increased privacy protection, it is suggested that x_R is shared and jointly generated by more than one revocation authority (RA) to prevent misuse. As argued in Section 6.1, anonymity revocation is a critical operation susceptible to single points of failure. Revocation is disclosure of information which cannot be reversed. In the following, we will briefly summarise the joined computation with shared keys as described in [9].

Step *b)*, denoted *Type-I Revocation*, is done by way of the stored values $S\sharp, \{u_i\}, R_{ID}$. The certificate is identifiable by way of c_b encrypted in u_i and can be recovered by knowledge of the revocation key x_R. We will not describe this part here because it runs along the same lines as Step *a)*.

Assuming there exists an algorithm for joint decryption D using a shared private key $x_R \equiv \sum_{i=1}^{k} x_i \mod \varphi(N)$, where the shares x_i are distributed over k revocation authorities. Then, the objective of anonymity revocation is to identify the certification process in which a given certificate $(c_b, c_o, s, \mathcal{T})$ was created. The members of a Quorum-RA collaborate in joint asymmetric decryption D_{x_R} on the encrypted revocation anchors v_i in $\mathcal{T}$ of a given certificate.

By deciphering all $v_i = E_{y_R}(S\sharp \,||\, \beta_i)$ in $\mathcal{T}$ the RA gets to know a set

$$r_{II} = \{D_{x_R}(v_i) : (\alpha_i, v_i) \in \mathcal{T}\}\,. \tag{6.3}$$

If certificate is valid values c_b within r_{II} are equal. The RA decrypt v_i by first truncating c_o off from the end of v_i, applying joint decryption and truncating the embedded α, which is summarised above as function D_{x_R}.

The RA-quorum, knowing $(S\sharp, c_o, \{\tilde{m}_i, \mathcal{S}, \{(r_i, u_i, \beta_i) : i \notin \mathcal{S}\}, \tilde{s})$ from the CA, can now find some $D_{x_R}(v_i) = S\sharp||\beta_i$. $S\sharp$ then identifies a certification protocol trace. In collaboration with the signing CA, this reveals the identity of the requester.

In Appendix C protocols for distributed RSA, usable to decrypt $E_{y_R}(c_b||\alpha)$ during a revocation process. It has been introduced by Boneh in [9].

6.4 Security Discussion

In this section, some attacks on the certification protocol described above are discussed. In the following the three classes of attack as defined in Section 6.1.3 are analysed. *Attacks against authorisation* mean that an attacker is able to gain an IVC certificate without being entitled, or to otherwise inject messages in the communication network that are accepted by other participants. *Attacks on privacy* denote any means to gain knowledge of the relation between certificates and vehicles without being authorised. *Circumvention of revocation* describes security of the protocol with respect to preventing an authorised requester from preventing revocation.

6.4.1 Authorisation Attacks

As mentioned before, IVC certificates are needed to participate in inter-vehicular communication insofar as messages are not accepted by vehicles if they are not authenticated by valid credentials. This is necessary to prevent injection of harmful IVC messages. Assuming that the signature algorithm is used correctly, an attacker has to get hold of valid IVC certificates to inject messages.

CA are trusted to verify the authenticity and permissions of a requester. An untrustworthy CA is able to sign any certificate and to pass arbitrary identifications to the RA. Thus a CA, or persons employed at the CA, are able to create certificates that have not been requested.

Furthermore, it is possible for a certification authority to generate certificates and bind them to arbitrary, known vehicle identifiers. A CA fakes a protocol run and sends $(S\sharp, u_i \in L, R_{ID})$ with an arbitrary R_{ID} to the revocation authorities. The CA then is able to take part in the network as if

it is, from the revocation authorities' point-of-view, the requester R_{ID}. This could be used by an untrustworthy CA to falsely accuse arbitrary vehicles.

This problem is basically an information injection problem. A certification authority is able to inject arbitrary information into the RA storage. There are multiple ways to work around this problem: foe example, introduction of a blinded identity statement in the certificate, encrypted again with the revocation key, introduction of peer-control, or inclusion of proofs of freshness of the data sent to the RA. Development of these protocols must be omitted here for the sake of keeping this work focused.

If an unauthorised requester requires a certificate, he needs to show valid credentials to the CA or to gain a valid certificate from another requester. This means the attacker either has to collaborate with another requester or to break the authentication scheme. We assume that the latter is not feasible in practise.

Multiple requesters might cooperate, probably by manipulating the vehicles OBU, to share certificates. Essentially, this would provide *plausible deniability* for malicious inter-vehicle communication. The protocol itself provides no protection against this. Trusted computing technology has already been discussed for use in this context, and the Trusted Platform Model has been proposed for software protection in general [58]. Trusted computing could well be used to prevent untrusted hard- or software to handle certificates. However, standard TPM Hardware may still be too inefficient to be used in vehicular communication [64].

The above attacks lead to unauthorised entities being able to inject messages into the inter-vehicle communication. Without a valid certificate, an attacker has to break the signature scheme used to sign the messages, which is out of scope of this work.

6.4.2 Attacks on Privacy

The main distinguishing feature of our approach is the focus on privacy of certificate-requesters. As described in Section 6.1, single authorities in general are considered potential attackers. Obviously, external entities must be considered as well.

A privacy attack in this context is any method for deriving information on the relation between certificates and vehicles without being authorised to do so. The only authorised way is through revocation authorities (RA).

Unauthorised Anonymity Revocation

Assuming that the encryption mechanism that is used to encrypt revocation anchors is secure, it is sufficient to discuss attacks by certification authorities.

Because, aside from RA, CA have the most information potentially usable for anonymity revocation, thus, in the following, the case of an malicious CA or an attacker with CA knowledge is analysed. If a certification authority is not able to link a given certificate $c = (c_b, c_o, s, \mathcal{T}), \mathcal{T} = \{(a, v)\}$ to an vehicle's identity, it is at least as difficult as for an external attacker.

A malevolent CA, or employee at a CA, might have implemented means to store protocol traces and might gain access to all information disclosed to the CA. A CA knows, for all certificates it signed, a *protocol trace* s consisting of information

$$s = (\{\tilde{m}_i : 1 \leq i \leq K\}, \{(r_i, u_i, \beta_i), i \notin \mathcal{S}\}, c_o, S\sharp, R_{ID}),$$

i. e., all blinded parts of the certificate $\tilde{m}_i$, the unblinded elements from the verification step (r_i, u_i, β_i), the public part of the certificate c_o, the related revocation anchor $S\sharp$ and requester identity R_{ID} known from authentication.

Under the assumption that certificate c has been signed by a CA itself, this CA knows that one of the stored protocol traces matches it. But assuming that blinding is secure, the CA cannot directly determine which trace is related to c. In the following, unlinkability of individual values in the trace is discussed.

Values $\{(r_i, u_i, \beta_i), i \notin \mathcal{S}\}$ are not included in the certificates and thus provide no information on relations between certificates. The vehicle's identity R_{ID} is, as well, not included in the certificate. Information from $\{\tilde{m}_i : 1 \leq i \leq K\}$ has been incorporated in the certificate, but assuming that r_i has been selected uniformly at random, no relations can be drawn between $\tilde{m}_i$ and c.

The revocation anchor $S\sharp$ is encrypted in each v_i from c and is included in the protocol trace. The malicious certification authority may either break the encryption of v_i, or the CA may try to encode each $S\sharp$ from its stored traces using the known encryption key y_R and compare the result with each v_i in the certificate. This would involve correctly guessing the value of β_i which is concatenated to $S\sharp$ before encryption. Assuming that the used encryption scheme is secure, the minimal complexity is that of guessing β_i or guessing x_R. Note that a CA is able to run a known plaintext attack against x_R because some β_i are revealed during each certification.

Concluding that the attacks described above are not feasible with CA-knowledge, this leaves the knowledge of the open,i. e., not blinded during certification, certificate part c_o, both available to the CA in stored traces and the certificate.

The *anonymity set* A_c of an certificate c with respect to a set of certification protocol traces S consists of all possible requester identities[5] contained

[5]Obviously there is a bijection between S and all granted certificates.

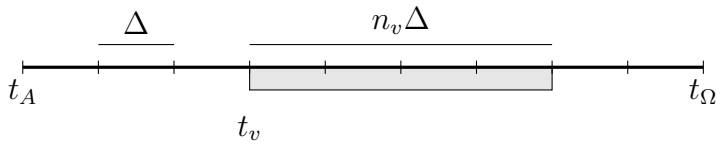


Figure 6.8: Validity Time Granularity/Encoding Scheme

in a trace $s \in S$ where the open data in that trace $c'_o \in s$ equals the open data of the certificate $c_o \in c$,

$$A_c(S) := \{R_{ID} : c'_o \in s \in S, c_o \in c, c_o = c'_o\}.$$

Cardinality of $A_c(S)$ depends on the number of traces that contain values of c'_o that are equal to c_o. For any given certificate c, the anonymity set size represents an measure of the anonymity[6] of the requester using this certificate. The distribution of the anonymity-set-size of all certificates from each requester, provides an expectation of anonymity of the system. This distribution is commanded by the distribution of open data values in the requested certificates.

As the requester chooses the actual value for c_o, the CA cannot influence this distribution directly, e. g., for marking a certain requester by assigning distinguishable values. Assuming that that requesters have no interest in being distinguishable, their choice for c_o would be as uniform as possible.

Assuming for simplicity reasons that the number of needed certificates is uniformly distributed over time and equal for all requesters, then the only parameter that controls $|A_c|$ is the *granularity* of the data format, i. e., the number of different values that can be assigned within a certain period of time.

Assuming for simplicity, that c_o contains only the validity dates for the certificate. Other contents are excluded here, but influence the granularity in a similar way. The following data model as depicted in Figure 6.8 can be used. Under the assumption that c_o is represented by a fixed length data field, the encoding restricts the notion of time to a given epoch starting with t_A and ending with t_Ω. Validity time is thus represented at an interval starting at t_v with length $n_v\Delta$, with Δ being the length of a single time unit.

Denoting the number of time-points, i. e., possible values for t_v, as $N = (t_\Omega - t_A)/\Delta$, the number of all possible time intervals is the sum of all numbers from one to N, i. e., $(N(N+1))/2$, which is in $O(N^2)$.

For an given cardinality of S, the expected size of the anonymity set is directly related to the number of possible time intervals. From a privacy

[6]See Section 5.4.1 for different anonymity measures.

point of view, it is thus useful to reduce the number of possible time intervals. This may be done by reducing the space of either t_v or n_v. In the vehicular scenario, it might be sensible to allow for only one possible interval length n_v because IVC certificates are all used in the same manner.

It is rather obvious that the identity of an vehicle is compromised if it is the only requester of certificates with certain distinguishable c_o. The optimum (global) anonymity is provided by a uniform distribution, similar to the PROB-channel optimum distribution of delay time developed in [66].

The ability of a CA to match an observed certificate is equal, assuming that the cryptographic concepts used are sound, to the ability of an CA to single out a distinguishable open certificate value c_o. As the requester produces this date, the CA can only indirectly — by way of denying certification — manipulate the distribution of c_o values. The other remaining problems are similar to problems known from mix-networks [18] with the CA providing a single-node mix-network.

6.4.3 Circumvent Revocation

A malicious requester may try to inject an $S\sharp$ into the certificate that is different from the one provided by the CA. His objective is to obtain a certificate that cannot be linked to the certification process. In this section we discuss the effects of protocol parameters K, k, and l on the success probability of such an attacker.

In the original cut-and-choose protocol [61] by Stadler, only one protocol parameter is used. The Stadler-scheme is equal to our protocol with parameters $K = 2k$ and $l = K - k$. The security parameters K and k control the probability with which a cheating requester is detected in the cut-and-choose inspection phase. In the original Stadler scheme, this probability grows exponentially with $\binom{2k}{k}^{-1} \approx 2^{-2k}\sqrt{\pi k}$.

In the protocol variant described above, the third parameter l controls how many certificate parts are selected into the signature. Parameter l is a lower bound for n_a,the number of certification parts an attacker manipulates in order to gain a non-revocable certificate. In Figure 6.9, the involved parameters are depicted on a scale from zero to K. One may note that there is a gap between l and $K - k$ which is area of indecision where certification parts are neither inspected nor included in the certificate. In the following, it will also be shown that the optimum for prevention of requester-cheating is to reduce this gap to zero.

A cheating requester — the attacker — provides blinded certificate parts $\tilde{m_1}, \ldots, \tilde{m_k}, \tilde{m_{k+1}}, \ldots, \tilde{m_K}$. Instead of signing all $\tilde{m_i}$ with $i \notin \mathcal{S}$, the CA chooses only $l \leq |\mathcal{S}| = K - k$ elements and signs only them. To create an unrevocable certificate, the attacker has correctly guess exactly those values

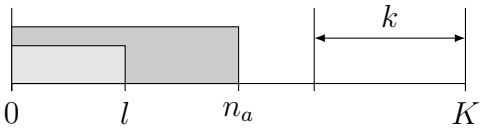


Figure 6.9: Protocol parameters K, k, l, and number of faked certification parts n_a. (Ordered, for simplicity reasons, on the interval $[0, K]$.)

$\mathcal{S}\sharp$ that are included in this signature. All values $\mathcal{S}\sharp$ that are checked by the CA have to be unmodified. He has to guess correctly, and chooses values in $n_a : l \leq n_a \leq K - l$ certificate parts $\tilde{m}_i$ that he modifies.

The protocol parameters in Figure 6.9 lead directly to the definition of possible outcomes of a protocol run. The result of a requester-cheating attempt may either be detection of the attempt (interval $[K-k, K]$), successful reception of a faked certificate (interval $[0, l]$) or reception of a broken certificate (interval $[0, n_a]$, excluding $[0, l]$). These three outcomes are produced by the two distinguished phases, detection and signature, of the protocol.

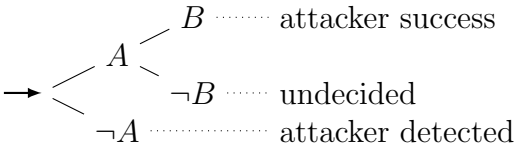


Figure 6.10: Event tree for a protocol run with cheating requester.

In Figure 6.10, the protocol's outcome is modelled as stochastic, two-phase experiment. Event A denotes success of an attacker in the first phase, where a cheating requester is not detected in the k out of K "Cut-and-Choose"-scheme. B denotes the event where an attacker guessed $n \geq l$ certificate parts correctly, i. e., that the received certificate contains only faked parts. From Figure 6.10, one can see that B or $\neg B$ can only occur if A has happened, because in the case of detection — event $\neg A$ — the protocol is stopped and no certificate is issued.

Event $A \cap B$ describes that a cheating requester has not been detected (Event A) and has received a certificate that consists solely of faked $\tilde{m}_i$ (Event B). Event $A \cap \neg B$ describes that a cheating requester has not been detected but has received a certificate where faked and non-faked $\tilde{m}_i$ have been included, which renders the certificate useless.

The probability of an attacker to successfully receive a non-revocable certificate is the product of the independent probabilities of not being detected by the cut-and-choose detection and having the CA choose only modified $\tilde{m}_i$ into the certificate.

The original protocol in [61] analyses the success probability of an attacker for the case $k = \lceil K/2 \rceil$. This case provides the optimum detection rate with $l = K - k$. In the following, the general case of free parameter choice is analysed to support complexity trade-off.

Cheating Probabilities

The probability of a cheating requester being detected is denoted $P(A)$. The probability of a successfully faked certificate, conditioned on A, denoted $P(B|A)$ can be modelled by a hyper-geometric probability distribution h[7],

$$P(A) = h(0|K, n, k) \tag{6.4}$$
$$P(B|A) = h(l|K - k, n, l) \tag{6.5}$$

Event B is conditioned on A because a cheating requester first has to avoid detection before the certificate parts for the signature are tried.

The event that a cheating requester receives a completely faked certificate, without being previously detected, in a single protocol run is denoted *p.succ* and given by $P(A \cap B)$ By applying the definition of conditional probability and Equations (6.4) and (6.5), we have

$$\begin{aligned} p.succ(K, k, n, l) :=& P(A \cap B) \\ =& P(B|A)P(A) \\ =& \frac{\binom{n}{l}\binom{K-k-n}{l-l}}{\binom{K-k}{l}} \cdot \frac{\binom{n}{0}\binom{K-n}{k-0}}{\binom{K}{k}} \\ =& \frac{n!(K-n)!(K-k-l)!}{K!(n-l)!(K-k-n)!}. \end{aligned} \tag{6.6}$$

Best-n_a Attack Strategy

Remember that the value of n_a is determined by the attacker. Values of K, k, and l are given by the system, i. e., the certification and revocation authorities. To reflect this, it is assumed that the attacker follows an optimum strategy. The optimum choice for n_a, considering *p.succ*, is denoted $n_a.max$ and — for given K, k and l — provided by

$$\begin{aligned} n_a.max(K, k, l) := n'_a \quad & \text{with } p.succ(K, k, n'_a, l) \geq p.succ(K, k, n_a, l), \\ & \text{for all } n_a : K - k \geq n_a \geq l. \end{aligned} \tag{6.7}$$

[7] With $h(k|N, M, n) = \frac{\binom{M}{k}\binom{N-M}{n-k}}{\binom{N}{n}}$, where k is the number of successes, N is the population size, M is the number of element with the targeted property, and n is the sample size.

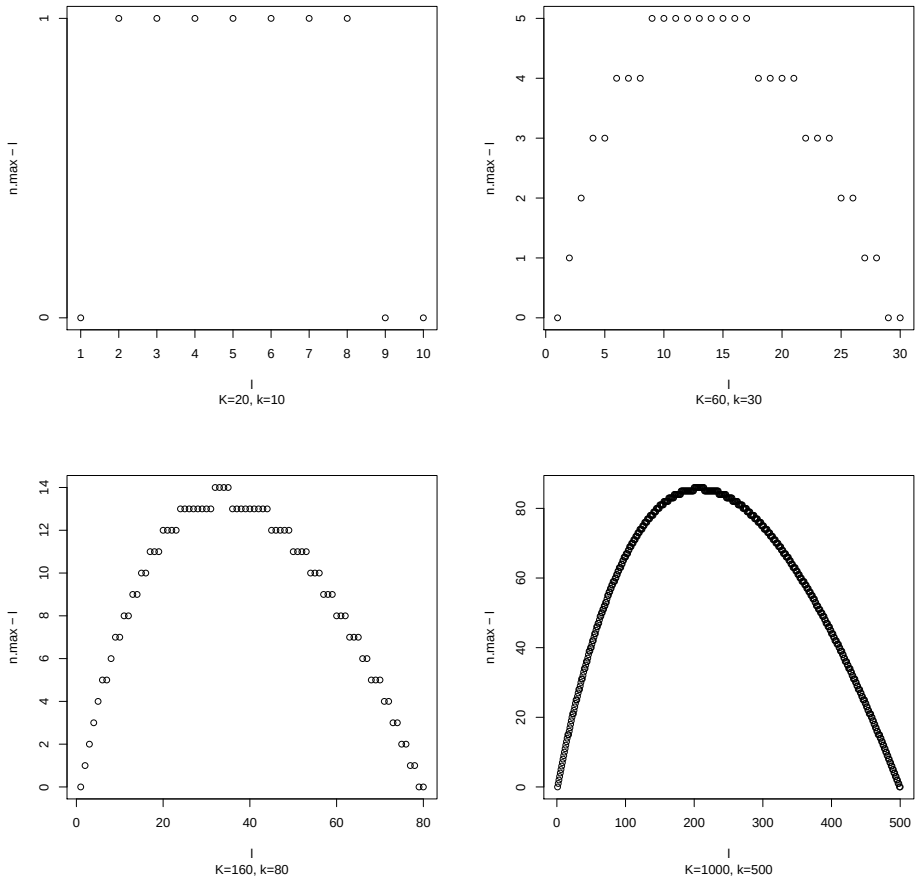


Figure 6.11: Number of surplus of modified $\mathcal{S}\sharp$: $n_a.max(K,k,l) - l$ for the optimum choice $n_a = n_a.max(K,k,l)$ given $K \in \{20, 60, 1600, 1000\}, k = K/2$.

In Figure 6.11, the difference between l and $n_a.max$ for $K = 80, k = K/2$. The shape becomes more precise for larger values of K as the number of possible values for l grows. Interestingly the shape is not symmetric to any parallel of the y-axis but resembles a slightly skewed inverse parabola.

This shape is the result of two opposed incentives. As the probability for a valid certificate (Event B) increases with a larger choice for n_a an attacker is inclined to maximise n_a. This on the other hand increases the probability of detection, which provides an incentive for reducing n_a. In the following, it is assumed that an attacker optimises n_a for the maximum probability for success in a single certification attempt and chooses $n_a = n_a.max$, unless explicitly noted otherwise.

Attacker Success/Detection Probability

We refer to one single run of the certification protocol as *round*. An attacker, who is neither detected, nor successful in a round (event in $A \cap \not B$) could try as many rounds, until he is detected (event in $\not A$) or receives a valid certificate with modified revocation anchors (event in $A \cap B$). Assume that one success is sufficient for the attacker. Likewise one detection is sufficient to stop further attacks. The question at hand is; what is the probability that an attacker has at least one successful round before he is detected? And what is the probability of an attacker being detected (at least once) before he is successful?

The probability of an attacker having exactly one success in the i-th round, denoted $p.si$, is modelled by a geometric probability distribution and calculated as follows:

$$\begin{aligned} p.si(i, K, k, n_a, l) &= (P(\neg B \cap A))^{i-1} P(B \cap A) \\ &= (h(0|K, n_a, k)(1 - h(l|k, n_a, l)))^{i-1} h(0|K, n_a, k) h(l|k, n_a, l) \\ &= h(0|K, n_a, k)^i (1 - h(l|k, n_a, l))^{i-1} h(l|k, n_a, l). \end{aligned} \tag{6.8}$$

The other case, where a cheating requester is detected in the i-th round without previously being successful, is denoted by the probability mass function $p.di$ defined as:

$$\begin{aligned} p.di(i, K, k, n, l) &= P(\neg A) \left(P(\neg B \cap A)\right)^{i-1} \\ &= P(\neg A) \ \left(P(\neg B|A) P(A)\right)^{i-1} \\ &= (1 - h(0|K, n, k)) \ \left((1 - h(l|k, n, l)) h(0|K, n, k)\right)^{i-1}. \end{aligned} \tag{6.9}$$

The probability of a cheating requester to be successful before he is detected up to the i-th round is the cumulative geometric probability distribution $P.s$. The opposite, the probability to be detected before success up to the i-round is given by $P.d$. The probability that neither event has occurred, the indecisive event, is given by the inverse probability of the sum of $P.s$ and $P.d$.

$$P.s(X \leq i) = \sum_{j=1}^{i} p.si(j, K, k, n, l) \tag{6.10}$$

$$P.d(X \leq i) = \sum_{j=1}^{i} p.di(j, K, k, n, l) \tag{6.11}$$

$$P.o(X \leq i) = 1 - (P.det + P.succ) \tag{6.12}$$

This finally leads to the question that is crucial for optimisation of the certificate validation time: How does l influence the success probability and risk of an attacker? For what values of K, k and l does sufficient security follow for the application? As $P.s$, $P.d$, and $P.o$ are monotonous, the answer is the bound of $P.s$, $P.d$, and $P.o$ for $i \to \infty$.

Because we assume that in the undecided case $A \cup \not{B}$ the attacker tries a new round, and because of monotony, the bounds can be interpreted as the final probability of success, respective detection, for the cut-and-choose scheme. With probability $\lim_{i\to\infty} P.s$ the cheating requester is successful without being previously detected. With probability $\lim_{i\to\infty} P.d$ a CA detects a cheating requester before he is successful.

$$\begin{aligned}
\lim_{i\to\infty} P.s &= \lim_{i\to\infty} \sum_{j=1}^{i} p.si \\
&= \lim_{i\to\infty} \sum_{j=1}^{i} P(B \cap A)(P(\neg B \cap A))^{j-1} \\
&= \frac{P(B \cap A)}{1 - P(\neg B \cap A)} \\
&= \frac{P(B|A)P(A)}{1 - P(A) + P(B|A)P(A)}
\end{aligned}$$

The limit $\lim_{i\to\infty} P.s$ provides us with the information about the overall attacker success probability that is needed to calculate the trade-off between complexity and security in choosing K, k and n.

$$\begin{aligned}
\lim_{i\to\infty} P.d &= \lim_{i\to\infty} \sum_{j=1}^{i} p.di \\
&= \lim_{i\to\infty} \sum_{j=1}^{i} P(\neg A)(P((\neg B) \cap A))^{j-1} \\
&= \frac{P(\neg A)}{1 - P(\neg B \cap A)} \\
&= \frac{1 - P(A)}{1 - P(A) + P(B|A)P(A)}
\end{aligned} \tag{6.13}$$

The limit $\lim_{i\to\infty} P.d$ provides the overall probability for the risk an attacker runs of being detected before gaining a faked certificate. As before, this allows for the protocol to be tuned to a given scenario.

One very interesting case is the probability for an attacker never being detected and never gaining a valid certificate. This is described by the limit $\lim_{i\to\infty} P.0$.

$$\begin{aligned}\lim_{i\to\infty} P.o &= \lim_{i\to\infty} 1-(P.d+P.s)\\ &= 1-\frac{P(\neg A)+P(B\cap A)}{1-P(\neg B\cap A)}\\ &= 1-\frac{1-P(A)+P(A)P(B|A)}{1-P(A)+P(A)P(B|A)}\\ &= 0\end{aligned} \tag{6.14}$$

Summarising, we define functions defined on the domains of K, k, n_a, and l as the limits of $P.s, P.d$, and $P.o$ w. r. t. $i \to \infty$ as follows:

$$P.succ(K,k,n_a,l) := \lim_{i\to\infty} P.s(X \le i) \tag{6.15}$$

$$= \frac{P(B|A)P(A)}{1-P(A)+P(B|A)P(A)} \tag{6.16}$$

$$P.det(K,k,n_a,l) := \lim_{i\to\infty} P.d(X \le i) \tag{6.17}$$

$$= \frac{1-P(A)}{1-P(A)+P(B|A)P(A)} \tag{6.18}$$

$$P.0(K,k,n_a,l) := \lim_{i\to\infty} P.o(X \le i) \tag{6.19}$$

$$= 0. \tag{6.20}$$

The limit above shows that the probability of an attacker being neither detected nor successful converges to zero. The success or detection probabilities depend solely on the choice of K, k, n, and l which now can be balanced to the complexity of certification, verification and security against certificates with faked revocation anchors.

Analysis of Success Probability

The bounds for $i \to \infty$ provide the means for finding optimal values for K, k, and l to balance security and complexity. The value for n_a is determined by the attacker, thus in the following the optimum attacker $n_a.max$ is chosen.

In [61] $k = K/2$, $l = k$ has been chosen, with the recommendation to set $K \ge 20$ to achieve sufficient security. These parameters are not usable in the vehicular scenario described in Chapter 3, mostly because message verification has to be very fast. Thus, the objective here is to reduce l while keeping a sufficient level of security against cheating requesters.

What is a *sufficient* level of security to prevent manipulation of revocation anchors? Assuming that communication with a CA is considered authenticated and integrity protected, false revocation anchors can only be explained by very serious defects in critical components of the vehicle's computer or active manipulation by the owner. Thus, it can be assumed by the authorities, that manipulations are most probably made with the intention of avoiding prosecution. Faked revocation anchors are a serious offence and we assume that authorities take action upon detection. Detection of a cheating requester in the vehicular scenario should have at least the effect, that the requester's vehicle is denied any certificates in the future.

Keeping in mind the serious consequences of detected cheating in a certificate request, we assume that a cheating requester has completely lost the game once he has been detected. Evading prosecution would require sufficient measures for hiding the owner's identity during registration of the vehicle, with every vehicle involved in the production of a single non-revocable certificate. Considering the physical components in the cheating procedure, we assume that attacker success probabilities below 10^{-5} provide sufficient security levels. The expectation for the attacker to receive his first usable certificate is after 10^5 tries. This means the attacker must also provide the resources for that many vehicles and identities, while facing the risk of being caught[8].

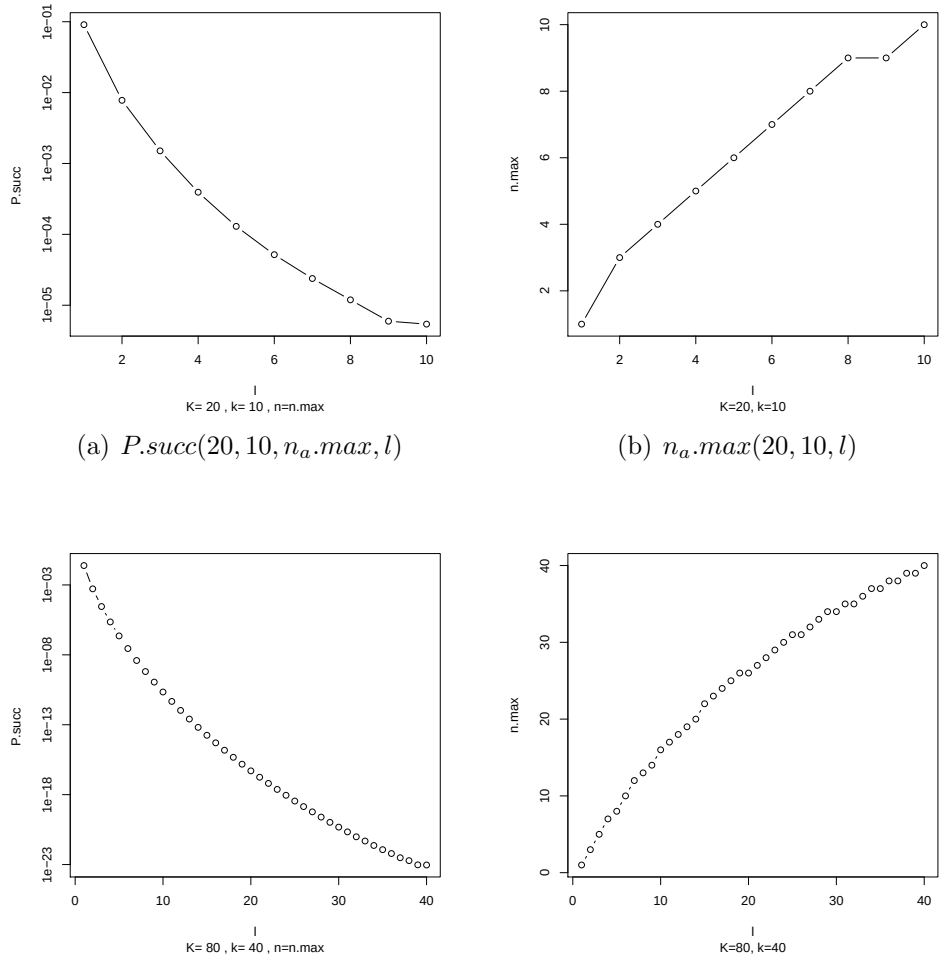
Regarding above assumptions, we first examine $k = K/2$-schemes with the intention of finding a minimum l that is secure. In Figur 6.12 success probabilities and attack strategies as functions of l are plotted for two settings of K. In Figure 6.13(a) and Figure 6.13(c) graphs of success probabilities *P.succ* are shown for (a) $K = 20$ and (c) $K = 80$. The graph in (a) reaches 10^{-5} only for maximum values of l. The graph in (c) is at probability 2.2610^{-06} for $l = 4$. But considering the transfer of 80 certificate parts for each certificate, this solution seems unusable.

The curves in Figure 6.13(a) and Figure 6.13(c) are rather smooth and strictly monotonously falling.

Success probabilities plotted as function of k are not similarly monotonous. Observe Figure 6.13(a) parallel to Figure 6.13(b) and Figure 6.13(c) parallel to Figure 6.13(d). One may recognise that the "jumps" in the success-probability-graphs correspond to jumps in the choice of $n_a.max$ plotted over k.

If plottet over k, *P.succ* behaves non-monotonous as depicted in Figure 6.13(a) and Figure 6.13(c). As *P.succ* is, for constant n_a monotonous,

[8]It can further be assumed that authorities will realise much further that someone is systematically gaming the system and actively begin to search for the attacker, increasing the difficulty of escaping prosecution. A the discussion about the required security level is beyond the scope of this work, we stop at this point.



(a) $P.succ(20, 10, n_a.max, l)$ (b) $n_a.max(20, 10, l)$

(c) $P.succ(80, 40, n_a.max, l)$ (d) $n_a.max(80, 40, l)$

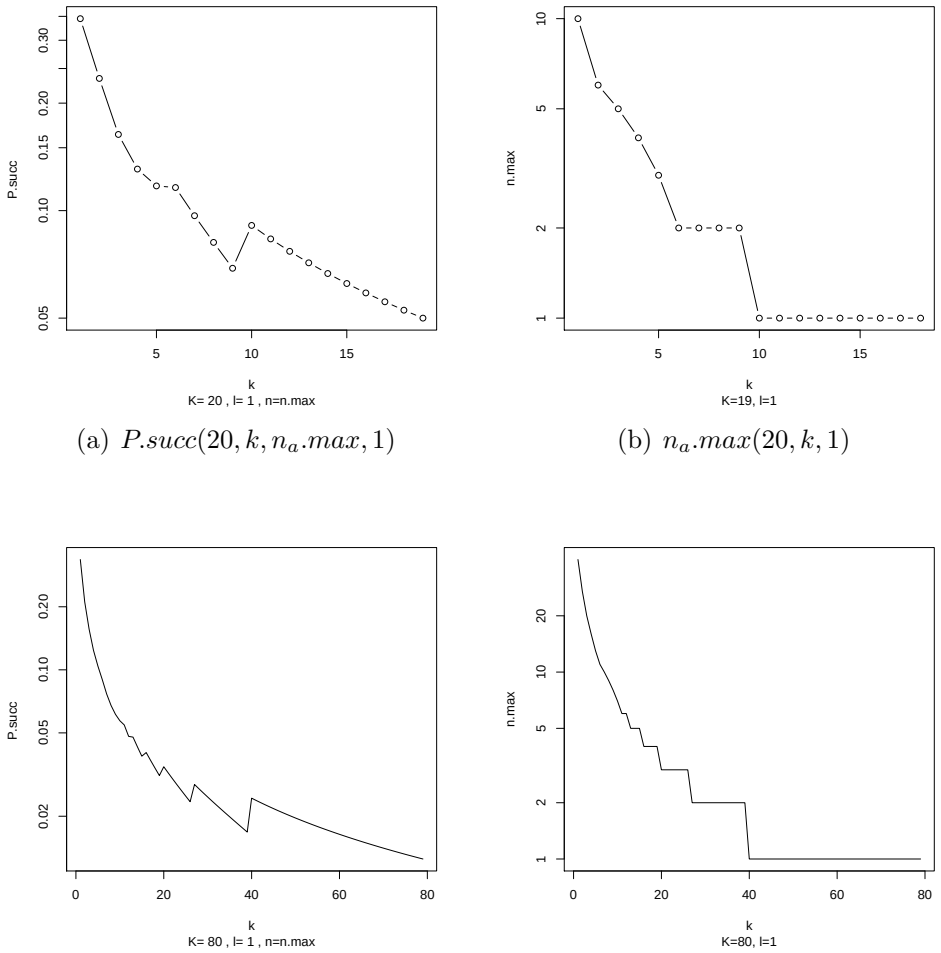
Figure 6.12: Logarithmic plots of $P.succ(K, k, n_a, l)$ with constant $k = K/2$.

the local spikes in $P.succ$ are caused by changes in $n_a.max$.

Having considered the behaviour of $P.succ$ over k and l individually, we now take a look at the whole picture to show the minimum of the attacker's success probability over the two main security parameters k and l. In Figure 6.14, the cheating requester's success probability is plotted over k and l for $K = 20$.

Figure 6.14 shows the lowest attacker success probability for parameters k and l is along the line $k + l = K$. The minimum obviously is found in the middle of that line at $l = k$, which coincides with the choice of values in the protocol scheme by Stadler.

But the objective here is not to find the global maximum, but to balance



(a) $P.succ(20, k, n_a.max, 1)$ (b) $n_a.max(20, k, 1)$

(c) $P.succ(80, k, n_a.max, 1)$ (d) $n_a.max(80, k, 1)$

Figure 6.13: Logarithmic plot of $P.succ$ with $l = 1, K \in \{20, 80\}, n = n - max$. The sawtooth-shape of $P.succ$ is correlated to $n_a.max$ over k.

security and complexity of certificate verification. The next step towards this goal is finding the optimum by considering only values of k and l with $k + l = K$. In Figure 6.15, graphs for $K = 20$ and $K = 80$ are depicted. For $K = 20$ it is possible to reach 10^{-5} with approximately $5 \leq k \leq 15$.

Considering that choices of k and l are best if $k + l = K$ and that complexity constraints at the verification we require a low value for l. Although the certification process is not as constraint as the verification process, we also want to achieve a reasonable low value for K, while meeting the security requirement in terms of low $P.succ$.

In Figure 6.16, graphs for $l \in \{1, \ldots, 5\}$ are plotted over $K \in \{10, \ldots, 180\}$.

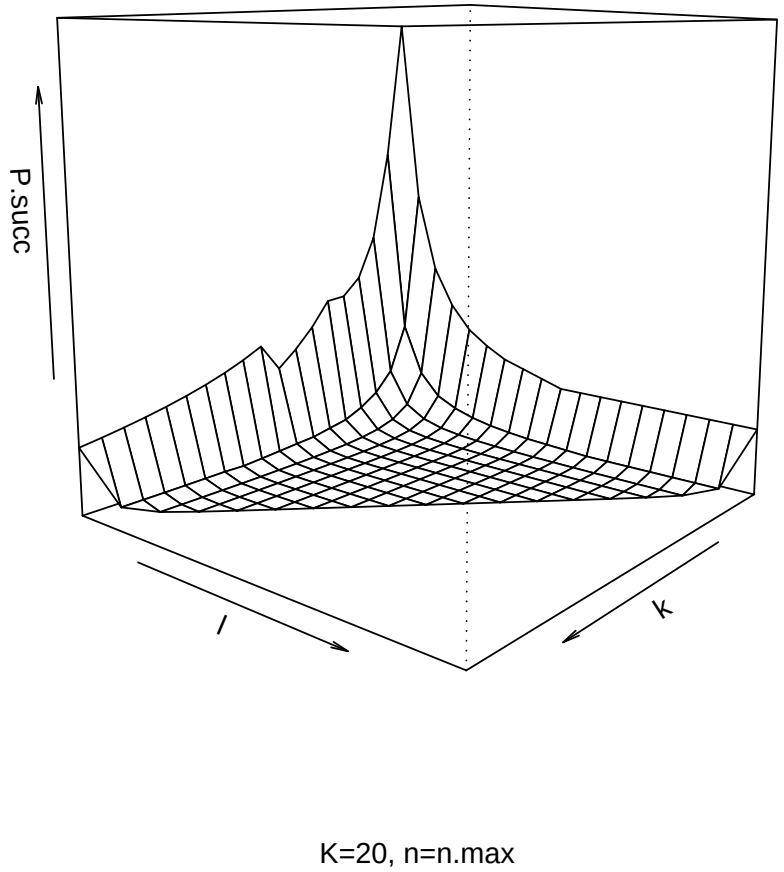
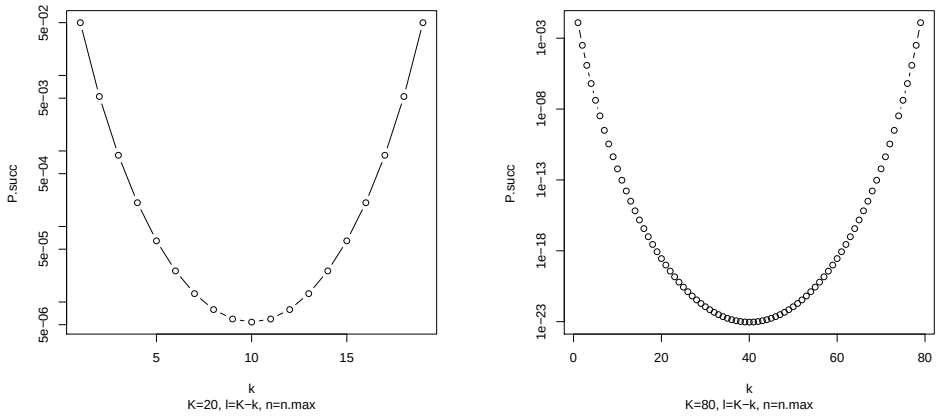


Figure 6.14: *P.succ* plotted over l and k. ($1 \leq l, k \leq 19$, $n = n_a.max(K, k, l)$)

For complexity reasons, solutions have to be found within this space if they exist at all. One may observe that the curves for $l \in \{1, 2\}$ do not fall below 10^{-5} in this domain. For $l = 3$, sufficient security is reached with $K \geq 90$, for $l = 4$ it is $K \geq 45$ and for $l = 5$ it is $K \geq 30$.

Finally, a conclusion is reached as we can now examine the minimum K (w. r. t. to l) and minimum l (w. r. t. to K) that ensure a defined security level (aka. attacker success probability), given that $l + k = K$. In Figure 6.17, we plot minimum values for parameters l and K that provide $P.succ \leq 10^{-5}$. In Figure 6.17(a) minimum values of l for $P.succ \leq 10^{-5}$ are plotted over $K \in \{2, \ldots, 180\}$. Similarly in Figure 6.17(b) minimum values of K for $P.succ \leq 10^{-5}$ are plotted over $l \in \{1, \ldots, 179\}$. The obvious sweet



(a) $P.succ(20, k, n_a.max, l), l = K - l$ (b) $P.succ(80, k, n_a.max, l)$

Figure 6.15: $P.succ$ for $l = K - k, n = n_a.max$.

spot is found in the right graph at the bottom of the plotted wedge where both K and l have relatively small values.

In this section, a thorough analysis of the cut-and-choose algorithm has been provided, with emphasis on unrestricted parameter choice. Obviously, a stochastic method like cut-and-choose will always leave a chance for error, especially as resources are restricted. However, considering that both defender and attacker have restricted resources, it is essential to weight the odds. In the vehicular scenario, for example, we may assume that any detected cheating might trigger an examination. A considerable chance of being caught before receiving even one certificate thus might deter sophisticated attackers as well.

6.5 Ressource Consumption

In vehicular communication, time is essential. A traffic-warning message is of little use if it is slower than nowadays break-lights. Requirements for time and overhead are strict. Vehicles' have only very limited resources for computation. In combination, this makes the computational complexity and memory overhead of communication protocols a crucial hindrance on the way to implementation of vehicular communication.

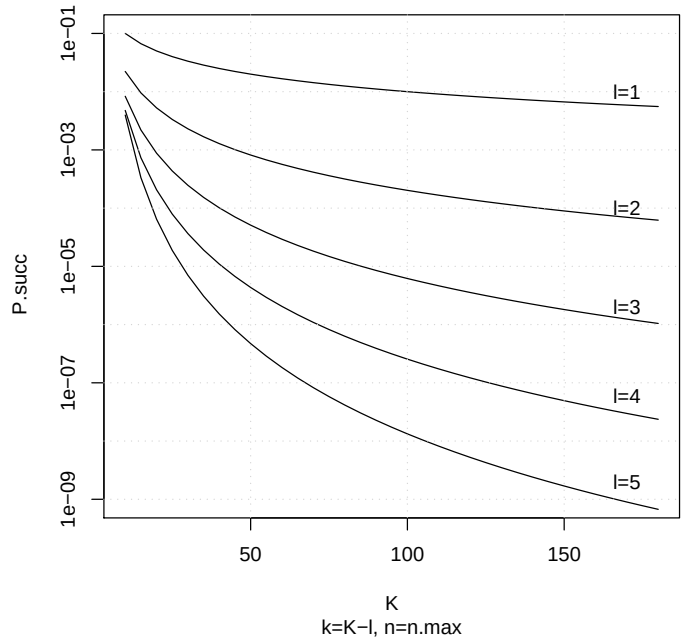


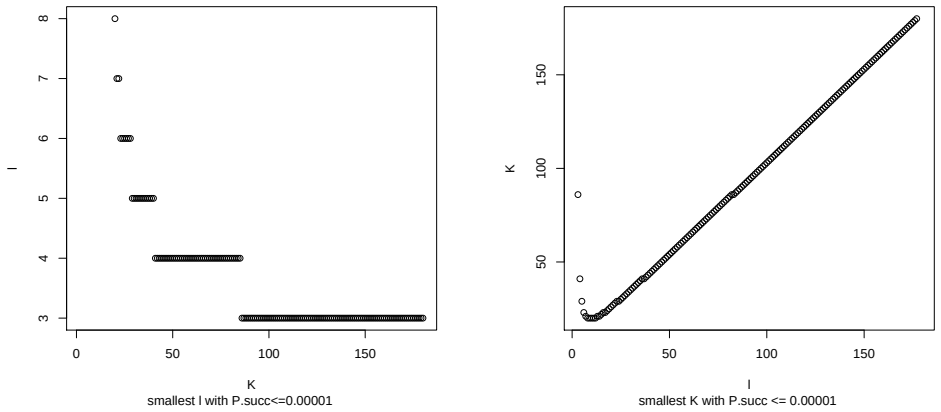
Figure 6.16: *P.succ* over K for $l \in \{1, \ldots, 5\}, k = K - l$.

6.5.1 Inter-Vehicle Communication Complexity

Reaction time in vehicular communication is a crucial parameter. We noted in Section 6.1.4 that the allowable latency, depending on the application, has to be below 100 milliseconds to 1 second. This time includes processing and marshaling of sensor data, sender side security measures (i. e., appending signature and credentials) as well as the receiver side un-marshaling, verification of security measures and interpretation. This section provides a rough analysis of the time needed to run the protocol from Section 6.3.

At the sending vehicle, the communication process consists of message generation and message-signing with the private key c_b^{-1}. (For simplicity, it is assumed that the blinded part of the certificate is equivalent to a public key.) Furthermore, marshaling of the IVC message and data adds further delays to emission. The message is then transferred, together with a certificate. A receiver has to verify the signature of the certificate and the signature on the message with the key c_b contained in the certificate.

The computational complexity of certificate verification is linearly dependent on l. The verification of a certificate consists of l hash and encryption operations using the encryption function and keys $E^a_{y_R}$. These values have to be combined in $l-1$ multiplications in $\mathbb{Z}/n\mathbb{Z}$, as given in Equation (6.2). Ver-



(a) smallest K for $10^{-5} \geq P.succ(K, k = K - l, n_a.max, l)$

(b) smallest K for $10^{-5} \geq P.succ(K, k = K - l, n_a.max, l)$

Figure 6.17: Smallest l, respectively K, for $P.succ \leq 10^{-5}$. We considered only K with $2 \leq K \leq 180$ and $l \leq K$.

ification of the message's signature usually requires calculation of one hash and one signature operation, depending on the specific hash and signature algorithms used.

To consider feasibility of our scheme in reality, we examine the timing of encryption algorithms. In the protocol, RSA is used for encryption of revocation information. In Table 6.1, timings of 1024-bit RSA, taken from [11], are shown. The implementation was executed on a Pentium II processor, and timings are for handling of a single data block. Assuming that one message contains not more than 100 bytes of payload [14], a message would completely fit into one data block. 1024-bit, depending on the algorithm used for IVC messages signatures, might be sufficient for the public key stored in c_b.

The modular nature of the protocol proposed in this chapter would allow easy exchange of the encryption algorithm used for revocation anchors and blinded certificate parts.

	key generation	encrypt	decrypt	sign	verify
RSA 1024-bit	2,740.87	5.34	67.32	66.56	1.23

Table 6.1: 1024-bit modulus RSA timings on a Pentium II in Milliseconds [11].

Table 6.1 shows that using RSA for encryption of c_b fulfils well the objective for strict time constraints during communication. During normal

operation, encryption function $E^a_{y_R}$ is often used, while the corresponding decryption function $D^a_{x_R}$ is only used for revocation. Encryption in RSA is much cheaper than decryption (factor ≈ 13). In Section 6.4.3 it has been shown that even small values for l provide sufficient security. Depending on the hashing algorithms, the proposed protocol seems usable for IVC communication.

Following [54] and [10], calculation of a message's or certificates hash is fast enough to be unimportant for these text sizes of ≈ 100 Byte (e. g., there are implementations that provide RIPEMD with throughput of 82.7 Mbit/s).

For values of $l = 2$ the overall verification takes roughly ≤ 15 milliseconds, which is reasonably acceptable, if aiming for an overall latency around 100 milliseconds.

6.5.2 Certification Complexity

During a single communication session between CA and requester, an arbitrary number of certificates may be requested — in theory. Depending on the "refill"-strategy, certificates with either short or longer validity time nearer or farther in the future may be requested[9]. In [43], two types of *pseudonym refill strategies* are defined. In Strategy 1 a vehicle requests as many certificates as possible at one point in time. Contacts to the CA then are rare, but many certificates have to be stored and the vehicle thus is authorised for a longer period of time. In Strategy 2 certificates are only granted with a validity time in the near future. Communication with the CA is more frequently but the vehicle needs less storage, and revocation lists (if needed) may be much smaller.

We distinguish between computational complexity at the certification authorities and at the side of requesters. In vehicular scenarios, a requester obviously has lower computational resources than a CA. The computational resources of the CA are easily improved by further division of the space of responsibility, by assigning multiple CA to highly frequented regions, or by enlarging bandwidth and servers at the CA.

For maximum unlinkability of messages, each message has to be signed with a new certficate. Although certificates can be requested in advance, the high number of certificates needed, and the complexity of potential revocation of a high number of certificates, most likely does not permit a supply of certificates sufficient for the whole expected lifetime of a vehicle. (Assuming 1024-bit certificates used only once, one message per second[10] and ten years

[9] Our protocol does not restrict the validity time. Discussion of the effects of different policies is out of scope of this work.

[10] Actually position beacons might be send at higher frequency.

lifespan ≈ 37 gigabytes memory would be needed.) The need for revocation is more critical with a large supply of certificates. Providing a vehicle with a lifetime supply implies huge revocation lists that must be kept (and used) for a similarly long time. Thus, a constant stream of fresh certificates for every vehicle would be needed.

Each certificate used has to be required beforehand. Denoting change-frequency of certificates by f_c, the time t_r needed for a single certification request defines an upper bound $f_c \leq f_r = 1/t_r$ to f_c. t_r is governed by the computational strength of the CA and requester, as well as the communication bandwidth between them.

Aside from other tasks, e. g., using the requested certificates, t_r is determined by the following numbers.

For every new certificate, a requester has to calculate

$3 \cdot 2K$	random numbers,
$2 \cdot 2K$	RSA encryptions $E^a_{y_R}$,
$2K$	hashes,
$2K - 1$	multiplications in $\mathbb{Z}/n\mathbb{Z}$,
$l - 1$	multiplications in $\mathbb{Z}/n\mathbb{Z}$.

This amounts to $4K$ encryptions $+6K$ random $+(2K + l - 2)$ multiplications $+2K$ hashes.

For every new certificate, a CA has to compute:

1	random $S\sharp$,
k	verifications = encryptions, hashes and multiplications in $\mathbb{Z}/n\mathbb{Z}$,
$l - 1$	multiplications in $\mathbb{Z}/n\mathbb{Z}$,
1	exponentiation in $\mathbb{Z}/n\mathbb{Z}$.

This amounts to $K \cdot$(encryption + hash + multiplication) + exponentiation + random $+(l - 1)$ multiplication.

Omitting the single exponentiation, the multiplications, and random number generation for the sake of simplicity, the determining parameter is K. The workload for a requester is about a factor 4 higher than that of a CA. Computations of both parties have to run sequentially, rendering $t_r \approx 5K \times$ encryption.

Thus, taking the recommendations from Section 6.4.3, and choosing $K = 20$ and the same assumptions on the certificate size, one certification takes ≈ 27 milliseconds on the mentioned Pentium II.

6.6 Related Work

The vehicular communications community has widely accepted privacy as one of the main security research topics, see [1][55][20][27][39]. In the following, important works from the body of privacy enhanced protocols are discussed, including protocols not from the vehicular communications domain.

Anonymity-preserving certification protocols can be distinguished first by having the ability to provide revocation of anonymity under defined circumstances or by being anonymous without exception. An example of the latter case is *Direct Anonymous Attestation* (DAA) by Anna Lysyanska and Jan Camenisch.[42] Examples of revocable anonymous protocols have been proposed by the author himself in [64]. Furthermore, we are able to distinguish between protocols by the attributes of their revocation mechanism, complexity and scalability of certification, verification, and revocation, as well as their modularity in terms of exchange of base cryptographic mechanisms.

In [23], we have already criticised the approach used by the Vehicular Safety Consortium in 2005. [14, 67] In WAVE, anonymity revocation is done by exhaustive search. The search space is chosen to be feasible to brute-force the anonymity of an individual certificate, but is too costly for revocation of a large number of certificates. This specification obviously violates the first of Kerckhoffs' principles [35].

In the following, we have proposed Secure Revocable Anonymous Authentic Certification (SRAAC) [23] and the improved Trusted-SRAAC (T-SRAAC) [64] using trusted computing hardware. In both works, we have emphasised the objective of reducing the necessary trust in certification authorities to handle anonymity revocation. Different from other works in this field that either provided complete anonymous certification or left the certification authorities in possession of all information on relations between identities and certificates, this work introduced peer-controlled revocation.

In [13] a certification scheme for vehicular communication is proposed that is based on *Baseline Pseudonyms* (BP) and a group signature scheme. Both schemes are combined in a *hybrid scheme* wherein each vehicle can individually sign pseudonyms with a group signature key. Only a CA is able to reveal a signer's identity and create a revocation list if malicious behaviour is detected. This complete CA-trust is the fundamental difference of our threat model as compared to others. Direct application to only partially trusted CA or revocation authorities is not discussed. Furthermore, BP omit validity times, creating the need for revocation lists which contain single entries for each revoked vehicle.

Ma, Kargl and Weber [43] propose a *pseudonym-on-demand* protocol to allow for frequent pseudonym changes, with a focus on efficiency. Their paper motivates and explains the advantages of a hierarchy of CA similar to the

concept we have briefly introduced in Section 6.2.3. The main difference is that their approach entrusts certification authority both with certification and revocation powers.

In [26] a protocol is introduced that uses shared symmetric keys, provided by road-side units (RSU), with dedicated mix-zones. The intention is to have all vehicles use one identity while driving within this zone and thus mitigate the ability of an attacker to trace individual vehicles. The main drawback, compared to revocable anonymity schemes like the one introduced in this work, is that again, either everyone or nobody is able to derive the identity of message emitters.

Efficient protocols for revocable-anonymous certification are a prerequisite for pseudonym changes. Related to the usability of any certification scheme is the question of pseudonym change strategies. In [22] simulations of vehicle re-encounters are used to estimate the length of silence periods during pseudonym changes.

6.7 Conclusion

In this chapter, a protocol for secure revocable anonymous certification has been defined and analysed. The main objective of this chapter is the introduction of separation of privilege and adaptability of the protocol with respect to complexity constraints at given communication phases. Seperation of privilege reduces the amount of trust that has to be put into single entities. Given the mandatory participation as it could be imagined for safety communication in vehicular networks, such an approach is well motivated.

The introduced protocol is a generalised variant of the blind-signature protocol by Stadler and Camenish [62] with additional un-blinded certificate parts and adaptable certificate parameters. These modifications allow for the certification authorities to verify the validity time and other critical parameters in the (otherwise blind) signed certificate. The introduced protocol parameters allow reduction of the complexity of the certificate-verification during communication. The latter attribute is crucial for latency-restricted communication, for example, in the vehicular communication scenario described in Chapter 3.

The security of the protocol has been analysed with a focus on cheating requesters that try to manipulate the revocation anchor or open data of the certificate. Furthermore, the anonymity of certificates with respect to the open certificate data c_o has been discussed.

Considering the vehicular scenario and similar applications, both reduced trust in single authorities, as well as techniques like requester-determined open data and anonymity revocation seems wise. Prototypical implementa-

tions are needed to validate that the protocols adapted complexity is usable in vehicular communication.

That particular odyssey is now over. My mind is now at rest.

Andrew Wiles

7 Conclusion

Privacy is a growing area of research, connected to a wide field of related knowledge. This work contributes to the relatively young field of *unlinkability measures*. Unlinkability herein is understood as the inability of an attacker to correctly solve an *unlinkability problem*. Unlinkability problems are general anonymity problems, thus every anonymity problem can be mapped onto an unlinkability problem. While anonymity problems concern mappings from a set of *identification anchors* (IA) to a set of *items of interest* (IOI), unlinkability problems concern partitions of a set of IOI. The focus of this work is quantification of the amount of unlinkability, i. e., quantification of the inability of an attacker to correctly determine the true relation between IOI. (See Chapter 2)

The vehicular communication scenario presented in Chapter 3 is used as a motivating example. The emphasis was put on restricted resources, tight latency, and security requirements. In Chapter 4, a graph-model of unlinkability and anonymity problems, based on certification infrastructures as used in the scenario, was introduced.

In Chapter 5, a new unlinkability measure has been motivated and explained. The proposed *expected distance unlinkability measure* captures a notion of *consistency* of an *attacker's probability mass assignment*. It has been shown by example that previously-known entropy-based measures, e. g., the *degree of unlinkability*, do not sufficiently discriminate between different attacker's assignments.

Unlinkability measures are the subject of ongoing research. A final answer how unlinkability can be measured efficiently and expressive being still

outstanding. The main contribution of this work is to introduce the notion of *inner structure* of unlinkability hypotheses and *consistency* of attacker assignments. Aside from providing motivation for the expected distance unlinkability, these notions provide criteria that can be applied to unlinkability measures in general. The understanding developed in this process about that and why the same inner structure is not applicable to anonymity problems improves the understanding of privacy problems in general.

Our definition of the *expected distance unlinkability* is based on three criteria, derived from inner and outer structure. These criteria can be considered, as well as the resulting measure, to be a step towards a quantitative understanding of unlinkability. Previous discussions about anonymity metrics have often been focused on attacks which were not recognised by known measures, the criteria-based approach positively defines attack-quality.

Due to exponential cardinality of the unlinkability hypotheses space, our measure provides only a theoretical notion of the amount of unlinkability. However, the aforementioned entropy-based measures also have the same problem. In practise, neither measure is computable for reasonably large sets of IOI and non-trivial attacker assignments. Considering that the complexity is caused by the cardinality of the hypotheses space, heuristics are needed to approximate the measure. In this work, it has been shown by example how properties of set partitions can be exploited to calculate sums over the hypotheses space. Still, generalisation of this example remains work in progress. The proposed measure could be seen as a defining measure for establishing the semantic of unlinkability by way of unlinkability measure criteria. It is ongoing work to apply these criteria to other related measures.

Knowing that anonymity problems are a sub-problem of unlinkability problems, unlinkability measures must be applicable to anonymity as well. Therefore, a relation between anonymity measures and unlinkability measures must exist. Furthermore, a single privacy scenario is often a combination of unlinkability and anonymity problems. Application of numerous different measures is unsatisfactory, and cannot be expected from average users for weighting different measures against each other. Thus, the average user will probably not use privacy measures because the benefit is not made clear to him. A unification of privacy measures is needed to provide information to users of privacy enhancing technology.

A minor contribution of this work is the *certification protocol* proposed in Chapter 6 and adapted from a blind-signature protocol by Stadler and Camenisch. The protocol is *secure revocable anonymous* and utilises *seperation of privilege* to reduce the amount of trust that has to be put into single authorities to not misuse knowledge on relations between certificates and requesters. This design is motivated by the observation, that even trustworthy

certification authorities consist of unknown individuals, and that one single untrustworthy is sufficient to irrevocably compromise privacy.

The vehicular communication scenario used herein sets tight constraints, especially for the complexity of certificate verification during communication. To meet these requirements, new parameters of the protocol's cut-and-choose algorithm had to be introduced, and an unblinded part of the certificate had to be included. An analysis of the generalised cut-and-choose algorithm showed that a combination of parameters that provides sufficient security levels does exist.

This work provided a complete analysis that allows the user to weight complexity, and thus costs, against security. The precise choice of parameters is up to the user. Although a structure of authorities has been sketched in this work, security might be enhanced by combination of our scheme with advanced PKI schemes, e. g., as proposed in [43]. The choice of protocol parameters, especially with respect to reduction of network load on the vehicular communication broadcast medium, and its applicability on vehicle computers has to be researched in the future.

Privacy is currently a field of political controversy. Especially because privacy always is a subjective choice, privacy enhancing technology has to be ready available, usable and adaptable to the individual choice. Precise definitions for privacy-related terms and measurements, approved by the scientific community, are crucial to enable decisions. The topics in this work — unlinkability-measures and scalable privacy protocols — are a part of ongoing research to grasp the complex notions of privacy. This knowledge might help to anchor privacy considerations into the design process for future information technology.

Bibliography

[1] *CAR 2 CAR Communication Consortium Manifesto*, v1.1 edition, 2007.

[2] M. Abadi and R. Needham. Prudent engineering practice for cryptographic protocols. *IEEE Transactions on Software Engineering*, 22(1):6–15, 1996.

[3] D. Agrawal and D. Kesdogan. Measuring anonymity: the disclosure attack. *Security & Privacy, IEEE*, 1(6):27– 34, Nov.-Dec. 2003.

[4] A. Aijaz, B. Bochow, F. Dötzer, A. Festag, M. Gerlach, R. Kroh, and T. Leinmüller. Attacks on intervehicle communication systems - an analysis. In *Proceedings of the 3nd International Workshop on Intelligent Transportation*, Hamburg, Germany, March 2006.

[5] J. Benaloh (Cohen). Secret sharing homomorphisms: keeping shares of a secret secret. In *Advances in Cryptology — Crypto '86*, number 263 in LNCS, pages 251–260, January 1987.

[6] A. R. Beresford. *Location privacy in ubiquitous computing*. PhD thesis, University of Cambridge, Robinson College, April 2004. updated TecReport available.

[7] A. R. Beresford and F. Stajano. Location privacy in pervasive computing. *Pervasive Computing*, 2(1):46 – 55, Jan-Mar 2003.

[8] A. R. Beresford and F. Stajano. Mix zones: User privacy in location-aware services. In *PERCOMW '04: Proceedings of the Second IEEE Annual Conference on Pervasive Computing and Communications Workshops*, page 127, Washington, DC, USA, 2004. IEEE Computer Society.

[9] D. Boneh and M. Franklin. Efficient generation of shared rsa keys. In *proceedings of Crypto '97*, volume 48 of *Journal of the ACM (JACM)*, pages 702–722, July 2001.

[10] A. Bosselaers, R. Govaerts, and J. Vandewalle. Fast hashing on the pentium. In N. Koblitz, editor, *Advances in Cryptology - CRYPTO 1996*, number 1109 in LNCS, pages 298–312. Springer, 1996.

[11] M. Brown, D. Cheung, D. Hankerson, J. L. Hernandez, M. Kirkup, and A. Menezes. PGP in constrained wireless devices. In *Proceedings of the 9th USENIX Security Symposium*, pages 247–262, 2000.

[12] L. Buttyán, T. Holczer, and I. Vajda. On the effectiveness of changing pseudonyms to provide location privacy in vanets. In *Proceedings of the Fourth European Workshop on Security and Privacy in Ad hoc and Sensor Networks (ESAS2007)*. Springer, 2007.

[13] G. Calandriello, P. Papadimitratos, J.-P. Hubaux, and A. Lioy. Efficient and robust pseudonymous authentication in vanet. In *Proceedings of VANET'07*, Montréal, Québec, Canada, September 2007. ACM.

[14] "CAMP Vehicle Safety Communications Consortium" consisting of BMW and Daimler-Chrysler, Ford, GM, Nissan, Toyota, and VW. Vehicle safety communications project task 3 final report identify intelligent vehicle safety applications enabled by dsrc. Technical report, National Highway Traffic Safety Administration U.S. Department of Transportation, March 2005.

[15] D. Chaum. The dining cryptographers problem: Unconditional sender and recipient untraceability. *Journal of Cryptology*, 1(1):65–75, January 1988.

[16] D. L. Chaum. Untraceable electronic mail, return addresses, and digital pseudonyms. *Communications of the ACM*, 24(2), February 1981.

[17] J. DeCew. Privacy. In E. N. Zalta, editor, *The Stanford Encyclopedia of Philosophy*. Metaphysics Research Lab, CSLI, Stanford University, fall 2008 edition, 2008.

[18] C. Díaz and A. Serjantov. Generalising mixes. In *Proceedings of Privacy Enhancing Technologies workshop (PET 2003)*, LNCS 2760, March 2003.

[19] C. Díaz, S. Seys, J. Claessens, and B. Preneel. Towards measuring anonymity. In *Designing Privacy Enhancing Technologies*, LNCS 2482, pages 54–68, April 2002.

[20] F. Doetzer. Privacy issues in vehicular ad hoc networks. In *Workshop on Privacy Enhancing Technologies*, Cavtat, Croatia, May 2005.

[21] M. Edman, F. Sivrikaya, and B. Yener. A combinatorial approach to measuring anonymity. pages 356–363, May 2007.

[22] S. Eichler. Strategies for pseudonym changes in vehicular ad hoc networks depending on node mobility. In *Intelligent Vehicles Symposium*, pages 541–546. IEEE, 2007.

[23] L. Fischer, A. Aijaz, C. Eckert, and D. Vogt. Secure revocable anonymous authenticated inter-vehicle communication (SRAAC). In *ESCAR 2006 - Embedded Security in Cars*, 2006.

[24] L. Fischer and C. Eckert. 50_6 ways to track your lover. In *proceedings of the WiVeC Symposium 2008*, 2008.

[25] M. Franz, B. Meyer, and A. Pashalidis. Attacking unlinkability: The importance of context. In *Proceedings of the Privacy Enhancing Technologies 2007*, 2007.

[26] J. Freudiger, M. Raya, M. Félegyházi, P. Papadimitratos, and J.-P. Hubaux. Mix-zones for location privacy in vehicular networks. In *Proceedings of WiN-ITS 2007*, Vancouver, British Columbia, Canada, 2007. ACM.

[27] M. Gerlach. Assessing and improving privacy in vanets. In *4th Workshop on Embedded Security in Cars (ESCAR 2006)*, November 2006.

[28] B. Gierlichs, C. Troncoso, C. Diaz, B. Preneel, and I. Verbauwhede. Revisiting a combinatorial approach toward measuring anonymity. In *Workshop on Privacy in the Electronic Society 2008*, volume ACM, page 5, Alexandria,VA,USA, 2008. ACM.

[29] D. Gusfield. Partition-distance: A problem and class of perfect graphs arising in clustering. *Information Processing Letters*, 82(3):159–164, May 2002.

[30] D. Huang. *Traffic analysis-based unlinkability measure for IEEE 802.11b-based communication systems*, pages 65–74. ACM, New York, NY, USA, 2006.

[31] A. K. Jain and R. C. Dubes. *Algorithms for Clustering Data*. Prentice Hall Advanced Reference Series : Computer Science. Prentice Hall, March 1988.

[32] A. Jøsang, E. Gray, and M. Kinateder. Analysing Topologies of Transitive Trust. In T. Dimitrakos and F. Martinelli, editors, *Proceedings*

of the First International Workshop on Formal Aspects in Security & Trust (FAST2003), pages 9–22, Pisa, Italy, Sept. 2003.

[33] F. Kargl, Z. Ma, and E. Schoch. Security engineering for vanets. In *4th Workshop on Embedded Security in Cars (ESCAR 2006)*, November 2006.

[34] D. J. Kelly, R. A. Raines, M. R. Grimaila, and R. O. Baldwin. A survey of state-of-the-art in anonymity metrics. In *Proceedings of the 1st ACM workshop on Network Data Anonymization*, pages 31–40, 2008.

[35] A. Kerckhoffs. La cryptographie militaire. *Journal des sciences militaires*, IX:5–38, Jan 1883.

[36] D. Kesdogan. Evaluation of anonymity providing techniques using queuing theory. In *The 26th Annual IEEE Conference on Local Computer Networks (LCN 2001), Tampa, Florida*, November 2001.

[37] D. Kesdogan, J. Egner, and R. Büschkes. Stop-and-go-mixes providing probabilistic anonymity in an open system. In *Workshop on Information Hiding*, volume 1525 of *Lecture Notes on Computer Science*, 1998.

[38] J. Kogan. *Introduction to Clustering Large and High-Dimensional Data*. Cambridge University Press, 2007.

[39] T. Kosch. Technical concept and prerequisites of car-to-car communication. 2005.

[40] H. W. Kuhn. The hungarian method for solving the assignment problem. *Naval Research Logistics*, Quart. 2:83–97, 1955.

[41] N. Li, T. Li, and S. Venkatasubramanian. t-closeness: Privacy beyond k-anonymity and -diversity. In *IEEE 23rd International Conference on Data Engineering*, 2007.

[42] A. Lysyanskaya. A signature scheme with efficient protocols. In *In SCN 2002, volume 2576 of LNCS*, pages 268–289. Springer, 2002.

[43] Z. Ma, F. Kargl, and M. Weber. Pseudonym-on-demand: A new pseudonym refill strategy for vehicular communications. In *Vehicular Technology Conference, 2008. VTC 2008-Fall*, pages 1 – 5. IEEE, September 2008.

[44] A. Machanavajjhala, J. Gehrke, D. Kifer, and M. Venkitasubramaniam. ℓ-diversity: Privacy beyond κ -anonymity. *TKDD*, 1(1), 2007.

[45] B. Malin. k-unlinkability: A privacy protection model for distributed data. *Data Knowl. Eng.*, 64(1):294–311, 2008.

[46] D. M. Martin. *Local Anonymity in the Internet*. PhD thesis, Boston University, Graduate School of Arts and Sciences, 1999.

[47] A. J. Menezes, P. C. van Oorschot, and S. A. Vanstone. *Handbook of Applied Cryptography*. CRC Press, 1997.

[48] P. Papadimitratos, L. Buttyan, J.-P. Hubaux, F. Kargl, A. Kung, and M. Raya. Architecture for secure and private vehicular communications. In *Telecommunications, 2007. ITST '07. 7th International Conference on ITS*, pages 1–6, June 2007.

[49] P. Papadimitratos, V. Gligor, and J.-P. Hubaux. Securing vehicular communications - assumptions, requirements, and principles. In *4th Workshop on Embedded Security in Cars (ESCAR 2006)*, November 2006.

[50] A. Pashalidis. Measuring the effectiveness and the fairness of relation hiding systems. In *Proceedings of the 2008 IEEE Asia-Pacific Services Computing Conference (APSCC 2008)*, pages 1387–1394, December 9.12 2008.

[51] A. Pfitzmann and M. Hansen. Anonymity, unlinkability, unobservability, pseudonymity, and identity managment - a consolidated proposal for terminology. Technical Report v0.31, TU-Dresden, February 2008.

[52] A. Pfitzmann and M. Hansen. A terminology for talking about privacy by data minimization: Anonymity, unlinkability, undetectability, unobservability, pseudonymity, and identity management. Technical Report v0.33, TU-Dresden, April 2010.

[53] F. Porikli. Trajectory distance metric using hidden markov model based representation, 2004.

[54] B. Preneel. Cryptographic hash functions. In W. Wolfowicz, editor, *Proceedings of the 3rd Symposium on State and Progress of Research in Cryptography*, pages 161–171. Fondazione Ugo Bordoni, 1993.

[55] M. Raya and J.-P. Hubaux. The security of vehicular ad hoc networks. In *SASN '05: Proceedings of the 3rd ACM workshop on Security of ad hoc and sensor networks*, pages 11–21, New York, NY, USA, 2005. ACM Press.

[56] M. Raya, P. Papadimitratos, I. Aad, D. Jungels, and J.-P. Hubaux. Eviction of misbehaving and faulty nodes in vehicular networks. *Selected Areas in Communications, IEEE Journal on*, 25(8):1557 – 1568, Oct. 2007.

[57] K. Sampigethaya, M. Li, L. Huang, , and R. Poovendran. Amoeba: Robust location privacy scheme for vanet. In *Journal on Selected Areas in Communications (JSAC),*, volume Special issue on Vehicular Networks. IEEE, Oct 2007.

[58] M. Scheibel, C. Stüble, and M. Wolf. Design and implementation of an architecture for vehicular software protection". In *4th Workshop on Embedded Security in Cars (ESCAR 2006)*, 2006.

[59] A. Serjantov and G. Danezis. Towards an information theoretic metric for anonymity. In *Workshop on Privacy Enhancing Technologies*, LNCS 2482, April 2002.

[60] C. E. Shannon. A mathematical theory of communication. *The Bell System Technical Journal*, 27:379–423, July 1948.

[61] M. Stadler. *Cryptographic Protocols for Revocable Privacy.* PhD thesis, Swiss Federal Institute of Technology Zürich, 1996.

[62] M. Stadler, J.-M. Piveteau, and J. Camenisch. Fair blind signatures. In *Advances in Cryptology – EUROCRYPT '95*, volume 921 of *LNCS*, pages 209–219. Springer, 1995.

[63] S. Steinbrecher and S. Köpsell. Modelling unlinkability. In R. Dingledine, editor, *Proceedings of Privacy Enhancing Technologies Workshop, PET 2003, Dresden, Germany, March 26–28*, volume 2760 of *LNCS*, pages 32–47. Springer, March 2003.

[64] F. Stumpf, L. Fischer, and C. Eckert. Trust, security and privacy in vanets – a multilayered security architecture for c2c-communication. In *VDI/VW-Gemeinschaftstagung: Automotive Security*, Wolfsburg, Germany, November 2007.

[65] G. Tóth and Z. Hornák. Measuring anonymity in a non-adaptive, real-time system. In *Proceedings of Privacy Enhancing Technologies workshop (PET 2004)*, volume 3424 of *Springer-Verlag, LNCS*, pages 226–241, 2004.

[66] G. Tóth, Z. Hornák, and F. Vajda. Measuring anonymity revisited. In S. Liimatainen and T. Virtanen, editors, *Proceedings of the Ninth*

Nordic Workshop on Secure IT Systems, pages 85–90, Espoo, Finland, November 2004.

[67] US Department of Transportation, National Highway Traffic Safety Administration (NHTSA). *Vehicle Safety Communications Project: Final Report, Appendix H: WAVE/DSRC Security*, 2005.

[68] M. van Kreveld and J. Luo. The definition and computation of trajectory and subtrajectory similarity. In *GIS '07: Proceedings of the 15th annual ACM international symposium on Advances in geographic information systems*, pages 1–4, New York, NY, USA, 2007. ACM.

[69] S. von Solms and D. Naccache. On blind signatures and perfect crimes. *Comput. Secur.*, 11(6):581–583, 1992.

[70] S. D. Warren and L. D. Brandeis. The right to privacy. *Harvard Law Review*, 4(5):193–220, Dec. 15 1890.

Index

A

Mapping Anonymity onto Unlinkability

Assume that an attacker is unable to solve anonymity problems, but owns an oracle that perfectly solves unlinkability problems. The attacker can express every anonymity problem by constructing the disjoint union $M_U := M \cup^* U$, where M is a set of IOI and U is a set of IA. The oracle then solves the problem to determine the real partition from Π_{M_U}.

Additionally the attacker may even reduce the complexity of the unlinkability problem by using the hint-class "breach of unlinkability" from [25]. This class describes the situation where an unlinkability-attacker gets to know a set of elements that pairwise are in different equivalence classes. We denote the set of set partitions $\Pi_{M_U}(\mathcal{H}_U)$ of a set M as conditioned by the hint $\mathcal{H}_U$. The hint $\mathcal{H}_U$ describes, that no two elements in of the set U are in the same cluster. The hypotheses space then is defined in [25] by

$$\Pi_{M_U}(\mathcal{H}_U) := \{\pi \in \Pi_{M_U} : \forall\{m, m'\} \subseteq U \Rightarrow m \nsim_\pi m'\}.$$

Where $m \nsim_\pi m'$ denotes that m and m' are not in the same equivalence class with respect to partition π.

Knowing the subject identifiers U hint $\mathcal{H}_{|U|}$, which defines that the number of clusters is equal to $|U|$ can be applied. This hint reflects the common global anonymity scenario where all subject identifiers are known and each item of interest has to be related to exactly one subject. This can be modelled as hint "number of equivalence classes". Combining both hints, a restricted

unlinkability hypotheses space of set partitions can be defined as

$$\Pi_{M_U}(\mathcal{H}_U, \mathcal{H}_{|U|}) = \{\pi \in \Pi_{M_U} : |\pi| = |U| \text{ and } \forall\{m, m'\} \subseteq U \Rightarrow m \nsim_\pi m'\}. \tag{A.1}$$

Where $|\pi|$ denotes the number of clusters, i. e., equivalence classes, in a set partition π.

Even without these hints the oracle would produce the real partition. By re-identifying the IA in M_U the attacker thus is able to derive the mapping from IOI to IA and thus has solved his original anonymity problem.

Thus, we have shown that any anonymity problem can be mapped to an unlinkability problem. Obviously the other direction is not possible because, as the above construction shows, the set of hypotheses in anonymity is but a subset of the unlinkability hypotheses set.

B

Example Implementation of *Ed*

The following source implements partition distance in R, a programming language for statistic computing. This code does not show the implementation of `matrix.trajectory.dist` which is used for computation of trajectory distance and has to be implemented according to the scenario.

```
### partition-distance
##################
## number of elements that have to be moved to create
## equal partitions.
## Based on the algorithm described in Gusfield 2002

dist.partition <- function(pa, pb) {
  if(length(pa) != length(pb))
    stop("set partitions differ in size of base set")
  if(clusternum(pa) > clusternum(pb)) {
    ## prevents false result
    ## This is a kludge
    warning("pb has more clusters, pa,pb have been switched")
    tmp <- pa
    pa <- pb
    pb <- tmp
  }

  ## maximum optimal assignment
```

```
## uses hungarian algorithm
optimal.assignment.max <- function(M) {
  ## turning into minimization problem
  m <- max(M) - M
  # m interpretation: NA means assigned
  check.solution <- function(m) {
    j <- sapply(1:nrow(m), function(i){
      ## get already decided fields
      j <- which(is.na(m[i,]))
      if(length(j)>0)
        return(j[1])
      ## check which 0 to use (there is at least 1)
      j <- which(m[i,] == 0)
      ## TODO check for more than one zero in col,
      ## which one to choose?
      j <- j[which(sapply(j,
                          function(j)
                            (any(is.na(m[,j])))==FALSE))]
      # filter NA-blocked j
      if(length(j) <= 0)
        return(NA)
      else
        return(j[which.min(sapply(j,
                                  function(j)
                                    length(which(m[,j]==0))))])
    })
    return(j)
  } # end:check.solution

  ## reduce rows/cols until optimal solution found
  res <- c()
  ## rowwise zeroing (kludge <<- operator)
  sapply(1:nrow(m), function(i) {
    m[i,] <<- m[i,] - min(m[i,],na.rm=TRUE)})
  res <- check.solution(m)
  sapply(which(!is.na(res)),
                  function(i){
                    m[i,res[i]] <<- NA})
  if(!any(is.na(res))) {
    return(sum(sapply(1:length(res),
                      function(i) M[i,res[i]])))
```

```
    } else {
      ## column wise zeroing
      sapply(1:ncol(m), function(i) {
        m[,i] <<- m[,i] - min(m[,i],na.rm=TRUE)})
      res <- check.solution(m)
      sapply(which(!is.na(res)), function(i){
        m[i,res[i]] <<- NA})
      if(!any(is.na(res))) {
        return(sum(sapply(1:length(res),
        function(i) M[i,res[i]])))
      } else {
        stop("todo: not implemented yet")
      }
    }
  } # end:optimal.assignment

  ## M: matrix of numbers of
  ##    elements in intersection of clusters
  ## a cluster is element of a set partition here
  M <- matrix.trajectory.dist(pa,pb)

  ## return number of elements that have to be deleted
  ## until both partitions are equal.
  return(sum(M)-optimal.assignment.max(M))
}
```

APPENDIX B. EXAMPLE IMPLEMENTATION OF *ED*

C

Shared Revocation Keys

In the following we describe the shared RSA-key generation and joint decryption as introduced in [9]. The described procedures can be used to implement a quorum-revocation authority in the certification protocol described in Chapter 6. The obvious advantage of this particular scheme is that the shared keys are not generated by a single (trusted) dealer who would be — again — the single point of failure we are avoiding here.

The described protocol is used to decrypt $E_{y_R}(c_b||\alpha)$. For simplicity reasons, we denote the ciphertext as c in this section. The decrypted plaintext is denoted m.

Private key shares x_i are calculated from the public key y_R and pre-shared key-generators p, q through three intermediate steps. First, each RA calculates a value φ_i from its share of p and q. The first step is to calculate

$$\varphi_i \quad = \quad \begin{cases} n_R - p_1 - q_1 + 1, & i = 1 \\ -p_i - q_i, & i \in \{2, \ldots, k\}. \end{cases}$$

There is one distinguished "first" among the RA which has to add $n_R + 1$ to its value.

In the second step, φ is used to produce ζ using l in a secret joint computation[1] without revealing the individual values of $\varphi_1, \ldots, \varphi_k$. Each member

[1] Using Benaloh's protocol [5] as described in [9].

of the quorum finally knows

$$\begin{aligned} l &= \varphi \mod y_R \\ \zeta &= l^{-1} \mod y_R. \end{aligned}$$

Now each RA generates its share of the private key x_i.

$$x_{Ri} = \left\lfloor \frac{-\zeta \cdot \varphi_i}{y_R} \right\rfloor.$$

This allows for joint decryption of cipher text c to

$$m = c^{x_R} \equiv c^r \prod c^{x_{Ri}} \mod n_R, \tag{C.1}$$

where r is known from a first trial decryption undertaken by the first RA and added to one of the x_{Ri}, e. g., x_{R1}. The plaintext m includes the necessary data to identify the certification run and, by this, the identity of the requester.

D

Curriculum Vitæ

Employment History

2009 – Scientific Advisor IT-Security, Specialist Division IT-Security, DIS AG (Branch Leipzig)

2003 – 2008 Research Assistant with Prof. Eckert, Working Group IT-Security, Department of Computer Science, Technische Universität Darmstadt

2001 – 2002 Student Assistant, Working Group Computer Networks, Department of Computer Science, University Bremen

1998 – 2001 Student Tutor, Basic Courses on Computer Logic, Operating Systems and Computer Networks, University Bremen, Department of Computer Science

1997 – 2003 Student of Computer Science, University Bremen

1996 in-between Jobs: Waiter, Dock Worker

1995 – 1996 Civil Service at Multiple-Sclerosis Nursing Home "Alten- und Pflegeheim Anna Stiegler, Haus Blumenkamp"

1995 Assistant IT-System Administrator (3 Month), Central Hospital "Sankt-Jürgen-Straße", Bremen

Education

2003 Graduate Degree (*Diplom*) in Computer Science, University Bremen, Department of Computer Science

1995 University-entrance Diploma (*Abitur*), Gymnasium an der Hamburger Straße, Bremen, Germany, Focus on Mathematics and Physics

Volunteer Work

2000 – 2003 Teaching Ju Jutsu and Modern Arnis (Selfdefense for Children and Adults), Polizeisportverein Bremen

since 2003 Deutscher Alpenverein, Sektion Darmstadt-Starkenburg (German Alpine Club)

Student Projects

2001 – 2002 ANIMA, 4D Syntactical Image Generation Lab. responsible for Networking Stack and Object Language

2000 GraphFruit, Graph Visualization Tool, Responsible for Object Network Interface (ONI)

Teaching Activities

Teaching Assistant

WS 2003	Operating Systems
WS 2004	IT-Security I
SS 2005	IT-Security II
WS 2005	Grundlagen der Informatik 3 [Basics of Computer Science 3]

Teaching

SS 2004	Advanced Seminar: Security of Ad Hoc Networks
SS 2006	Practical Course: Hacker Contest
WS 2006	Advanced Seminar: Topological Adressing
SS 2007	Practical Course: Hacker Contest
WS 2007	Seminar: Security in Car2Car-Communication
SS 2008	Seminar: Secure Vehicular Communication