On the Provisioning of Vehicle-Based Public Sensing Services

Sherin Abdelhamid  
School of Computing  
Queen’s University  
Kingston, ON, Canada  
sherin@cs.queensu.ca

Hossam S. Hassanein  
School of Computing  
Queen’s University  
Kingston, ON, Canada  
hossam@cs.queensu.ca

Glen Takahara  
Department of Mathematics & Statistics, Queen’s University  
Kingston, ON, Canada  
takahara@mast.queensu.ca

ABSTRACT
Thanks to their abundant on-board resources, ubiquity, and mobility, smart vehicles can be considered major candidates for providing pervasive information services. With the diversity of in-vehicle sensors along with abundant storage, processing, and communication capabilities, smart vehicles can bring a wide scope of applications into action under the public sensing paradigm outstripping other candidate mobile resources such as smartphones. In this paper, we propose the vehicular public sensing (VPS) platform that aims at utilizing the abundant resources of smart vehicles for providing ubiquitous public sensing services. The VPS platform encompasses underlying components that address the recruitment, communication, sensing, reporting, and data analytics functionalities of a typical public sensing process. Taking into account different environmental and practical set-ups, the VPS platform provides potential adjustments and different approaches for the operation of each component. We anticipate that by engaging smart vehicles in providing ubiquitous sensing-based services, a plethora of information services and applications will be unleashed bringing a new era of service provisioning.

Categories and Subject Descriptors
H.4.3 [Information Systems Applications]: Communications Applications; C.2.3 [Computer-Communication Networks]: Network Operations

General Terms
Design, Management, Services

Keywords
Smart vehicles; Public sensing; Service provisioning

1. INTRODUCTION
With the ubiquitous availability of smartphones, a data collection and service provisioning paradigm has emerged in the past few years to utilize such mobile resources for providing sensing-based services. Such a paradigm is known as “public sensing” as it utilizes public owned devices and assistance for providing sensory services [9][10]. Although smartphones have achieved high benefits under the public sensing paradigm, their use suffers from limitations due to the scarcity of their on-board resources especially the energy resource. Concurrently, smart vehicles have been gaining wide interest from researchers and service providers as major ubiquitous resource providers with their ample sensing, storage, processing, and communication capabilities [2]. With the abundance of these on-board resources in smart vehicles engaged with the vehicular ubiquity and mobility, such vehicles can be considered major candidates for providing public sensing services that can go beyond the array of services provided by other mobile resources such as smartphones.

In [4], we presented a categorization of the public sensing applications that can be provided by smart vehicles into instant sensing and on-move sensing applications. The instant sensing applications are those that require only instantaneous readings such as reporting weather conditions and pollution levels. Contrarily, the on-move sensing applications require continuous reporting by vehicles while moving in an area of interest (AoI). Examples of this category are reporting road anomalies, traffic conditions, and parking availability. Such applications/services can provide sensory data for the vehicle’s own use, other vehicles on roads, and/or third parties such as municipalities and data centers.

Generally speaking, the architecture of a public sensing system consists of three main elements: 1) data contributors/participants such as vehicles and smartphones, 2) data consumers/end-users, and 3) a service provider (SP) that is responsible for managing the whole process, obtaining data from the participants, and providing it to the end-users (see Figure 1). The public sensing process consists of three main stages: a) an SP interested in collecting data from a specific area of interest sends sensing requests to candidate participants, b) after generating the data of interest, the participants send it to the designated SP, c) after performing data analytics on the collected data, the SP publishes it to the end-users as a part of a subscribed service. The process can be adapted to handle situations when an end-user initiates the communication asking for specific information of interest.

Motivated by the hype of public sensing, the abundance of the hard-to-neglect vehicular resources, and the wide array of services that can be provided by smart vehicles under the
umbrella of the public sensing paradigm, in this paper we propose the Vehicular Public Sensing (VPS) platform that aims at utilizing vehicular resources of sensing, storage, processing, and communication for provisioning sensing-based information services. The platform encapsulates four components that collaboratively manage the whole process providing different approaches under different environmental setups.

The remainder of this paper is organized as follows. In Section 1, we discuss the general architecture of the VPS platform along with its functional scope. Section 3 delineates the functions of the underlying components of the proposed platform, the interactions among these components considering some practical and environmental setups, along with different functional approaches under each component. We summarize the discussion and present our future work in Section 4.

2. THE VEHICULAR PUBLIC SENSING PLATFORM – SCOPE AND ARCHITECTURE

Towards provisioning sensing-based vehicular information services, the proposed VPS platform targets providing an inclusive architecture that handles the full vehicular public sensing process starting from participant selection and task assignment until having the collected data ready for publishing/usage. To achieve this objective, the platform consists of four underlying components each of which is responsible for handling a stage of the process while interacting with the other components for a fully-managed public sensing service. These components are: 1) the recruitment component, 2) the communication component, 3) the sensing and reporting component, and 4) the data analytics component.

Under each component, the platform addresses challenges that may face the service provisioning process. For example, some challenges such as participant selection, incentive assignment, and the quality of information (QoI) diversity of participants may face an SP while handling participant recruitment. Data delivery challenges may arise while handling the communication between SPs and participants. Data assessment is a potential challenge when it comes to data analytics. The VPS platform covers some approaches that can be used by the platform components for handling such challenges.

The core architecture of the platform including the four underlying components and the interaction among them is shown in Figure 2. Details about the components and corresponding approaches under each component are discussed in the next section.

3. THE PLATFORM COMPONENTS

In this section, we delve into the functions of each component of the proposed VPS platform, its adaptation according to practical considerations, its interaction with the other components in the platform, and some approaches that can be employed for handling the operation of such a component.

3.1 The Recruitment Component

This component is mainly responsible for selecting a set of participants among all the possible candidates to handle a sensing task in a way that satisfies the application and cost requirements. A main input to the recruitment component is the recruitment policy which is used to define different requirements related to the task of interest and participant selection. These requirements include all thresholds related to the sensing attributes and the corresponding pecuniary issues. They also include all restrictions on the selected participants such as excluding participants who have not participated before, or confining the selection to participants owning a specific vehicle make/model.

A vital element of any recruitment policy is the incentive given to participants as a reward for the service they provide and to encourage them to keep engaged for future participation. In general, incentives can be of three types: 1) a participant’s willingness to serve the public, 2) receiving a service in return, or 3) receiving monetary rewards in return. Among these types, it has been shown that the incentives based on monetary values are the most encouraging ones. Monetary incentives can be in the form of pecuniary returns, parking passes, express route passes, or vouchers. Regardless of how that monetary value is represented, the SP paying such rewards to participants is in need of minimizing such paid rewards especially with having a budget cap controlling the recruitment process.

We remark that the average traffic density of an environment directs how participant recruitment can be handled. In that regard and based on the density factor, we distinguish between recruitment in urban and rural environments as follows.

a) Urban Environments

One of the major challenges that face an SP during the recruitment process in urban environments is the wide availability of candidate participants in an AoI. Due to the aforementioned incentive perspective, those available participants cannot all be recruited for a sensing task as they all will have to be paid for their participation. Therefore, efficient recruitment schemes should be used to select a sufficient number of trusted participants that can provide a required level of coverage for the intended AoI taking into account the following design considerations:

1. An SP should consider the availability of participants to limit the selection to those that are spatiotemporally available in the AoI.
2. With the diversity of participants’ behavior and the quality of vehicular resources and reported data, a re-
Recruiting SP has to consider participants’ reputation and QoI in the selection process to maximize the benefit and quality out of the recruited participants.

3. Since in practical scenarios a sensing task would have a budget limit that controls the amount of rewards paid out to the recruited participants, the recruitment process has to take this budget constraint into account.

Different approaches can be used to handle the underlying functionalities of the recruitment schemes. Such approaches can be categorized based on different aspects as follows.

– The considered availability scope.

One approach is to consider the instantaneous availability of participants to support applications that require instantaneous coverage. An example of a recruitment framework that uses such an approach is the one proposed by Reddy et al. in [16]. This framework is proposed mainly for the recruitment of smartphones for public sensing services but it can be adapted to support the recruitment of vehicles for instantaneous coverage applications. It targets selecting a set of participants that maximizes the coverage of an AoI under a budget limit. To achieve this objective, the authors consider the budgeted maximum coverage problem [11].

The other approach is to consider the on-move availability of participants for supporting the on-move sensing applications. Examples of recruitment schemes employing such an approach are proposed in [4][1]. These scheme are specifically targeting the recruitment of smart vehicles in urban environments. They utilize the trajectories of candidate participants as indicators of their on-move availability. The trajectory-based recruitment (TBR) scheme [4] focuses on finding the minimum number of participants that achieve a desired coverage to an AoI applying a minimal-cover greedy algorithm for selection. In [1], the authors present an optimal reputation-aware, trajectory-based recruitment framework that builds on the availability concept introduced in [4]. This framework is formulated as an integer linear programming (ILP) optimization problem to present performance benchmarks for two different recruitment objectives.

– The redundancy requirement.

One approach is to minimize redundancy/overlapping when selecting participants to cover an AoI. The motivation for the use of such an approach is to avoid the drawbacks of redundant reporting represented in the unnecessary cost incurred by the recruiting SP and the undesired waste for the communication bandwidth. This approach is used in [1] through its underlying ‘maximum coverage with minimum overlap’ recruitment objective.

Another approach is to accommodate a level of data redundancy upon demand. This approach is for handling cases when multiple readings from vehicles monitoring the same area are required to achieve a certain level of reliability especially when monitoring critical events on roads. In [4], the authors present a solution that employs such an approach.

– The used pricing model.

Many approaches can be used for formulating the pricing model and computing a recruitment cost for each participant. One simplistic, albeit not efficient approach is to assign an identical cost for all participants. This approach is used by the recruitment framework proposed in [16]. Such an approach is easy to use; however, it is unfair for the participants’ side.

Another approach is to employ a dynamic model that takes participant-related parameters into account. For example, the scheme proposed in [1] involves an underlying pricing model that takes participants’ reputation into consideration along with their availability in computing their recruitment cost.

One more approach is the reverse auction [12] in which the candidate participants get the flexibility of bidding for their data. The recruitment scheme in [8] is an example that utilizes such an approach.

b) Rural Environments

Due to the scarcity of candidate participants in rural environments, there is no need to use a participant selection scheme for recruitment. In rural environments, a simplistic
scheme that we call the ‘naive’ scheme can be used in which an SP tasks whatever participants are available in the AoI during the determined sensing period.

3.2 The Communication Component

The main function of the communication component is linking SPs and participants. It handles sending sensing requests/tasks by SPs to participants and sending the sensed data back. It is also responsible for delivering any control messages needed for the operation of the other components (e.g., vehicle trajectories needed for recruitment as highlighted in the previous section).

The operation of the communication component differs based on the availability of broadband connectivity. This is discussed in the following and illustrated through examples shown in Figure 3 where vehicle $S$ needs to send a packet to a service provider $P$.

a) Broadband-Connected Environments

In environments with ubiquitous availability of broadband communication infrastructure, the Internet can be adopted to be the dominant communication backbone supporting real-time communication between SPs and participants. For example, in Figure 3(a), the environment is broadband-connected, therefore vehicle $S$ manages to communicate directly with $P$ to deliver the packet using the backbone network.

b) Broadband-Restricted Environments

It may happen that the SP moderating the sensing and data collection process cannot use the Internet for communication with participants to avoid the cost of accessing broadband networks. Such restrictions of the use of broadband connectivity are encountered also in rural environments where it is not feasible to depend on the Internet for connectivity because of the lack/difficulty of using a broadband infrastructure.

Therefore, to accommodate the aforementioned practical setups, the delivery of both interest and data reply packets (i.e., sensing requests and sensed data, respectively) between the SPs and participants should be handled through wireless multi-hop communication between vehicles and, when applicable, between vehicles and roadside assistants. For example, in the broadband-restricted environment shown in Figure 3(b), $S$ has to deliver the packet using vehicle-to-vehicle (V2V) and vehicle-to-roadside (V2R) multi-hop communication to an on-road data collector $D$. Examples of data collectors are mentioned later in the reporting module. The data can be kept at $D$ for later retrieval by $P$, or $D$ can be supported with a limited Internet access to report the collected data periodically to $P$.

Generally speaking, two main approaches can be used to support the data delivery paradigm discussed above:

- Autonomous vehicular data delivery.

This approach depends solely on communication between vehicles for packet forwarding. Each packet-carrying vehicle is responsible for selecting its next-hop forwarding vehicle from its neighboring vehicles. Many data delivery/gathering schemes that follow this approach are proposed in the literature and they can be adopted under this component. For example, the delay-bounded vehicular data gathering (DB-VDG) scheme [14] supports the retrieval of vehicular information through directing requests to areas of interest and retrieving resolved replies over vehicular communication. To reduce the communication overhead and aggregate data from multiple sources, vehicles in DB-VDG decide on either forwarding the packets immediately or carrying them while moving based on a delay-bound.

- Roadside-assisted vehicular data delivery.

Vehicles in this approach get assistance from roadside nodes, deployed usually at intersections, for more efficient forwarding/data delivery. The roadside nodes can either be used for just forwarding assistance or they can be utilized as well for adding some processing intelligence/optimization such as data aggregation and filtering. Such an approach has been employed by some schemes proposed in the literature that can fit in our platform under this component. One of the recent schemes is the caching-assisted data delivery (CADD) scheme [3] that is specifically proposed to handle the delivery of both interest and data reply packets between SPs and participants in vehicular public sensing. To reduce the round-trip communication delay and the cost of accessing vehicular resources for each sensing request, CADD employs an on-
road caching concept to introduce cache hits of previously collected data (see Figure 4). Vehicles in CADD get both caching and forwarding assistance from light-weight road caching spots (RCSs) deployed at intersections on electric poles or traffic lights. The end-users/requesters communicate with the vehicular network through Internet-connected gateways deployed one at each main area.

### 3.3 The Sensing and Reporting Component

After getting an assigned sensing task through the communication component, it is the responsibility of the sensing and reporting component from selected vehicles to perform this task and report data back to the SP.

#### 3.3.1 The Sensing Module

After getting tasked, a vehicle starts sensing the phenomenon and generating data periodically as long as it is in the area of interest defined in the sensing request or until the required sensing period expires, whichever comes first.

After generating the required sensing data, the sensing module passes this data to the reporting module.

#### 3.3.2 The Reporting Module

The role of the reporting module is getting data out of the participating vehicle through the communication component to be delivered to the SP. A vehicle should report the collected data based on a reporting period, and one of the main functions of the reporting module is determining that period.

![Figure 4: The CADD scheme components and benefit of caching [3].](image)

A replica is found at $I_1$. The reporting period is controlled by two parameters; the data type (data criticality) and the communication paradigm, as follows.

- **Data type:**
  - Delay-critical data should be reported once sensed using the available broadband connectivity (reporting rate = sensing rate).
  - Delay-tolerant data has some flexibility in its reporting rate. Such data can be stored and aggregated to be reported at a later time depending on the connectivity paradigm (reporting rate < sensing rate).

- **Connectivity paradigm:**
  - If the collected data has to be reported using an onboard broadband connectivity and this data is delay-critical, it should be reported once sensed as mentioned above. If the data is delay-tolerant and has to be sent also using broadband connectivity, aggregated data can be reported at the end of the sensing task.
  - If data has to be delivered to an on-road data collector (e.g., a wireless roadside unit or sink at an intersection), data can be stored and carried until the participating vehicles come in contact with the dedicated data collector. If the data-carrying vehicle will not encounter the data collector on its trajectory, the vehicle can utilize the data-relaying capabilities of other vehicles on the road to send the data to the collector using multi-hop V2V communication.
  - In cases where data is to be delivered through opportunistic connectivity, aggregated data should be stored and carried until the carrying vehicle reaches a connectivity opportunity.

### 3.4 The Data Analytics Component

The role of the data analytics component is to receive reported data and apply necessary data aggregation, filtering, and validation functionalities. In addition, the component builds a QoI index for each participant through assessing the quality of retrieved data and the behavior of its corresponding participant. Such QoI indices can be used as inputs for the recruitment component to guide the selection process as discussed earlier.

#### 3.4.1 Data Aggregation and Filtering

Many data aggregation and filtering techniques have been proposed in the literature in the past decade for the use by the data-centric paradigms such as wireless sensor networks. State of the art mechanisms handling such functionalities can be borrowed and adopted in our platform [6][15].

#### 3.4.2 Data Validation

Data validation aims at deciding on the correctness of the reported data and detecting the untrusted and malicious participants. This functionality is to ensure the reliability of the data to be published to the users and the overall trustworthiness of the sensing-based service to be offered.
A research direction towards ensuring data validity is to assess the trustworthiness of candidate participants and confine the interactions with only the trusted participants. Many schemes are proposed in the literature under this topic [18].

Another direction is to come up with validation/verification techniques to assess the trustworthiness of the received data. Since the location where the reported data has been generated is a very crucial attribute when it comes to data analytics/visualization, the validation of a reporting vehicle’s location is an essential functionality in the proposed platform. Many schemes have been proposed in the literature for location verification and validation in vehicular networks and such schemes can be utilized by the platform to handle this functionality. For example, the vehicular misbehavior detection framework proposed in [5] includes a location verification scheme utilizing the periodic beacons exchanged between vehicles. The scheme depends on the integration of road and map information along with the previously received location updates for estimating a vehicle’s trajectory. The vehicle’s announced location is compared to an expected plausible geographic area on the estimated trajectory to decide on the consistency of the announced location.

### 3.4.3 Quality of Information Assessment

As one of its main functions, this component assesses the QoI of the contributions/readings provided by participants and maintains a QoI index of each contributor/participant accordingly. This function encompasses two main tasks: a) analyzing the received data and computing a QoI rating for each contribution, and b) maintaining the ratings of the contributions and using them to compute an overall QoI index for each participant. To carry out these tasks, our platform can encompass the following two modules under this component: 1) a watchdog module, and 2) a QoI aggregation module. In addition, a database would be used to maintain the history of a participant through storing the ratings of the latest $n$ corresponding contributions and the most recent QoI index. In the following, we discuss these two modules highlighting some approaches that can be adopted for their use.

#### a) The Watchdog Module

The raw data received by the platform is passed directly to the watchdog module. This module is responsible for assessing the QoI of each contribution and computing a corresponding rating in the range $[0,1]$. To handle this task, two different approaches can be used based on the availability of redundant data:

- **Redundancy-dependent rating**: This approach depends on correlated readings from other participants reporting the same event to apply a consensus-based technique such as outlier detection [7][17]. The input to this technique is the set of correlated readings from different contributors and it measures the distance of each reading value to a common value (e.g., average of the correlated readings). The computed distance of each reading is translated to its rating after normalizing it to the $[0,1]$ range.

- **Redundancy-independent rating**: In cases where there is no other correlated contributions reported, QoI assessment can depend on metrics that consider features associated with the reported data and the contributor’s behavior such as the completeness of the data and the contributor’s commitment.

The assessment values of these metrics are computed and all combined into a single value representing the overall rating of the contribution in the $[0,1]$ range.

#### b) The QoI Aggregation Module

This module is responsible for computing a QoI index for each participant and maintaining the QoI history. It receives as inputs the ratings of participants’ contributions from the watchdog module upon encountering new contributions. It utilizes the ratings of the latest $n$ corresponding contributions stored in the database along with the received rating of the most recent contribution for computing an aggregated QoI index. Such ratings can be aggregated using a simple average function or more advanced weighted mechanisms.

Figure 5 summarizes the QoI assessment modules and the interaction among them.

### 4. SUMMARY AND FUTURE WORK

Motivated by the rising interest in public sensing and the abundant resources of smart vehicles, we proposed the Vehicular Public Sensing (VPS) platform that utilizes such vehicular resources for the purpose of provisioning vehicle-based public sensing services. The platform consists of four main components that handle the sensing and data collection process, starting from participant selection to having the collected data ready for use. Our platform handles participant selection through an underlying recruitment component. In addition, the proposed platform manages data delivery through underlying communication, sensing, and reporting components. The VPS platform also takes into consideration the final analytical stage needed by the public sensing process through a data analytics component. We discussed how the different components can be adapted to handle different practical setups (e.g., environmental densities and available connectivity), and how they interact with one another.

In our future work, we will explore the adaptation of the proposed platform to support dynamic reuse of the vehicular resources and incorporation with neighboring transient resources offered by other network paradigms. The applicability of the resilient paradigm proposed in [13] to be adopted into our platform will be investigated.

### 5. REFERENCES


