PEPSI: Privacy-Enhanced Participatory Sensing Infrastructure*

Emiliano De Cristofaro[1], Claudio Soriente[2]

[1] University of California, Irvine

[2] Universidad Politécnica de Madrid

Abstract

Participatory Sensing combines the ubiquity of mobile phones with the sensing capabilities of Wireless Sensor Networks. It targets the pervasive collection of information, e.g., temperature, traffic conditions, or medical data. Users produce measurements from their mobile devices, thus, a number of privacy concerns – due to the personal information conveyed by reports – may hinder the large-scale deployment of participatory sensing applications. Prior work has attempted to protect privacy in participatory sensing, but it relied on unrealistic assumptions and achieved no provably-secure guarantees. In this paper, we introduce PEPSI: Privacy-Enhanced Participatory Sensing Infrastructure. We explore realistic architectural assumptions and a minimal set of formal requirements aiming at protecting privacy of both data producers and consumers. We also present an instantiation that attains privacy guarantees with provable security at very low additional computational cost and almost no extra communication overhead. Finally, we highlight some problems that call for further research in this developing area.

1 Prelude

Participatory sensing is an emerging paradigm that targets the seamless collection of data from a large number of user-carried devices. By embedding a sensor to a mobile phone, participatory sensing (also called *opportunistic* or *urban* sensing) enables harvesting dynamic information about environmental trends, such as ambient air quality [39], urban traffic patterns [33], monitoring Wi-Fi access points for place discovery applications [25], health-related information [28], parking availabilities [32], sound events [29], earthquakes [47], data reported by automobiles [23], sharing consumer pricing information in offline market [16], etc. (For more information on participatory sensing initiatives, please refer to our project page [1].)

To allow large-scale deployment, researchers are proposing platforms for application developers [14] and devising innovative business models, based on incentive mechanisms, for the capitalization on sensed data [27, 41]. In the last few years, participatory sensing initiatives have multiplied, ranging from research prototypes [32, 25, 33, 37] to deployed systems [26, 39, 44, 34].

Participatory sensing combines the ubiquity of mobile phones with sensing capabilities typical of Wireless Sensor Networks (WSNs). However, it differs in several aspects. *Sensors* are high-end mobile devices, such as smartphones, with much greater resources than traditional WSN sensors. Their batteries can be easily recharged and production cost constraints are not as tight. They are extremely *mobile*, as they leverage the ambulation of their carriers. Moreover, in traditional WSNs, the network operator is assumed to own and query all sensors, while this assumption does not apply to most participatory sensing scenarios. Indeed, mobile devices are *tasked* to participate into gathering and sharing local knowledge; thus, different entities co-exist and might not trust each other.

A typical participatory sensing infrastructure involves (at least) the following parties:

*A preliminary version of this paper appears in the Proceedings of **ACM WiSec 2011**. This is the full version.

- **Sensors:** Installed on smartphones or other wireless-enabled devices, they emit data reports and form the basis of the participatory sensing infrastructure.
- **Carriers:** Usually envisioned as the people carrying their smartphones, they could also be vehicles, animals or any other entity carrying the mobile sensing device. In the rest of the paper, we refer to a sensor and its carrier as a ***Mobile Node***.
- **Network Operators:** They manage the network used to collect and deliver reports, e.g., maintaining the WiFi, GSM, or 3G network infrastructure.
- **Queriers:** They subscribe to specific information collected in a participatory sensing application (e.g., "*temperature readings from all sensors in Irvine, CA*") and obtain corresponding data reports.

Motivation. The number and the heterogeneity of entities involved in participatory sensing prompts a range of new challenges. Unlike in WSNs, sensing devices are no longer "dull" gadgets, owned by the network operator; they are personal devices that follow users at all time, and their reports often expose personal information. Thus, not only traditional security but also privacy issues must be taken into account, as concerns on personal information disclosure are a fundamental obstacle to large-scale deployment. Participatory sensing can pave the way to novel distributed computing scenarios and new business models. However, its success is strongly related to the number of users willing to devote device resources to sensing applications. Thus, relying on large and ubiquitous user participation [6], participatory sensing will become effective only if it can protect privacy of participating entities.

Contribution. Prior work has focused on privacy concerns in participatory sensing and proposed a few solutions to protect privacy of user locations and reports. that, however, introduce unrealistic network assumptions and provide no provably-secure guarantees. On the contrary, we aim at a *cryptographic treatment* of privacy protection in participatory sensing. We investigate realistic architectural assumptions and a minimal set of formal requirements intended to protect privacy of both data producers (i.e., mobile nodes) and data consumers (i.e., queriers). Finally, we provide an instantiation that attains privacy guarantees with provable security at very low additional computational cost and almost no extra communication overhead.

Organization. The rest of the paper is organized as follows. Section 2 reviews previous privacy-enhancing solutions and highlights their limitations. Section 3 presents the PEPSI infrastructure and its privacy requirements, while Section 4 yields an efficient instantiation with provable security. Finally, Section 5 concludes the paper with a list of open problems. Appendix A reviews the Identity-Based Encryption cryptosystem leveraged by our solutions, while Appendix B presents the details of our privacy analysis.

2 Related Work

In the last years, research interest in participatory sensing has ramped-up. Many researchers have highlighted security and privacy challenges [46], [24], [11], but without proposing actual solutions.

Recent proposals in [12] and [22] are—to the best of our knowledge—the only results to address privacy-related problems, hence, they are most related to our work. They aim at protecting anonymity of users, using Mix Network techniques [10], and provide either k-anonymity [49] or l-diversity [31]. They rely on statistical methods to protect privacy and do not achieve provably-secure guarantees. They also provide report integrity using group signatures (i.e., all sensor share the same group keys to sign reports). Also, they only provide limited confidentiality, as they reports are encrypted under the public key of a *Report Service* (RS), a trusted party responsible for collecting reports and distributing them to queriers. That is, the RS learns both sensors' reports and queriers' interests.

Additional research work focuses on somewhat related problems. [8] argues that privacy issues can be addressed if each user has access to a private server (e.g., a virtual machine hosted by a cloud service) and uses it as a proxy between her *sensors* and the application requesting her data. However, given the number of contributors in a participatory sensing application, the requirement of *per-user* proxies would severely limit the feasibility of this approach. [45] studies privacy-preserving data aggregation, e.g., computation of sum, average, variance, etc. Similarly, [18] 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). Other proposals, such as [17] and [19], aim at guaranteeing integrity and authenticity of user-generated contents, by employing Trusted Platform Modules (TPMs).

Finally, we remark that effective results have been obtained in somewhat related areas in the context of wireless (sensor) networks and mobile communications, e.g., for location privacy [38, 51], query privacy [15, 50, 9], secure routing [21, 2], authentication [40, 43, 36], or key management [3, 48, 35].

2.1 Limitations of prior work

As discussed above, the research community has devised a few solutions to address privacy issues in participatory sensing applications; We now discuss in details limitations and open problems of prior work on security and privacy in participatory sensing.

Assuming an ubiquitous WiFi infrastructure. One common feature of existing proposals is the assumption of an ubiquitous WiFi infrastructure used to collect and deliver reports [12, 22, 45]. In particular, [12, 22] use standard MAC-IP address recycling techniques to guarantee user unlinkability between reports with respect to WiFi access points. While there exists extensive research on privacy, anonymity, and unlinkability in WiFi networks [7], Such an assumption imposes severe limitations on the scope of participatory sensing applications, as an ubiquitous presence of open WiFi networks is not realistic today nor anticipated in the next future. Therefore, such an assumption would heavily limit applications' availability and accuracy.

Actually, the majority of existing participatory sensing applications operate from smartphones and use the cellular network to upload reports [42, 39, 32]. Thus, one cannot use WiFi-based anonymization techniques and, in particular, cannot leverage MAC-IP address recycling to guarantee unlinkability with respect to the access point. In cellular networks, devices are identified through their International Mobile Subscriber Identity (IMSI), and ID recycling—besides being impossible with current technologies—would lead to denial of service (e.g., the device would not receive incoming calls for its original ID). Thus, it seems not possible to protect privacy of user locations with respect to the network operator. In fact, the regular usage of cellular networks (e.g., including incoming/outgoing phone calls), as well as heartbeat messages exchanged with the network infrastructure, irremediably reveal device's location to the operator.

Using Mix Networks. Another limitation of prior work, such as [12, 22], concerns the use of Mix Networks [10] – anonymizing channels used to de-link reports submitted by sensors before they reach the applications. In other words, Mix Networks act as proxies to forward user reports only when some system-defined criteria are met. Several metrics, such as k-anonymity [49] or l-diversity [31] have been defined to characterize privacy through Mix Networks. Observe that a Mix Network may wait to receive k reports before forwarding them to the application, e.g., to guarantee k-anonymity. However, the anonymity level directly depends on the number of reports received and "mixed" by the Mix Network. They rely on statistical methods to protect privacy and do not guarantee provably-secure privacy. Moreover, there could be scenarios where a relatively long time could pass before the desired level of anonymity is reached (when "enough" reports have been collected). As a result, Mix Networks may remarkably decrease system throughput and cannot be used in settings where *timely* reports are required.

Multiple Semi-Trusted Parties. Available techniques to protect privacy in participatory sensing often involve many semi-trusted independent parties, that are always assumed not to collude. The solution in [12], besides

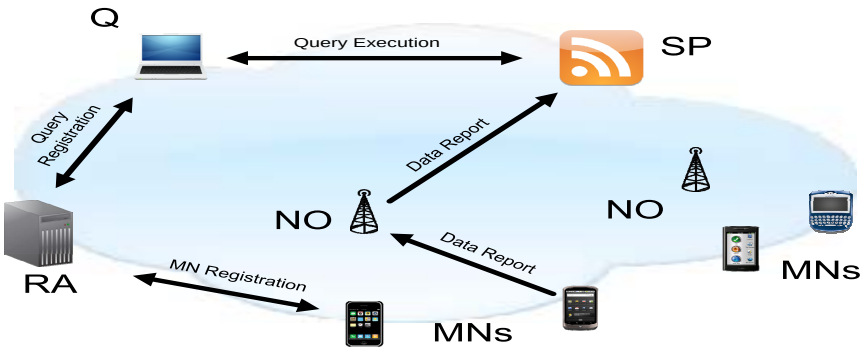

Figure 1: Privacy-Enhanced Participatory Sensing Infrastructure: Mobile Nodes (MNs) register to the Registration Authority (RA) and, subsequently, report sensed data to a Service Provider (SP). Queriers, after registering to RA, subscribe to queries offered by the SP and receive corresponding reports.

Mobile Nodes, Registration Authority, and WiFi Access Points, also assumes the presence and the non-collusion of a Task Service (used to distribute tasks to users), a Report Service (to receive reports from sensors), and several Mix Network nodes (i.e., a trusted anonymizing infrastructure). The assumption of multiple non-colluding parties raises severe concerns regarding its practicality and feasibility. It appears difficult to deploy all of the parties in a real world setting where entities provide services only in exchange of some benefit. For instance, it is not clear how to deploy the Task and the Report services as two separate entities having no incentive to collude. Whereas, we aim at minimizing the number of needed semi-trusted parties (and, in general, the number of involved entities), and propose a participatory sensing infrastructure that can be deployed with formal privacy guarantees.

3 Preliminaries

Following a cryptographic treatment of privacy in participatory sensing, in this section, we formalize: (i) the entities involved in a privacy-enhanced participatory sensing infrastructure, (ii) involved operations, and (iii) privacy requirements.

3.1 Infrastructure

We envision a participatory sensing infrastructure composed by the following entities:

Mobile Nodes (MNs). They are computing devices with sensing capabilities (i.e., equipped with one or more *sensors*) and with access to a cellular network. They are carried by people or attached to mobile entities. We assume that MNs run on smartphones and that users voluntarily engage into participatory sensing. We denote with $\mathcal{N}$ a generic mobile node of a participatory sensing application.

Queriers. Queriers are end-users interested in receiving sensor reports in a given participatory sensing application. A generic querier is denoted with $\mathcal{Q}$.

Network Operator (NO). The Network Operator is responsible for the communication infrastructure. We assume that the NO maintains, and provides access to, a cellular network infrastructure (e.g., GSM or 3G).

Registration Authority (RA). The Registration Authority handles the application setup, as well as the regis-

tration of participating parties. In our solutions, the RA also contributes to privacy protection, by generating cryptographic public parameters, handling the registration of MNs, and managing queriers' subscription.

Service Providers (SP). The Service Provider acts as an intermediary between the nodes reporting readings and queriers subscribed to them. We envision one or more SPs running participatory sensing applications that offer different query types. (For example, a national service provider might run a pollution monitoring application and define queries to retrieve reports of pollution levels in different cities). Service provider's duties may include listing available sensing services, micropayment, data collection, and notification to queriers.

3.2 Operations

We now describe the common operations performed within participatory sensing applications.

Setup. In this phase, the RA generates all public parameters and its own secret key.

MN Registration. Users register their sensor-equipped device to the RA and install participatory sensing software.

Query Registration. Queriers approach the appropriate RA and request an *authorization* to query the participatory sensing application to obtain a specific type of data reports, e.g., *"Pollution level in Madrid, Spain"*. (A public list of available sensing services and query syntax may be available from the SP or the RA). Next, they may subscribe to one or more (authorized) queries, by submitting a request to SP and awaiting for the responses containing the desired readings. Ideally, only queriers authorized by the RA should receive the desired reports. Also, no information about query interests should be revealed to the SP.

Data Report. MNs report to the SP their readings, using the network access provided by the NO. Ideally, this operation should not reveal to the SP, the NO, or unauthorized queriers any information about reported data, such as type of reading (e.g., pollution) or quantitative information (e.g., $35mg/m^3$ carbon oxide). Also, the SP and any querier should not learn the identity of the source MN.

Query Execution. With this operation, the SP matches incoming data reports with query subscriptions. Ideally, this should be done *blindly*, i.e., the SP should learn nothing beyond the occurrence of an (unspecified) match, if any.

In Figure 1, we illustrate our participatory sensing infrastructure. In the depicted scenario, one may envision that a phone manufacturer (e.g., Nokia) acts as the RA and embeds a given type of sensor (e.g., air pollution meter) in one or more of its phone models, operated by smartphone users, i.e., the MNs. A service provider (such as Google, Microsoft, Yahoo, or a non-profit/academic organization) offers participatory sensing applications (used, for instance, to report and access pollution data), and acts as an intermediary between queriers and mobile nodes. Finally, queriers are users or organizations (e.g., bikers) interested in obtaining readings (e.g., pollution levels).

Note that—similar to related work—we do not address the problem of encouraging mobile phone users to run participatory sensing applications, nor we focus on business incentives for phone manufacturers or for service providers. Nonetheless, it is reasonable to envision that queriers are willing to pay small fees (or receive advertisement) in return to obtaining measurements of interest.

3.3 Privacy Requirements

Before entering the details of our privacy requirements, observe that the main purpose of a participatory sensing application is to allow queriers to obtain MNs reports. While our main goal is to protect the privacy of both data producers and consumers, entities registered as a querier should still be able to subscribe to a query

and receive desired measurements, thus, techniques to identify legitimate parties before registering them are beyond the scope of our work.

We now define the requirements of a Privacy-Enhanced Participatory Sensing Infrastructure (PEPSI). Our definitions here are concise, whereas, formal adversarial games can be found in Appendix B.

Soundness. We say that PEPSI is *sound* if, upon subscribing to a query, a querier in possession of the appropriate authorization obtains the desired readings (if any).

Node Privacy. We say that PEPSI is *node-private* if neither the NO, the SP, nor any unauthorized querier, learn any information about the type of reading or the data reported by the MN. Also, other MNs should not learn any information about a given node's reports. In other words, *only queriers in possession of the corresponding authorization obtain MN's readings*.

Query Privacy. We say that a PEPSI is *query-private* if neither the NO, the SP, nor any mobile node or any other querier, learn any information about the query subscribed by a querier. We leave as part of future work how to guarantee query privacy also w.r.t. the RA.

Report Unlinkability. Report unlinkability prevents any party from linking two or more reports as originating from a given mobile node. PEPSI provides report unlinkability if different reports originating by the same MN cannot be linked to the source by the RA, the SP, the queriers or any other MN[1].

As discussed earlier, it seems unlikely to provide node location privacy w.r.t. NO, as the NO knows phone's position at any time. We still consider node location privacy as for the RA, the SP, or unauthorized queriers.

4 PEPSI

We now present our PEPSI instantiation, in accordance to the architectural design in Figure 1 and that complies with privacy definitions of Section 3.3. We also discuss arguments of its provable security and estimate its performance.

4.1 PEPSI Instantiation

In PEPSI, data reports are always labeled using keywords that identify the nature of the information announced by MNs. Similarly, queriers subscribe to given queries by specifying the corresponding keywords. In the rest of the paper, we use the term *identifier*, and the notation ID (or ID^*) to identify the data report/query type. Examples of such identifiers include: "*Temperature in Irvine, CA*" or "*Pollution in Madrid, Spain*". The list of identifiers – depending on the application – can be obtained either from the SP or the RA. In particular, the RA defines which services (i.e., queries) will be available for MNs to contribute and for users to query. However, as these identifiers can be public, they can be downloaded from the SP or any bulletin board. For ease of presentation, in the rest of the paper, we assume that query identifiers are available at the RA.

4.1.1 High Level Description

One of the main goals of PEPSI is to hide reports and queries to unintended parties. Thus, those cannot be transmitted *in-the-clear*, but need to be encrypted. In the rest of this section, we discuss how to achieve, at the same time, (1) secure encryption of reports and queries, and (2) efficient and oblivious matching performed at the Service Provider.

[1]Observe that we cannot guarantee user anonymity and report unlinkability with respect to the NO, thus, we do not consider the NO in our definition.

Report/Query Encryption. One naïve possibility is to let each querier and each mobile node share a unique pairwise secret key and employ a symmetric-key cipher, such as AES [13]. This approach requires queriers and MNs to interact and establish a shared secret. On the contrary, participatory sensing requires no contact (nor mutual knowledge) between them: that is, MNs provide reports obliviously of (any) potential receiver. Similarly, queries subscribe to measurements without knowing the identity of MNs producing reports potentially matching their interests. Even if we allow interaction between each mobile node and queriers, we would still need MNs to encrypt reports under each key shared with queriers (recall that MNs do not know which queriers are interested in their reports). This would generate a number of ciphertexts quadratic in the number of measurements. Alternatively, we could use a public key encryption scheme and provide MNs with the public keys of the queriers. Still, scalability would be an issue as each report would be encrypted under the public key of each querier.

Using Identity-based Encryption. PEPSI's main building block is Identity-Based Encryption (IBE) [4], specifically, the construction given by Boneh and Franklin in [4], which we review in Appendix A. The main advantage in using IBE, as opposed to standard public-key cryptography, is to enable non-interactivity in our query protocol design. This is crucial in participatory sensing scenarios, where MNs and queriers have no direct communication nor mutual knowledge. that enables asymmetric encryption using any string ("identity") as a public key. Recall that, In IBE, anyone can derive public keys from some unique information about recipient's identity. Private decryption keys are generated by a third-party, called the Private Key Generator (PKG).
Our idea is to use labels (i.e., keywords) that define the type of reports as identities in an IBE scheme. For examples labels "Temperature" and "Central Park, New York" can be used to derive a unique public encryption key, associated to a secret decryption key. Mobile Nodes encrypt sensed data using report's labels as the (public) encryption key. Query registration then consists in obtaining the private decryption keys corresponding to the labels of interest. Decryption keys are obtained, upon query registration, from the Registration Authority – which acts like a PKG.

PEPSI Overview. PEPSI works as follows. The RA runs the *Setup* algorithm to generate public parameters and secret keys. In order to pose a query, e.g., identified by ID, queriers first need to register to the RA and obtain the corresponding authorization (*Query Authorization*). Then, they subscribe their queries to the SP (*Query Subscription*): in PEPSI, this process reveals nothing about queriers' interests. At the same time, before starting reporting data, MNs need to authenticate to the RA, and obtain: (i) the identifier ID corresponding to the type of their reports, and (ii) a token that allows them to announce data (*MN Registration*).

The *on-line* part of PEPSI includes two operations: *Data Report* and *Query Execution*. With the former, MNs upload encrypted reports to the SP. In the latter, the SP *blindly* matches received reports with queries and forwards (matching) reports to all subscribed queriers. Only authorized queriers obtain query responses, can decrypt data reports, and retrieve original measurements. Finally, we let the RA periodically run a *Nonce Renewal* procedure to evict malicious MNs from the participatory sensing application.[2] This procedure is run periodically (e.g., once a week or once a month) and the new nonce is securely delivered to honest MNs using broadcast encryption [5].

4.1.2 Algorithms Specification

Setup. The Registration Authority (RA), given a security parameter λ, generates a prime q, two groups $\mathbb{G}_1, \mathbb{G}_2$ of order q, a bilinear map $e : \mathbb{G}_1 \times \mathbb{G}_1 \rightarrow \mathbb{G}_2$.[3] Next, a random $s \in \mathbb{Z}_q^*$ and a random generator $P \in \mathbb{G}_1$, are chosen; Q is set such that $Q = P^s$. (P, Q) are public parameters. s is RA's private master key.

[2]Techniques to identify malicious MNs are beyond the scope of this work.
[3]Recall that the map e is *bilinear* if $e(U^a, V^b) = e(U, V)^{ab}$.

Also, a nonce $z \in \mathbb{Z}_q^*$ is selected and R is set such that $R = P^z$. Finally, three cryptographic hash function, $H_1 : \{0,1\}^* \rightarrow \mathbb{G}_1$, $H_2 : \{0,1\}^{\mathbb{G}_2} \rightarrow \{0,1\}^\lambda$, and $H_3 : \{0,1\}^{\mathbb{G}_2} \rightarrow \{0,1\}^\lambda$ are chosen.

MN Registration. The MN registration is run between MN and RA. We assume that, after being identified by RA, (e.g., using the IMSI number), the Mobile Node $\mathcal{N}$ obtains the pair (z, ID), where z is the nonce generated by RA during setup, and ID identifies the nature of the readings for which $\mathcal{N}$ provides reports.

Query Registration. The query registration is as follows:

1. `Query Authorization`: Querier $\mathcal{Q}$ registers to RA to get an authorization to receive readings for a given query, identified by ID^*. $\mathcal{Q}$ obtains:
$$sig = H_1(ID^*)^s.$$
2. `Query Subscription`: Querier $\mathcal{Q}$—subscribing to readings identified by ID^*—sends to the SP:
$$T^* = H_2[e(R, sig)].$$

Data Report. Mobile node, $\mathcal{N}$. periodically submits data reports to SP, using NO's infrastructure. In our protocol, to upload a reading, $\mathcal{D}$, related to query ID, $\mathcal{N}$ sends SP (using NO's infrastructure) the pair $\langle T, CT \rangle$, such that:
$$T = H_2[e(Q, H_1(ID)^z)]$$
$$CT = ENC_k(\mathcal{D})$$
for $k = H_3[e(Q, H_1(ID)^z)]$.

Query Execution. The query execution is as follows:

1. `Blind Matching`: The Service Provider matches T with stored T^*'s and returns $\mathcal{Q}$ all matching T's, alongside associated CT's.
2. `Notification`: On receiving (T, CT), $\mathcal{Q}$ computes $k^* = H_3[e(R, sig)]$ and obtains:
$$\mathcal{D} = Dec_{k^*}(CT)$$

Nonce renewal. We assume a dynamic set of subscribed MNs where new sensors can register and malicious ones are evicted. In order to ban misbehaving sensors, the RA periodically generates and distributes a fresh z to sensors and $R = P^z$ to queriers. The former can be securely distributed to honest sensors using broadcast encryption [5].

4.2 Privacy Analysis

We now consider privacy properties of PEPSI. We hereby provide concise proofs, whereas, the reader can find detailed proofs in Appendix B. We assume that the system is immune to eavesdropping. In fact, 3G networks encrypt communication between mobile phones and the network operator. Communication between other parties (i.e., $RA \leftrightarrows MN$, $Querier \leftrightarrows SP$, etc.) are encrypted using standard techniques, e.g., using SSL.

Soundness. Our PEPSI solution is *sound*, since: for any (ID^*, sig) held by a querier $\mathcal{Q}$, and ID held by a node $\mathcal{N}$, if: (1) $sig = H_1(ID^*)^s$, where s is RA's secret key, and (2) $ID^* = ID$, we obtain :
$$\begin{aligned} T &= H_2[e(Q, H_1(ID)^z)] = H_2[e(P^z, H_1(ID^*)^s)] \\ &= H_2[e(R, sig)] = T^* \end{aligned}$$

and, similarly, also $k = k^*$. Therefore, (1) SP correctly matches $\mathcal{Q}$'s (authorized) request T^* with the appropriate sensor report $\langle T, CT \rangle$, and (2) $\mathcal{Q}$ can correctly decrypt CT and recover $\mathcal{D}$.

Node Privacy. Our PEPSI solution is *node-private* since only authorized queriers in possession of a valid sig can learn any information about the report $\langle T, CT \rangle$. Privacy w.r.t. the NO, the SP, other MNs and non-authorized queriers, stems from the security of the underlying Identity-based Encryption scheme [4]. The main intuition is that this kind of adversary could obtain information about a node's report only if she was in possession of the appropriate sig or—assuming that ID's have low entropy—the nonce z. Assuming a CPA-secure and key-private IBE system, the resulting PEPSI scheme is trivially *node-private* w.r.t. the NO, the SP, and non-authorized queriers. Indeed, one can easily sketch a proof of this claim by contradiction. (Detailed proof is deferred to Appendix B.) Assuming our claim is not true, then there exists a polynomial-bounded adversary $\mathcal{A}$ that violates node privacy of PEPSI. $\mathcal{A}$ is given ID and the IBE-encryption of $\mathcal{D}$ under the key ID but not the corresponding $sig = H_1(ID)^s$. If $\mathcal{A}$ decrypts $\mathcal{D}$ with non-negligible probability, then we can construct a polynomial-bounded adversary $\mathcal{B}$ which uses $\mathcal{A}$ to break the CPA-security of IBE. This contradicts our assumption.
Note that the RA could use its secret key to "test" an arbitrary ID^* against an eavesdropped report $\langle T, CT \rangle$. That is, the RA could learn whether $ID^* = ID$ and violate node privacy. However, since we assume that reports $\langle T, CT \rangle$ are encrypted under SP's public key, the RA can access nodes' reports only if it colludes with the SP.

Query Privacy. PEPSI is *query-private* since neither the NO, the SP, other queriers, nor any mobile node, learn any information about query interests of a querier $\mathcal{Q}$. Query privacy stems from the security of the underlying Identity-based Encryption scheme [4]. Arguments behind this claim mirror those outlined above for node privacy, thus, we do not repeat them here. Again, detailed proof is reported in Appendix B. Observe that the privacy of a querier $\mathcal{Q}$, subscribed to ID^* (i.e., in possession of $sig = H_1(ID^*)^s$), could be violated by a malicious party, subscribed to ID' (i.e., in possession of $sig' = H_1(ID')^s$), only if: (1) she obtains $T = H_2[e(R, sig)]$ sent from $\mathcal{Q}$ to the SP during *Query Execution* and (2) $ID^* = ID'$. Since the communication between $\mathcal{Q}$ and the SP is encrypted, (1) happens only if such a malicious party colludes with the SP.

Report Unlinkability and Location Privacy. As argued above, it is not possible to guarantee report unlinkability with respect to the network operator. However, one could trust the NO to remove privacy-sensitive metadata from each report (such as mobile nodes' identifiers, the cell from which the report was originated, etc.), before forwarding it to the SP. Nonetheless, this would not require the NO to act as a Mix Network. Also, the NO never delays message forwarding, e.g., until "enough" reports to protect privacy are collected, but forwards "the payload" of each report (i.e., $\langle T, CT \rangle$) as soon as it is received. More details are reported Appendix B.

Trust Assumptions. The security of PEPSI only relies on the assumption that the SP is not colluding with either the RA or queriers — on the contrary, prior work assumed the presence of *several* non-colluding and/or fully-trusted parties [12, 22, 17, 19]. Specifically, if the RA and the SP colluded, they could violate node privacy using RA's secret key, s, and pairs $\langle T, CT \rangle$ received by the SP. Also, recall that any party registered as a querier could potentially collude with the SP and try violating query privacy: it could test a given sig (obtained during the *Query Authorization*) against messages sent by the victim querier to the SP (during *Query Subscription*). We argue that assuming a non-colluding SP is realistic since in participatory sensing, SPs often capitalize on the services they provide, thus, they have no incentive to deviate from an honest-but-curious behavior.

Furthermore, PEPSI needs to trust the NO to remove sensitive MN information from reports before forwarding to the SP. Recall that this assumption is essential since anonymity w.r.t the NO is not achievable in 3G networks.

4.3 Performance Evaluation

Even if resources in participatory sensing are not as constrained as in WSNs, we aim at minimizing the overhead incurred at mobile nodes. This section provides preliminary figures on the cost of cryptographic operations used to achieve intended privacy features.

We implemented protocol operations executed by MNs on a Nokia N900 (equipped with a 600 MHz ARM processor and 256 MB RAM) running the `libpbc` cryptographic library [30]. We selected Type-A pairings and 160-bit prime q. Computation overhead is due to the computation of T, the encryption key k, and the encrypted report CT. Note that the first two values can be computed off-line, independently of the sensed data. Communication overhead is merely due to the transmission of T, which is the output of a hash function (e.g., SHA-1), and can be as small as 160-bit. Indeed, using available symmetric-key cryptosystems (e.g., AES), the length of CT is almost the same as a reading $\mathcal{D}$.

Without leveraging off-line precomputation, we measured the time to compute and transmit $\langle T, CT \rangle$, using integers as data reports. Over 100 experiments, we experienced an average time of $93.47ms$ to compute $\langle T, CT \rangle$ and around $80ms$ for transmission over the 3G network. Note that a naïve (non-private) solution would save in computation (since data would not be encrypted) but would spend roughly the same transmission time to send the report. Finally, remark that the SP incurs no communication nor computational overhead: its task is limited to forwarding and hash comparisons. Similarly, the only additional operation that queriers perform during query execution is the symmetric decryption of received readings, which incurs a negligible overhead.

5 Conclusion

The Participatory sensing paradigm bears an irrefutably great potential. However, its success depends on the number of users willing to report measurements from their mobile devices. Clearly, a wide-scale user participation is bound to effective protocols that preserve privacy of both data producers (i.e., mobile nodes) and data consumers (i.e., queriers). In this paper, we have highlighted shortcomings of previous solutions and we embarked toward a cryptographic treatment of privacy in participatory sensing. To this aim, we analyzed which are the privacy features that can be guaranteed with provable security and introduced a participatory sensing protocol that attains them. Finally, we provided figures of the incurred overhead at mobile nodes.

As often happens, deploying actual solutions based on our proposal requires addressing additional (potential) security issues, such as authentication, data integrity, DoS prevention, active attacks, Sybil attacks, etc. Our next step is to deploy testing applications using the PEPSI infrastructure, as well as to devise a large-scale evaluation of its global overhead. Our future work also includes extending the protocols to efficiently support query privacy w.r.t. the RA (i.e., queriers can register without the RA learning their interests). Interesting open challenges remain in how to provide location privacy with respect to cellular network operator, addressing potential collusion between different parties, and supporting more complex queries (e.g., aggregate and conjunctive queries).

Acknowledgments. This research has been partially funded by US Intelligence Advanced Research Projects Activity (IARPA) under grant FA8750-09-2-0071, the Madrid Regional Council – CAM under project CLOUDS (S2009TIC-1692), the Spanish Research Agency – MICINN under project CloudStorm (TIN2010-19077), and the European Commission under projects MASSIF (FP7-257475) and STREAM (FP7-216181). We are also grateful to Nokia for the devices used in our experiments.

References

[1] Participatory Privacy: Enabling Privacy in Participatory Sensing. `http://sprout.ics.uci.edu/PEPSI/`.

[2] G. Acs, L. Buttyan, and I. Vajda. Provably secure on-demand source routing in mobile ad hoc networks. *IEEE Transactions on Mobile Computing*, 5(11), 2006.

[3] N. Asokan and P. Ginzboorg. Key agreement in ad hoc networks. *Computer Communications*, 23(17), 2000.

[4] D. Boneh and M. Franklin. Identity-based encryption from the Weil pairing. In *Crypto*, 2001.

[5] D. Boneh, C. Gentry, and B. Waters. Collusion resistant broadcast encryption with short ciphertexts and private keys. In *Crypto*, 2005.

[6] J. Burke, D. Estrin, M. Hansen, A. Parker, N. Ramanathan, S. Reddy, and M. Srivastava. Participatory sensing. In *World Sensor Web Workshop*, 2006.

[7] L. Buttyán and J.-P. Hubaux. *Security and cooperation in wireless networks*. Cambridge University Press, 2007.

[8] R. Cáceres, L. P. Cox, H. Lim, A. Shakimov, and A. Varshavsky. Virtual individual servers as privacy-preserving proxies for mobile devices. In *MobiHeld Workshop*, 2009.

[9] B. Carbunar, Y. Yu, W. Shi, M. Pearce, and V. Vasudevan. Query privacy in wireless sensor networks. *ACM Transactions on Sensor Networks*, 6(2), 2010.

[10] D. L. Chaum. Untraceable electronic mail, return addresses, and digital pseudonyms. *Communications of ACM*, 24(2), 1981.

[11] D. Christin, M. Hollick, and M. Manulis. Security and Privacy Objectives for Sensing Applications in Wireless Community Networks. In *IEEE ICCCN*, 2010.

[12] C. Cornelius, A. Kapadia, D. Kotz, D. Peebles, M. Shin, and N. Triandopoulos. AnonySense: Privacy-aware people-centric sensing. In *Mobisys*, 2008.

[13] J. Daeman and V. Rijmen. AES proposal: Rijndael. 1999.

[14] T. Das, P. Mohan, V. Padmanabhan, R. Ramjee, and A. Sharma. PRISM: Platform for Remote Sensing using Smartphones. In *Mobisys*, 2010.

[15] E. De Cristofaro, X. Ding, and G. Tsudik. Privacy-preserving querying in wireless sensor networks. In *ICCCN*, 2009.

[16] Y. Dong, S. Kanhere, C. Chou, and N. Bulusu. Automatic collection of fuel prices from a network of mobile cameras. In *DCOSS*, 2008.

[17] A. Dua, N. Bulusu, W. Feng, and W. Hu. Towards trustworthy participatory sensing. In *HotSec*, 2009.

[18] R. Ganti, N. Pham, Y. Tsai, and T. Abdelzaher. PoolView: stream privacy for grassroots participatory sensing. In *SenSys*, 2008.

[19] P. Gilbert, L. Cox, J. Jung, and D. Wetherall. Toward trustworthy mobile sensing. In *HotMobile*, 2010.

[20] M. Green and S. Hohenberger. Blind identity-based encryption and simulatable oblivious transfer. In *Asiacrypt*, 2007.

[21] Y.-C. Hu, A. Perrig, and D. B. Johnson. ARIADNE: A secure on-demand routing protocol for ad hoc networks. *Wireless Networks*, 11(1-2), 2005.

[22] K. Huang, S. Kanhere, and W. Hu. Preserving privacy in participatory sensing systems. *Computer Communications*, 33(11), 2010.

[23] B. Hull, V. Bychkovsky, Y. Zhang, K. Chen, M. Goraczko, A. Miu, E. Shih, H. Balakrishnan, and S. Madden. CarTel: a distributed mobile sensor computing system. In *SenSys*, 2006.

[24] A. Kapadia, D. Kotz, and N. Triandopoulos. Opportunistic sensing: Security challenges for the new paradigm. In *COMNETS*, 2009.

[25] D. Kim, J. Hightower, R. Govindan, and D. Estrin. Discovering semantically meaningful places from pervasive RF-beacons. In *UbiComp*, 2009.

[26] S. Kinkade and K. Verclas. Wireless technology for social change: Trends in mobile use by NGOs. `www.mobileactive.org/files/MobilizingSocialChange_full.pdf`, 2008.

[27] J. Lee and B. Hoh. Sell Your Experiences: A Market Mechanism based Incentive for Participatory Sensing. In *PerCom*, 2010.

[28] B. Longstaff, S. Reddy, and D. Estrin. Improving activity classification for health applications on mobile devices using active and semi-supervised learning. In *PervasiveHealth*, 2010.

[29] H. Lu, W. Pan, N. Lane, T. Choudhury, and A. Campbell. SoundSense: scalable sound sensing for people-centric applications on mobile phones. In *Mobisys*, 2009.

[30] B. Lynn. PBC: The Pairing-Based Cryptography Library. `http://crypto.stanford.edu/pbc/`, Last Accessed, 2011.

[31] A. Machanavajjhala, D. Kifer, J. Gehrke, and M. Venkitasubramaniam. l-diversity: Privacy beyond k-anonymity. *ACM Transactions on Knowledge Discovery from Data (TKDD)*, 1(1), 2007.

[32] S. Mathur, T. Jin, N. Kasturirangan, J. Chandrasekaran, W. Xue, M. Gruteser, and W. Trappe. ParkNet: drive-by sensing of road-side parking statistics. In *MobiSys*, 2010.

[33] P. Mohan, V. Padmanabhan, and R. Ramjee. Rich monitoring of road and traffic conditions using mobile smartphones. In *Sensys*, 2008.

[34] M. Mun, S. Reddy, K. Shilton, N. Yau, J. Burke, D. Estrin, M. Hansen, E. Howard, R. West, and P. Boda. Peir, the personal environmental impact report, as a platform for participatory sensing systems research. In *Mobisys*, 2009.

[35] D. Nilsson, T. Roosta, U. Lindqvist, and A. Valdes. Key management and secure software updates in wireless process control environments. In *ACM WiSec*, 2008.

[36] P. Ning, A. Liu, and W. Du. Mitigating DoS attacks against broadcast authentication in wireless sensor networks. *ACM Transactions on Sensor Networks*, 4(1), 2008.

[37] Nokia Research Center. SensorPlanet. `http://research.nokia.com/research/projects/sensorplanet`, 2010.

[38] C. Ozturk, Y. Zhang, and W. Trappe. Source-location privacy in energy-constrained sensor network routing. In *SASN*, 2004.

[39] E. Paulos, R. Honicky, and E. Goodman. Sensing atmosphere. In *SenSys Workshops*, 2007.

[40] A. Perrig, R. Szewczyk, V. Wen, D. Culler, and J. D. Tygar. SPINS: Security protocols for sensor networks. *Wireless Networks*, 8(5), 2002.

[41] S. Reddy, D. Estrin, and M. Srivastava. Recruitment Framework for Participatory Sensing Data Collections. *Pervasive Computing*, 2010.

[42] S. Reddy, K. Shilton, G. Denisov, C. Cenizal, D. Estrin, and M. Srivastava. Biketastic: sensing and mapping for better biking. In *CHI*, 2010.

[43] K. Ren, S. Yu, W. Lou, and Y. Zhang. Multi-user broadcast authentication in wireless sensor networks. *IEEE Transactions on Vehicular Technology*, 58(8), 2009.

[44] A. Santanche, S. Nath, J. Liu, B. Priyantha, and F. Zhao. Senseweb: Browsing the physical world in real time. In *IPSN*, 2006.

[45] J. Shi, R. Zhang, Y. Liu, and Y. Zhang. PriSense: Privacy-Preserving Data Aggregation in People-Centric Urban Sensing Systems. In *Infocom*, 2010.

[46] K. Shilton. Four billion little brothers?: Privacy, mobile phones, and ubiquitous data collection. *Communications of the ACM*, 52(11), 2009.

[47] Stanford University. Quake-Catcher Network. `http://qcn.stanford.edu`, 2010.

[48] M. Steiner, G. Tsudik, and M. Waidner. Key agreement in dynamic peer groups. *IEEE Transactions on Parallel and Distributed Systems*, 11(8), 2002.

[49] L. Sweeney. k-Anonymity: A model for Protecting Privacy. *Int. J. Uncertain. Fuzziness Knowl.-Based Syst.*, 2002.

[50] R. Zhang, Y. Zhang, and K. Ren. DP2AC: Distributed privacy-preserving access control in sensor networks. In *INFOCOM*, 2009.

[51] G. Zhong and U. Hengartner. A distributed k-anonymity protocol for location privacy. In *Percom*, 2009.

A Boneh and Franklin's IBE

We recall that Boneh and Franklin's IBE [4] is composed by four algorithms: `setup`, `extract`, `encrypt`, `decrypt`:

- `Setup`, given a security parameter k, is used to generate a prime q, two groups $\mathbb{G}_1, \mathbb{G}_2$ of order q, a bilinear map $e : \mathbb{G}_1 \times \mathbb{G}_1 \to \mathbb{G}_2$. Then a random $s \in \mathbb{Z}_q^*$, a random generator $P \in \mathbb{G}_1$, P are chosen and Q is set such that $Q = sP$. (P, Q) are public parameters. s is the private master key. Finally, two cryptographic hash function, $H_1 : \{0,1\}^* \to \mathbb{G}_1$ and $H_2 : \{0,1\}^n \to \mathbb{G}_2$ for some n, are chosen.
- `Extract`, given a string $ID \in \{0,1\}^*$, is used to compute the corresponding private key $H(ID)^s$.
- `Encrypt` is used to encrypt a message M under a public key ID: for a picked random $r \in Z_q^*$ the ciphertext is set to be $C = \langle rP, M \oplus H_2(e(Q, H_1(ID)^r)) \rangle$.
- `Decrypt` is used to decrypt a ciphertext $C = \langle U, V \rangle$, by computing $M = V \oplus H_2(e(U, H(ID)^s)$.

B Detailed PEPSI Privacy Analysis

In this section, we formally analyze the privacy properties of PEPSI. Observe that PEPSI's intuition lies in the application of a key-private IBE system—where query identifiers are used as identities—to protect privacy in the participatory sensing setting. Therefore, its privacy requirements rely on the security and the key-privacy of the underlying IBE system.

Recall that we assume communication over cellular networks, thus, the system is immune to external adversaries. (In 3G/4G networks, communication betweeen mobile phones and the network operator is encrypted). Further, communication between other parties (i.e., $RA \leftrightarrows MN$, $Querier \leftrightarrows SP$, etc.) is encrypted using SSL.

B.1 Soundness

PEPSI is *sound* if, at the end of *Query Execution*, querier $\mathcal{Q}$ outputs $\mathcal{D}$, given that:
(1) $\mathcal{Q}$ registered query ID^* to the RA.
(2) $\exists$ node N such that N reports $(ID, \mathcal{D})$.
(3) $ID^* = ID$.

Our PEPSI solution is *sound*, since, for any registered query (ID^*, sig) held by querier $\mathcal{Q}$, and ID reported by a node N, if: (1) $sig = H_1(ID^*)^s$, where s is RA's secret key, and (2) $ID^* = ID$, we obtain:

$$\begin{aligned} \mathbf{T} &= H_2[e(Q, H_1(ID)^z)] = H_2[e(P^z, H_1(ID^*)^s)] \\ &= H_2[e(R, sig)] = \mathbf{T}^* \end{aligned}$$

and, similarly, also $k = k^*$. Therefore, (1) SP correctly matches $\mathcal{Q}$'s (authorized) request T^* with the appropriate node report $\langle T, CT \rangle$, and (2) $\mathcal{Q}$ can correctly decrypt CT and recover $\mathcal{D}$. □

B.2 Node Privacy

Informally, PEPSI is *node-private* if neither the NO, the SP, nor any unauthorized querier, learn any information about the type of reading or the data reported by the MN. Also, other mobile nodes in the infrastructure should not learn any information about a given node's reports. That is, only queriers in possession of the corresponding authorization obtain MN's readings.

Formally, privacy of node N, running on input $(ID, \mathcal{D})$, is guaranteed if no information about $(ID, \mathcal{D})$ is leaked to unauthorized parties.

We distinguish between node privacy w.r.t. the NO and the SP and w.r.t. unauthorized queriers.

Node Privacy w.r.t. the NO and the SP. Privacy is considered as the probabilistic advantage that an adversary $\mathcal{A}$ gains from obtaining encrypted reports. We say that PEPSI is *node-private* w.r.t. the NO/SP if no polynomially bounded adversary $\mathcal{A}$ can win the following game with non-negligible probability above $\frac{1}{2}$. The game is between $\mathcal{A}$ and a challenger Ch:

1. Ch executes setup operations and computes public parameters $(e, q, \mathbb{G}_1, \mathbb{G}_2, P, Q, R, H_1(\cdot), H_2(\cdot), H_3(\cdot))$ and private parameters (s, z).
2. $\mathcal{A}$, on input the public parameters, selects two inputs $((ID_0, \mathcal{D}_0), (ID_1, \mathcal{D}_1))$ and gives them to Ch.
3. Ch picks a random bit $b \in_R \{0, 1\}$ and interacts with $\mathcal{A}$ executing the role of the node N, on input the public parameters and private input $(ID_b, \mathcal{D}_b)$.
4. $\mathcal{A}$ outputs b' and wins if $b' = b$.

Assuming that the underlying IBE system is semantic secure against a chosen-plaintext attack (CPA) and key-private (such as Boneh-Franklin IBE [4]), PEPSI is trivially node-private w.r.t. the NO and the SP in the Random Oracle Model (ROM). Assuming that H_2 and H_3 are modeled as a random oracle, if our claim is not true then there exists a polynomial-bounded adversary $\mathcal{B}$ that breaks the CPA-security of IBE.

Node Privacy w.r.t. unauthorized queriers. Privacy is considered as the probabilistic advantage that an adversary $\mathcal{A}$ gains from submitting queries to the SP. We say that PEPSI is *node-private* w.r.t. unauthorized queriers if no polynomially bounded adversary $\mathcal{A}$ can win the following game with non-negligible probability above $\frac{1}{2}$. The game is between $\mathcal{A}$ and a challenger Ch:

1. Ch executes setup operations and computes public parameters $(e, q, \mathbb{G}_1, \mathbb{G}_2, P, Q, R, H_1(\cdot), H_2(\cdot), H_3(\cdot))$ and private parameters (s, z)

2. $\mathcal{A}$, on input the public parameters, adaptively queries Ch a number n of times on a set of identifiers $L = \{ID_0, \ldots, ID_n\}$. For every ID_i, Ch responds by giving $\mathcal{A}$ a signature $sig_i = H(ID_i)^s$.

3. $\mathcal{A}$ announces two new identifier strings $(ID_0^*, ID_1^*) \notin L$ and generates a data record $\mathcal{D}^*$.

4. Ch picks a random bit $b \in_R \{0, 1\}$ and interacts with $\mathcal{A}$ executing the role of the node N, on input the public parameters and private input $(ID_b^*, \mathcal{D}^*)$.

5. $\mathcal{A}$ outputs b' and wins if $b' = b$.

Assuming an underlying IBE system semantically secure against a chosen-plaintext attack (CPA) and key-private, the resulting PEPSI scheme is trivially *node-private* w.r.t. unauthorized queriers, in the random oracle model. Indeed, to win the above game, $\mathcal{A}$ needs to forge signature on ID_0^* or ID_1^*. Again, if this happens, then there exists a polynomial-bounded adversary $\mathcal{B}$ that breaks the CPA-security of IBE.□

Remark. Observe that the RA could use its secret key, s, to "test" an arbitrary ID^* against a report $\langle T = H_2[e(Q, H_1(ID)^z)], CT = ENC_k(\mathcal{D})\rangle$. That is, the RA could learn whether $ID^* = ID$ and violate node privacy. However, assuming that reports $\langle T, CT\rangle$ are super-encrypted under SP's public key, the RA can access nodes' reports only if it colludes with the SP.

B.3 Query Privacy

Informally, PEPSI is *query-private* if neither the NO, the SP, other queriers, nor any mobile node, learn any information about query interests of a querier $\mathcal{Q}$. Query privacy w.r.t. the NO, any mobile node, and any other querier, is trivially guaranteed as none of them obtains any cleartext message from $\mathcal{Q}$. Thus, we focus on privacy against a malicious SP, described as the probabilistic advantage that SP gains from obtaining queries.

Formally, PEPSI is *query-private* if no polynomially bounded adversary $\mathcal{A}$ can win the following game with probability non-negligibly over 1/2. The game is between $\mathcal{A}$ and a challenger Ch:

1. Ch executes setup operations and computes public parameters $(e, q, \mathbb{G}_1, \mathbb{G}_2, P, Q, R, H_1(\cdot), H_2(\cdot), H_3(\cdot))$ and private parameters (s, z)

2. $\mathcal{A}$, on input the public parameters, chooses two strings ID_0^*, ID_1^*.

3. Ch picks a random bit $b \in \{0, 1\}$ and interacts with $\mathcal{A}$ playing the role of the querier on input the public parameters and private input (ID_b^*).

4. $\mathcal{A}$ outputs b' and wins if $b' = b$.

Assuming that the underlying IBE system is semantic secure against a chosen-plaintext attack (CPA) and key-private (such as Boneh-Franklin IBE [4]), PEPSI is trivially query-private in the Random Oracle Model (ROM). Assuming that H_2 and H_3 are modeled as a random oracle, if our claim is not true then there exists a polynomial-bounded adversary $\mathcal{B}$ that breaks the CPA-security of IBE.

Remark. Note that we do not address in this version, query privacy w.r.t. the RA. However, we anticipate that this can be achieved if the query registration is performed blindly, using Blind-IBE [20]. Also, observe that we do not address at this stage potential collusion between authorized queriers or RA and the SP.

B.4 Report Unlinkability and Location Privacy

Report unlinkability prevents any party from linking two or more reports as originating from a given mobile node. We say that PEPSI provides report unlinkability since different reports originating by the same MN cannot be linked to the source by the RA, the SP, the queriers or any other MN.

We do not guarantee user anonymity and report unlinkability with respect to the network operator, given the intristic nature of the underlying cellular network. Similarly, as discussed in our related work section, it seems impossible to provide node location privacy w.r.t. the NO, since the NO knows phone's position at any time. Hence, PEPSI does not pursue MN location privacy.

Nonetheless, it is possible to modify our protocol to provide report unlinkability and location privacy w.r.t. all other parties, if we assume that the NO removes privacy-sensitive metadata from each report (such as mobile nodes' identifiers, the cell from which the report was originated, etc.), before forwarding it to the SP. Note that this would not need the use of MixNetworks, i.e., the NO does not have to delay message forwarding (e.g., until "enough" reports to protect privacy are collected) but forwards "the payload" of each report (i.e., $\langle T, CT \rangle$) as soon as it is received.