Trustworthy People-Centric Sensing: Privacy, Security and User Incentives Road-Map

Thanassis Giannetsos, Stylianos Gisdakis, Panos Papadimitratos
Networked Systems Security Group
KTH Royal Institute of Technology
Stockholm, Sweden
{*athgia, gisdakis, papadim*}@kth.se

***Abstract*—The broad capabilities of widespread mobile devices have paved the way for *People-Centric Sensing* (PCS). This emerging paradigm enables *direct* user involvement in possibly large-scale and diverse data collection and sharing. Unavoidably, this raises significant privacy concerns, as participants may inadvertently reveal a great deal of sensitive information. However, ensuring user privacy, e.g., by anonymizing data they contribute, may cloak faulty (possibly malicious) actions. Thus, PCS systems must not only be privacy-preserving but also *accountable* and *reliable*. As an increasing number of applications (e.g., assistive healthcare and public safety systems) can significantly benefit from people-centric sensing, it becomes imperative to meet these seemingly contradicting requirements. In this work, we discuss security, user privacy and incentivization for this sensing paradigm, exploring how to address all aspects of this multifaceted problem. We critically survey the security and privacy properties of state-of-the-art research efforts in the area. Based on our findings, we posit open issues and challenges, and discuss possible ways to address them, so that security and privacy do not hinder the deployment of PCS systems.**

I. Introduction

The emergence of resource-rich mobile devices has changed the landscape of mobile sensing. Today's smart-phones tend to become the main user computing and communication platform, incorporating multiple embedded sensors, e.g., accelerometers, gyroscopes, GPSs, cameras, etc. With more than 6 billion mobile subscriptions worldwide [1], these sensors (collectively) can be used to sense the environment and gather valuable data of unprecedented quality and quantity, practically from everywhere. This new sensing paradigm, *People-Centric Sensing* (PCS) [2, 3, 4], makes individuals and user communities the focal point of the sensing infrastructure; with their mobile handsets, users collect targeted information about their daily patterns and interactions [5].

In fact, PCS complements earlier efforts in location-based services by leveraging the duality of a user's role; each user is a data *contributor* as well as a data (service) *consumer*. Therefore, increasing voluntary participation, to provide a sufficient and continuous influx of contributions (the "*chasm of critical mass*" [6]), is key to the success of any PCS campaign. However, as mobile devices gather sensor data from the user's immediate environment, privacy concerns (i.e., understanding, choosing, and controling *what* kind of information users share, with *whom* and for *how long*) are rightfully raised. Sharing such sensed data tagged with spatio-temporal information (e.g., time and/or location) could reveal many personal attributes, such as a user's personal activities and health condition among others [7]. Furthermore, due to the strong correlation with the user's current context (e.g., whether they are at home or at work, walking or driving, etc.), there is a significant risk of *indirectly* inferring their daily routines or habits [8, 9]. As people realize the serious consequences of disclosing their sensitive information, it is vital to address these privacy challenges so that users do not opt out of PCS. Especially in the light of recent revelations of mass surveillance [10] that aggravate such user anxieties.

Privacy protection is necessary to motivate user participation, but it is not (by itself) a sufficient condition. What is needed, actually, is to provide *incentives* to engage as many people as possible, notably with diverse backgrounds, interests and availability. Indeed, relying only on the altruistic behavior of contributing participants [11] may not be adequate. This is why the research community has identified various types of incentives and ways to materialize them, such as reputation systems [12], service quotas [13] and monetary rewards [14]. However, it is necessary to provide such incentives in a privacy-preserving manner. For example, users should be rewarded with credits for their contributions without revealing what kind of data they shared or the task they participated in.

On the other hand, protecting user privacy by hiding any identifying information, can threaten the *trustworthiness* of the system. In fact, PCS applications could suffer from incorrect contributions due to their inherently open nature [15]; participants may inadvertently or deliberately submit falsified data. Therefore, full anonymity may tempt malicious user behavior, compromising the reliability of the entire sensing campaign. To thwart such behavior, we need fine-grained protocols that provide some form of *accountable anonymity*. This way offending users can be tied to their actions, without necessarily disclosing their identity.

As more and more applications could greatly benefit from the incorporation of people-centric sensing, it becomes imperative to address all these issues in a concrete way that will allow PCS to achieve its full potential. The importance of user participation, user privacy and system trustworthiness has been made clear thanks to numerous research efforts. However, current solutions have been only moderately successful, as they address facets of the problem at hand; they focus on privacy and security without considering either accountability [16, 17, 18, 19, 20] or the user's dual role (i.e., producing and consuming information) [21, 22]; or they try to enable mass participation by linking incentives to users' contributions without considering privacy implications [12, 13, 14].

There is clearly a need to shed more light on the underpinnings of the necessary solutions that will not only ensure the system's trustworthiness and protect users' privacy, but will also balance their interplay. Simply put, *how can we design massive (in terms of user participation) PCS systems that are, at the same time, reliable, accountable and privacy-preserving?*

978-1-4799-5258-8/14/$31.00 ©2014 IEEE

Device	**Inertial**	**Compass**	**GPS**	**Mic.**	**Camera**	**Proximity**	**Light**	**Gyro.**	**Acceler.**	**Magneto.**	**Barometer**	**Heart Rate**
iPhone 5s	✓	✓	✓	✓	✓	✓	✓	✓	✓	-	-	-
Samsung Galaxy S5	✓	✓	✓	✓	✓	✓	✓	✓	✓	-	✓	✓
HTC One M8	✓	✓	✓	✓	✓	✓	✓	✓	✓	-	✓	
Sony Xperia Z2	✓	✓	✓	✓	✓	✓	✓	✓	✓	-	-	-
Nokia Lumnia	✓	✓	✓	✓	✓	✓	✓	-	✓	✓	-	-
Google Nexus 5	✓	✓	✓	✓	✓	✓	✓	✓	✓	-	✓	-

TABLE I. SENSORS ON VARIOUS MOBILE HANDSETS

Future research should leverage well-understood solutions towards comprehensive architectures that systematically address all aforementioned challenges. Nevertheless, such a holistic treatment mandates a deep reflection on the current status of people-centric sensing.

Contributions: The aim of this study is to explore, in detail, the factors that influence security, user privacy and incentives in PCS. In our analysis, we consider the security and privacy of participants (i.e., *protecting users from the system*), as well as the trustworthiness of the PCS campaigns (i.e., *protecting the system from the users*). We start by identifying the involved stakeholders and actors and we craft an application-agnostic system model (Sec. II). We compile a set of security and privacy requirements (Sec. III) that includes user incentivization and accountability properties. This leads to a refined adversarial model that incorporates possible *collusion* between PCS entities (Sec. IV); such enhanced attacker capabilities are highly relevant to the complex multi-user PCS environment. Then, we critically survey the state-of-the-art research efforts in the area, with respect to their security and privacy (Sec. V). Based on our findings, we posit open issues and challenges (Sec. VI), and discuss possible ways to address them; so that security and privacy vulnerabilities do not become inextricable impediments as PCS evolves towards commercial deployment.

II. PEOPLE-CENTRIC SENSING

People-centric sensing lies at the intersection of several research domains, including sensor networking, ubiquitous computing, machine learning and social networking. The literature has identified different approaches for PCS: *mobile crowd-sensing* [5], *urban sensing* [3], *participatory* [2] or *opportunistic sensing* [4]. While all these terms are closely related, they can be different by system aspects such as the *nature* of collected data (e.g., environmental data, health data, etc.) and the *degree of user involvement* in the sensing process. For example, active user engagement in participatory systems versus automated data collection without user involvement in opportunistic sensing campaigns. Nevertheless, the common denominator is the incorporation of people in the sensing loop (as *custodians of sensing devices*), which leads to a mixture of human and machine intelligence, enabling new classes of collective applications [23].

Indeed, this kind of pervasive computing is proving its usability in various domains ranging from environmental monitoring [24, 25] and urban sensing [26, 27, 28] to intelligent transportation systems [29, 30], assistive healthcare [31, 32] and public safety [33]. A new breed of sensing applications is motivated by the ever increasing number of online *social network* users. Such systems try to exploit social networks (formed by groups of users) in an attempt to create a contextual picture of their surroundings, improving for example the classification of locations as points of interest [34, 35].

Despite this broad gamut of emerging practices, there is little consensus on the underlying system architectures with different people-centric sensing applications assuming different system models. In what follows, we derive a generic design (including stakeholders and architectural components) to capture all key features of this complex environment.

A. Stakeholders and System Model

The actors involved in people-centric sensing can be broadly classified as (i) *users*, (ii) *task initiators* and (iii) *infrastructure components* (Figure 1). Users are participants operating their mobile devices (e.g., smart-phones, tablets, smart vehicles and wearable platforms), equipped with multiple embedded sensors (Table I) 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) and cellular networks (3rd Generation (3G) and Long Term Evolution (LTE)).

Task initiators (or campaign administrators) are defined [36] as organizations, public authorities or, even, individuals that initiate targeted data collection campaigns, by recruiting users and distributing sensing tasks to them. Based on the adopted sensing approach, the recruitment depends on the desired degree of user involvement in the sensing process [37]. For instance, participatory sensing assumes that volunteers opt in to sensing tasks and contribute information out of personal interest or altruism [38]. Such a modus operandi requires users to (periodically) *pull* the list of active tasks and select the ones that they wish to participate in. The other strategy, i.e., opportunistic sensing, diminishes user control on the data collection process. Mobile devices will automatically join a sensing task as long as the *explicit task requirements* (described later in the section) are met. This variant employs a *push* discipline according to which the task providers, or the underlying infrastructure, forward tasks to participating devices.

There is no standard (to the best of our knowledge) governing the structure of a sensing task or the information that should be included in its description [39]. Nevertheless, based on our discussion on various PCS applications, such a specification should (at least) contain the following attributes:

- **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 statistical property of a sensing modality (e.g., minimum, maximum, average, median temperature). It might 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 [40].
- **Area of Interest**: The locality within which the mobile devices must contribute sensed data. The area of interest for a task 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 the sensing task is active.
- **Reporting Frequency**: Definition of the periodicity and the conditions under which mobile devices submit the values of their sensing modalities. For example, a task might request devices to submit data every $1\,\mathrm{min}$ or whenever the value of the sensed phenomenon is within (or not) some predefined range (e.g., noise pollution exceeds $140\ dB$).
- **Eligibility Criteria**: People-centric sensing heavily relies on mass user participation. Nevertheless, sensing-capable devices are, still, fragmented when considering their sensing and computational capabilities (for example not all phones have a barometer). Furthermore, sensing tasks might require contributions from specific user groups (e.g., cyclists). As a result, a task description must contain the eligibility criteria that define the conditions under which user participation is allowed or desired.
- **Rewards**: Description of the incentives and the rewards that users shall receive for their data reports.

The back-end infrastructure is responsible for supporting the life-cycle of a sensing task. Some works [36, 17, 18] suggest, or assume, the existence of a *centralized infrastructure* component to which users submit their data. Usually, such a centralized server is run by the campaign administrators so that they can have direct access to the results of the sensory tasks (i.e., reported sensor readings). As shown in Figure 1, the campaign administrator initiates sensing tasks by either uploading the corresponding specifications to a central repository (Step 1) or by directly distributing them to the targeted pool of users (Step 2). Users can then start participating in the sensing process and upload their contributions to the central server (Step 3); which may perform some type of aggregation on the submitted reports, depending on the application scenario. The collected (and possibly analyzed) sensor readings are finally released in various forms, e.g., attached to maps or as statistics, made available to the participants or to a larger public.

Another approach that has recently started gaining momentum leverages a *decentralized* set-up [16, 22, 41]: the campaign administrator directly *tasks* and queries participating devices for sensing data (Step 4). Nevertheless, such sensing systems still require some form of centralized infrastructure to enable the interaction of campaign administrators and users' devices. Regardless the approach (i.e., centralized or decentralized), campaign administrators can offer rewards to users for their participation (Step 5). As mentioned, providing incentives is necessary to attract large numbers of users and, thus, ensure a continuous influx of high quality contributions (more details can be found in Sec. V).

III. Security and Privacy Requirements

Seeking to design PCS successful systems, one has to cater to the security and privacy requirements of all involved actors and stake-holders. 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). Achieving these two objectives is not straight-forward. Thus, a necessary first step is a clear definition of security and privacy requirements PCS systems must meet (illustrated in Figure 2):

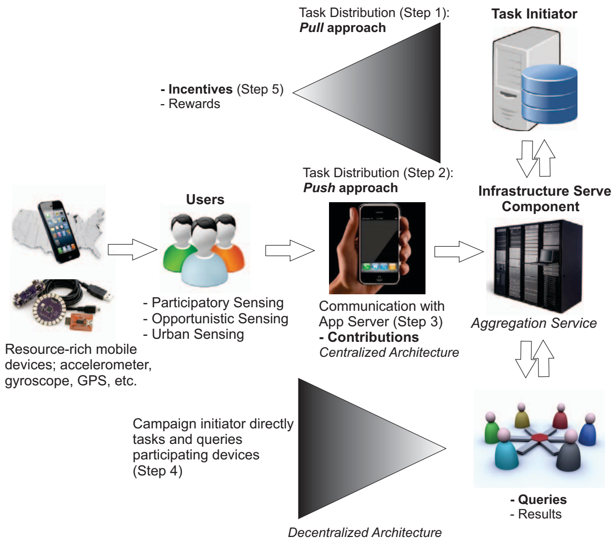
Task Distribution (Step 1):
Pull approach
Task Initiator
- Incentives (Step 5)
- Rewards
Task Distribution (Step 2):
Push approach
Infrastructure Server Component
Users
- Participatory Sensing
- Opportunistic Sensing
- Urban Sensing
Resource-rich mobile devices; accelerometer, gyroscope, GPS, etc.
Communication with App Server (Step 3)
- Contributions
Centralized Architecture
Aggregation Service
Campaign initiator directly tasks and queries participating devices (Step 4)
- Queries
- Results
Decentralized Architecture

Fig. 1. Architectural Overview of People-Centric Sensing

- **R1: Privacy Preserving Participation**: Users should participate in sensing tasks 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).

Furthermore, user participation must be *unobservable*: no external (e.g., cellular providers or ISPs) or internal (i.e., PCS infrastructure components or users) observers should be able to deduce that an anonymous user has (or will) participated in a task. The level of the achieved anonymity and unobservability strongly depends on the task's *anonymity set* (i.e., the number of users that participate in a given task). More specifically, a user is unidentifiable and indistinguishable from the rest of it's anonymity set.

Nonetheless, such notions of anonymity and unobservability do not preclude inference attacks against user privacy. Indeed, observers, leveraging filtering techniques [42], could still infer, with relatively high probability, the identity and the actions of anonymous users. Towards this, a PCS system should provide *strong probabilistic privacy*: From the perspective of the observer, any action (e.g., report submission, joining or quitting tasks) could have been performed by any user belonging to the anonymity set of a task [43].

In PCS, 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 [42] 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. To mitigate such attacks, *report unlinkability* is necessary. More specifically, it should be hard for an observer to link together reports originating from the same user and, thus, reconstruct (parts of) her whereabouts.

- **R2: Privacy-Preserving, Resilient Incentive Mechanisms and Fairness**: Users should receive credits and rewards for their participation without associating themselves with the data or the tasks 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.

Ideally, as users are expected to offer (i.e., contribute measurements) and receive services (e.g., query the results of a sensing task), the rewards they get should correspond to the extent of their contributions. Nevertheless, selfish users might attempt to receive disproportional, compared to their contributions, services or refrain from engaging with the system. To this end, the PCS system should discourage, detect and isolate such selfish behavior.

- **R3: Communication integrity, confidentiality and authentication**: All PCS entities should be authenticated and their communications should be protected from any alteration and disclosure to unauthorized parties.
- **R4: Authorization and Access Control**: Participating users should act according to the policies specified by the sensing task, defined by campaign administrators. To enforce such policies, *access control* and *authorization* services must be in place.
- **R5: Data-Centric Trust**: The PCS system should provide the means and mechanisms to assess the trust-worthiness and the validity of user submitted data.
- **R6: 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. The PCS system 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.

IV. Threats to PCS

Adversaries in the PCS context can be broadly categorized as *external* and *internal*. The former are unauthorized actors, i.e., not PCS system entities, that try to degrade the performance of the system and disrupt its operation. They can eavesdrop, modify and forge messages. Furthermore, they can target the availability of the system by launching clogging, e.g., D(D)oS, attacks. Since external adversaries have no association with the PCS system, their disruptive capabilities are relatively limited.

Internal adversaries can launch more sophisticated attacks (with far more grave implications) that cannot be easily mitigated. Any PCS entity can pose as an internal adversary: users, task administrators or, even, the PCS infrastructure itself (see Sec. II). Users might be *malicious* or *selfish*. Malicious users (acting with compromised devices) can exhibit arbitrary behavior, completely deviating from the expected functionality and communication protocols. They might also attempt to impersonate other PCS users or even try to simultaneously pose as multiple, authorized, ones (i.e., act as a *sibyl*). Malicious users (devices) could contribute faulty but seemingly valid reports, thus, *polluting* the data collection process and degrading the usefulness of the collected data. Such orchestrated pollution attacks, usually, target the overall result of a sensing task which is 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., CO2, CO) in an attempt to manipulate the aggregative result and avoid further consequences (e.g., pollution fines).

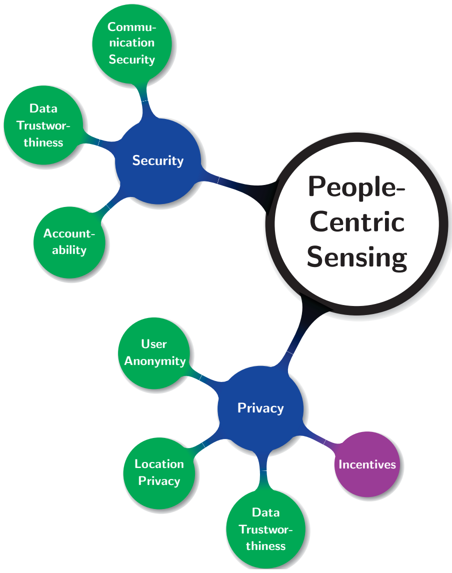

Fig. 2. Security and Privacy Requirements for PCS

Selfish users aim at minimizing their *effort-to-utility* ratio. Such a behavior is relevant to tasks and sensing campaigns 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. We broadly categorize such behavior as follows:

- **Task Leeching:** This term originates from P2P systems and describes users that exhibit a *hit-and-run* approach; they try to benefit with minimal (or non-existent) contributions. In the context of PCS, leechers would 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).
- **Low Quality Contributions:** Instead of honestly contributing the values of required sensing modalities, selfish users might deliberately submit low quality data (e.g., previously acquired, non-fresh or random sensor readings among others). This behavior resembles the data pollution attacks; the difference being that pollution attacks target the overall result of a sensing task and not single data measurements.
- **Double-spending Redeemed Quotas:** This is a combination of abusing and selfish behavior where users try to exploit the incentive mechanisms in order to reclaim already awarded rewards (e.g., monetary rewards).

Adversarial behavior is not limited only to users; campaign administrators and PCS infrastructure components can also misbehave. It is in the benefit of campaign administrators to ensure high added value for the sensing tasks they manage. Nevertheless, this does not preempt attacks against user privacy. More specifically, campaign administrators might leverage PCS specific actions to deanonymize users [17]. To reduce the size of the anonymity set, and, thus, track and deanonymize users, campaign administrators can initiate tasks with strict eligibility criteria satisfiable by only a few (or even a single) devices. In that case, even the strongest privacy-protection mechanisms cannot offer adequate privacy.

Furthermore, components of the PCS infrastructure can be *honest-but-curious*: They may implement the correct communication and authentication protocols but they are curious to learn private user data. Multiple curious such entities might *collude* to de-anonymize users. Finally, PCS infrastructure entities can be also malicious and exhibit arbitrary behavior.

V. Current Status of Security, Privacy and User Incentives

In this section we survey the state-of-the-art literature for secure and privacy-preserving PCS. We refrain from presenting stand-alone security solutions and privacy-enhancing technologies and we, instead, focus on architectures that address subsets of the requirements defined in Sec. III.

AnonySense [17] was one of the first works to propose a general-purpose security and privacy architecture for PCS. It tessellates geographical areas to achieve *statistical k-anonymity* [44]; individuals cannot be identified within a set of k users assumed to reside in the same area at a given moment in time. This approach prevents inference attacks aiming to link reports back to users. As a second layer of protection, AnonySense aggregates (at least λ) user reports before sending them to the campaign administrator. To achieve user anonymity, AnonySense leverages *group signatures*; cryptographic schemes that enable users to anonymously sign their reports. Anonymity is conditional in the sense that, in case it is needed, a signature can be opened to reveal the real identity of the user. This property can be used to evict misbehaving users from the system. Nevertheless, filtering out past and faulty contributions of such offending users requires also the de-anonymization of benign reports. Furthermore, due to the way that AnonySense employs group-signatures it is vulnerable to sibyl attacks (see Sec. IV).

PoolView [45] 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 and, thus, does not cover aspects such as security, accountability and data trustworthiness.

PEPSI [18] 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 [46], 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 [47].

TAPAS [20] presents a participatory sensing framework that enables privacy-preserving user contributions. Additionally, it considers data trust-worthiness 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).

	PCS Requirements					
Work	**R1**	**R2**	**R3**	**R4**	**R5**	**R6**
AnonySense [17]	✓	✗	✓	✗	✗	✓
PoolView [45]	✓	✗	✓	✗	✗	✗
Pepsi [18]	✓	✗	✓	✓	✗	✗
TAPAS [20]	✓	✗	✓	✗	✓	✗
[48]	✗	✗	✓	✗	✓	✗
PEPPeR [22]	✓	✗	✓	✓	✗	✗
[16, 49]	✓	✗	✓	✗	✗	✗
[50, 51]	✓	✗	✓	✗	✓	✗
[52]	✓	✓	✗	✗	✗	✗

TABLE II. Comparative Analysis of State-of-the-art

In [48] 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 sibyl attacks. Nevertheless, the use of TPM cannot protect the privacy of participating users.

PEPPeR [22] 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 [19] 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 [16] and [49] 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 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 [50]. Similarly, in [51] 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 PCS.

Significant efforts have been made on the provision of *incentives* to stimulate user participation [12, 13, 14, 53, 54]. 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 [52] 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 II contains a comparative analysis of all the research efforts discussed in this section. We evaluate each one of the presented solutions with respect to the requirements defined in Sec. III. Despite the intense research interest in the area of secure and privacy-preserving PCS systems, we still lack a comprehensive solution that can meet all the key requirements. While it is feasible to combine some of the aforementioned works, it is not an easy task.

VI. Road-Map and Future Prospects for People-Centric Sensing

A. Data Trustworthiness: Can we Trust Users?

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 (see Sec. V) touch upon such aspects of PCS, we are far from conclusive and convincing answers. Data-trustworthiness turns out to be a research challenge well beyond the boundaries of people-centric sensing. 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 straight-forward task.

Meeting this challenge requires work on three different manifestations of trust [55]; (i) *user (default) trust*, (ii) *task-specific trust* and (iii) *dynamic trust*. The default user trust depends on the attributes of a user's mobile device. 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 task-specific trust is relevant. Simply put, devices without accelerometers and/or inertia sensors cannot be trusted for, e.g., tasks related to traffic congestion monitoring. Bring Your Own Device (BYOD) solutions are of interest for such trust assessment.

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?* Again, solutions addressing the security of BYOD systems could help. Furthermore, the trustworthiness of the data, reported by a device, should decrease as the device moves away from the point of interest. Finally, compromised or modified (e.g., jail-broken) devices should not be trusted.

In addition, how should the PCS system cope with unreliable users (and their devices), especially if such a detection and identification of misbehaving devices/users is made? Fortunately, such aspects have been extensively considered in related areas, e.g., in vehicular communication systems; various eviction schemes have been proposed [56, 57, 58, 59, 60] and can serve as a starting point for PCS.

B. User Empowerment: Putting Users in the Privacy Equation

The design of PCS systems must take into consideration the main stake-holder: users should be explicitly informed about the privacy implications their participation entails. The notion of privacy is highly individual and depends on the user's views and understanding. Nevertheless, this user empowerment should not presume that users are privacy-experts. What is needed, actually, is to include the participants in the privacy equation, but also reduce their *friction* (i.e., effort) to understand the underlying complex configurations. For example, users might not understand that providing the values of their accelerometer can reveal the location of their home [61]. To this end, PCS applications should work on two directions: (*i*) provide users with privacy recommendations and (*ii*) protect the user whenever possible.

The latter requirement is well understood and has been intensively addressed by the research community (see Sec. V). The former remains an open challenge as it requires Machine Learning (ML) algorithms that not only understand the possible privacy leakage risks, relating to a task, but can also translate them in a user friendly manner. This research direction is currently gaining momentum and various works are exploring such advanced privacy aspects [8, 9].

C. 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 [62]. The user's personal context that can be inferred from such multimodal sensor streams [61] 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 [16, 41]. 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 trustworthiness of PCS systems. Such adaptive solutions allow us to better cope with the multi-party nature of people-centric sensing environments and meet the requirements of all involved PCS actors and stakeholders.

VII. Conclusions

Ensuring user privacy and system trustworthiness is, perhaps, the most fundamental requirement for people-centric sensing applications. In this work, we discussed key aspects of security and privacy for PCS systems and surveyed the state-of-the-art literature. Our investigation showed that existing solutions focus on facets of the problem at hand. Furthermore, by taking into consideration the salient characteristics of PCS along with the requirements of the involved stake-holders, we identified a number of open research challenges. It is our strong belief that if these challenges are tackled now while PCS is still at an early stage, then, this emerging paradigm can reach its full potential.

REFERENCES

[1] International Communication Union. *"The World in 2013: ICT Facts and Figures"*. July 2013. URL: http://www.itu.int/ITU-D/ict/facts/material/ICTFactsFigures2013.pdf.

[2] A. T. Campbell, S. B. Eisenman, N. D. Lane, E. Miluzzo, R. A. Peterson, H. Lu, X. Zheng, M. Musolesi, K. Fodor, and G.-S. Ahn. "The Rise of People-Centric Sensing". In: *IEEE Internet Computing* 12.4 (July 2008), pp. 12–21.

[3] D. Cuff, M. Hansen, and J. Kang. "Urban Sensing: Out of the Woods". In: *Communications ACM* 51.3 (Mar. 2008), pp. 24–33.

[4] R. K. Ganti, F. Ye, and H. Lei. "Mobile crowdsensing: current state and future challenges." In: *IEEE Communications Magazine* 49.11 (2011), pp. 32–39.

[5] N. D. Lane, E. Miluzzo, H. Lu, D. Peebles, T. Choudhury, and A. T. Campbell. "A Survey of Mobile Phone Sensing". In: *Communications Magazine* 48.9 (Sept. 2010), pp. 140–150.

[6] M. Brereton, P. Roe, M. Foth, J. M. Bunker, and L. Buys. "Designing Participation in Agile Ridesharing with Mobile Social Software". In: *Proceedings of the 21st Annual Conference of the Australian Computer-Human Interaction Special Interest Group*. OZCHI. Melbourne, Australia, 2009.

[7] I. Krontiris, F. Freiling, and T. Dimitriou. "Location privacy in urban sensing networks: research challenges and directions". In: *IEEE Wireless Communications* 17.5 (Oct. 2010), pp. 30–35.

[8] M. Götz, S. Nath, and J. Gehrke. "MaskIt: Privately Releasing User Context Streams for Personalized Mobile Applications". In: *Proceedings of the 2012 ACM SIGMOD International Conference on Management of Data*. Scottsdale, Arizona, USA, 2012.

[9] A. Parate, M.-C. Chiu, D. Ganesan, and B. M. Marlin. "Leveraging Graphical Models to Improve Accuracy and Reduce Privacy Risks of Mobile Sensing". In: *Proceeding of the 11th Annual International Conference on Mobile Systems, Applications, and Services*. MobiSys. Taipei, Taiwan, 2013.

[10] G. Greenwald. *NSA Prism Program Taps in to User Data of Apple, Google and Others*. June 2013. URL: http://www.theguardian.com/world/2013/jun/06/us-tech-giants-nsa-data.

[11] P. Kollock. "The economies of online cooperation: gifts and public goods in cyberspace". In: *Communities in the cyberspace*. 1st ed. Routledge, Feb. 1999, pp. 259–262.

[12] E. Androulaki, S. G. Choi, S. M. Bellovin, and T. Malkin. "Reputation systems for anonymous networks". In: *Privacy Enhancing Technologies*. 2008.

[13] T. Luo and C.-K. Tham. "Fairness and social welfare in incentivizing participatory sensing." In: *IEEE SECON*. 2012.

[14] I. Krontiris and A. Albers. "Monetary incentives in participatory sensing using multi-attributive auctions". In: *International Journal of Parallel, Emergent and Distributed Systems* 27.4 (2012), pp. 317–336.

[15] K. L. Huang, S. S. Kanhere, and W. Hu. "Are You Contributing Trustworthy Data?: The Case for a Reputation System in Participatory Sensing". In: *Proceedings of the 13th ACM International Conference on Modeling, Analysis, and Simulation of Wireless and Mobile Systems*. MSWIM. Bodrum, Turkey, 2010.

[16] I. Boutsis and V. Kalogeraki. "Privacy preservation for participatory sensing data". In: *Pervasive Computing and Communications (PerCom), IEEE International Conference on*. 2013.

[17] C. Cornelius, A. Kapadia, D. Kotz, D. Peebles, M. Shin, and N. Triandopoulos. "Anonysense: privacy-aware people-centric sensing". In: *Proceedings of the 6^{th} international conference on Mobile systems, applications, and services*. Breckenridge, USA, 2008.

[18] E. De Cristofaro and C. Soriente. "Extended Capabilities for a Privacy-Enhanced Participatory Sensing Infrastructure (PEPSI)". In: *IEEE Transactions on Information Forensics and Security* 8.12 (2013), pp. 2021–2033.

[19] T. Das, P. Mohan, V. N. Padmanabhan, R. Ramjee, and A. Sharma. "PRISM: platform for remote sensing using smartphones". In: *Proceedings of the 8^{th} international conference on Mobile systems, applications, and services*. San Francisco, USA, 2010.

[20] L. Kazemi and C. Shahabi. "TAPAS: Trustworthy privacy-aware participatory sensing". In: *Knowledge and Information Systems* 37.1 (2013), pp. 105–128.

[21] L. Kazemi and C. Shahabi. "Towards preserving privacy in participatory sensing". In: *Pervasive Computing and Communications (PerCom) Workshops*. 2011.

[22] T. Dimitriou, I. Krontiris, and A. Sabouri. "PEPPeR: A querier's Privacy Enhancing Protocol for PaRticipatory sensing". In: *Security and Privacy in Mobile Information and Communication Systems*. Springer, 2012, pp. 93–106.

[23] S. Tilak. "Real-World Deployments of Participatory Sensing Applications: Current Trends and Future Directions". In: *Int. Scholarly Research Notices for Sensor Networks* (2013).

[24] M. v. Kaenel, P. Sommer, and R. Wattenhofer. "Ikarus: Large-scale Participatory Sensing at High Altitudes". In: *Proceedings of the 12^{th} Workshop on Mobile Computing Systems and Applications*. Phoenix, USA, 2011.

[25] D. Mendez, A. Perez, M. Labrador, and J. Marron. "P-Sense: A Participatory Sensing system for air pollution monitoring & control". In: *IEEE International Conference on Pervasive Computing and Communications*. 2011.

[26] E. Miluzzo, N. D. Lane, K. Fodor, R. Peterson, H. Lu, M. Musolesi, S. B. Eisenman, X. Zheng, and A. T. Campbell. "Sensing Meets Mobile Social Networks: The Design, Implementation and Evaluation of the CenceMe Application". In: *Proceedings of the 6th ACM Conference on Embedded Network Sensor Systems*. SenSys. Raleigh, NC, USA, 2008.

[27] L. Deng and L. P. Cox. "LiveCompare: Grocery Bargain Hunting Through Participatory Sensing". In: *Proceedings of the 10th Workshop on Mobile Computing Systems and Applications*. HotMobile. Santa Cruz, California, 2009.

[28] E. Miluzzo, M. Papandrea, N. D. Lane, A. M. Sarroff, S. Giordano, and A. T. Campbell. "Tapping into the Vibe of the City Using VibN, a Continuous Sensing Application for Smartphones". In: *Proceedings of 1st International Symposium on From Digital Footprints to Social and Community Intelligence*. SCI. Beijing, China, 2011.

[29] 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: *Proceedings of the 4^{th} international conference on Embedded networked sensor systems*. Boulder, USA, 2006.

[30] A. Thiagarajan, L. Ravindranath, K. LaCurts, S. Madden, H. Balakrishnan, S. Toledo, and J. Eriksson. "VTrack: Accurate, Energy-aware Road Traffic Delay Estimation Using Mobile Phones". In: *Proceedings of the 7^{th} ACM Conference on Embedded Networked Sensor Systems*. Berkeley, USA, 2009.

[31] T. Giannetsos, T. Dimitriou, and N. R. Prasad. "People-centric sensing in assistive healthcare: Privacy challenges and directions". In: *Security and Communication Networks* 4.11 (Nov. 2011), pp. 1295–1307.

[32] N. Lane, M. Mohammod, M. Lin, X. Yang, H. Lu, S. Ali, A. Doryab, E. Berke, T. Choudhury, and A. Campbell. "BeWell: A Smartphone Application to Monitor, Model and Promote Wellbeing". In: *5th International ICST Conference on Pervasive Computing Technologies for Healthcare*. Apr. 2012.

[33] J. Ballesteros, M. Rahman, B. Carbunar, and N. Rishe. "Safe cities. A participatory sensing approach". In: *IEEE 37th Conference on Local Computer Networks*. 2012.

[34] I. Krontiris and F. C. Freiling. "Integrating people-centric sensing with social networks: A privacy research agenda." In: *Pervasive Computing and Communications (PerCom) Workshops*. IEEE, 2010.

[35] D. Peebles, H. Lu, N. D. Lane, T. Choudhury, and A. T. Campbell. "Community-Guided Learning: Exploiting Mobile Sensor Users to Model Human Behavior." In: *Proceedings of the Twenty-Fourth AAAI Conference on Artificial Intelligence*. 2010.

[36] D. Christin, A. Reinhardt, S. S. Kanhere, and M. Hollick. "A survey on privacy in mobile participatory sensing applications". In: *Journal of Systems and Software* 84.11 (2011), pp. 1928 –1946.

[37] N. D. Lane, S. B. Eisenman, M. Musolesi, E. Miluzzo, and A. T. Campbell. "Urban Sensing Systems: Opportunistic or Participatory?" In: *Proceedings of the 9th Workshop on Mobile Computing Systems and Applications*. HotMobile. 2008.

[38] J. Goncalves, D. Ferreira, S. Hosio, Y. Liu, J. Rogstadius, H. Kukka, and V. Kostakos. "Crowdsourcing on the Spot: Altruistic Use of Public Displays, Feasibility, Performance, and Behaviours". In: *Proceedings of the 2013 ACM International Joint Conference on Pervasive and Ubiquitous Computing*. UbiComp. Zurich, Switzerland, 2013.

[39] A. Kapadia, D. Kotz, and N. Triandopoulos. "Opportunistic Sensing: Security Challenges for the New Paradigm". In: *Proc. of the International Conference on COMmunication Systems And NETworks*. Bangalore, India, 2009.

[40] J. Burke, D. Estrin, M. Hansen, A. Parker, N. Ramanathan, S. Reddy, and M. B. Srivastava. "Participatory sensing". In: *In: Workshop on World-Sensor-Web: Mobile Device Centric Sensor Networks and Applications*. 2006.

[41] I. Krontiris and T. Dimitriou. "Privacy-Respecting Discovery of Data Providers in Crowd-Sensing Applications". In: *IEEE International Conference on Distributed Computing in Sensor Systems* 0 (2013).

[42] R. E. Kalman. "A New Approach to Linear Filtering & Prediction Problems". In: *Transactions of the ASME–Journal of Basic Engineering* 82.Series D (1960), pp. 35–45.

[43] P. Papadimitratos, V. Gligor, and J.-P. Hubaux. "Securing Vehicular Communications - Assumptions, Requirements, and Principles." In: *Proceedings of 4th Workshop on Embedded Security in Cars (ESCAR)*. Berlin, Germany, 2006.

[44] L. Sweeney. "k-anonymity: A Model for Protecting Privacy". In: *International Journal of Uncertainty, Fuzziness and Knowledge Based Systems.* 10.5 (Oct. 2002), pp. 557–570.

[45] R. K. Ganti, N. Pham, Y.-E. Tsai, and T. F. Abdelzaher. "PoolView: Stream Privacy for Grassroots Participatory Sensing". In: *Proceedings of the 6th ACM Conference on Embedded Network Sensor Systems*. SenSys '08. Raleigh, NC, USA, 2008.

[46] D. Boneh and M. Franklin. "Identity-Based Encryption from the Weil Pairing". In: Springer-Verlag, 2001, pp. 213–229.

[47] S. Al-Riyami and K. Paterson. "Certificateless Public Key Cryptography". In: *Advances in Cryptology - ASIACRYPT 2003*. Vol. 2894. Lecture Notes in Computer Science. Springer Berlin Heidelberg, 2003, pp. 452–473.

[48] A. Dua, N. Bulusu, W.-C. Feng, and W. Hu. "Towards Trustworthy Participatory Sensing". In: *Proceedings of the 4th USENIX Conference on Hot Topics in Security*. HotSec. Montreal, Canada, 2009.

[49] S. Gao, J. Ma, W. Shi, G. Zhan, and C. Sun. "TrPF: A Trajectory Privacy-Preserving Framework for Participatory Sensing". In: *IEEE Transactions on Information Forensics and Security* 8.6 (2013), pp. 874–887.

[50] H. Ahmadi, N. Pham, R. Ganti, T. Abdelzaher, S. Nath, and J. Han. "Privacy-aware Regression Modeling of Participatory Sensing Data". In: *Proceedings of the 8th ACM Conference on Embedded Networked Sensor Systems*. SenSys. 2010.

[51] K. Xing, Z. Wan, P. Hu, H. Zhu, Y. Wang, X. Chen, Y. Wang, and L. Huang. "Mutual privacy-preserving regression modeling in participatory sensing". In: *INFOCOM, Proceedings IEEE*. 2013.

[52] B. O. Holzbauer, B. K. Szymanski, and E. Bulut. "Socially-aware Market Mechanism for Participatory Sensing". In: *Proceedings of the First ACM International Workshop on Mission-oriented Wireless Sensor Networking*. MiSeNet. Istanbul, Turkey, 2012.

[53] J.-S. Lee and B. Hoh. "Dynamic pricing incentive for participatory sensing." In: *Pervasive and Mobile Computing* 6.6 (2010), pp. 693–708.

[54] S. Reddy, D. Estrin, M. Hansen, and M. Srivastava. "Examining Micro-payments for Participatory Sensing Data Collections". In: *Proceedings of the 12th ACM International Conference on Ubiquitous Computing*. Copenhagen, Denmark, 2010.

[55] M. Raya, P. Papadimitratos, V. Gligor, and J.-P. Hubaux. "On Data-Centric Trust Establishment in Ephemeral Ad hoc Networks". In: *Proceedings of the 28th IEEE Conference on Computer Communications (INFOCOM)*. Phoenix, AZ, USA, 2008.

[56] T. Moore, M. Raya, J. Clulow, P. Papadimitratos, R. Anderson, and J.-P. Hubaux. "Fast Exclusion of Errant Devices from Vehicular Networks". In: *Proceedings of the Fifth IEEE-CS Conference on Sensor, Mesh and Ad Hoc Communications and Networks (SECON)*. San Francisco, CA, USA, 2008.

[57] P. Papadimitratos, G. Mezzour, and J.-P. Hubaux. "Certificate Revocation List Distribution in Vehicular Communication Systems". In: *ACM MobiCom VANET*. San Francisco, CA, USA, 2008.

[58] M. Raya, P. Papadimitratos, I. Aad, D. Jungels, and J.-P. Hubaux. "Eviction of Misbehaving and Faulty Nodes in Vehicular Networks". In: *IEEE Journal on Selected Areas in Communications* 25.8 (2007), pp. 1557–1568.

[59] S. Gisdakis, M. Lagana, T. Giannetsos, and P. Papadimitratos. "SEROSA: SERvice Oriented Security Architecture for Vehicular Communications". In: *Vehicular Networking Conference (VNC), IEEE*. Boston, USA, 2013.

[60] N. Alexiou, M. Laganà, S. Gisdakis, M. Khodaei, and P. Papadimitratos. "VeSPA: Vehicular Security and Privacy-preserving Architecture". In: *Proceedings of the 2nd ACM Workshop on Hot Topics on Wireless Network Security and Privacy*. HotWiSec. Budapest, Hungary, 2013.

[61] A. Santos, L. Tarrataca, J. Cardoso, D. Ferreira, P. Diniz, and P. Chainho. "Context Inference for Mobile Applications in the UPCASE Project". In: *MobileWireless Middleware, Operating Systems, and Applications*. Vol. 7. Lecture Notes of the Institute for Computer Sciences, Social Informatics and Telecommunications Engineering. Springer Berlin Heidelberg, 2009, pp. 352–365.

[62] R. Shokri, P. Papadimitratos, G. Theodorakopoulos, and J.-P. Hubaux. "Collaborative Location Privacy". In: *IEEE 8th International Conference on Mobile Adhoc and Sensor Systems (MASS)*. 2011.