Proceedings of the 2014 Winter Simulation Conference
A. Tolk, S. Y. Diallo, I. O. Ryzhov, L. Yilmaz, S. Buckley, and J. A. Miller, eds.

PRIVACY ASSESSMENT IN VEHICULAR NETWORKS USING SIMULATION

Isabel Wagner

University of Hull
Cottingham Road
Hull, HU6 7RX, UK

David Eckhoff

University of Erlangen
Martensstrasse 3
91058 Erlangen, GERMANY

ABSTRACT

Vehicular networks are envisioned to play an important role in the building of intelligent transportation systems. However, the dangers of the wireless transmission of potentially exploitable information such as detailed locations are often overlooked or only inadequately addressed in field operational tests or standardization efforts. One of the main reasons for this is that the concept of privacy is difficult to quantify. While vehicular network algorithms are usually evaluated by means of simulation, it is a non-trivial task to assess the performance of a privacy protection mechanism.

In this paper we discuss the principles, challenges, and necessary steps in terms of privacy assessment in vehicular networks. We identify useful and practical metrics that allow the comparison and evaluation of privacy protection algorithms. We present a systematic literature review that sheds light on the current state of the art and give recommendations for future research directions in the field.

1 INTRODUCTION

Equipping vehicles with WiFi technology to exchange information among each other or infrastructure nodes is a promising approach to increase road safety, improve traffic flow, and bring new types of comfort applications to the driver. Vehicles will periodically broadcast their state – position, speed, heading, etc. – over an interface very similar to consumer wireless LANs. This information is sent unencrypted and can be received by anyone close enough to the sender – typically in a maximum distance of 200 m to 1000 m. This raises privacy concerns, as it would not only allow to easily track vehicles but also to build a fully automated traffic surveillance system (Eckhoff and Sommer 2014).

The need for effective privacy protection has been understood from the beginning (Hubaux, Čapkun, and Luo 2004). Upcoming standards describe the use of changing, pseudonymous identifiers instead of static addresses. However, concrete privacy measures, for example, how or when these pseudonyms are changed, have not yet been recommended in these documents. One of the major reasons is that the abstract concept of privacy is hard to grasp and cannot be easily put in numbers, leading to the fact that protection measures are also often neglected in field operational tests.

One of the ways to assess how well a vehicular network preserves the drivers' privacy is by means of simulation. The system is evaluated using tailored, often very complex metrics in specific scenarios with various assumptions regarding mobility and the adversary to protect against, and a vast number of parameters. Publications often do not give complete information about their simulations, reducing their reproducibility, comparability, and validity of results.

By performing a systematic literature review, we investigate the current state of the art in privacy simulation and aim to identify reasons why privacy protection mechanisms are not yet established. We discuss current trends, benefits and shortcomings of different methodologies (i.e., metrics, adversary models, evaluation) and give detailed recommendations on how to assess privacy in vehicular networks.

The remainder of this paper is organized as follows: In Section 2 we discuss the general concept of privacy in vehicular networks, followed by a description of our literature review method (Section 3). We summarize our taxonomy of the field (Section 4) and, based on our classifications, present our results in Section 5. We give detailed recommendations for the simulation of privacy protection in vehicular networks in Section 6 and discuss future work and open topics (Section 7).

2 PRIVACY IN VEHICULAR NETWORKS

One reason why the current state of privacy and privacy-enhancing technologies leaves much to be desired could be that privacy is a somewhat nebulous concept.

One of the early definitions of privacy defines privacy as "the ability of an individual to control the terms under which personal information is acquired and used." (Westin 1967). Twenty years later, the EU privacy directive (European Parliament 1995) defines "personal data" as "any information relating to an identified or identifiable natural person [...]; an identifiable person is one who can be identified, directly or indirectly, in particular by reference to an identification number or to one or more factors specific to his physical, physiological, mental, economic, cultural or social identity."

Reducing the vagueness of these definitions, Nissenbaum (2004) defines privacy in terms of contextual integrity, which "ties adequate protection for privacy to norms of specific contexts, demanding that information gathering and dissemination be appropriate to that context." This means that norms determine which information is appropriate to reveal in a given context (e.g., a hospital visit), and also how that information may be distributed to other parties. Information is therefore always associated with a context that governs how it may be used. For example, the re-purposing of personal data collected for a specific purpose will often lead to a violation of contextual integrity. Even though contextual integrity is a compelling concept, it is not technical, and as such difficult to measure.

Thus, to engineer better privacy protection solutions, better quantitative measures of privacy are needed. All kinds of analyses and automated algorithms – and simulations – depend on numeric values. Without a measure it is therefore difficult to compare different privacy-enhancing technologies and judge which better serves its purpose.

Applying contextual integrity to the area of vehicular networks, Zimmer (2005) argues that existing norms in this context allow for the sharing of visually observable information, such as a general description of the occupants in a vehicle, or its license plate number, but generally only by persons who happen to be nearby. The introduction of vehicular networks for safety purposes has the potential to expand the range of data to be shared, to increase the data's precision and accuracy, to allow data to be collected without human intervention, without line of sight, and to store and analyze this information on a large scale and across large geographic areas.

One of the first things that come to mind when looking at the concept of vehicles periodically broadcasting detailed location information is a possible violation of location privacy – the ability to track vehicles, and to create and share detailed driver profiles. In fact, most studies almost exclusively focus on location privacy, since many other types of privacy, such as social privacy or the privacy of behavior and action (Finn, Wright, and Friedewald 2013), can be derived from a person's location. To prevent tracking, vehicles are envisioned to use pseudonymous identifiers that are frequently changed and, in a best case scenario, cannot be linked to each other or to the identity of the driver by anyone but the provider. Current efforts to make vehicular networking capabilities legally mandated are especially worrisome in the context of road traffic: it would allow for the deployment of a fully automated traffic surveillance system (Eckhoff and Sommer 2014), where the user, unlike with cell phones or public WiFi, does not have the choice to turn it off.

3 LITERATURE REVIEW METHOD

A systematic literature review is a process following a fixed set of steps. The goal is to ensure that the review encompasses as much of the relevant literature as possible, and at the same time reduce the influence

of biases the authors may have – in essence, to make sure the review is thorough and fair. Kitchenham (2004) adapted the approach to Computer Science from its origin in medical research. The three phases constituting a systematic review are planning, conducting, and reporting the review.

The planning phase results in a set of research questions, as well as a literature review protocol that specifies the search engines and keywords to be used, and a set of inclusion/exclusion criteria. In the conducting phase, the literature search is carried out, and papers are classified according to a taxonomy, including an assessment of paper quality. This phase also includes data synthesis and answering the research questions. The reporting phase is then concerned with writing up and publishing the results.

Research Questions The main aim of our review is to study how privacy is assessed in simulations of vehicular networks. This leads to the following research questions:

1. What is the state of the art in evaluating privacy in vehicular network simulations?
2. How can we classify privacy metrics used in vehicular network simulations?
3. What patterns, gaps, challenges – and strengths/weaknesses of individual metrics – can be inferred from this?

Literature Review Protocol In order to find the relevant literature to answer our research questions, we decided conduct keyword-based searches using several major search engines. We used the three keywords *(simulation and privacy and vanet)* joined by a Boolean AND, and collected results from Google Scholar, the ACM Digital Library, IEEE Xplore, ScienceDirect, Springer Link, and Web of Science. To these search results, we added a number of papers we knew to be relevant but were missing one or more of our keywords from their titles or abstracts. After de-duplication of results, our search had produced around 180 papers. To align these search results with the focus of our study, we defined additional inclusion criteria: we selected only those papers that use discrete-event simulation to evaluate privacy. This resulted in a final list of 48 papers.[1]

4 TAXONOMY

When considering how to classify literature with regard to the evaluation of privacy protection mechanisms, three main dimensions have to be taken into account.

The first dimension concerns privacy properties – the specific aspects or types of privacy that are analyzed. The second dimension regards the adversary and her capabilities. This includes the question against whom privacy should be protected, for example, a local adversary or one that has access to all data within the system. The third dimension is the privacy metric, the method employed to measure and quantify the level of privacy. This could be counting among how many persons an individual is indistinguishable, or estimating the probability of an adversary's success.

In the following sections, we shed light on each of these aspects; we introduce how privacy properties, adversaries, and privacy metrics are commonly classified; and we discuss the strengths and weaknesses associated with each. Finally, we present criteria to assess quality and repeatability of simulation studies.

4.1 Privacy Properties

There is general agreement in the field (Pfitzmann and Hansen 2010; Deng et al. 2011) that five privacy properties can be enforced by technical means.

Anonymity describes the property that an adversary cannot distinguish the target from a set of other subjects, the anonymity set. In vehicular networks, the anonymity set usually describes a number of vehicles among which the adversary is not able to single out the target vehicle.

Unlinkability describes the adversary's inability to link two or more subjects, actions, or locations. In vehicular networks, this property is often used to describe whether the adversary is able to link a vehicle's pseudonyms through (a series of) pseudonym changes.

[1] A spreadsheet containing full details for all 48 papers as well as our classification is available online at https://hydra.hull.ac.uk/resources/hull:8478

Undetectability describes an adversary's inability to discern whether an item of interest exists or not. Undetectable messages, for example, would be indistinguishable from static noise.

Plausible deniability, or repudiation, describes that a subject is able to deny having performed an action. In the context of vehicular networks, this can refer to users being able to deny having visited a specific location.

Confidentiality refers to an adversary's inability to access the content of data. In vehicular networks, this could consist of encrypting messages to keep the content of queries (and their replies) hidden.

4.2 Adversary Models

As explained by Díaz (2006), the evaluation of anonymity depends on the specific capabilities of the adversary. This statement can be generalized to all of the above mentioned privacy properties. Naturally, the results of the privacy evaluation will change based on the adversary's type which can be classified along five dimensions.

Internal vs. external describes whether the adversary is part of the system – a participating vehicle, for example, or a road-side unit – or outside of the system, for example operating her own (network of) eavesdropping equipment.

Local vs. global refers to the geographic extent of the adversary's operations. A local adversary, for example, could be operating WiFi sniffers at a single intersection, while a global adversary is assumed to have access to the entire system's communications. Real-world scenarios might also feature mixtures between the two, such as distributed, connected local adversaries.

Active vs. passive describes whether the adversary is trying to infer information by passively observing the system, or by actively participating in it, for example by spoofing road-side units or triggering responses from specific vehicles.

Static vs. adaptive describes whether the adversary's strategy and behavior is static – fixed from the outset – or whether the adversary is able to adapt her behavior to whatever information she is able to learn.

Prior knowledge describes all the bits of information that may be at the adversary's disposal before she starts attacking the system. This knowledge has to be differentiated between general information about the system itself (such as WiFi frequencies or boundaries of vehicular movement) and to properties of the specific situation, like the initial identities of nodes, city layouts, or statistics about traffic flow. In this paper, when we refer to prior knowledge, we mean the latter type.

4.3 Privacy Metrics

There is a large number of different privacy metrics to be found in the literature. For the purpose of this paper, we focus on metrics that are used to evaluate privacy through discrete-event simulation. These metrics can be grouped into five main categories: anonymity set size, entropy, adversary's tracking success rate, statistics on pseudonym changes, and maximum tracking time.

Anonymity Set Size Anonymity set size describes among how many other vehicles it is not possible to distinguish the target vehicle. The advantages of this metric lie in its simplicity and ease of calculation. However, being an absolute number, the anonymity set size depends on the total number of vehicles in a scenario, making results difficult to compare among studies. In addition, the metric measures only anonymity, and disregards the other privacy properties.

Anonymity set size is similar to the concept of k-anonymity which describes that a specific database record is indistinguishable from k other records. The literature has shown that records can be de-anonymized even if k-anonymity is fulfilled (Shokri et al. 2010), and also that k-anonymity for location privacy, i.e. the anonymity set size, is insufficient for similar reasons.

Finally, Serjantov and Danezis (2003) discuss two further problems with the anonymity set size. First, there is no way to represent an adversary's prior knowledge. Second, the metric assumes that all vehicles in the anonymity set are equally likely to be the target. However, the adversary may make observations

that make some vehicles (much) more likely than others. This cannot be adequately described using the anonymity set size.

Entropy To alleviate this last problem, many authors have turned to entropy. Entropy is a concept from information theory expressing the uncertainty in a random variable. The entropy of an anonymity set can therefore represent the adversary's beliefs about the likelihood of individual vehicles. Formally, entropy is expressed as $H(X) = -\sum_{i=1}^{N} p_i \log_2(p_i)$, where N denotes the number of vehicles in the anonymity set and p_i usually refers to the adversary's estimation of the probability of i being the target vehicle.

Unfortunately, calculating the entropy of the anonymity set also only evaluates the anonymity dimension of privacy, disregarding other privacy properties.

Similar to the anonymity set size, entropy also depends on the absolute number of vehicles in the anonymity set, making it difficult to compare entropy values between different studies. A remedy for this is the degree of anonymity which has been proposed by Díaz et al. (2003). It normalizes the entropy using the maximum possible entropy value $H_M = \log_2(N)$, resulting in a range of $[0,1]$ for the degree of anonymity $d = \frac{H(X)}{H_M}$.

Another drawback of entropy was discussed by Hoh and Gruteser (2005). They argue that entropy will have a high value if an adversary cannot decide between two vehicles. But if those two vehicles are in fact standing next to each other, the adversary has successfully inferred their location. This low level of privacy is then not reflected in the entropy value.

In addition to computing the entropy of an anonymity set, entropy can also be used to represent the uncertainty in other quantities, for example the adversary's uncertainty in assigning trips to individuals. Some authors also look at cumulative entropy, adding up entropy values for each successful pseudonym change, or at the difference between entropies before and after a pseudonym change.

Adversary's Success Rate The adversary's success rate measures to what extent an adversary is able to achieve her goal, without specifying what exactly that goal is. As such, it can be used to measure any of the privacy properties, depending on the specific way it is defined.

This flexibility is also the main disadvantage of measuring the adversary's success rate. Since the adversary's goal is not specified, researchers adapt the metric to their own needs and scenarios, leading to a huge variety of subtly different metrics that make comparisons between studies difficult, if not impossible.

Maximum Tracking Time The maximum tracking time is similar to the adversary's success rate in that it measures how well the adversary is able to succeed. The maximum tracking time is much more specific as it specifies exactly what the adversary's goal is, namely, tracking vehicles for as long as possible. As such, the metric focuses on the unlinkability property.

The metric assumes that the adversary will eventually be completely confused by pseudonym changes and measures the time until this happens. Its usage is problematic if the adversary draws probabilistic conclusions, or if she is able to re-link vehicles at a later time.

Statistics on Pseudonym Changes Metrics calculating statistics on pseudonym changes only give an indirect estimation of privacy because they only estimate how well a particular mechanism – pseudonym changing – is working. The only privacy property evaluated in the surveyed papers using this metric was unlinkability. Statistics on pseudonym changes are straightforward to compute since they directly concern the functioning of the privacy-preserving mechanism. However, this restricts their applicability to pseudonym changing strategies, and also limits comparability because their value depends on the functioning of the mechanism.

Other metrics Several metrics used in the surveyed papers do not fit into the above classification, but were not used frequently enough to merit their own category.

Some of them are quite complicated and hard to understand which may have hindered their adoption. Examples include the location privacy gain (a combination of anonymity set size, the probability of pseudonym change, and how important privacy is to the user), user-centric privacy loss (taking into account the entropy of the anonymity set and a user-specific time decay), or the metric introduced by Ma, Kargl, and Weber (2010) which combines the entropy of trip probabilities with information accumulated by the

adversary. Others are very specific to a particular mechanism and as such have limited applicability. Examples include the quiet time to be observed after a pseudonym change, or the number of destination locations revealed.

Location privacy metrics have also been investigated outside the context of simulation studies, mainly for analytic evaluations. A good survey on the topic is Shokri, Freudiger, and Hubaux (2010). As an example, Fischer, Katzenbeisser, and Eckert (2008) argue that entropy-based measures do not suffice to measure unlinkability and therefore introduce the expected distance unlinkability measure, that accounts for the "inner structure," i.e., the similarity between the adversary's choices, and the robustness of these choices.

Another example is the expectation of the adversary's distance error introduced by Hoh and Gruteser (2005). This metric captures how well an adversary is capable of estimating a user's position. This could be a useful metric in combination with an uncertainty-based metric such as the entropy.

Many metrics from other application areas have not yet been applied to location privacy, for example mutual information (Chen and Pang 2012) and differential privacy (Dwork 2008). Mutual information is an information theoretic concept that could be useful in a vehicular network context to measure the commonalities between a user's real location trace and the adversary's estimation of it. Differential privacy was originally developed for use in statistical databases, but has since been adapted to many other areas. It is a mechanism that can guarantee a level of privacy to the user. It could be used to allow a vehicle to gauge whether the data it transmits will violate the user's privacy level.

4.4 Quality Criteria

In addition to the dimensions introduced above, we defined criteria to indicate the quality and repeatability of simulation studies. In decreasing order, the criteria for the quality of studies are whether the authors reported on confidence intervals; quantiles or standard deviation as measures for statistical error; replications; and averages. The importance of reporting confidence intervals and measures of statistical error has been analyzed in the seminal paper by Pawlikowski, Jeong, and Lee (2002), which states that simulation studies lacking these features risk their credibility. We furthermore investigated the repeatability of the surveyed simulation studies, that is, whether the authors use publicly available models, list all relevant parameters, and describe the simulation package that was used.

5 SURVEY RESULTS AND ANALYSIS

Of the 48 papers in our survey, there were 30 conference papers and 18 journal publications. Figure 1 shows how the publication dates are distributed between the earliest papers in 2005 and the latest in 2014.

5.1 Privacy Properties

Figure 2 shows that the only privacy properties addressed by papers in our survey are unlinkability and anonymity. Undetectability, plausible deniability, and confidentiality are not investigated by any of the papers. While unlinkability and anonymity are certainly the most pressing privacy properties to protect in vehicular networks, we believe neglecting the other three properties will be detrimental to the overall quality of privacy research. This is an area that should be investigated in future research.

In addition, most papers claim to investigate location privacy. This is usually mapped to anonymity, unlinkability, or both, but in some cases the exact meaning is left unspecified.

Only about a quarter of the papers surveyed investigate both unlinkability and anonymity. These can be very distinct properties – concerning the linkage of successive pseudonyms or locations in one case, and the connection of an individual to locations in the other. Therefore, it may often be necessary to analyze both properties simultaneously, which is a gap in current research.

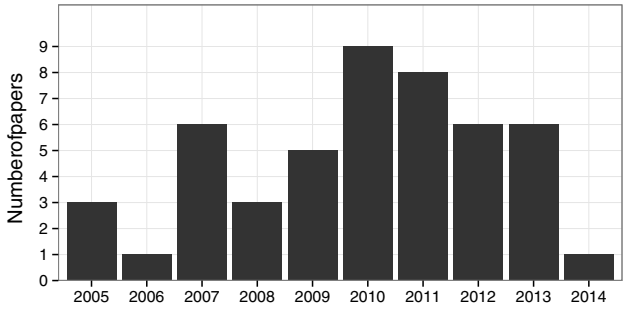

Figure 1: Surveyed papers by year.

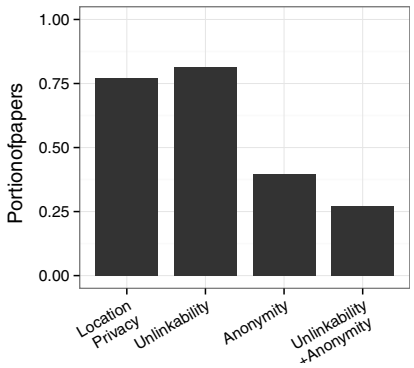

Figure 2: Privacy properties.

5.2 Adversary Models

Figure 3 clearly shows that research focuses on global (70%), passive (80%) adversaries, that is, eavesdroppers who can listen in to all of the communication in the network. The relationship between an external, eavesdropping adversary and an internal adversary who might have access to some of the cryptographic material (e.g., participating vehicles or providers of road-side units), or an active adversary who might be able to alter communications, is not immediately clear. This means that there is no obvious ordering in the strength of an adversary such that it would suffice to only investigate the strongest possible adversary. Investigating the effects of other adversary types is clearly a gap in existing research.

Only about one in five papers considered the effect of prior knowledge available to an adversary. However, in the age of big data, it is becoming increasingly likely that an adversary would have access to data that she might be able to correlate with her own observations. This may either increase her chances of success, or allow her to draw surprising new conclusions. Investigations in this direction seem to be promising areas for future research.

Another point to note is that none of the papers classified their adversary models along the static/adaptive dimension. Since adaptivity would certainly have been described, we conclude that the adversaries considered were all static. However, real-world adversaries would in all likelihood be able to learn and adapt. We believe it would be worthwhile to reflect this in simulation studies, even if it is more difficult to model.

5.3 Privacy Metrics

Figure 4 shows an overview of the frequencies with which the five broad classes of privacy metrics were used by the surveyed papers. Most authors favor the more general metrics – anonymity set size, entropy, and adversary's success rate – while metrics specific to vehicular networks are significantly less popular.

Anonymity Set Size About one third of the papers in our survey use the anonymity set size as a metric to evaluate privacy. Given the criticisms that have been directed at this metric for more than a decade now (and the number of viable alternatives in the literature) it is somewhat surprising that it is still in such widespread use. As the first group of bars in Figure 5 shows, there is no visible decline in the use of this metric over the years.

Entropy The vast majority of papers we classified as using entropy computed the entropy of the anonymity set, taking the probability over the adversary's belief whether a specific vehicle is the target. In most papers, the probability distribution depends on the adversary's strength. However, in almost 25% of papers the probability distribution was assumed to be uniform, in which case entropy does not carry any more information than the anonymity set size.

Surprisingly, none of the papers in our study used the degree of anonymity instead of entropy, although this would make it easier to evaluate and compare results, even if a reader is not entirely familiar with the characteristics and typical value range of entropy.

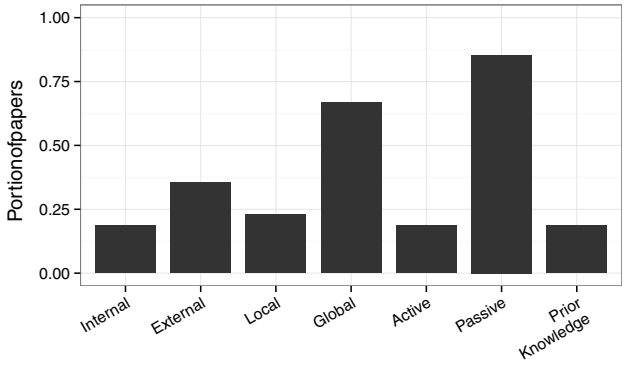

Figure 3: Adversaries.

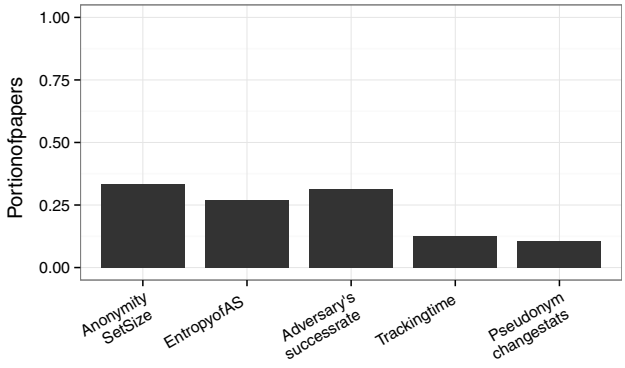

Figure 4: Metrics.

Adversary's Success Rate Of the 15 papers in our survey that looked at the adversary's success rate in some form, we identified eight different variants, a combination of ratios, probabilities and chances relating to successful attacks, adversary guesses, vehicle identification, tracking, pseudonym mapping, and anonymity set size. Taking into account that some papers were authored by the same group, this means that almost every group defines their own version of the adversary's success rate. This is clearly not helping to make studies more comparable, and it might even be actively confusing to a reader who does not delve into the detailed definition of the metric in every single paper.

Maximum Tracking Time According to the fourth group of bars in Figure 5, the maximum tracking time seems to have fallen out of use recently, the latest paper using it dating from 2011. As discussed in the taxonomy section, the metric does have shortcomings in terms of the assumptions necessary for its use. However, its strength lies within its public outreach capabilities as it is easy to understand and market. Depending on the publication venue, the maximum tracking time can be the right choice to advertise the effect of privacy protection mechanisms to a wider audience.

Statistics on Pseudonym Changes The same statements also apply to statistics on pseudonym changes. Six of the papers in the survey employed this type of metric using a total of five different variants, a combination of total numbers and rates of pseudonyms received, successful and failed pseudonym changes, and changes per minute.

5.4 Quality of Simulation Studies

Only about one quarter of the papers report either confidence intervals or a measure of statistical error (quantiles or standard deviation). This means that the validity of simulation results is questionable for three quarters of the papers (Pawlikowski, Jeong, and Lee 2002). This is clearly a situation that needs to be improved upon. One in five derives results from a single simulation run – effectively drawing conclusions based on a dataset with size $N = 1$. Given the random nature of simulation inputs, this is highly problematic. Most of the papers – 80% – only report averages. The reporting of averages is problematic because it does not give information about the result distribution. If the result distribution is bipolar, e.g., consisting of a group of cars with perfect privacy, and another group with a much reduced level of privacy, then the averaged result would wrongly indicate that everybody has the same average level of privacy. Of those papers reporting averages, 60% report that they have conducted independent replications. For the remaining papers, it remains unclear how exactly these averages are derived.

Concerning repeatability, about 80% of the papers in our survey report on the simulation package used, but only one of the papers indicates that their models are publicly available. However, the URL given in this paper is no longer working. This highlights one issue that may be keeping authors from publishing their models: once a specific URL is published, someone has to be responsible for maintaining it and making sure that it remains available.

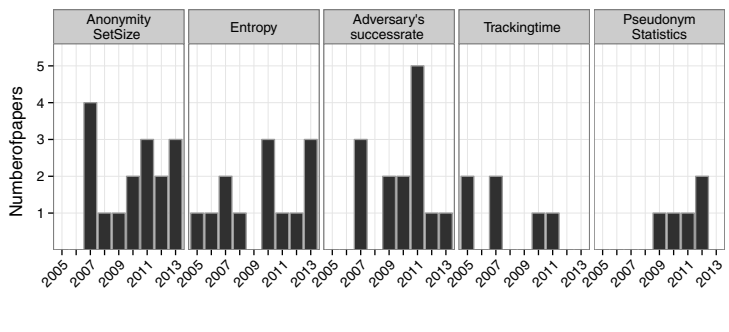

Figure 5: Metrics by year.

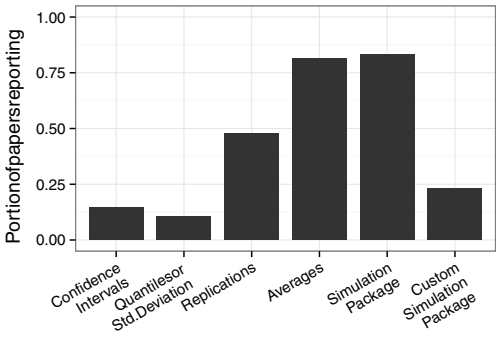

Figure 6: Quality of simulation studies.

Apart from one, all papers report on simulation parameters in at least some degrees of detail. However, as the right-hand side of Figure 6 shows, almost one in four papers use a custom simulation package that is not publicly available (e.g., all those developed in Matlab). Therefore, only little more than half the papers have made use of a publicly available simulation package.

This leads to the alarming conclusion that none of the surveyed simulation studies are in fact easily repeatable for non-involved authors.

6 RECOMMENDATIONS FOR SIMULATIVE ASSESSMENT OF PRIVACY

In this section we discuss the challenges faced by simulation studies aiming to assess privacy and give recommendations on how these challenges can best be overcome. A discussion of general problems and recommendations for scientific programming can be found in Merali (2010).

Reproducibility The biggest challenge in the simulation of privacy protection methods is reproducibility. If the research community is unable to repeat simulations or confirm results presented in a publication the meaningfulness of the proposed approach suffers. However, explaining an algorithm in detail is not sufficient as privacy simulations rely on a large number of input parameters.

The setup for privacy simulations is quite consistent throughout the reviewed literature: vehicles move through a road network and emit location information that is collected by an adversary. Using this information the adversary tries to track the vehicles using some kind of tracking algorithm. The deployed privacy mechanism, e.g., pseudonym changing, influences how well the tracking algorithm works. The quality of this protection mechanism is then linked to the success rate of the adversary, often by using metrics that reflect the uncertainty of the adversary. This evaluation method contains a vast number of parameters, making it difficult or even impossible to properly reproduce results if only one of them is not described. The growing availability of digital online repositories at universities[2] could help alleviate this problem and allow researchers easy long-term sharing of their simulation scenarios and setups.

Mobility Another important factor is the mobility of vehicles. Microscopic traffic simulators have replaced unrealistic mobility models such as the random way point model, however, they still need to be tweaked to provide fully realistic road traffic. For example, some of the investigated papers use mobility models that did not account for lane changing or overtaking. This is problematic because real drivers are less predictable than their simulated counterparts. Car following and lane change models need to account for this diversity as the individual behavior of a driver has a large impact on the success of a tracking algorithm. It seems that currently the best option is the use of sophisticated state-of-the-art traffic simulators such as SUMO (Krajzewicz et al. 2002) or VisSim (Fellendorf and Vortisch 2010).

Another possibility is the use of mobility traces which, when collected properly, have the highest degree of realism. Their availability, however, is problematic as public traces of private vehicles raise privacy concerns. To overcome this problem, researchers often use traces from public vehicles, such as buses,

[2]Similar, for example, to the *Hydra* repository at the University of Hull, where the data underlying this paper are published.

taxis, or public safety vehicles. Since these vehicles have distinct mobility patterns, a privacy protection algorithm evaluated using a taxi trace cannot be said to also protect the privacy of private persons.

Scenarios Even when two protection mechanisms are evaluated using the same simulator, it can be impossible to compare the presented results when the simulation scenarios differ. For example, highway traffic is very different from traffic in an urban environment. Scenarios for highway traffic often use a straight line with several lanes, with or without on/off-ramps. Urban environments are simulated using Manhattan Grid-like scenarios or, giving considerably different results, using real map data (Sommer, Eckhoff, and Dressler 2010). Regardless of the scenario, researchers should always explain their general setup in detail or make their scenarios publicly available. Ideally, the research community would agree on a collection of default scenarios to be used in simulative privacy assessment (Eckhoff and Sommer 2014, Uppoor and Fiore 2011). These scenarios would consist of various settings that cover the common mobility patterns (e.g., freeway, urban, intersection, rotary, etc.).

Adversaries Collecting location information from vehicles always depends on the adversary model. A global, omniscient adversary is often used as a very strong passive adversary, for example to derive upper bounds. If the collected location information is accurate enough, it seems to be almost impossible to confuse a modern tracking system ((Wiedersheim et al. 2010); Blackman and Popoli 1999) requiring researchers to also study more realistic, local adversaries. While the exploitation of data by passive, external adversaries is certainly a threat, internal or even active attacks can be just as or even more harmful. New privacy protection algorithms should therefore also be evaluated with respect to these kinds of attacks, especially also considering adaptive adversaries. Articles should always include a detailed adversary model as the effectiveness of privacy protection algorithms will depend on the adversary they try to protect against. They should also discuss against which adversary types the presented scheme does *not* protect.

The tracking algorithm itself is a very important element of the adversary model and will significantly influence the outcome of the simulation. Implementing a state-of-the-art tracking system such as the ones presented by Blackman and Popoli (1999) is complicated and error-prone. Unfortunately, researchers can only rarely resort to established and reviewed libraries as both their availability and applicability is often insufficient. If newly implemented, authors should always describe the building blocks of their tracking system accurately enough for others to reproduce.

Metrics The metrics to illustrate the effectiveness of a privacy preserving mechanism should be selected carefully. Their applicability depends on the used scenario, adversary model, and on the proposed mechanism itself. Ideally, we would use metrics that measure the right privacy properties with sensible accuracy and reproducibility but are still easy to understand and apply. Given the advantages and disadvantages of the metrics we discussed in Section 4, it seems that this is unlikely to be achieved by any single metric. In general, we would therefore opt to use a combination of metrics instead of only a single one. Among them, this combination should measure all relevant privacy properties; consider the characteristics of the adversary; be sufficiently accurate from a scientific viewpoint; and be easy enough to understand to be usable in public outreach activities. We believe that the degree of anonymity would be a good uncertainty-based metric to measure the anonymity property. The expected distance could serve as an unlinkability measure, and the maximum tracking time could be used to target wider audiences. We also recommend to take inspiration from the wider literature on privacy metrics.

Common Framework To facilitate the sharing of simulation parameters, models and algorithms, we believe the research community would benefit from a common simulation framework that includes a set of scenarios, adversary models, tracking algorithms, metrics, and a list of sound, predefined parameters, released under a permissive open source license such as the GPL. Articles could then simply refer to this framework, name the adversary and scenario, and only list parameters that were different from the default configuration. Ideally, new privacy protection mechanisms can then be released as a module for this framework to allow for the easy comparison and reproduction of results.

7 CONCLUSION

We conducted a systematic literature review concerning the assessment of privacy in simulations of vehicular networks. We introduced a taxonomy along the dimensions of privacy properties, adversary models, privacy metrics, and quality of simulation studies, and classified 48 papers accordingly. Our three main findings were first, that there was a distinct lack of quality concerning the statistical analysis of simulation results. Second, some privacy properties (undetectability, repudiation, confidentiality) and some types of adversaries (internal, adaptive) are investigated rarely or not at all. Third, there is a huge variety of subtly different privacy metrics, with common ground centering on metrics that are not as meaningful as others. Based on these findings, we formulated a set of recommendations for conducting privacy assessments in simulation studies. The most important are focusing on proper statistical analysis and proper reporting; enhancing reproducibility through supplementary material in digital repositories; using a combination of metrics to measure different facets of privacy; and a common framework with predefined mobility models, scenarios, adversaries, and metrics. Throughout the paper, we highlighted potential areas for future investigation. Most importantly, we believe that more research is necessary toward meaningful and comprehensible privacy metrics, taking inspiration from other areas. This will ensure that the nebulous concept of privacy becomes easier to grasp, paving the way for significant improvements in privacy protection mechanisms.

REFERENCES

Blackman, S., and R. Popoli. 1999. *Design and Analysis of Modern Tracking Systems*. Artech House Boston.

Chen, X., and J. Pang. 2012. "Measuring query privacy in location-based services". In *Proceedings of the 2nd ACM Conference on Data and Application Security and Privacy*, 49–60. San Antonio, TX, USA: ACM.

Deng, M., K. Wuyts, R. Scandariato, B. Preneel, and W. Joosen. 2011. "A privacy threat analysis framework: supporting the elicitation and fulfillment of privacy requirements". *Requirements Engineering* 16 (1): 3–32.

Díaz, C. 2006. "Anonymity Metrics Revisited". In *Anonymous Communication and its Applications*, edited by S. Dolev, R. Ostrovsky, and A. Pfitzmann, Number 05411 in Dagstuhl Seminar Proceedings. Dagstuhl, Germany: Internationales Begegnungs- und Forschungszentrum für Informatik.

Díaz, C., S. Seys, J. Claessens, and B. Preneel. 2003. "Towards Measuring Anonymity". In *Privacy Enhancing Technologies*, edited by R. Dingledine and P. Syverson, Volume 2482 of *Lecture Notes in Computer Science*, 54–68. Springer Berlin Heidelberg.

Dwork, C. 2008. "Differential Privacy: A Survey of Results". In *Theory and Applications of Models of Computation*, edited by M. Agrawal, D. Du, Z. Duan, and A. Li, Volume 4978 of *Lecture Notes in Computer Science*, 1–19. Springer Berlin Heidelberg.

Eckhoff, D., and C. Sommer. 2014. "Driving for Big Data? Privacy Concerns in Vehicular Networking". *IEEE Security and Privacy* 12 (1): 77–79.

European Parliament 1995. "Directive 95/46/EC". *Official Journal* L 281:0031–0050.

Fellendorf, M., and P. Vortisch. 2010. "Microscopic Traffic Flow Simulator VISSIM". In *Fundamentals of Traffic Simulation*, edited by J. Barceló, Volume 145 of *International Series in Operations Research & Management Science*, 63–93. Springer.

Finn, R. L., D. Wright, and M. Friedewald. 2013. "Seven Types of Privacy". In *European Data Protection: Coming of Age*, edited by S. Gutwirth, R. Leenes, P. de Hert, and Y. Poullet, 3–32. Springer.

Fischer, L., S. Katzenbeisser, and C. Eckert. 2008. "Measuring unlinkability revisited". In *Proceedings of the 7th ACM Workshop on Privacy in the Electronic Society*, 105–110. Alexandria, Virginia, USA: ACM.

Hoh, B., and M. Gruteser. 2005. "Protecting Location Privacy Through Path Confusion". In *Proceedings of the 1st International Conference on Security and Privacy for Emerging Areas in Communications Networks*, 194–205.

Hubaux, J.-P., S. Čapkun, and J. Luo. 2004. "The Security and Privacy of Smart Vehicles". *IEEE Security and Privacy* 2 (3): 49–55.

Kitchenham, B. 2004. "Procedures for Performing Systematic Reviews". Technical Report TR/SE-0401, Keele University.

Krajzewicz, D., G. Hertkorn, C. Rössel, and P. Wagner. 2002. "SUMO (Simulation of Urban MObility); An Open-source Traffic Simulation". In *Proceedings of the 4th Middle East Symposium on Simulation and Modelling*, 183–187. Sharjah, United Arab Emirates.

Ma, Z., F. Kargl, and M. Weber. 2010. "Measuring long-term location privacy in vehicular communication systems". *Elsevier Computer Communications* 33 (12): 1414–1427.

Merali, Z. 2010. "Computational science: Error – why scientific programming does not compute". *Nature* 467 (7317): 775–777.

Nissenbaum, H. 2004. "Privacy as contextual integrity". *Washington Law Review* 79 (1): 119–158.

Pawlikowski, K., H.-D. Jeong, and J.-S. R. Lee. 2002. "On Credibility of Simulation Studies of Telecommunication Networks". *IEEE Communications Magazine* 40 (1): 132–139.

Pfitzmann, A., and M. Hansen. 2010. "A terminology for talking about privacy by data minimization: Anonymity, Unlinkability, Undetectability, Unobservability, Pseudonymity, and Identity Management". v0.34, http://dud.inf.tu-dresden.de/literatur/Anon_Terminology_v0.34.pdf.

Serjantov, A., and G. Danezis. 2003. "Towards an Information Theoretic Metric for Anonymity". In *Privacy Enhancing Technologies*, edited by R. Dingledine and P. Syverson, Volume 2482 of *Lecture Notes in Computer Science*, 41–53. Springer Berlin Heidelberg.

Shokri, R., J. Freudiger, and J.-P. Hubaux. 2010. "A unified framework for location privacy". In *Proceedings of the 3rd Symposium on Hot Topics in Privacy Enhancing Technologies*, 203–214. Berlin, Germany.

Shokri, R., C. Troncoso, C. Díaz, J. Freudiger, and J.-P. Hubaux. 2010. "Unraveling an old cloak: k-anonymity for location privacy". In *Proceedings of the 9th Annual ACM Workshop on Privacy in the Electronic Society*, 115–118. Chicago, Illinois: ACM.

Sommer, C., D. Eckhoff, and F. Dressler. 2010. "Improving the Accuracy of IVC Simulation using Crowd-sourced Geodata". *Praxis der Informationsverarbeitung und Kommunikation* 33 (4): 278–283.

Uppoor, S., and M. Fiore. 2011. "Large-scale Urban Vehicular Mobility for Networking Research". In *Proceedings of the 3rd IEEE Vehicular Networking Conference*, 62–69. Amsterdam, Netherlands.

Westin, A. 1967. *Privacy and freedom*. Atheneum.

Wiedersheim, B., Z. Ma, F. Kargl, and P. Papadimitratos. 2010. "Privacy in Inter-Vehicular Networks: Why simple pseudonym change is not enough". In *Proceedings of the 7th International Conference on Wireless On-Demand Network Systems and Services*, 176–183.

Zimmer, M. 2005. "Surveillance, Privacy and the Ethics of Vehicle Safety Communication Technologies". *Ethics and Information Technology* 7 (4): 201–210.

AUTHOR BIOGRAPHIES

ISABEL WAGNER is a Lecturer in Computer Science at the University of Hull, United Kingdom. She holds a M.Sc. and Ph.D in Computer Science from the University of Erlangen, Germany. Her research focuses on metrics to quantify the effectiveness of privacy protection mechanisms, as well as on privacy-enhancing technologies in vehicular networks, smart grids, and tele-health. Her email address is i.wagner@hull.ac.uk.

DAVID ECKHOFF received his M.Sc. in Computer Science (Dipl.-Inf. Univ.) from the University of Erlangen in July 2009, graduating top of his class. He is currently pursuing his Ph.D. degree at the Chair for Computer Networks and Communication Systems in Erlangen. His research interests include privacy concerns in vehicular networks, ITS simulation, lower layer modeling, and safety applications in IVC environments. He coordinates the study program "Computer Science in Automotive Engineering" and lectures on Vehicular Communication. His email address is David.Eckhoff@fau.de.