Towards privacy-driven design of a dynamic carpooling system

Jesús Friginal, Sébastien Gambs, Jérémie Guiochet, Marc-Olivier Killijian

Towards Privacy-Driven Design of a Dynamic Carpooling System

Jesús Friginal[a], Sébastien Gambs[b],
Jérémie Guiochet[a], Marc-Olivier Killijian[a]

[a] *LAAS-CNRS, 7 Avenue du Colonel Roche, 31400 Toulouse Cedex, France*
[b] *Université de Rennes 1, Avenue du Général Leclerc 35042 Rennes Cedex, France*

Abstract

Dynamic carpooling (also known as instant or ad-hoc ridesharing) is a service that arranges one-time shared rides on very short notice. This type of carpooling generally makes use of three recent technological advances: (i) Navigation devices to determine a driver's route and arrange the shared ride; (ii) smartphones for a traveller to request a ride from wherever she happens to be; and (iii) social networks to establish trust between drivers and passengers. However, the mobiquitous environment in which dynamic carpooling is expected to operate, raises several privacy issues. Among all the personal identifiable information, learning the location of an individual is one of the greatest threats against her privacy. For instance, the spatio-temporal data of an individual can be used to infer the location of her home and workplace, to trace her movements and habits, to learn information about her centre of interests or even to detect a change from her usual behaviour. Therefore, preserving location privacy is a major issue to be able to leverage the possibilities offered by dynamic carpooling. In this paper we use the principles of privacy-by-design to integrate the privacy aspect in the design of dynamic carpooling, henceforth increasing its public (and political) acceptability and trust.

Keywords: Carpooling, privacy, mobility, location-based systems

Email addresses: `jesus.friginal@laas.fr` (Jesús Friginal), `sebastien.gambs@irisa.fr` (Sébastien Gambs), `jeremie.guiochet@laas.fr` (Jérémie Guiochet), `marco.killijian@laas.fr` (Marc-Olivier Killijian)

1. Introduction

The automotive and transport sector present an enormous potential for the exploitation of mobiquitous systems, as seen for example in the research around Vehicular Ad Hoc Networks (VANETS) [1]. According to recent studies [2], this sector is expected to generate business opportunities of $202 Billion until 2020. A lot of such opportunities turn around to the proposal of novel and innovative transports enhancing the user's experience on vehicles while mitigating its impact on the environment.

Carpooling, also known as ridesharing, is a solution that could enable private cars to become part of the public transportation system, thus benefiting both users and environment [3]. Passengers benefit by having an alternative when their usual transportation mode is unavailable, and by possibly eliminating the need for an additional car for occasional use. Drivers benefit by having someone to share the cost of the trip or to gain enough passengers to qualify for high occupancy vehicle lanes. Finally, the environment benefits from a reduction of greenhouse gases emissions. Although different implementations can be found[1], carpooling is typically characterised by four stages: a system registration where some personal information is required, (ii) a negotiation where users finally agree sharing the same car for a determined itinerary, (iii) the trip execution, and (iv) the assessment of the carpooling experience. While first commercial carpooling solutions are already a reality on the Internet, there are privacy aspects that affect the trust of users, consequently limiting their definitive takeoff.

Privacy, defined as the state or condition of being free from being observed or disturbed by other people [4], has become a real concern of users in the last years. Indeed, after the several espionage incidents involving the National Security Agency (NSA) [5], users have started to gain awareness of the importance of their *privacy assets*, that is, the sensible information related to their daily lives. Carpooling systems are not alien to this mistrust. So, their successful exploitation will depend on the ability of system providers to offer not only functionally appealing but also trusty and privacy-aware carpooling solutions.

To date, carpooling systems present two main characteristics:

- Static nature: Trips are scheduled several days in advance, but no interaction between users (such as picking up a new passenger on the

[1] http://www.carpooling.com
http://www.blablacar.com

fly) is possible once the trip has started.

- Centralised infrastructure: Applications rely on fixed preestablished Trusted Third Parties (TTP) in charge of collecting and storing sensitive information from carpooling users (such as their identity, location, usual trips).

Although the static nature of carpooling remains effective, it does not allow the continuous interaction between users. Conversely, a dynamic approach would offer passengers a more flexible ride in just a few minutes. However dynamic carpooling presents a problem of privacy, as it implies a major data exchange between driver and passenger. This increases the risk that attackers may infer private information from both participants.

Regarding the second characteristic, TTPs act as guarantors for trust between users. However, the security of carpooling systems may be compromised by those attackers that are able to gain access to centralised TTPs [6]. The adoption of a distributed architecture would make more difficult for an attacker to find the information about users, as it would be scattered around the network. However the development of distributed TTPs architectures is still a challenging problem that limit the trust of users.

The goal of this paper is to promote the concept of a dynamic and distributed carpooling system, taking into account the non-functional requirements of privacy and trust. In consequence, this paper focuses on answering challenging questions such as (i) what is the critical information required by dynamic carpooling systems to work, (ii) how to exchange such informations between drivers and passengers in an infrastructureless context while offering them trust in the information they receive? and (iii) how to protect the privacy of carpooling users from potential attackers? By addressing such issues, this paper aims at integrating the principles of the privacy-by-design [7] for dynamic carpooling systems.

The rest of this paper is structured as follows: Section 2 presents current challenges in dynamic carpooling systems, highlighting the importance of privacy. Section 3 identifies the most important privacy assets from the viewpoint of users as well as the security and privacy properties that need to be addressed to protect them. Section 4 identifies the most important functional entities of dynamic carpooling and how they relate one another. Then, Section 5 proposes the mechanisms to protect the system from a privacy viewpoint. Section 6 discusses about the need of covering the gap between design and final prototypes. Finally, Section 7 concludes the paper.

2. Challenges of dynamic carpooling

Carpooling was early introduced in 1970s. However it is in the last 5 years when it has gained momentum in our society, specially in countries with high population density. For example, recent studies carried out in China [8], state the increasing interest of society in carpooling solutions, highlighting the cost saving and congestion reduction as the most important benefits. However, to address its dynamic decentralised deployment it is necessary to face some issues. The AMORES project[2] [9] is an initiative that addresses the challenges of dynamic carpooling from a cooperative and distributed way. In overall, they involve improving the performance and the comfort of the carpooling experience while protecting the user privacy.

2.1. Scheduling the meeting point

One of the main bottlenecks of carpooling is scheduling the meeting point. In traditional carpooling this decision is taken jointly between driver and passenger days before the trip. However, in practice, users are rarely warned in case the other party delays since there is not a continuous monitoring of user location. This problem, identified as complex for carpooling [8], has been explored in European projects such as Eureka-Celtic WiSafeCar[3] from a dynamic perspective. This project focused on studying the real-time problems of carpooling such as the scheduling and automated allocation of users, taking into account unforeseen events. In particular, some works have tackled this issue by designing multi-agent-based platforms to perform an optimised and distributed assignment of vehicles to users' queries [10, 11]. However, this solution does not solve the problem of cheating users where for example Alice is aware of arriving 30 minutes late, but just notifies Bob a delay of 5 minutes to force him to wait. To avoid these annoying situation, it would be necessary that Bob could validate Alice's real position to really trust her. Furthermore, this validation would limit the delays caused by the selection of confusing meeting points. The challenge for dynamic carpooling consists in providing mechanisms so that scheduling the meeting point becomes a more reliable task.

2.2. Privacy aspects

The work in [8] revealed privacy risks as an important drawback of current carpooling systems. The fact that the collection and transmission of

[2] http://projects.laas.fr/AMORES
[3] http://wisafecar.gforge.uni.lu

personal data can also be used against the privacy of users, either at the transmission time (e.g., to send unwanted advertisement), or later in the future, is seen by users as an important threat. This problem becomes specially important in dynamic carpooling as messages use the wireless medium and attackers may be equipped with eavesdropping capacities. Currently, there is no universal location privacy mechanism that has reached a consensus in the privacy community. Recently, some cooperative schemes for neighbour position verification were proposed. From the point-of-view of privacy, authors of [12] ensure certain degree of anonymity of communication by relying on a generator of MAC address during the discovery phase in order to obfuscate the identity of the users. However, this protocol assumes that the verifier is trusted (i.e., honest). Indeed, the verifier is granted the privilege to decide which entities are really in her proximity without requiring the help of an external trusted entity to verify the correctness of this proximity map. However, even if the user knows which entity should in principle be responsible for keeping her data private, she has no guarantee other than the promises of this entity that her location data will not be disclosed to other entities (e.g., for instance to a marketing company for a profiling purpose or to nearby shops for targeted advertising). The work in [13] proposes a privacy-preserving location verification mechanism called APPLAUS. However the solution relies on a centralised TTP. The challenge for dynamic carpooling consists in developing distributed TTPs that can replace centralised ones to protect user's privacy.

2.3. Trust between users

The potential disputes between users is a significant problem identified in [8]. This problem has been addressed from a prevention viewpoint in traditional carpooling, i.e., by proposing matching mechanisms to enhance the compatibility between driver and passenger. However, to date, there is a lack of trust mechanisms that may protect users in case of needing to reclaim legal responsibilities to the other party involved in the carpooling activity. For example, if during a police investigation Bob denies having carpooled with her, Alice should be able to prove that she was with Bob during a certain period. Dynamic carpooling has the potential to collect location proofs that could be used as evidences by offended users. However, how to do it while respecting the principle of privacy is a challenge that has not been addressed before from a distributed and collaborative way in the domain of carpooling.

3. Towards private-by-design carpooling systems

Previous challenges limit the commercial exploitation of carpooling. In this paper we propose to address them through the concept of privacy-by-design. Privacy-by-design [7] is a notion that aims to ensuring privacy protection and gaining personal control over one's own information. Among all the benefits it presents, the following are very interesting for carpooling: (i) anticipating and preventing privacy invasive events before they happen; (ii) addressing the privacy of users as strong privacy defaults; (iii) designing privacy as a core functionality, and not as an add-on; and (iv) ensuring all these requirements throughout the entire lifecycle of the data involved.

However, privacy-by-design has been generally criticised [14] due to the lack of methodologies to guide the definition of Privacy Enhancing Technologies (PETs) that match previous requirements.

To address this problem, we pose a private-by-design methodology encompassing the following steps within the domain of dynamic carpooling:

- Identifying the assets that need to be protected
- Matching assets with key privacy and security properties
- Specifying functional and non-functional aspects of the system
- Designing the system architecture interweaving non-functional with functional aspects

3.1. Identifying key assets

This step addresses the identification of the private information (assets) used by carpooling systems. It is obvious that, when asking users in general about their privacy, they will probably agree on its importance for them. Contrarily, it is rather difficult to find a widely-accepted definition of privacy, since it largely depends and affects the personal perspective of each person in a different way. Privacy can be a malleable notion involving a gradation of needs and trust levels. Thus, to identify basic privacy assets, users' preferences need to be identified taking into account a social perspective [15]. Questionnaires can be useful tools to capture the user's viewpoint.

Following this principle, we designed a set of questions intended to help design teams through specific privacy issues in the domain of carpooling. For the sake of representativeness, the questionnaire was designed to be answered by non-expert users, thus trying to capture the opinion of conventional participants, which compound the widest range of users in the domain

of mobile applications. After being introduced to carpooling through the description in Table 1, they were invited to express their opinion about the implications to personal privacy in case any attacker was able to disclose different informations about them. In particular they were asked about their level of concern (not concerned, not very concerned, somewhat concerned and very concerned) of 14 different personal data in case they were leaked. Such data correspond to the first column in Table 2.

Table 1: Description of carpooling provided in the questionnaire.

Carpooling means sharing car journeys, so that more than one person travels in a car. By having more people using one vehicle, carpooling reduces each person's travel costs such as fuel costs, tolls, and the stress of driving. Real-time ridesharing is a service that arranges one-time shared rides on very short notice. This type of carpooling generally makes use of recent technological advances like smartphones with GPS capabilities. Such a system could be the target of attacks, leading to retrieve some of your personal data.

The population of the experiment consisted in 64 volunteers from our laboratory, of which 33 (51.6%) were men and 31 (48.4 %) were women. All of them were in a range from 23 to 50 years old. From the response of users, shown in Table 2, we were able to fusion the 14 personal data into the following 8 basic privacy assets:

- *Identity.* It is any subset of attributes which sufficiently identifies a person within any set of persons. E.g., the full name, pseudonym used in system (nickname), IP address, MAC Address (unique identifier of your device), Personal preferences (Smoker, rock addicted, what car you have, animal lover, etc).

- *Location.* This refers to a particular place or position. Location could be referred as a physical and symbolic information. For example, GPS provides physical location. In contrast, symbolic location encompasses abstract ideas of where something is: e.g., in the city centre.

- *Social relations* between users by considering for instance that two users that are in contact during a non-negligible amount of time share some kind of social link. This information can also be derived from mobility traces by observing that certain users are in the vicinity of each other on a frequent basis.

- *Itinerary.* This can be defined as a detailed plan for a journey, especially a list of places to visit. An itinerary could be referred to as

Table 2: Analysis of the user responses to our survey.

	Not concerned	Not very concerned	Somewhat concerned	Very concerned
One of your past location	29,7% (19)	**35,9% (23)**	26,6% (17)	7,8% (5)
Home address	6,3% (4)	14,1% (9)	39,1% (25)	**40,6% (26)**
List of persons you usually travel using this system	14,3% (9)	**36,5% (23)**	31,7% (20)	17,5% (11)
Movement semantics (e.g. Every Thursday you go to gym)	11,1% (7)	22,2% (14)	**44,4% (28)**	22,2% (14)
Full Name	11,1% (7)	17,5% (11)	**39,7% (25)**	31,7% (20)
Itinerary used for carpooling	11,1% (7)	**42,9% (27)**	33,3% (21)	12,7% (8)
MAC Address (unique identifier of your device)	9,4% (6)	18,8% (12)	**43,8% (28)**	28,1% (18)
Current location (if you are using the system)	1,6% (1)	27,0% (17)	**49,2% (31)**	22,2% (14)
Personal preferences (Smoker, rock addicted, what car you have, animal lover, etc)	23,8% (15)	**38,1% (24)**	27,0% (17)	11,1% (7)
Predict your movements	12,7% (8)	22,2% (14)	**41,3% (26)**	23,8% (15)
Work address	7,9% (5)	31,7% (20)	**39,7% (25)**	20,6% (13)
Name of the persons you travelled with (shared your car)	17,5% (11)	**38,1% (24)**	36,5% (23)	7,9% (5)
Pseudonym used in system (ex: fast_but_safe, etc)	22,2% (14)	**38,1% (24)**	27,0% (17)	12,7% (8)
List of your past itineraries	14,1% (9)	**46,9% (30)**	28,1% (18)	10,9% (7)

both physical and symbolic information because it contains a starting point and an ending point for the journey but also complementary information (deviation, seats, possible preferences).

- *Important places*, called *Points Of Interests* (POIs), which characterise the interests of an user. A POI may be for instance the home or place of work of an user. Revealing the POIs of a particular user is likely to cause a privacy breach as this information may be used to infer sensitive information such as hobbies, religious beliefs, political preferences or even potential diseases.

- *Mobility patterns of an user*. A trajectory described by different locations in time to infere future information. From the mobility patterns, it is possible to deduce other informations such as the mode of trans-

port, the age or even the lifestyle[4].

- *Semantics of the mobility behaviour of an user* from the knowledge of her POIs and mobility patterns. For instance, some mobility models such as *semantic trajectories* [16] do not only represent the evolution of the movements of an user over time but also attach a semantic label to the places visited.

- *Linkable records of the same user*, which can be contained in different geolocated datasets or in the same dataset, either anonymised or under different pseudonyms. For example, the association of the mobility of Alice's car (contained for instance in dataset A) with the tracking of her cell phone locations (recorded in dataset B).

From the previous list, identity and location can be defined as primary assets, since the rest can be derived or composed from them.

3.2. Key properties of security and privacy

Protecting the privacy of user's identity and location is very important for the confident use of dynamic carpooling applications. To properly address this question, it is necessary to reinforce dynamic carpooling with different security and privacy properties for the assets previously identified.

According to [17], *authentication*, *confidentiality* and *non-repudiation* are key properties of security that have a special significance in the domain of privacy. Ensuring identity and location authentication and confidentiality should prevent malicious adversaries from monitoring and tracing the carpooling user's activity, thus improving the level of trust on the system. Additionally, The fulfilment of non-repudiation on the dynamic carpooling scenario should assist third parties, such as the police, to demonstrate the participation of users in dynamic carpooling sessions. From the strict viewpoint of privacy, *unlinkability* is also a desired property complementing previous ones. This property prevents attackers from reconstructing derived assets from basic informations. These properties are listed in Table 3.

3.3. Functional and non-functional aspects of dynamic carpooling

Once identified the key assets of carpooling and the properties to protect them, the goal we pose is how to interweave the PETs required to

[4]See for instance `http://www.sensenetworks.com`.

Table 3: Properties to enhance privacy on dynamic carpooling.

Property	**Description**
Authenticated positioning (AP)	The user's location has been verified. The user is where she claims to be.
Confidential positioning (CP)	The user's location is only revealed to authorised users.
Authenticated identity (AI)	The user's identity is verified. The user is who claims to be.
Confidential identity (CI)	The user's identity is only revealed to authorised users.
Non-repudiation (NR)	Users cannot deny their participation in a carpooling activity.
Unlinkable identity-location (UIL)	An attacker cannot sufficiently distinguish whether an identity is related to a location or not.

protect the system without affecting the functional aspects offered by dynamic carpooling system. To properly address this problem, the rest of the paper is structured as follows. For simplicity, Section 4 will focus on the basic functional aspects of a dynamic carpooling system, while Section 5 will introduce non-functional ones in the design of the system following the principles stated in this Section.

4. Functional specification of dynamic carpooling

The system specification is in charge of detailing the bounds of our dynamic carpooling approach, which implies defining the entities concerned, the role they will play as well as the processes in which they are involved. This Section presents the functional specification of dynamic carpooling to tackle this problem.

For the sake of clarity, the application resulting from applying our specification should be able to run in different embedded devices: from vehicle applications to mobile devices.

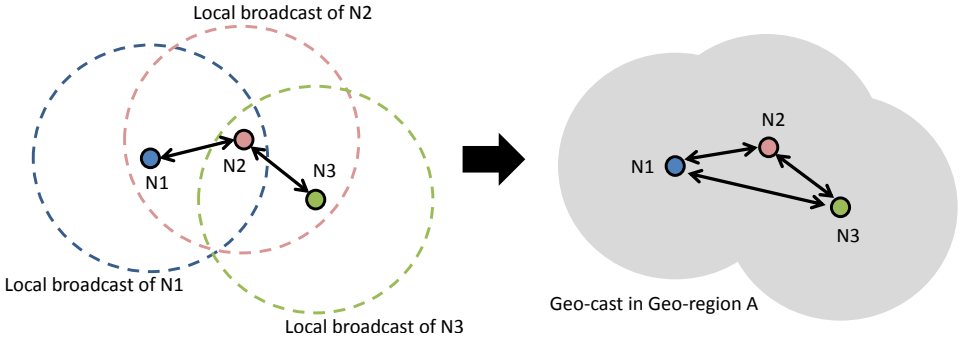
4.1. Basic geo-primitives

The infrastructureless nature required for dynamic carpooling leads us to tackle the complexity of communications between distant users. So, the definition of primitives seems essential for the conception of the system.

Such primitives, which are a geo-located variant of group communication in the domain of classical distributed systems [18], are addressed to ensure the consistency of the group membership views (i.e., who currently belongs to the group), and how nodes join and leave the group. Accordingly, we introduce the notion of *geo-casting* to define a service such that, at the

moment of the group creation, nodes located in the same vicinity or within a particular delimited area communicate with one another to create a original group structure or *geo-region*. As Figure 1 shows, *geo-casting* is built on top of traditional local broadcast and routing protocols to enable a user to spread a message to a particular spatial neighbourhood. By using this primitive, a user can send a message along a particular direction such as towards her destination. Thus, longer trips might be facilitated using "multihop" matches in which passengers change cars to reach their final destination.

Figure 1: Abstraction of geo-casting. Although nodes N1 and N3 are out of range, they can geo-cast messages to each other through node N2.



The notions of *geo-casting* and *geo-region* will be, as we will seen in further sections, a very useful building block for privacy-by-design dynamic carpooling systems.

4.2. Functional carpooling entities

Regardless the final implementation of the system, the intrastructureless nature of dynamic carpooling, requires the consideration of three main entities.

- Driver: The driver corresponds to an active user of the carpooling application that possesses a vehicle and is willing to carpool with another user on some part of her itinerary. A user declares himself as a driver by activating the corresponding mode on the carpooling application. By extension in the description of this use case, the concept of driver also refers to the personal device owned by the driver that is in charge of running the carpooling application.
- Passenger: The passenger is an active user of the carpooling application that does not use his own vehicle and would like to be match

with a driver whose itinerary matches her own mobility desiderata. A user declares himself as a passenger by activating the associated mode on the carpooling application. By extension in the description of this use case, the concept of passenger also refers to the personal device owned by the passenger that is in charge of running the carpooling application.

- Intermediary: The intermediary is a passive user in charge of forwarding messages from one geo-region to another, consequently enabling distant users to communicate. The key role of intermediary nodes is to improve the connectivity of the carpooling network. The intermediary node is implicitly activated with the application, in such a way that, a user can be involved in a trip negotiation as a driver, while helping other nodes to communicate in the background.

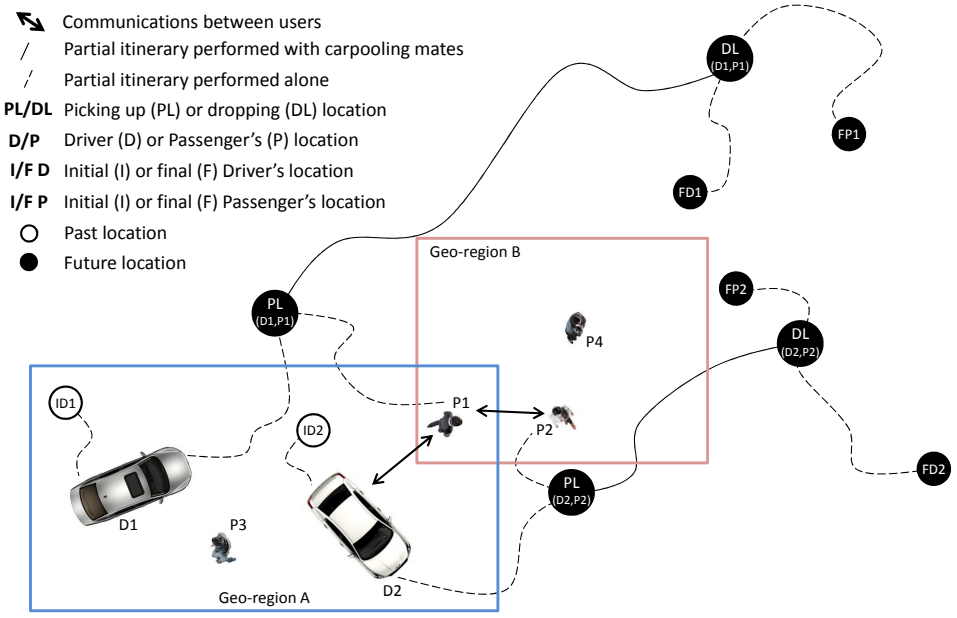
4.3. Dynamic interactions

Previously introduced entities must participate in an interactive dynamic process that enables users to deploy the carpooling activity. Given the social dimension of the problem, this is a flexible process that may present different (but valid) versions. In this Section we propose a basic approach that could be used as a model to inspire future designers.

Figure 2 presents an example of scenario showing the dynamic interactions between the different entities of the system. In this scenario we find different participants located at different geo-regions. Driver $D1$, Driver $D2$ and Passenger $P3$ are positioned at geo-region A whereas Passenger $P2$ and Passenger $P4$ are located at geo-region B. Finally, Passenger $P1$ is located at an intersection of both geo-regions.

Let us focus on the interaction between $D2$ and $P2$ to illustrate the general behaviour of dynamic carpooling. $D2$ has planned an itinerary in the carpooling system to go from $ID2$ (her initial position) to $FD2$ (her final position). From that moment, the application geo-casts messages requesting carpooling mates. At a given moment, such a request arrives to $P1$, which belongs to geo-regions A and B. However, although she is not interested in such a proposal, as she is already involved in a trip negotiation with $D1$, she geo-casts the message for other concerning users in geo-region B. At that time, Passenger $P2$ wishes to travel to $FP2$ and geo-casts a trip request. Such a message arrives to $P1$, who as stated before, is also located at geo-region B. Accordingly, $P1$ can process if there is a potential matching between users $D2$ and $P2$ (i.e., if they share the same destination and trip preferences, e.g., they are non-smokers). Given the successful matching

Figure 2: Example of dynamic carpooling.



between both users, $P1$ geo-casts this itinerary proposal towards $D2$ and $P2$. Once received, users $D2$ and $P2$ must consider accepting or not such a proposal and acknowledge directly to the other part their choice. The acceptance of the trip involves performing a common itinerary for $D2$ and $P2$ that could satisfy both parts. Such itinerary requires $D2$ to pick up $P2$ at a given location ($PL(D2, P2)$) and dropping her at another ($DL(D2, P2)$). Thus, if both users accept sharing this trip, they should meet at the location proposed. When the journey in common finishes, $D2$ drops $P2$ and they can continue alone toward their final destination ($FD2$ and $FP2$ respectively). At that moment, they could rate the experience with their carpooling mates to recompute their reputation for future carpooling activities.

5. A privacy-by-design proposal for dynamic carpooling

Protecting the privacy of the process shown in the prior Section is essential for the confident use of dynamic carpooling. In this paper we propose the use of protection mechanisms against generic attackers (either malicious users or intruders) with wireless packet eavesdropping capabilities. The provision of inter-nodes-cooperation incentives and trust mechanisms for small

mobile devices is a real challenge when designing PETs against this attacker model.

5.1. On protecting the privacy of assets

One of the first questions that naturally arises when dealing with dynamic carpooling is how a particular user can convince others about the validity of its current position while preserving her privacy. Without the aim of being exhaustive, hereafter we present a list of PETs that could assist designers to reinforce the privacy properties of the system:

- Asymmetric cryptography: a cryptographic system requiring two separate keys, one of which is secret (or private) and one of which is public. Although different, the two parts of the key pair are mathematically linked. One key locks or cyphers the message, and the other unlocks or uncyphers the cyphered text. The use of the private keys to cypher messages ensures confidentiality, while signing messages through public key ensures authentication.

- Static distribution of cryptographic keys: the static distribution of cryptographic keys, typically done by centralised authorities, prevents the digital identity of users from being compromised during the carpooling activity. Separating distribution from usage reinforces the properties of authentication and confidentiality.

- Challenge-response handshaking: this is a negotiation process to validate the authentication between two entities, before normal communication over the channel begins.

- Dynamic use of watchdog nodes: watchdogs are passive nodes in charge of monitoring the activity of the network. However, they can also participate as third parties in authentication processes to legitimate the authorship of messages.

- Random pseudonyms: the use of random pseudonyms is a way to reinforce unlinkability. All the random pseudonyms belonging to a user are generated from the same seed, so that the user's identity can be verified by adequate parties while it remains confidential for other parties.

Table 4 matches previous PETs with the properties they may reinforce. However, it is necessary to articulate the way in which they may collaborate together. The rest of this Section introduces such PETs within the design of the carpooling system.

Table 4: PETs for dynamic carpooling systems.

Protection mechanism	**Authenticated positioning (AP)**	**Confidential positioning (CP)**	**Authenticated identity (AI)**	**Confidential identity (CI)**	**Non-repudiation (NR)**	**Unlinkable identity-location (UIL)**
Asymmetric cryptography keys	✓	✓	✓	✓	✓	
Static distribution of cryptographic keys	✓	✓	✓	✓		
Challenge-response handshaking		✓				
Dynamic use of watchdog nodes	✓		✓			
Random pseudonyms				✓		✓

5.2. Non-functional carpooling entities

To preserve user privacy, our dynamic carpooling proposal relies on the separation of trust. This paradigm requires dislocating the storage of sensitive assets to avoid having a single point of failure. In other words, the separation of trust ensures that compromising just one entity node will not be sufficient to have access to link the user's identity with her location. This paradigm neutralises our attacker model. For the sake of this property, we have considered the definition of the following non-functional entities:

- Prover: A user who wants to convince others of her current location while preserving her anonymity behind an unique pseudonym.
- Witness: Also known in the literature as watchdog node, is a user in the local broadcast vicinity of a prover who acknowledges her legitimate location. Such an acknowledgement, or location proof, is signed with a private key and can be read using a group public key shared by all the carpoolers.
- Verifier: A user who checks the location announced by a prover in a distant geo-region and acknowledged by their witnesses.
- Certification authority (CA): A trusted third party to assign credentials to new users of the carpooling application. The role of this entity

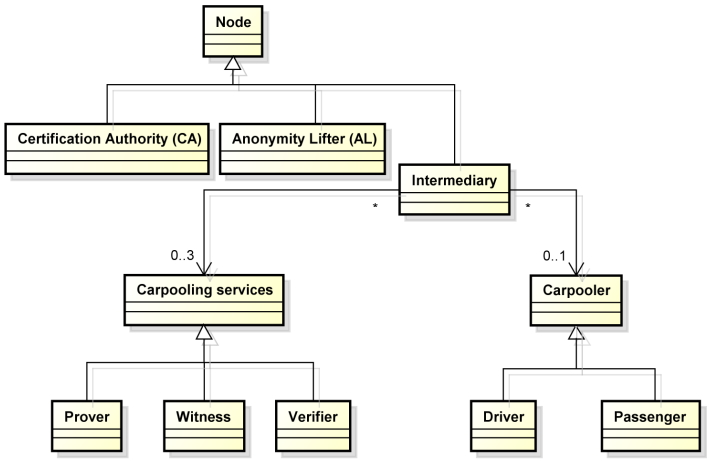
is limited to the registration time, but ensures the property of authenticated identity during the execution of carpooling activities.

- Anonymity lifter (AL): This entity is a trusted third party used to lift the anonymity of users when legally required (e.g, to sue a given user). This is a passive entity in the sense that it does not participate in carpooling activities. However, the AL is essential to ensure non-repudiation.

Following the example initiated in Section 4.2, $D2$ would play the role of the prover and $P2$ would be the verifier in the interaction $D2 \rightarrow P2$ previously shown in Figure 2. Conversely, in the interaction $P2 \rightarrow D2$, $P2$ would be the prover and $D2$ the verifier. In both cases, $P1$ would be the witness of $D2$ and $P2$.

The consideration of non-functional entities in dynamic carpooling plays an adequate role against malicious behaviours. On one hand, since witnesses only accept to acknowledge the location of users in their vicinity, malicious drivers or passengers announcing to be located in a different geo-region will not be able to obtain an acknowledgement from them. So, both prover and witness should be necessarily compromised for the success of such an attack. On the other hand, a malicious verifier could not re-use the location proof generated by a witness to maliciously acknowledge a different user, since, for this, it is necessary to obtain the private key of the witness. Figure 3 represents all the entities involved in dynamic carpooling. This class diagram presents all dynamic nodes as intermediary nodes.

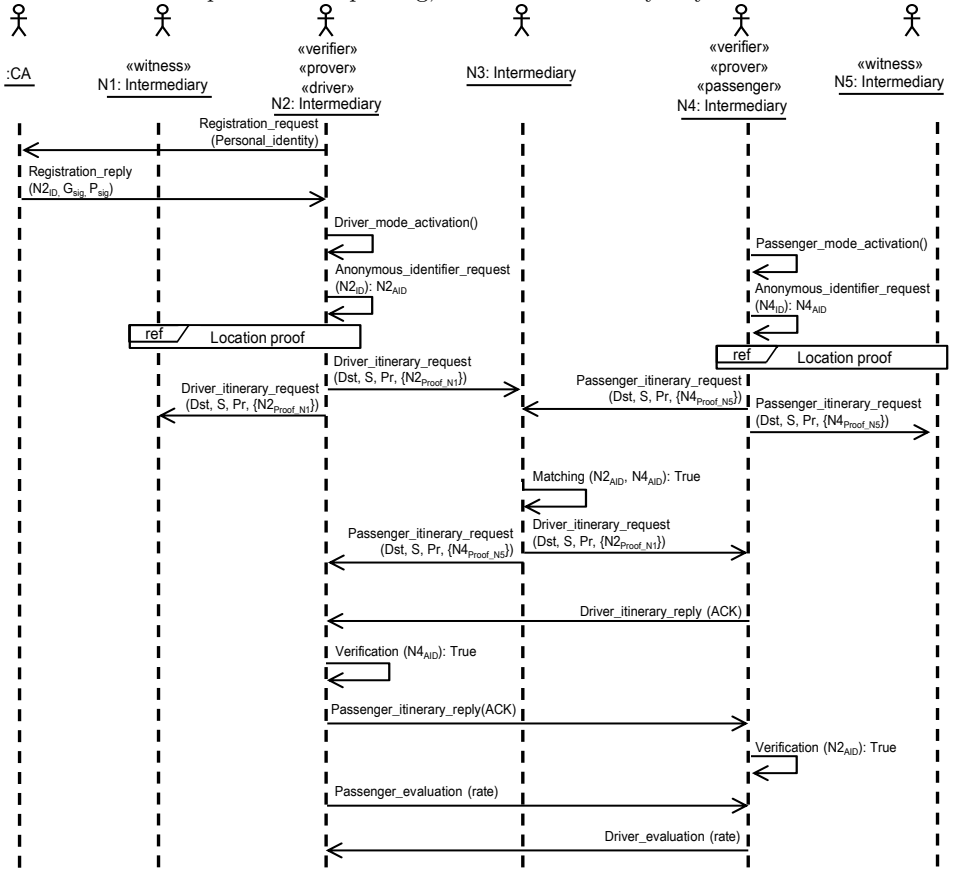
Figure 3: UML class diagram of dynamic carpooling entities.



5.3. Carpooling process

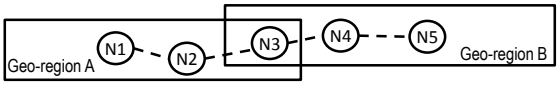
Figure 4 presents a sequence diagram describing the interactions for all the (functional and non-functional) entities involved.

Figure 4: Sequence diagram of privacy-by-design dynamic carpooling. Since we mainly focus on the online process of carpooling, the role of the anonymity lifter has been omitted.



The topology considered to instantiate our sequence diagram is depicted in Figure 5. After being registered in the system, the user receives an identifier and a pair of private-group signatures [19]. The system identifier, e.g. $N2_{ID}$ in the case of node N2, can be understood as a seed to generate multiple pseudonyms unequivocally linked to N2, which reinforces the property of *unlinkability*. The pair of private-group signatures (P_{Sig}, G_{Sig}) is a version of traditional public-private asymmetric model. Yet, in this case, the public key (P_{Sig}) is shared by the group of carpooling users. The use of

Figure 5: Topology considered in our sequence diagram.

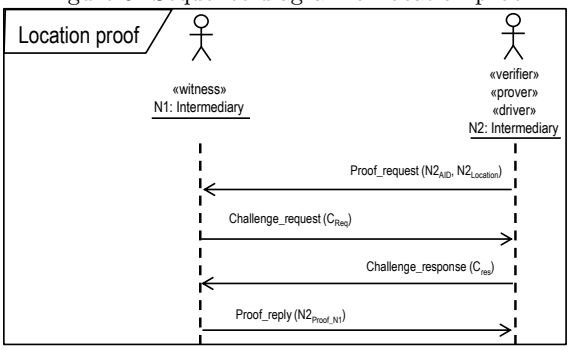


these signatures ensures the properties of *authenticated identity and identity* and *confidential identity and location*.

Once users registered, the online activity of the carpooling application begins. First of all, it is necessary to select the user mode. Since this process is similar for both drivers and passengers let us focus on the case of the driver. After N2 activates the "driver mode", the application solicits a new anonymous pseudonym for the user ($N2_{AID}$). Before being able to send any itinerary request, N2 must obtain a proof of her location.

The location proof process is addressed in Figure 6. This process concerns N2 as a prover and N1 as a witness. The process is initiated by N2, which broadcasts a location proof request to the nodes in her vicinity. N1, which is within the same radio range of N2 receives such a request. Such a request, signed with N2's private key, includes $N2_{AID}$ as well as her current location ($N2_{Location}$). When such a request is received by N1, the witness can disclose the content of the packet using the group signature and verify that the announced location corresponds to a position in her vicinity. After verifying the real existence of N2 through a challenge-response handshaking, N1 sends N2 her location proof $N2_{Proof\_N1}$. Such a proof is the result of signing both $N2_{AID}$ and $N2_{Location}$ with the public key of N1. Since various nodes may have replied to N2's location proof request, N2 may concatenate them to obtain a more consistent proof of her location. In no case the personal identity of provers is revealed to witnesses.

Figure 6: Sequence diagram of location proof.



After processing her location proofs, N2 can geo-cast an itinerary request to geo-region A composed of her current location proof, her destination (Dst) as well as the number of seats available (S) and her preferences to share the car (Pr). In our example, the arrival of packet N1 and N3 happens in parallel to the itinerary request made by passengers. In this case, the itinerary request of N4 is received by N3 and N5. When the itinerary requests of N2 and N4 are processed by N3, which plays the role of intermediary node, the system runs a matchmaking algorithm between those drivers and passenger with similar preferences. After that, N3 contacts each pair of compatible driver and passenger, in our case the pair N2-N4. After receiving the itinerary proposal, N2 and N4 verify the location proof of the other party. So, they decrypt the proof using the group key. This process, executed locally, determines if the anonymous pseudonym associated to the delivered location is correct or not. If it is correct, the user will decide whether accepting the proposed trip or not. If both driver and passenger accept the trip, the system sends an acknowledgement packet to both of them (possibly with some additional information) to close the deal.

Finally, N2 and N4 send their feedback to recompute the reputation of users for further trips.

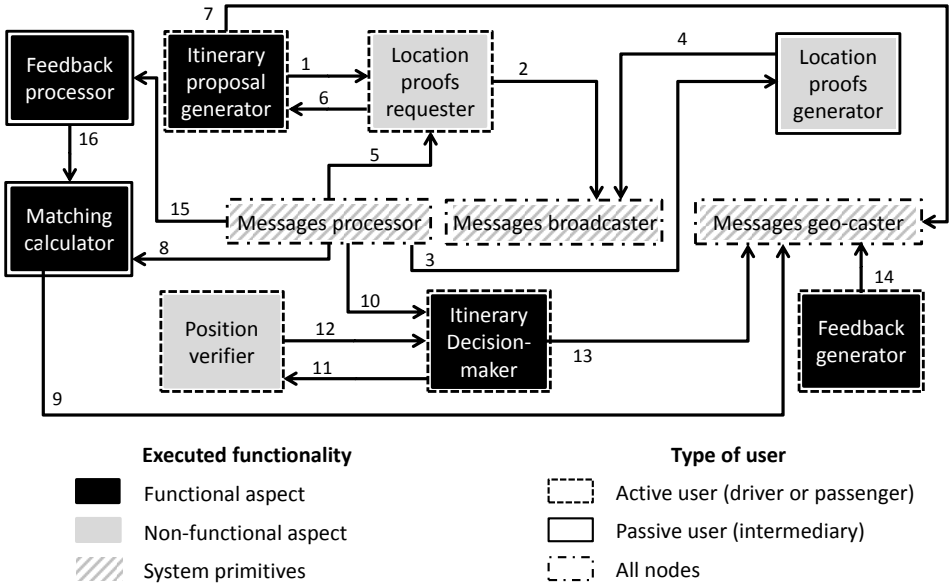
5.4. A specification-compliant architecture

To finish our privacy-by-design proposal, it is necessary to map the carpooling process into a specification-compliant architecture that can be implemented. Figure 7 provides an abstract viewpoint of such an architecture encompassing different modules, which basically can be characterised according to two different perspectives.

Regarding their functionality, architecture modules can represent functional aspects (if they address the visible performance pursued by the application, such as the exchange of itinerary requests or the matching computation) or non-functional ones (if they deal with properties, such as those related to trust and privacy). Additionally, it is possible to find primitives that will be used by both types of modules, like those related to network management and packet sending and processing.

On the other hand, modules can be characterised by the type of entities in charge of executing them. In this case, we classified them according to the type of functional entity (see Section 4.2). Thus, we have considered active users (if they play the role of drivers or passengers), passive ones (intermediary nodes) or all of them, if both types of nodes are concerned. Taking these concepts into account, it is easy to characterise architecture modules as follows.

Figure 7: Dynamic carpooling architecture.



- The *itinerary proposal generator*, is the point through which active users access the application. This is a functional module in charge of building the trip proposals. For that, this module calls the *location proofs requester* (see interaction #1) and the *location proofs generator* (see interaction #7).

- The *location proofs requester* is a module executed by active users with a non-functional purpose. Its goal is to build a location proof request. It calls the *message broadcaster* (see interaction #2) and the *itinerary proposal generator* (see interaction #6).

- The *messages broadcaster* is a communication primitive responsible for spreading packets in the node's vicinity to exchange information concerning the location proof. It is used by all nodes.

- The *messages processor* is a communication primitive in charge of dispatching incoming carpooling packets to concerned modules. Concretely, it receives and handles carpooling information addressed to the *location proofs generator* (see interaction #3), the *location proofs requester* (see interaction #5), *the matching calculator* (see interaction #7), the *itinerary decision maker* (see interaction #8) and the

feedback processor (see interaction #13). It is used by all nodes.

- The *location proofs generator* is a non-functional module used by passive users to forge location proofs. Once generated, the proof is sent using the *messages broadcaster* (see interaction #4).
- The *messages geo-caster* is a communication primitive used by all the nodes to send and forward itinerary requests between the distant nodes of the carpooling network.
- The *matching calculator* is a functional module executed by passive users to determine the geo-spatial matching and the social affinity between active users. If this rate is high enough, the *messages broadcaster* is invoked to notify concerned users of the matching (see interaction #8).
- The *itinerary decision maker* is a functional module used by active users to take a choice about the acceptance or rejection of a carpooling trip. The execution of this module requires the invocation of the *position verifier* (see interaction #9).
- The *position verifier* is a non-functional module executed by active users to authenticate the received itinerary proposals. Once the authentication determined, it calls back the *itinerary decision maker* (see iteration #10).
- The *feedback generator* is a functional module used by active users to notify the system about their opinion of the trip for the computation of further trips. The result of this feedback is geo-casted using the *messages geo-caster* (see interaction #12).
- The *feedback processor* is a functional module used by passive users to update the reputation of active users. Such reputation will be used in the *matching calculator* (see interaction #14).

6. Future work

The future of dynamic carpooling requires to pay attention to more practical aspects combining privacy, resilience and trust issues with performance aspects. In consequence, it is necessary to find a feasible trade-off between the different dimensions of the problem taking into account the resource-limited nature of small mobile devices. This point is essential so that the

application becomes usable, and thus attractive for people. Studying the technological challenges behind the implantation of our methodology is one of our future axis. So, the provision of a middleware to implement proposed techniques and algorithms is necessary to evaluate its usability in real scenarios. Currently, thanks to the AMORES project, we are in contact with some companies in the domain of Location-Based Services (LBS) development that are interested in the practical deployment of dynamic carpooling. We think that this type of synergy is positive for the goal we pursue.

Going beyond, we argue that dynamic carpooling would be an excellent gateway to cover the gap between private cars and public transportation, thus approaching the notion of intermodal mobility to integrate public transport, commercial carriers and peer-to-peer services (including both carpooling, trains, buses, bikes, etc.). The goal is thinking of dynamic carpooling not as a competitor for public or B2C transportation systems but rather as the perfect complement. Thus, we ambition at extending the specification and design addressed in this paper to include the intermodal paradigm, thus exploring the privacy-by-design concerns from social cars to the future of social travelling. Following the same strategy as in this paper, it would be necessary to pose to what extent the assets identified for carpooling match to different types of transports, which are the privacy properties concerning such assets or how to define the interfaces between the different transports in such a way the privacy-by-design concept is preserved.

7. Conclusions

Given their autonomous capability to freely move, vehicles in general, and cars in particular will play an important role in the development of socially-inspired mobility services in a near future.

This paper has explored the potential of cars to inspire new types of social interactions. Dynamic carpooling is a novel social-inspired service offering drivers and passengers the possibility to easily share a car. Thus, this paper has proposed a privacy-by-design approach enabling users to interact in a proactive, private and trusty way. To the best of our knowledge, there is no such type of model. We strongly believe that, in the social car model, with plenty of interacting moving nodes, the mobility and the geographical distribution should be explicitly taken into account for the design of privacy-aware systems.

The mobiquitous environment, in which the devices are mobile and geo-located and where the services are location-based, raises several privacy

issues due to the fact that the geo-located device belongs usually to an individual (or a group of person such as a family) and as such its location corresponds to the location of its owner(s). Therefore, preserving location privacy is a major challenge limiting the possibilities offered by the mobiquitous setting to provide efficient and trusted geo-services. Our approach, which follows the privacy-by-design principle, integrates the privacy aspect in the design of the system, henceforth increasing its public (and political) acceptability and trust.

This aspect becomes even more important if we take into account that nowadays, over half the world's population already lives in urban areas and by 2050 this is expected to reach 70% [20]. With an increasing population, the emphasis to research new alternatives to efficiently reduce traffic jams and carbon emissions in cities is only going to grow. As our planet becomes more populated, our cities need to evolve smartly. Vehicles sharing is not only going to be a convenient way of travelling in the future, it is going to become a necessity.

Acknowledgements

This work is partially supported by the ANR French project AMORES (ANR-11-INSE-010) and the Intel Doctoral Student Honour Programme 2012.

References

[1] Saleh Yousefi, Mahmoud Siadat Mousavi, and Mahmood Fathy. Vehicular ad hoc networks (VANETs): challenges and perspectives. In *6th International Conference on ITS Telecommunications Proceedings*, pages 761–766. IEEE, 2006.

[2] Albrecht Fehske, Gerhard Fettweis, Jens Malmodin, and Gergely Biczok. The global footprint of mobile communications: The ecological and economic perspective. *IEEE Communications Magazine*, 49(8):55–62, 2011.

[3] Matthew Barth and Kanok Boriboonsomsin. Real-world carbon dioxide impacts of traffic congestion. *Transportation Research Record: Journal of the Transportation Research Board*, 2058(1):163–171, 2008.

[4] Yves Deswarte and Carlos Aguilar Melchor. Current and future privacy enhancing technologies for the internet. In *Annales des télécommunications*, volume 61, pages 399–417. Springer, 2006.

[5] Leslie Cauley. NSA has massive database of Americans phone calls. *USA today*, 11(06), 2006.

[6] Hector Marco-Gisbert and Ismael Ripoll. Preventing brute force attacks against stack canary protection on networking servers. In *International Conference on Network and Computer Applications (NCA)*, pages 243–250, 2013.

[7] Ann Cavoukian et al. Privacy by design: The 7 foundational principles. *Office of the Information and Privacy Commissioner*, 2011.

[8] Wenhui Xin, Shunying Zhu, Hong Wang, Yongfei Yan, and Jining Xiong. Analyzing early market potential and strategies for carpooling in china: A case study of wuhan. In *Management and Service Science, 2009. MASS '09. International Conference on*, pages 1–4, 2009.

[9] Christian Artigues, Yves Deswarte, Jérémie Guiochet, Marie-José Huguet, Marc-Olivier Killijian, David Powell, Matthieu Roy, Christophe Bidan, Nicolas Prigent, Emmanuelle Anceaume, Sébastien Gambs, Gilles Guette, Michel Hurfin, and Frédéric Schettini. Amores: an architecture for mobiquitous resilient systems. In *Proceedings of the 1st European Workshop on AppRoaches to MObiquiTous Resilience*, ARMOR '12, pages 7:1–7:6. ACM, 2012.

[10] M. Sghaier, S. Hammadi, H. Zgaya, and C. Tahon. An optimized dynamic carpooling system based on communicating agents operating over a distributed architecture. In *11th International Conference on Intelligent Systems Design and Applications (ISDA)*, pages 124–129, 2011.

[11] C. Bonhomme, G. Arnould, and D. Khadraoui. Dynamic carpooling mobility services based on secure multi-agent platform. In *Global Information Infrastructure and Networking Symposium (GIIS), 2012*, pages 1–6, 2012.

[12] M. Fiore, C. Ettore Casetti, C. Chiasserini, and P. Papadimitratos. Discovery and verification of neighbor positions in mobile ad hoc networks. *IEEE Transactions on Mobile Computing*, 12(2):289–303, 2013.

[13] Zhichao Zhu and Guohong Cao. Toward privacy preserving and collusion resistance in a location proof updating system. *IEEE Transactions on Mobile Computing*, 12(1):51–64, 2013.

[14] Seda Gürses, Carmela Gonzalez Troncoso, and Claudia Diaz. Engineering privacy by design. *Computers, Privacy & Data Protection*, 2011.

[15] P.E. Agre and M. Rotenberg. *Technology and Privacy: The New Landscape*. Mit Press, 1998.

[16] Sébastien Gambs, Marc-Olivier Killijian, and Miguel Núnez del Prado Cortez. Show me how you move and I will tell you who you are. In *Proceedings of the 3rd ACM SIGSPATIAL International Workshop on Security and Privacy in GIS and LBS*, SPRINGL '10, pages 34–41. ACM, 2010.

[17] Algirdas Avizienis, J-C Laprie, Brian Randell, and Carl Landwehr. Basic concepts and taxonomy of dependable and secure computing. *IEEE Transactions on Dependable and Secure Computing*, 1(1):11–33, 2004.

[18] Marc-Olivier Killijian, Raymond Cunningham, René Meier, Laurent Mazare, and Vinny Cahill. Towards group communication for mobile participants. In *Proceedings of ACM Workshop on Principles of Mobile Computing (POMC 2001)*, pages 1–8, 2001.

[19] Haeryong Park, Seongan Lim, Ikkwon Yie, Kitae Kim, and Junghwan Song. Strong unforgeability in group signature schemes. *Comput. Stand. Interfaces*, 31(4):856–862, June 2009.

[20] United Nations Expert Group Meeting on Population Distribution, Urbanisation, Internal Migration and Development. An Overview of Urbanisation, Internal Migration, Population Distribution and Development in the world. 2008.