Ensuring Privacy in Self Participatory Devices

Jyothi Karjagi¹, Mr. B.N. Veerappa², Dr. Mohamed Rafi³

¹Department of, Computer Science, UBDTCE Davangere, Visvesvaraya Technological University, India
²Department of, Computer Science, UBDTCE Davangere, Visvesvaraya Technological University, India
³Department of, Computer Science, UBDTCE Davangere, Visvesvaraya Technological University, India

1jyothikarjigi@gmail.com; 2bnveerappa@yahoo.co.uk; 3rafimohamed17@gmail.com

Abstract— Participatory Sensing is an emerging computing paradigm that enables the distributed collection of data by self-selected participants. It allows the increasing number of mobile phone users to share local knowledge acquired by their sensor-equipped devices, e.g., to monitor temperature, pollution level or consumer pricing information. While research initiatives and prototypes proliferate, their real-world impact is often bounded to comprehensive user participation. If users have no incentive, or feel that their privacy might be endangered, it is likely that they will not participate. In this article, focus on privacy protection in Participatory Sensing and introduce a suitable privacy-enhanced infrastructure. First, we provide a set of definitions of privacy requirements for both data producers (i.e., users providing sensed information) and consumers (i.e., applications accessing the data). Then, we propose an efficient solution designed for mobile phone users, which incurs very low overhead.

Keywords—“WSN”, “Participatory Sensing”, ”Cryptography”, ”PEPSI”, ”IBE”

I. Introduction

In the last decade, researchers have envisioned the outbreak of Wireless Sensor Networks (WSNs) and predicted the widespread installation of sensors, e.g., in infrastructures, buildings, woods, rivers, or even the atmosphere. This has triggered a lot of interest in many different WSN topics, including identifying and addressing security issues, such as data integrity, node capture, secure routing, etc. On the contrary, privacy has not really been a concern in WSNs, as sensors are usually owned, operated, and queried by the same entity. (For instance, the National Department of Transportation deploys sensors and collects traffic information related to national highways). On the other hand, the proliferation of mobile phones, along with their pervasive connectivity, has propelled the amount of digital data produced and processed every day. This has driven researchers and IT professionals to discuss and develop a novel sensing paradigm, where sensors are not deployed in specific locations, but are carried around by people. Today, many different sensors are already deployed in our mobile phones, and soon all our gadgets (e.g., even our clothes or cars) will embed a multitude of sensors (e.g., GPS, digital imagers, accelerometers, etc.). As a result, data collected by sensor-equipped devices becomes of extreme interest to other users and applications. For instance, mobile phones may report (in real-time) temperature or noise level; similarly, cars may inform on traffic conditions.
This paradigm is called Participatory Sensing (PS) – sometimes also referred to as opportunistic or urban sensing [3]. It combines the ubiquity of personal devices with sensing capabilities typical of WSN. As the number of mobile phone subscriptions exceeds 5 billions, PS becomes a cutting-edge and effective distributed-computing (as well as business) model. We argue that PS appreciably expands the capabilities of WSN applications, e.g., allowing effective monitoring in scenarios where the setup of a WSN is either not economical or not feasible.

However, its success is strongly related to the number of users actually willing to commit personal device resources to sensing applications, and thus, to associated privacy concerns. Observe that sensing devices are no longer “dull” gadgets, owned by the entity querying them. They are personal devices that follow users at all times, and their reports often expose personal and sensitive information. Consider, for instance, a PS application like http://www.gasbuddy.com/ where gas prices are monitored via user reports, and information announced by participants inevitably exposes their current and past locations, hence, their movements. If users have no incentive in contributing sensed data or feel that their privacy might be violated, they will (most likely) refuse to participate. Thus, not only traditional security but also privacy issues must be taken into account.

In this article, we focus on privacy protection in PS. We define privacy in this new context, present a privacy-enhanced PS infrastructure, and elaborate on a number of desirable features which constitute challenging research problems. Proposed privacy-protecting layer can be easily adopted by available PS applications to enforce privacy and enhance user participation.

II. Literature Survey

Participatory Sensing Projects, In the last few years, Participatory Sensing initiatives have multiplied, ranging from research prototypes to deployed systems. Due to space limitations we briefly review some PS application that apparently exposes participant privacy (e.g., location, habits, etc.). Each of them can be easily enhanced with our privacy-protecting layer. Interested readers may find a larger list of PS applications at [4]. Quake-Catcher [1] aims at building the world’s largest, low-cost strong-motion seismic network by utilizing accelerometers embedded in any internet-connected device. Kim et al. use the power of PS for meaningful places (e.g., home, office, etc.) discovery. PS has been shown to be an effective mean to monitor levels of air pollution, noise pollution and water quality. PS to aid health care providers in patient monitoring has been investigated in this paper.

Privacy, Only little attention has been paid to arising privacy issues in PS. The authors of [2] study privacy in participatory sensing relying on weak assumptions: they attempted to protect anonymity of Mobile Nodes through the use of Mix Networks. (A Mix Network is a statistical-based anonymizing infrastructure that provides k-anonymity – i.e., an adversary cannot tell a user from a set of k). However, Mix Networks are unsuitable for many PS settings. They do not attain provable privacy guarantees and assume the presence of an ubiquitous WiFi infrastructure used by Mobile Nodes, whereas, PS applications do leverage the increasing use of broadband 3G/4G connectivity. In fact, an ubiquitous presence of open WiFi networks is not realistic today nor anticipated in the next future. By contrast our work aims at identifying a minimal set of realistic assumptions and clear privacy guarantees to be achieved with provable security.

The work in studies privacy-preserving data aggregation (e.g., computation of sum average, variance, etc). Similarly, presents a solution for community statistics on time-series data, while protecting anonymity (using data perturbation in a closed community with a known empirical data distribution). Finally, aims at guaranteeing integrity and authenticity of user-generated contents, by employing Trusted Platform Modules (TPMs). The main technical challenge in providing provable privacy in participatory sensing infrastructure stems from the simultaneous presence of several mutually untrusted (and potentially unknown) entities, including data producers, data consumers, and Service Providers.
III. System Design

A. Architecture

PEPSI protects privacy using efficient cryptographic tools. Similar to other cryptographic solutions, it introduces an additional (offline) entity, namely the Registration Authority. It sets up system parameters and manages Mobile Nodes or Queries registration. However, the Registration Authority is not involved in real-time operations (e.g., query/report matching) nor is it trusted to intervene for protecting participants’ privacy. Figure 1 illustrates the PEPSI architecture. The Registration Authority can be instantiated by any entity in charge of managing participants registration (e.g., a phone manufacturer). A Service Provider offers PS applications (used, for instance, to report and access pollution data) and acts as an intermediary between Queries and Mobile Nodes. Finally, Mobile Nodes send measurements acquired via their sensors using the network infrastructure and Queries are users or organizations (e.g., bikers) interested in obtaining reports (e.g., pollution levels).

B. Use Case Diagram

Fig. 2: Use Case Diagram of Participant, Querier, Registration Authority and Service Provider.
C. Sequence Diagram

Fig. 3: Sequence Diagram of System

IV. PEPSI Operations

Fig. 4: System Operations
Figure 4 shows how PEPSI work. The upper part of the figure depicts the offline operations where the Registration Authority is involved to register both Mobile Nodes and Queries.

**Queries Registration:** In the example, Queries Q (the laptop on the right side) picks “Temp” among the list of available queries and obtains the corresponding decryption key (yellow key).

**Mobile Node Registration:** Similarly, Mobile Node M (the mobile phone on the left side) decides to report about temperature in its location and obtains the corresponding secret used for tagging (grey key). The bottom part of Figure 2 shows the online operations where the Service Provider is involved.

**Queries Subscription:** Q subscribes to queries of type “Temp” in “Irvine, CA” using these keywords and the decryption key acquired offline, to compute a (green) tag; the algorithm is referred to as TAG(). The tag leaks no information about Q’s interest and is uploaded at the Service Provider.

**Data Report:** Any time M wants to report about temperature, it derives the public decryption key (red key) for reports of type “Temp” (via the IBE() algorithm) and encrypts the measurement; encrypted data is pictured as a vault. M also tags the report using the secret acquired offline and a list of keywords characterizing the report; in the example M uses keywords “Temp” and “Irvine, CA”.

Our tagging mechanism leverages the properties of bilinear maps to make sure that, if M and Q use the same keywords, they will compute the same tag, despite each of them is using a different secret (M is using the grey key while Q is using the yellow one). As before, the tag and the encrypted report leak no information about the nature of the report or the nominal value of the measurement. Both tag and encrypted data are forwarded to the Service Provider.

**Report Delivery:** The Service Provider only needs to match tags sent by Mobile Nodes with the ones uploaded by Queries. If the tags match, the corresponding encrypted report is forwarded to the Queries. In the example of Figure 2 the green tag matches the blue one, so the encrypted report (the vault) is forwarded to Q. Finally, Q can decrypt the report using the decryption key and recover the temperature.

V. Results
Participatory Sensing is a novel computing paradigm that bears a great potential. If users are incentivized to contribute personal device resources, a number of novel applications and business models will arise. In this article we discussed the problem of protecting privacy in Participatory Sensing. We claim that user participation cannot be afforded without protecting the privacy of both data consumers and data producers. We also proposed the architecture of a privacy-preserving Participatory Sensing infrastructure and introduced an
efficient cryptographic solution that achieves privacy with provable security. Our solution can be adopted by current Participatory Sensing applications to enforce privacy and enhance user participation, with little overhead.

This work represents an initial foray into robust privacy guarantees in PS, thus, much remains to be done. Items for future work include (but are not limited to):

1. Protecting query privacy with respect to the Registration Authority. Recall, in fact, that Queries Alice needs to obtain the IBE decryption keys from the Registration Authority, which would then learn Alice’s query interests.

2. Protecting node privacy with respect to the Network Operator. Current technology does not allow to hide users’ locations and identities from to the Network Operator. Hence, it is an interesting challenge to guarantee node anonymity in broadband networks.

3. Addressing collusion attacks, where multiple entities might collaborate in order to violate the privacy of Mobile Nodes or Queries.

4. Improving the syntax of supported query types. In fact, PEPSI so far allows query/report matching based on the tags provided by both Mobile Nodes and Queries. However, PS applications might require more complex queries where Queries are interested in an aggregate of the reports (e.g., average or sum), or even complex query predicates (e.g., comparisons). While simple aggregate function evaluation over encrypted data is viable with available cryptographic techniques (e.g., homomorphic encryption), enabling efficient evaluation of complex predicates remains an open challenge.

References


