In-network Collaborative Mobile Crowdsensing

Yifan Du
Supervised by Valérie Issarny & Françoise Sailhan
Inria Paris, France
yifan.du@inria.fr

***Abstract*—Our work aims to make opportunistic crowdsensing a reliable means of detecting urban phenomena, as a component of smart city development. We believe that the optimal method for achieving this is by enforcing the cost-effective collection of high quality data. We then investigate a supporting middleware solution that reduces both the network traffic and computation at the cloud. To this end, our research focuses on defining a set of protocols that together implement "*context-aware in-network collaborative mobile crowdsensing*" by combining: (i) The inference of the crowdsensors' physical context so as to characterize the gathered data; (ii) The context-aware grouping of crowdsensors to share the workload and filter out low quality data; and (iii) Data aggregation at the edge to enhance the knowledge transferred to the cloud.**

***Index Terms*—Crowdsensing, Context Inference, D2D Collaboration, Edge Computing, Environment Sensing**

I. Introduction

Mobile crowdsensing [1] is a sensing paradigm that empowers ordinary citizens to contribute with data sensed or generated from their sensor-enhanced mobile devices (e.g., mobile phones, wearable devices, tablets). Crowdsensing supports two sensing strategies [2]: (1) *Participatory crowdsensing* that requires the proactive involvement of individuals who consciously contribute with sensing data; (2) *Opportunistic crowdsensing* that collects data in the background autonomously and does not necessitate any explicit action from the user. In that framework, opportunistic mobile crowdsensing appears as a scalable and cost-effective alternative to the deployment of static wireless sensor networks for the dense coverage of large areas, and especially urban areas. However, opportunistic crowdsensing faces two major challenges: (i) the low quality data due to the relatively low accuracy of mobile sensors and the background sensing, and (ii) high resource -hence financial- cost due to the reliance on the cloud for mass raw data collection and processing.

Our work addresses the above challenges in the context of opportunistic urban environmental monitoring where the mobile crowdsensing application contributes measurements related to the physical environment (e.g., ambient temperature, air pressure, ambient humidity, ambient light, sound level, magnetic field, location) using the embedded/connected sensors. Each new measurement necessarily comes with both spatial (e.g., location coordinates) and temporal (e.g., time stamps) metadata. The crowdsensors are further connected to the internet via cellular network or Wi-Fi so that they can send sensing data to the –often cloud-based– server.

II. Research Issues & Objective

Unlike professional wireless sensor networks where experts regularly calibrate, heterogeneous crowdsensing devices often introduce high bias in sensor measurements. The bias is not only due to the low accuracy of the contributing sensors but also the diversity of the sensing contexts. For example, the aggregation of crowdsensed noise measurements must distinguish the contributions of the devices that are out-pocket (i.e., free of friction) from the others; similarly, the processing of crowdsensed temperature measurements must distinguish the in-door contributions from the out-door ones (at same location). In a nutshell, combining uncorrelated crowdsensed data introduces significant errors in the resulting knowledge. Another critical issue arising with opportunistic crowdsensing relates to its cost for both the individual crowdsensors (especially in terms of battery consumption and network traffic fees) and the infrastructure (regarding the uploading to –and further processing at– the cloud of massive raw data). Moreover, such a cost includes the one associated with the useless gathering of low quality data. We further note that the device positioning using GPS and data uploading over cellular networks are the two main sources of power consumption for a crowdsensing user, while the one associated with the operation of the embedded environmental sensors is relatively negligible.

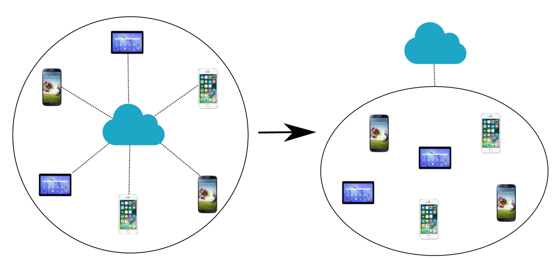

Fig. 1. From cloud-centric to in-network collaborative crowdsensing

The above leads us to argue that "*in-network collaboration among mobile crowdsensors*", which leverages device-to-device communication, may significantly contribute to enable cost-effective, high-quality crowdsensing by moving part of the filtering and (pre-)processing of the sensing data closer to the end user (Figure 1). For the purpose of increasing data quality while decreasing the crowdsensing cost –thereby contributing to enhancing the overall efficiency of crowdsensing–, our approach promotes *context-aware in-network collaboration*.

III. Published Research Work

Our approach to "*context-aware in-network collaborative mobile crowdsensing*" subdivides into three complementary steps: (1) Crowdsensor-specific context inference that accounts for the device attributes and user behavior; (2) Context-aware grouping so that the group members share position/internet access and perform selective sensing, to avoid redundant tasks and thereby decrease the overall resource consumption; (3) In-network data processing that reduces the data uploading and processing at the cloud. We have so far developed the solution -from design to implementation and evaluation- to the first two steps.

Personalized context inference: While the meaning of "*context*" is broad, our work focuses on the inference of the physical sensing context –beyond the geographical position in the Euclidean space– when gathering observations about the physical environment, that is, whether the smartphone is in-/out-pocket, in-/out-door and upper-/under-ground. Considering the sensors embedded in today's smartphones, we select in [3] the most relevant set of features to characterize the specific context information, according to the features' high information gain ratio. We also consider the availability of features depending on the device and the preference of the user who decides which embedded components are switched on/off. We then introduce an online learning approach to support a local inference of the sensing context that can evolve over time according to the environment in which it takes place and the available features. The challenge is to devise a classifier that accounts for the diversity in: the characteristics of the contributing devices, the behavior of the contributing users, and even the usage scenarios. For this purpose, we personalise the classifier so as to overcome the disparity of the classification performance. While the personalized inference of the sensing context is running on the device, feedback is requested to the end user so as to assess the correctness of the inference result and update the current learning model accordingly. Our approach detailed in [4] specifically features a hierarchical algorithm for inference that limits the number of opportunistic feedback requested to the user, while increasing the accuracy of the context inference per user. The inferred context information is a parameter of interest for the *group-based crowdsensing* framework.

Group-based crowdsensing: Traditional crowdsensing requires that each individual device provides the spatio-temporal sensing data to the cloud server, which processes it. Device-to-device wireless networks, such as Bluetooth and Wi-Fi Direct, bring the ability to collaborate using short range communication. To support cooperation among crowdsensors, we introduce a context-aware and collaborative grouping strategy in [5], in which a group is maintained in an autonomous and distributed way for the purpose of monitoring a physical phenomenon that is of interest. The user activity (e.g., static *vs* mobile) and the context (e.g., in-door *vs* out-door) are used to create homogeneous groups of crowdsensors that share the same sensing environment. Indeed, grouping together the crowdsensors that are co-located and that behave alike facilitates the collection of measurements that relate to a common physical phenomena and that henceforth can be aggregated locally. The grouping strategy designates a *group owner* that manages the group and handles advanced capabilities so as to enhance the resource-efficiency and the data quality of the crowdsensing system, both locally and globally. In particular, the group owner registers the services offered by the crowdsensors that belong to the group and assigns the crowdsensing tasks as needed. In order to optimize the task assignment and to ensure that only relevant group members perform sensing, the crowdsensor contexts such as in-/out-pocket, the sensor accuracy and remaining power are used to estimate the utility of the crowdsensor, which defines to which extent a member can provide the requested service. Overall, our group-based strategy introduces a novel context-aware clustering strategy and task allocation scheme that enhance the local and global crowdsensing efficiency.

IV. On-going Research

The final step of our approach is focused on the collaborative processing of the crowdsensing data to reduce the resource-cost at the cloud.

In-network processing: Our on-going work is related to the in-network collaboration for processing the crowdsensor measurements. Indeed, large scale crowdsensing usually involves significant communication, computation and financial costs due to the dependence on the cloud for the post-processing of huge amount of raw sensing data. As an alternative, we investigate a distributed data processing approach running on the crowdsensors so as to pre-process the sensing data at the very edge, and thereby to both reduce the resource consumption and enhance the relevance of the knowledge collected, at the cloud. In particular, the work in [6] introduces a multi-hop, multi-party calibration approach that suits well our collaborative crowdsensing application and allows enhancing the quality of the measurements. We are currently investigating the integration of such a solution together with other data aggregation algorithms in our collaboration framework.

References

[1] B. Guo, Z. Wang, Z. Yu *et al.*, "Mobile crowd sensing and computing: The review of an emerging human-powered sensing paradigm," *ACM Computing Surveys*, vol. 48, no. 1, 2015.

[2] J. Phuttharak and S. W. Loke, "A review of mobile crowdsourcing architectures and challenges: Toward crowd-empowered internet-of-things," *IEEE Access*, vol. 7, 2018.

[3] Y. Du, V. Issarny, and F. Sailhan, "When the power of the crowd meets the intelligence of the middleware: The mobile phone sensing case," *ACM SIGOPS Operating Systems Review*, vol. 53, no. 1, 2019.

[4] ——, "User-centric context inference for mobile crowdsensing," in *ACM International Conference on Internet of Things Design and Implementation*, 2019.

[5] Y. Du, F. Sailhan, and V. Issarny, "Let opportunistic crowdsensors work together for resource-efficient, quality-aware observations," in *IEEE International Conference on Pervasive Computing and Communications*, 2020.

[6] F. Sailhan, V. Issarny, and O. Tavares-Nascimiento, "Opportunistic multi-party calibration for robust participatory sensing," in *IEEE International Conference on Mobile Ad Hoc and Sensor Systems*, 2017.