Abstract

The emergence of resource-rich mobile devices and smart vehicles has paved the way for Urban Sensing. In this new paradigm, users sense their environment and become part of an unprecedented large-scale network of sensors, with extensive spatial and temporal coverage, that enables the collection and dissemination of real-time information, practically, from anywhere. Urban sensing can facilitate the deployment of innovative applications that can address the ever-growing concerns for citizens’ well-being. Nevertheless, the openness of such systems (ideally anyone can participate) and the richness of the data users contribute unavoidably raise significant concerns for both the security of urban sensing applications and the privacy of the participating users.

In this thesis we consider different urban sensing application domains: vehicular communication networks, intelligent transportation systems and environmental monitoring applications. We begin with a detailed analysis of the security and privacy requirements of these applications domains. Our objective is to protect users from the system (by ensuring their anonymity and privacy) and urban sensing systems from malicious users (by holding malicious users accountable of their actions). This is not straight-forward; anonymity may tempt malicious user behavior, compromising the reliability of the entire urban sensing system.

Towards that, we design and implement secure and privacy-preserving identity management systems that can accommodate these requirements. We demonstrate their efficiency, practicality, and scalability through extensive experimental evaluations. Furthermore, we formally evaluate formally their security and privacy-preserving properties.
Acknowledgments

First, I would like to thank my supervisor Prof. Panos Papadimitratos for his continuous guidance and support. I am grateful for his efforts throughout these three years. Thank you for giving me the chance to work in Academia. I would also like to thank my colleague, and friend, Dr. Thanassis Giannetsos. Your help and support was more than important; you are the colleague I wished I would work with. I also want to thank all my colleagues at the Networked Systems Security group (NSS) and at the Laboratory of Communication Networks (LCN). I feel lucky to be part of such a stimulating, and friendly, academic environment. I am thankful to all my collaborators and especially to Nikos Alexiou, Marcello Laganà and Mohammad Khodaei for their help.

I am very grateful to my wife, Elena. It would require another thesis to express how much your understanding, support and, more importantly, your love helped me going. Thank you for always being there, for giving me the correct advice and for making my life beautiful. Finally, I would like to express my greatest gratitude to my family: Giorgos, Alexandra, Eleni, Giorgos, Alexandros and Marigina. You have been by my side from the first day of my (academic) life. I owe you everything.
Contents

1 Introduction 1
  1.1 Thesis Structure 2

2 Urban Sensing Application Domains 3
  2.1 Vehicular Communications (VC) and Intelligent Transport Systems (ITS) 3
  2.2 Other Urban Sensing Applications 4

3 Secure & Privacy Preserving Urban Sensing Systems: Requirements and Adversaries 7
  3.1 Threats & Adversaries 9

4 Current Status of Urban Sensing 11

5 Summary of Original Work 15
  5.1 SEROSA: SERvice Oriented Security Architecture for Vehicular Communications 15
  5.2 SPPEAR: Security & Privacy-Preserving Architecture for Mobile Crowd-Sensing Applications 16
  5.3 Secure and Privacy-Preserving Smart-phone based Traffic Information Systems 16
  5.4 Publications not included in this thesis 17

6 Future Research Directions 19

Bibliography 21
The emergence of resource-rich mobile devices and smart vehicles is changing the landscape of mobile sensing. These platforms are becoming the main user computing and communication platform, by incorporating multiple embedded sensors (e.g., accelerometers, gyroscopes, GPSs, cameras) and communication modules (e.g., 802.11p, 3/4G cellular networks). With more than 6 billion mobile subscriptions worldwide [1], and smart vehicles on the way [2], we can now sense our environment and gather valuable data of unprecedented quality and quantity, practically from everywhere. These facts set the way for the emerging paradigm of Urban Sensing where users become the focal point of the sensing infrastructure by collecting targeted information about their environment and interactions [3].

Extensive user participation is the key to the success of such systems as it can provide a sufficient and continuous influx of contributions. However, as mobile devices and smart vehicles collect data from the user’s immediate context, privacy concerns emerge; user contributed data tagged with spatio-temporal information (i.e., time and location) could reveal a user’s personal activities, home/work location, health condition [4] and daily routines [5], [6]. These concerns are aggravated in the light of the recent revelations of mass surveillance [7]. Towards that, users must be able to contribute to urban sensing systems without compromising their privacy.

Privacy protection is a necessary condition for user participation, but, it is not (by itself) a sufficient one. Indeed, the research community has identified the importance of incentivizing users so that they provide a continuous influx of contributions. However, it is necessary to provide incentives in a privacy preserving manner.

At the same time, the very openness of urban sensing systems (ideally any user operating sensing capable devices should participate [8]) renders them vulnerable to scores of inadvertent, or deliberate, contributions of faulty data (measurements) that threaten the trustworthiness [9] of the system. Detecting data falsification attacks is not straight-forward; anonymity may tempt malicious user behavior,
compromising the reliability of the entire sensing system. This sets a challenging trade-off: although users should be able to contribute to the urban sensing system in an anonymous manner, they should be held, at the same time, accountable of their actions. To thwart malicious, system-harming or privacy violating behavior, we need solutions that provide some form of accountable anonymity: users are tied to their actions, without disclosing their identity.

Security and privacy can facilitate urban sensing systems to reach their full potential [10]. However, we still lack comprehensive security and privacy solutions as state-of-the-art works address facets of the problem at hand (see Chapter 4 and [11]); they focus on privacy and security without considering either accountability [12]–[16] or the user’s dual role (i.e., producing and consuming information) [17]; or they try to enable mass participation by linking incentives to user contributions without considering the possible privacy implications [18]–[20].

This sets the scope and the motivation of this thesis: how can we design broad (in terms of user participation) urban sensing systems that are, at the same time, reliable, accountable and privacy-preserving?

1.1 Thesis Structure

The structure of this thesis is as follows: Chapter 2 presents use cases of urban sensing applications followed by a discussion on their security and privacy requirements (Chapter 3). Chapter 4 presents the current status of urban sensing systems with a focus on their security and privacy. Chapter 5 provides a summary of the papers in the context of this thesis and Chapter 6 concludes this thesis with a discussion on future research directions.
Urban Sensing Application Domains

Urban sensing lies at the intersection of various application domains [21]. Although the work presented in this thesis is largely independent of application domains, we use ITS\(^1\) and environmental monitoring systems as motivational use-cases.

2.1 VC and ITS

Automated vehicles featuring powerful embedded platforms and antennas paved the way for the deployment of a new trajectory of safety, traffic efficiency and comfort applications. Such applications are built by leveraging Cooperative Awareness Messages (CAMs) and Decentralized Environmental Notification Messages (DENMs) broadcasted by vehicles. With these messages vehicles can exchange information with respect to their status (e.g., location, velocity) and their environment (e.g., traffic congestion, road-hazards). VC render feasible such applications by enabling Vehicle-to-Vehicle (V2V) and Vehicle-to-Infrastructure (V2I) communications over cellular networks or the 5.9 GHz wireless channel (i.e., 802.11p).

The security and the privacy of VC have been extensively investigated not only by the academia but also by standardization bodies (e.g., the Car2Car Communication Consortium (C2C-CC) [22]) and the automotive industry. These efforts have converged to a comprehensive framework for the protection of V2V and V2I communications. More specifically, according to the Institute of Electrical and Electronics Engineers (IEEE) and European Telecommunications Standards Institute (ETSI) standards, each vehicle has a unique, long-term identifier, that is a public key and the corresponding private key [23], [24]. The public key is bound to the vehicle’s long-term identity by means of certificates. Each vehicle is also provided with a set of anonymous credentials, the pseudonyms, which correspond to ephemeral asymmetric key-pairs. In contrast to the vehicle’s long-term identity, pseudonyms

\(^1\)Note the distinction between ITS enabled by vehicle to infrastructure communication (e.g., cellular systems) and the more general technology of VC that also leverages vehicle-to-vehicle communications
contain no information that can identify the vehicle. The credential management of the different keys is performed by a number of Certification Authorities (CAs) organized into a Vehicular Public-Key Infrastructure (VPKI).

Intelligent Transport Systems also aim at enhancing transportation safety by collecting traffic data, producing traffic estimates, and providing drivers with feedback and location-based recommendations. The increasing smart-phone penetration, in combination with the wide coverage of cellular networks, defines an unprecedented large-scale network of sensors, with extensive spatial and temporal coverage, able to serve as traffic probes for ITS [25]. Nevertheless, as these systems require fine-grained location information, the privacy and the whereabouts of the contributing participants must be protected. This need for privacy is intensified in the context of smart-phone based traffic information systems; smart-phones already reveal a great deal of, possibly sensitive, information to the cellular operators (e.g., user identity, coarse grained location and calling/messaging actions among others) and thus, it is important that the introduction of smart phone based traffic information systems does not, under no circumstances, deteriorate the exposure of user’s private and sensitive information.

2.2 Other Urban Sensing Applications

Urban Sensing systems span well beyond transportation safety and efficiency. Indeed, scores of urban sensing applications, focusing on different aspects of urban life, have surfaced; environmental monitoring [26], [27], assistive health-care [28], [29] and public safety [30] applications among others.

In contrast to VC, where there is tangible industrial interest, such applications have emerged as the result of research efforts and, thus, they are not regulated by any standards; different works assume different system models and stake-holders. Nonetheless, they all converge on a basic set of actors: Users, campaign administrators and the supporting infrastructure (Figure 2.1).

Users are operators of mobile devices (e.g., smart-phones, tablets, smart vehicles and wearable platforms), equipped with multiple embedded sensors and navigation modules (GPS, GLONASS and Galileo). These platforms also possess transceivers that can communicate over wireless local area (i.e., 802.11a/b/g/n/p) and cellular networks (3rd Generation (3G) and Long Term Evolution (LTE)).

Campaign Administrators are organizations, public authorities or, even, individuals [31] that initiate targeted data collection campaigns, by recruiting users and distributing the descriptions of sensing tasks to them. Such descriptions should contain (at least) the following information:

- Sensing Modalities: Specification of the sensors that devices must employ within the scope of a given task. Participants may be requested to provide either raw sensor data (e.g., current temperature) or some statistic over the data (e.g., minimum, maximum, average, median temperature). It could also be the case that multiple modalities need to be combined for increasing the
utility of the contributed data. For example, it is easier to assess the quality of noise-level measurements if the orientation of the phone is known [32].

- **Area of Interest**: The locality within which the mobile devices must contribute sensed data. The area of interest can be defined either explicitly, by means of geographic fields (e.g., a set of coordinates forming a polygon on the map), or implicitly, by leveraging annotated geographic areas (e.g., city of Stockholm).

- **Task Duration**: The time interval during which users are expected to contribute data.

- **Reporting Frequency**: The periodicity and the conditions under which mobile devices submit the values of their sensing modalities. For example, a task could request devices to submit data every 1 min or whenever the value of the sensed phenomenon is within (or not) some predefined range (e.g., noise pollution exceeds 140 dB).

- **Eligibility Criteria**: Sensing-capable devices are, still, fragmented when considering their sensing and computational capabilities (for example not all phones have a barometer). Furthermore, sensing tasks could require contributions from specific user groups (e.g., cyclists). As a result, a task description must explicitly define the conditions under which user participation is allowed or desired.
• **Rewards**: Although not mandatory, the provision of rewards and incentives to the participating users is motivated by numerous research efforts [18]–[20], [33], [34]. In case a sensing task provides incentives or rewards to participating users, they should be mentioned in its description.

The back-end infrastructure (Fig. 2.1) is responsible for supporting the lifecycle of an urban sensing task by facilitating the recruitment and the registration of users and offering credential management, authentication, and access control services. Some works [13], [14], [31] require a centralized component that is responsible for the collection and the aggregation of user submitted data. Others [12], suggest a decentralized approach where mobile devices can be directly queried by campaign administrators instead of submitting their data to a central aggregation server.
Secure & Privacy Preserving Urban Sensing Systems: Requirements and Adversaries

To design trustworthy urban systems, one has to cater to the security and privacy requirements of all involved actors. In a nutshell, the general design goal should be to protect the users from the system (e.g., ensuring user privacy) and to protect the system from the users (e.g., by guaranteeing the security of the communications and ensuring the validity of user data) [11], [35]. A first step towards that is a clear definition of the security and privacy requirements such systems must meet (Fig. 3.1):

- **R1: Privacy Preserving Participation**: Users should participate in urban sensing systems, anonymously, i.e., without revealing their identity. We refer to both user-specific data (e.g., name, email address) and device identifiers, such as the International Mobile Subscriber Identity (IMSI) and the International Mobile Station Equipment Identity (IMEI), or, the vehicle’s registration number (in the case of vehicular networks).

  The level of the achieved anonymity strongly depends on the anonymity set (i.e., the number of users that participate in a given task or are within the same geographical area).

  During their participation in sensing campaigns, users are expected to contribute fine-grained location measurements. Successive (anonymous) location updates from user devices still reveal spatial and temporal correlations that can be used as indirect identifiers. Such correlations can be exploited by tracking techniques [36] to reconstruct a user’s whereabouts and, thus, infer frequently visited places, e.g., home or workplace. In such cases, user de-anonymization could be easy [37]. To mitigate such attacks, unlikability
is necessary. More specifically, it should be hard for an observer to link together actions performed by the same user and, thus, reconstruct (parts of) her whereabouts.

Furthermore, user participation must be unobservable: no external observer (e.g., cellular providers or ISPs) should be able to deduce that an anonymous user has (or will) participated in the urban sensing system.

- **R2. Privacy-Preserving, Resilient Incentive Mechanisms and Fairness**: Users should be able to receive rewards for their participation without associating themselves with the data they contributed. Such an incentive mechanism should be resilient: misbehaving or selfish users should not be able to exploit them to increase their utility without making the desired contributions.

- **R3. Communication integrity, confidentiality and authentication**: All system entities should be authenticated and their communications should be protected from any alteration and disclosure to unauthorized parties.

- **R4. Authorization and Access Control**: Users should act according
to the system’s specifications and policies. To enforce such policies, access control and authorization services must be in place.

- **R5. Accountability**: Offending entities (i.e., users, infrastructure components and campaign administrators) should be held accountable for actions that could disrupt the system operation or harm the users. Accountable urban sensing systems should provide the necessary means to shun misbehaving users and filter out their faulty contributions.

Ensuring the aforementioned properties separately is relatively straight-forward. Nevertheless, ensuring all of them at the same time is a challenge due to their inherent contradictions. For example, achieving anonymous and unobservable participation hardens the task of enforcing accountability. Similarly, it is hard to mediate and authorize anonymous participants.

### 3.1 Threats & Adversaries

Similarly to any networked system, adversaries in urban sensing systems can be classified based on whether they are part of the system (i.e., internal) or not (i.e., external).

External adversaries are principals that do not possess system credentials. Their aim is to degrade the performance of the system and disrupt its operation. They can eavesdrop, modify and replay messages\(^1\). Furthermore, they can target the availability of the system by launching clogging, e.g., jamming and D(D)oS attacks. Since external adversaries have no security associations (with the system), their disruptive capabilities are relatively limited.

In contrast, internal adversaries can launch more sophisticated attacks. Any system entity can pose as an internal adversary: users and the system infrastructure itself. Users might be malicious or selfish. Malicious ones can exhibit arbitrary behavior, completely deviating from the expected functionality and protocols. They might also attempt to impersonate other users or even try to simultaneously pose as multiple, authorized, ones (i.e., act as a sybil). Furthermore, users could contribute faulty but seemingly valid data, thus, polluting the data collection process and degrading the usefulness of the collected data. Such orchestrated pollution attacks target the overall result of a sensing task based on a statistical analysis of the contributed data. For example, users in an air pollution monitoring application may fake a series of sensor readings (e.g., CO\(_2\), CO) in an attempt to manipulate the aggregative result and avoid further consequences (e.g., pollution fines).

Selfish users aim at minimizing their effort-to-utility ratio. Such a behavior is relevant to participatory sensing applications that entail incentive and reward mechanisms. Selfish users may try to exploit such procedures to increase their utility without offering the required contributions to the system. More specifically,

\(^1\)Although such attacks can be easily mitigated by means of encryption and digital signatures.
they might *leech*\(^2\) a sensing task: offer the minimum contributions that suffice for being awarded a task’s reward. Of course, users might quit a task for various (non-selfish) reasons: They might move away from a task area of interest or run out of resources (e.g., battery).

Adversarial behavior is not limited only to users; infrastructure components can misbehave too. More specifically, infrastructure components can be *honest-but-curious*: They execute correctly the various protocols but they are curious to learn private user data. Multiple curious entities might *collude* to de-anonymize users. Finally, malicious infrastructure entities can exhibit arbitrary behavior.

\(^2\)This term is borrowed from Peer-to-Peer (P2P) systems and describes users that exhibit a *hit-and-run* approach and download content without sharing it.
Current Status of Urban Sensing

In this section we survey the state-of-the-art literature with respect to the application domains discussed in Chapter 2 and the requirements defined in Chapter 3.

For VC, pseudonyms are the de-facto standard for privacy protection [38]–[40]. Towards that, there has been a significant body of work that examines their effectiveness [41], efficiency [42]–[45] along with other aspects of credential management [46], [47]. Besides pseudonyms, the literature discusses alternative privacy protection schemes such as Group Signatures [48]–[51].

In the context of ITS, the Mobile Century team presented a privacy-preserving smart-phone based Traffic Information System (TIS) [52] based on a sampling scheme, known as Virtual Trip Lines, and a privacy-aware algorithm that defines the road points at which samples should be submitted. Their system uses different encryption keys for the participating entities: a client application on mobile phones, an ID-proxy server, the traffic server and a Virtual Trip Lines (VTL) generator. Communications between the mobile clients and the traffic server or the VTL generator are done through the ID-proxy server, which is responsible for user authentication. Each location update, submitted by the mobile clients to the traffic server, contains the location and the identity of the phone each encrypted with a different key. The identity of the device is encrypted with a symmetric key known to the ID-proxy. Similarly, location information is encrypted with the public key of the traffic server and thus, it is accessible only by it. These keys are established and pre-installed on the mobile during its initialization. The scheme achieves privacy under the assumption that the traffic and ID-proxy servers will not collude and it requires a third-party for the identity management. This point introduces an extra burden for deployment and it requires that such a third-party establishes trust relations with the clients participating in the traffic information system.

AnonySense [13] was one of the first works to propose a general-purpose security and privacy architecture for urban sensing systems. It tessellates geographical areas to achieve statistical $k$-anonymity [53]; individuals cannot be identified within a set of $k$ users assumed to reside in the same area at a given moment in time. This ap-
proach prevents inference attacks aiming to link reports back to users. As a second layer of protection, AnonySense aggregates (at least \( \lambda \)) user reports before sending them to the campaign administrator. To achieve user anonymity, AnonySense leverages group signatures. Nevertheless, filtering out past and faulty contributions of such offending users requires also the de-anonymization of benign reports.

PoolView [54] is a privacy-preserving architecture that enables mobile clients to perturb private measurements before sharing them. To thwart inference attacks that leverage the correlation of user data, the authors propose an obfuscation model. The novelty of this scheme is based on the fact that although private user data cannot be obtained, statistics over them can be accurately computed. PoolView considers only privacy of data streams without covering aspects such as security, accountability and data trustworthiness.

PEPSI [14] prevents unauthorized entities from querying the results of sensing tasks with provable security. It is based on a centralized solution that focuses on the privacy of data queriers; i.e., entities interested in sensing information. PEPSI does not consider aspects such as accountability and privacy preserving incentive mechanisms and it does not ensure privacy against cellular Internet Service Providers (ISPs). Furthermore, as PEPSI leverages Identity Based Cryptography, and more specifically the scheme presented in [55], it inherits its key escrow properties. This aspect, besides placing strong trust on the key generating entity, weakens the non-repudiation properties of the system [56].

TAPAS [16] presents a participatory sensing framework that enables privacy-preserving user contributions. Additionally, it considers data trustworthiness by employing redundancy; multiple users (termed as replicators) collect and report data from the same geographical areas. The more the users engage in the data collection process, the higher the trustworthiness of the collected data is (under the assumption that the majority of the nodes is benign).

In [9] the authors propose a system that leverages a Trusted Platform Module (TPM) to ensure the integrity of sensor readings. This approach renders the system resilient against malicious users that aim to (collectively) pollute the data collection process by submitting faulty measurements or by launching sybil attacks. Nevertheless, the use of TPM cannot protect the privacy of participating users.

PEPPeR [17] protects the privacy of the parties querying mobile nodes (and not of the mobile nodes themselves), by decoupling the process of node discovery from the access control mechanisms used to query these nodes. PRISM [15] focuses on the secure deployment of sensing applications and does not consider privacy. It follows the push model for distributing tasks to nodes: service providers disseminate tasks to mobile devices (according to criteria such as their location). This approach enables timely and scalable application deployment, but harms user privacy since service providers have knowledge of the device locations.

The works presented in [12] and [57] propose decentralized frameworks for storing data on user devices (instead of some central authority) and for privacy-preserving disclosure of user trajectory information. Since these works focus mostly
Security & Privacy Requirements

<table>
<thead>
<tr>
<th>Works</th>
<th>R1</th>
<th>R2</th>
<th>R3</th>
<th>R4</th>
<th>R5</th>
</tr>
</thead>
<tbody>
<tr>
<td>AnonySense [13]</td>
<td>✓</td>
<td>✗</td>
<td>✓</td>
<td>✗</td>
<td>✓</td>
</tr>
<tr>
<td>PoolView [54]</td>
<td>✓</td>
<td>✗</td>
<td>✓</td>
<td>✓</td>
<td>✗</td>
</tr>
<tr>
<td>Pepsi [14]</td>
<td>✓</td>
<td>✗</td>
<td>✓</td>
<td>✓</td>
<td>✗</td>
</tr>
<tr>
<td>TAPAS [16]</td>
<td>✓</td>
<td>✗</td>
<td>✓</td>
<td>✗</td>
<td>✓</td>
</tr>
<tr>
<td>[9]</td>
<td>✗</td>
<td>✗</td>
<td>✓</td>
<td>✓</td>
<td>✗</td>
</tr>
<tr>
<td>PEPPeR [17]</td>
<td>✓</td>
<td>✗</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>[12], [57]</td>
<td>✓</td>
<td>✗</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>[58], [59]</td>
<td>✓</td>
<td>✗</td>
<td>✓</td>
<td>✗</td>
<td>✓</td>
</tr>
<tr>
<td>[60]</td>
<td>✓</td>
<td>✓</td>
<td>✗</td>
<td>✗</td>
<td>✓</td>
</tr>
</tbody>
</table>

Table 4.1: Comparative Analysis of State-of-the-art

on location privacy, they do not consider aspects such as accountability, data-trustworthiness and user incentivization.

Ahmadi et. al. present a scheme for regression modeling that enables efficient and privacy-preserving transformation of user data [58]. Similarly, in [59] the authors present a framework for privacy-preserving collection, analysis and aggregation of user data. This approach also enables regression analysis over private data. As both works focus on data privacy, they do not consider the security and data-trustworthiness aspects of urban sensing.

Significant efforts have been made on the provision of incentives to stimulate user participation [18]–[20], [33], [34]. These works leverage mechanisms such as auctions, dynamic pricing, monetary coupons, service quotas and reputation accuracy. However, they do not consider user privacy and, thus, can leak sensitive information by linking the identity of users with the data they contribute. The approach presented in [60] differs from the aforementioned works as it considers user privacy. Nevertheless, instead of proposing privacy protection measures, the authors suggest that as user privacy exposure increases, users should receive better services (e.g., QoS-wise) and rewards as a compensation.

Table 4.1 presents a comparative analysis of all the research efforts discussed in this section with respect to the requirements presented in the previous chapter.
Summary of Original Work

5.1 SEROSA: SERvice Oriented Security Architecture for Vehicular Communications

Stylianos Gisdakis, Marcello Laganà, Thanassis Giannetsos, Panos Papadimitratos
Presented at: Conference on Vehicular Communication Networks (IEEE VNC), Boston, USA, 2013.

Summary

This work presents a secure and privacy-preserving service-oriented architecture for VC; SEROSA provides a comprehensive set of identity management and access control services while guaranteeing the security and privacy requirements of VC. Our approach emerges as a synthesis of VC (e.g., IEEE 1609.2) and Internet standards such as the Security Assertion Markup Language (SAML). We devise a comprehensive set of security and privacy requirements and a full implementation of our system. Additionally, we extensively assess its efficiency, practicality, and dependability. Overall, SEROSA significantly extends the state of the art and serves as a catalyst for the integration of vehicles into the vast domain of Internet-based services.

Contribution

The author of this thesis developed the framework jointly with the fourth author. He further elaborated the design and the implementation of the system, along with the performance analysis, with the help of the other authors. The paper was written by all authors.
5.2 SPPEAR: Security & Privacy-Preserving Architecture for Mobile Crowd-Sensing Applications

Stylianos Gisdakis, Thanassis Giannetsos, Panos Papadimitratos

Summary

In this paper we present SPPEAR: a comprehensive secure and privacy preserving architecture for Participatory Sensing systems, that systematically addresses all key PS aspects, i.e., privacy, security, accountability and incentives provision. More specifically, SPPEAR (i) is scalable, dependable and applicable to any type of PS application, (ii) guarantees user non-identifiability and offers stronger privacy protection, (iii) limits participation to legitimate users in a fully accountable manner, (iv) efficiently shuns out offending users without, necessarily, revealing their identity, (v) is resilient to compromised and colluding PS entities, and (vi) can support various incentive mechanisms in a privacy-preserving manner. We provide a full-blown implementation of our system, on real mobile devices, and extensively assess its efficiency and practicality. Furthermore, we present a formal analysis of the achieved security and privacy properties.

Contribution

The author of this thesis designed the system along with the second and the third authors. He also carried out the system implementation the formal analysis, the simulations and the performance evaluation. The article was written by all three authors.

5.3 Secure and Privacy-Preserving Smart-phone based Traffic Information Systems

Stylianos Gisdakis, Vasileios Manolopoulos, Sha Tao, Ana Rusu, Panos Papadimitratos
Submitted to: IEEE Transactions on Intelligent Transportation Systems (second revision).

Summary

In this work, we systematically address the traffic estimation and the security and privacy aspects of smart-phone based TISs. We present the first instance of smartphone based TIS and assess its accuracy through Global Positioning System (GPS) traces in the presence of traffic estimation errors and for different values of location reporting rates and accumulation frames. Furthermore, we leverage cellular
providers and existing telecommunication standards, and with the use of state-of-the-art cryptographic schemes, converge them into a comprehensive security and privacy preserving architecture resilient against offending users and misbehaving TISs entities. We formally assess the security and privacy properties of the system and demonstrate its efficiency and practicality through extensive evaluations.

**Contribution**

This work is a continuation of prior work [25], [61] of the second and the third authors. The author of this thesis enhanced the design and the implementation of the security and privacy architecture, jointly with the second and the last authors. He also carried out the formal analysis of the system, the simulations and the performance evaluation. The paper was written primarily by the author of this thesis and the last author of the paper.

**5.4 Publications not included in this thesis**


---

1 The introduction of this thesis is based on this publication.
Future Research Directions

The work presented in this thesis makes a step towards secure and privacy-preserving architectures for Urban Sensing applications. Nevertheless, there is still a plethora of research challenges that require further investigation. In the rest of this section we present our future research directions.

Trust in Open-Access Systems: Data-trustworthiness

In all open-access systems (i.e., systems where anyone that can get involved should contribute data) questions regarding the trustworthiness of user data are unavoidably raised. Although some works touch upon the data-trustworthiness aspects of information security, we are far from conclusive and convincing answers. Moreover, well known solutions from the quiver of the security and privacy research community, i.e., cryptography and privacy mechanisms, have proved to be insufficient. Indeed, assessing the trust of data generated by (anonymous) users, forming complex and ephemeral networks, is not a straightforward task.

Meeting this challenge requires work on three different manifestations of trust; (i) user (default) trust, (ii) context-specific trust and (iii) dynamic trust [62]. The default user trust depends on the attributes of a user’s computing platform (e.g., mobile device, smart vehicle). For example, users carrying devices updated with the latest firmware versions, are better protected (from attacks) and, thus, should be considered more trustworthy.

User devices should be trusted only for tasks that they can execute; this is where the notion of context-specific trust is relevant. For example, devices without accelerometers and/or inertia sensors cannot be trusted for, e.g., traffic congestion monitoring.

Users’ trustworthiness needs to be updated based on the overall quality (and quantity) of their contributions. Users with a “well-behaved” history can have a higher degree of trust (but not unconditionally). Nonetheless, what happens when trusted users operate untrusted devices?
Hybrid Approach: Providing Composable Security and Privacy Architectures

User privacy depends on how user data are collected, stored and used. This becomes even more important for data containing (and thus, revealing) personal information. Despite extensive privacy research, current works do not cope with all privacy aspects as they are, usually, tailored to specific application scenarios; they assume the use (in many cases) of trusted centralized entities that collect user data. However, the privacy implications of having a centralized repository hosting sensitive information are far from negligible. The user’s personal context that can be inferred from such multimodal sensor streams is still an open challenge.

To address these concerns, current research points towards user-centric approaches that enable users with full control over their own personal data. Thus, it is up to their discretion what statistics they are going to share and with whom. Apart from the data trustworthiness issues (discussed earlier), the downside of this decentralized set-up is the need for information discovery. More specifically, how can consumers discover the source of information they have to query for a specific dataset? This may require some form of centralized infrastructure.

Both approaches (i.e., centralized and decentralized) have distinct merits and operate under different assumptions. Nevertheless, their synthesis can yield numerous advantages; we can leverage decentralized architectures that enable large-scale dissemination of information in a peer-to-peer manner, in combination with centralized approaches that offer identity management and accountability, thus, ensuring the system’s trustworthiness.