# Towards a Practical Deployment of Privacy-preserving Crowd-sensing Tasks

Nicolas Haderer
University Lille 1, INRIA
nicolas.haderer@inria.fr

Vincent Primault
Université de Lyon, CNRS
INSA-Lyon
vincent.primault@liris.cnrs.fr

Patrice Raveneau
Université de Lyon, INRIA
patrice.raveneau@inria.fr

Christophe Ribeiro
University Lille 1, INRIA
christophe.ribeiro@inria.fr

Romain Rouvoy
University Lille 1, INRIA
romain.rouvoy@univ-lille1.fr

Sonia Ben Mokhtar
Université de Lyon, CNRS
INSA-Lyon
sonia.ben-mokhtar@liris.cnrs.fr

## ABSTRACT

Recent generations of mobile phones, embedding a wide variety of sensors, have fostered the development of open sensing applications, such as network quality or weather forecast applications. In this paper, we present a novel privacy-preserving crowdsourcing platform relying on two components: APISENSE and PRIVAPI. APISENSE is a distributed middleware platform that leverages the dynamic deployment of crowdsourcing tasks across a population of mobile phones. PRIVAPI is a middleware handling privacy-preserving publication of mobility data.

## Categories and Subject Descriptors

C.2.4 [**Distributed Systems**]: Distributed Applications; K.4.1 [**Computers and society**]: Public policy issues—*privacy*

## General Terms

Algorithms, Security

## Keywords

crowd-sensing, privacy, mobile crowdsourcing, middleware

## 1. INTRODUCTION

The increasing popularity of mobile devices has raised the opportunity to easily connect to a crowd of mobile users and to engage them in the completion of a variety of cyber-physical tasks. In particular, mobile crowdsourcing (or crowd-sensing) refers to the capability of lifting a large and diffuse group of users to delegate the task of retrieving trustable data from the field. The pervasive nature of mobile crowd-sourcing allows easy gathering of mobility data, which consists in all timestamped locations where a user has been during the experiment. This obviously raises some serious concerns about privacy, not taken in account by existing solutions like the Funf Open Sensing Framework[1]. As it has been demonstrated by several papers like [2] it is possible to use mobility data about individuals to infer a lot of knowledge like where they work, where they live and ultimately who they are. Mobility data can also be used to learn sensitive information like religious or political preferences.

APISENSE is a mobile crowdsourcing platform that makes easier the deployment of crowd-sensing experiment by taking care of the critical challenges in this area [1]. The novelty is therefore to provide a *Software-as-a-Service* platform where crowd-sensing experiments are described as scripts, which will be offloaded onto mobile devices in order to collect data on the field.

PRIVAPI is a privacy-preserving middleware that can be easily integrated on-top of APISENSE. Its goal is to pre-process gathered mobility data before it is released. Thanks to its knowledge on the whole dataset it can use an optimal anonymization strategy on mobility data while still offering a satisfactory level of utility. At the moment one novel strategy satisfying these two criteria is implemented in PRIVAPI.

In the remainder of the paper, we provide an overview of the APISENSE platform in Section 2) and the PRIVAPI middleware in Section 3), before concluding in Section 4).


*Middleware'14:Demos and Poster* Dec 08 - December 12 2014, Bordeaux, France
ACM 978-1-4503-3220-0/14/12.
http://dx.doi.org/10.1145/2678508.2678530

## 2. THE APISENSE PLATFORM

APISENSE builds on a distributed architecture. In its center sits the Hive service, that is responsible for managing the community of mobile users and publishing crowd-sensing tasks. These crowd-sensing tasks are uploaded on the Hive from Honeycomb endpoints, which are deployed and used by people interested in collecting specific datasets. The Honeycomb is therefore used to describe the crowd-sensing tasks as scripts (based on an extension of JavaScript) that are seamlessly offloaded onto mobile devices by the Hive. Once triggered by the mobile device, these scripts will automatically produce a dataset, which will be sent back to the Honeycomb to be processed and stored depending on experiments.

One of the benefits of building a common platform like APISENSE lies in the federation of communities of mobile users who are willing to contribute to mobile crowdsourcing

[1] `http://www.funf.org`

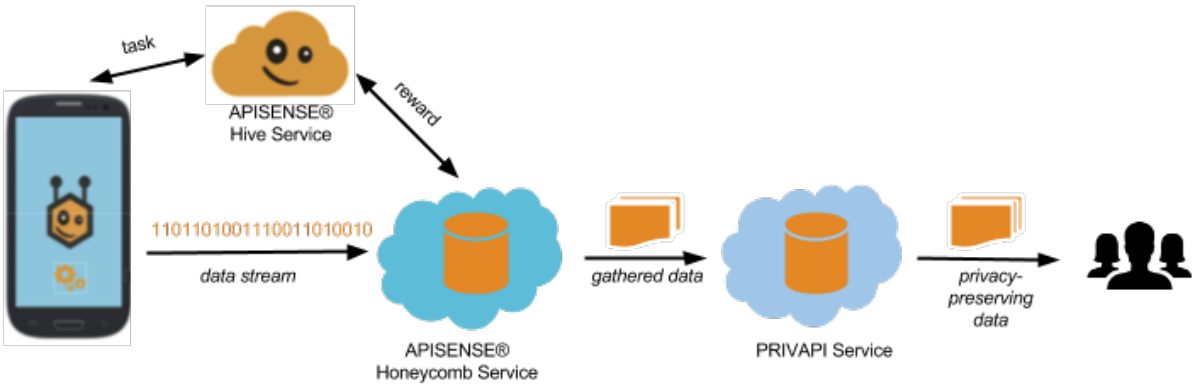


**Figure 1: Architecture of the data collection platform.**

experiments in order to ease their recruitment and therefore focus on the nature of the datasets to be collected. The APISENSE platform also implements the concept of virtual sensors as a mean to abstract the individual devices and therefore offer a set of additional services that self-organize a group of mobile devices to orchestrate the retrieval of datasets according to different strategies (e.g., round robin, energy-aware). The APISENSE platform supports the implementation of different incentive strategies, including user feedback, user ranking, user rewarding and win-win services. The selection of incentive strategies carefully depends on the nature of the crowdsourcing experiments.

Concerning privacy, a first layer is deployed on the mobile device and implements several algorithms to filter out and blur sensitive information (e.g., address book, location) depending on user preferences. The user keeps the control of her mobile phone to select the sensors to be shared, as well as when and where these sensors can be used by the platform.

## 3. THE PRIVAPI MIDDLEWARE

PRIVAPI is a generic middleware that can be integrated with any crowd-sensing platform. It focuses on protecting mobility data, which comes with specific threats against location privacy. We have previously studied in [3] the impact of points of interest, which are places where a user spends significant amounts of time like his home, his office, a cinema, etc. These places are highly sensitive because they convey rich semantic information. Moreover, they allow to almost uniquely identify individuals by studying their mobility patterns. We have shown that even a recent state-of-the art protection mechanism still allows to re-identify at least 60 % of the points of interest from a real-life dataset.

PRIVAPI leverages the global knowledge of the whole system to apply an optimal anonymization strategy and produce a privacy-preserving mobility dataset. Because published data will be used by researchers or industrials, it must guarantee both privacy and utility. A minimum level of privacy must be enforced, as parametrized by the users and/or the platform owner. In the same time, our middleware wants to be utility-driven. We believe there is not one unique anonymization strategy that always performs well but many from which we can choose the one that fits the best to the usage that will be done with the anonymized dataset.

Because of the threats related to points of interest, we have implemented an original anonymization strategy that focuses on hiding these places. To achieve that we use an algorithm that smoothes speed along a trajectory (typically one day of data) to guarantee that speed is constant. This still allows to analyze the trajectory of a user but prevents to find out places where he stopped during his day. First experiments show that under such a protection utility of our anonymized dataset remains high for useful data mining tasks such as finding out crowded places or predicting traffic.

## 4. CONCLUSION

We introduced a privacy-preserving crowd-sensing platform composed of two complementary components. APISENSE[2] as an open platform that can be used to quickly deploy a wide diversity of crowdsourcing tasks. PRIVAPI is a middleware protecting mobility data by producing privacy-preserving yet useful anonymized datasets.

## 5. ACKNOWLEDGMENTS

This work was supported by the LABEX IMU (ANR-10-LABX-0088) of Université de Lyon, within the program "Investissements d'Avenir" (ANR-11-IDEX-0007) operated by the French National Research Agency (ANR).

## 6. REFERENCES

[1] N. Haderer, R. Rouvoy, and L. Seinturier. Dynamic Deployment of Sensing Experiments in the Wild Using Smartphones. In *DAIS 2014*, pages 43–56.

[2] J. Krumm. Inference Attacks on Location Tracks. In *Pervasive 2007*, pages 127–143.

[3] V. Primault, S. Ben Mokhtar, C. Lauradoux and L. Brunie. Differentially Private Location Privacy in Practice. In *MOST 2014*.

[2]APISENSE online: `http://www.apisense.com`