Participatory Sensing: Crowdsourcing Data from Mobile SmartPhones in Urban Spaces

Salil S. Kanhere
School of Computer Science and Engineering
The University of New South Wales, Sydney, NSW 2052, Australia
Email: salilk@cse.unsw.edu.au

***Abstract*—The recent wave of sensor-rich, Internet-enabled, smart mobile devices such as the Apple iPhone has opened the door for a novel paradigm for monitoring the urban landscape known as participatory sensing. Using this paradigm, ordinary citizens can collect multi-modal data streams from the surrounding environment using their mobile devices and share the same using existing communication infrastructure (e.g., 3G service or WiFi access points). The data contributed from multiple participants can be combined to build a spatiotemporal view of the phenomenon of interest and also to extract important community statistics. Given the ubiquity of mobile phones and the high density of people in metropolitan areas, participatory sensing can achieve an unprecedented level of coverage in both space and time for observing events of interest in urban spaces. Several exciting participatory sensing applications have emerged in recent years. For example, GPS traces uploaded by drivers and passengers can be used to generate realtime traffic statistics. Similarly, street-level audio samples collected by pedestrians can be aggregated to create a citywide noise map. In this advanced seminar, we will provide a comprehensive overview of this new and exciting paradigm and outline the major research challenges.**

I. Introduction

In recent times, mobile phones have been riding the wave of Moore's Law with rapid improvements in processing power, embedded sensors, storage capacities and network data rates. The mobile phones of today have evolved from merely being phones to full-fledged computing, sensing and communication devices. It is thus hardly surprising that over 5 billion people globally have access to mobile phones. These advances in mobile phone technology coupled with their ubiquity have paved the wave for an exciting new paradigm for accomplishing large-scale sensing, known in literature as *participatory sensing* [1], [2] The key idea behind participatory sensing is to empower ordinary citizens to collect and share sensed data from their surrounding environments using their mobile phones.

Mobile phones, though not built specifically for sensing, can in fact readily function as sophisticated sensors. The camera on mobile phones can be used as video and image sensors. The microphone on the mobile phone, when it is not used for voice conversations, can double up as an acoustic sensor. The embedded GPS receivers on the phone can provide location information. Other embedded sensors such as gyroscopes, accelerometers and proximity sensors can collectively be used to estimate useful contextual information (e.g., is the user walking or traveling on a bicycle). Further, additional sensors can be easily interfaced with the phone via Bluetooth or wired connections, e.g., air pollution or biometric sensors.

Participatory sensing offers a number of advantages over traditional sensor networks which entails deploying a large number of static wireless sensor devices, particularly in urban areas. First, since participatory sensing leverages existing sensing (mobile phones) and communication (cellular or WiFi) infrastructure, the deployment costs are virtually zero. Second, the inherent mobility of the phone carriers provides unprecedented spatiotemporal coverage and also makes it possible to observe unpredictable events (which may be excluded by static deployments). Third, using mobile phones as sensors intrinsically affords economies of scale. Fourth, the widespread availability of software development tools for mobile phone platforms and established distribution channels in the form of App stores makes application development and deployment relatively easy. Finally, by including people in the sensing loop, it is now possible to design applications that can dramatically improve the day-to-day lives of individuals and communities.

A typical participatory sensing application operates in a centralized fashion, i.e., the sensor data collected by the phones of volunteers are reported (using wireless data communications) to a central server for processing. The sensing tasks on the phones can be triggered manually, automatically or based on the current context. On the server, the data are analyzed and made available in various forms, such as graphical representations or maps showing the sensing results at individual and/or community scale. Simultaneously, the results may be displayed locally on the carriers' mobile phones or accessed by the larger public through web-portals depending on the application needs.

In this advanced seminar we will provide a comprehensive overview of this exciting and new paradigm. Section II will provide an overview of innovative participatory sensing applications that have been proposed in literature. Section III will provide a detailed discussion on the key research challenges posed by participatory sensing and focus on novel approaches to deal with the same. Finally, Section IV will conclude the article.

DOI 10.1109/MDM.2011.16

II. Innovative Participatory Sensing Applications

The emergence of the participatory sensing paradigm has resulted in a broad range of novel sensing applications, which can be categorized as either *people-centric* or *environment-centric* sensing. People-centric applications mainly focus on documenting activities (e.g., sport experiences) and understanding the behavior (e.g., eating disorders) of individuals. In contrast, environment-centric sensing applications collect environmental parameters (e.g. air quality or noise pollution). In this section, we present an overview of a few representative applications within each category.

A. People-Centric Sensing Applications

People-centric sensing uses the sensor devices integrated in mobile phones to collect data about the user. In the following we present oe example

1) Personal Health Monitoring: In personal health monitoring, mobile phones are used to monitor the physiological state and health of patients/participants using embedded or external sensors (e.g., wearable accelerometers, or air pollution sensors). For example, DietSense [3] assists participants who want to lose weight by documenting their dietary choices through images and sound samples. The mobile phones are worn around the neck of the participants and automatically take images of the dishes in front of the users. The images document the participants' food selection and allow for an estimation of the food weight and waste on the plates. Moreover, the mobile phones capture the participants' context during their meals by recording time of day, location, and sound samples to infer potential relationships between the participants' behavior and their context (e.g., having lunch in a restaurant or eating chips late at night on the sofa). All captured data are uploaded to a personal repository, where the participants can review them to select/discard the information to be shared with their doctors and nutritionists.

2) Calculating Environmental Impact: PEIR (Personal Environmental Impact Report) is a system that allows users to use their mobile phone to determine their exposure to environmental pollutants [4]. A sensing module installed on the phone determines the current location of the user as well as information about the currently used mode of transportation (e.g., bus vs car), and transfers this information to a central server. In return, the server provides the users with information about the environmental impact of their traveling in terms of carbon and particle emissions. Additionally, the server estimates the participants' exposure to particle emissions generated by other vehicles and fast food restaurants while commuting. The latter may be useful for health conscious users who may want to avoid the temptation of stopping by such restaurants. The mode of transport is inferred using accelerometer readings, while the route travelled is extracted from the captured location traces. Additional input parameters and models are considered for determining the environmental factors, such as weather conditions collected by weather stations, road traffic flow models, and vehicle emission models.

3) Monitoring and Documenting Sport Experiences: BikeNet [5] presents a system for monitoring bicycling experiences of the participants. BikeNet draws a fine-grained portrait of the cyclist by measuring his current location, speed, burnt calories, and galvanic skin response. Multiple peripheral sensors are used to obtain this information: Microphone, magnetometer, pedal speed sensor, inclinometer, lateral tilt, GSR stress monitor, speedometer/odometer, and a sensor for CO_2 concentration. The peripheral sensors form a body area network and interact with the mobile phone over a wireless connection. The captured data can be reviewed by the cyclists themselves, but can also be merged with other participants' data or combined with additional parameters, such as air quality and traffic properties, in order to construct complete maps for the cycling community.

4) Enhancing Social Media: A large pool of applications utilizes data captured by sensors to enrich the contents shared in social media, such as blogs, social networks, or virtual worlds. CenceMe [6] integrates virtual representations of the participants' current state and context in social networks and virtual worlds. Based on multimodal information (acceleration, audio samples, pictures, neighboring devices, and location) captured by the mobile phone, context information is inferred in various dimensions, including the user's mood, location, and habits, as well as information about the currently performed activity and the environment. The inferred information is then posted as status message in social networks or translated into the virtual representation of participants in virtual worlds.

5) Price Auditing: PetrolWatch [7] presents a system for automatic collection of fuel prices using embedded cameras in phones. The mobile phone is mounted on the passenger seat of a car and faces the road to automatically photograph fuel price boards (using GPS and GIS) when the vehicle approaches service stations. The pictures are then uploaded to a central entity, which is responsible for image processing and price extraction. The brand of the service station is first inferred from the capture location in order to reduce the image processing complexity, as price boards of different brands differ in colors and dimensions. Assisted by this information, computer vision algorithms extract the fuel prices, and uploads them to the database. Users can query the system to determine the cheapest fuel that is available in their area of interest.

B. Environment-Centric Sensing Applications

In environment-centric scenarios, the mobile phones capture information via their embedded sensors and additional peripheral sensors about the surroundings of the participants. In contrast to most people-centric sensing scenarios, the captured data are mainly exploited at a community scale, e.g., to monitor the evolution of environmental parameters like air quality, noise, road and traffic conditions in cities, or to detect socially interesting events.

1) Air Quality Monitoring: In Haze Watch [8], mobile phones were interfaced to external pollution sensors, in order to measure the concentration of carbon monoxide, ozone, sulphur dioxide, and nitrogen dioxide concentration in the air. In comparison to meteorological stations, the mobile phones may collect less accurate measurements. However, their inherent mobility allows observing unpredictable events (e.g., accidental pollution), which cannot be detected by static stations and provide large spatial coverage. The mobile phones can thus complement static high-fidelity data captured by traditional meteorological stations by providing finer-grained readings. The timestamped and geotagged measurements are then uploaded to a server to build maps, which aggregate the readings of all participants and are accessible by the public. Individual measurements may also be displayed on the participant's mobile phone.

2) Monitoring Noise and Ambiance: Microphones in mobile phones can be configured to measure the surrounding noise level and give insights about the nature of contextual events. In EarPhone [9], the noise levels are used to monitor noise pollution, which can e.g. affect human hearing and behavior, and build representative pollution maps accessible to either specialists to understand e.g. the relationships between noise exposition and behavioral problems for the general public.

3) Monitoring Road and Traffic Conditions: The mobile phones can be exploited to document road and traffic condition. In Nericell [10], the embedded accelerometer, microphone, and positioning system (GPS or GSM radio) are used to detect and localize traffic conditions and road conditions, e.g., potholes, bumps, or braking and honking (which are both implicit indicators of traffic congestion). The application integrates the provided information about the surface roughness of the roads, the surrounding noise, and the traffic conditions into traffic maps, which are available to the public.

III. Key Research Challenges

Since participatory sensing relies on existing networking and sensing infrastructure, the key challenges are associated with handing the data, i.e., data analysis and processing, ensuring data quality, protecting privacy, etc. We now present a short overview of the key research challenges posed by this new and exciting paradigm. We also discuss some recent work in overcoming these challenges.

A. Dealing with Incomplete Samples

Since a participatory sensing system relies on volunteers contributing noise pollution measurements, these measurements can only come from the place and time where the volunteers are present. Furthermore, volunteers may prioritize the use of their mobile devices for other tasks. Or they may choose to collect data only when the phone has sufficient energy. Consequently, samples collected from mobile phones are typically randomly distributed in space and time, and are incomplete. The challenges thus is to recover the original spatiotemporal profile of the phenomenon being monitored from random and incomplete samples obtained via crowdsourcing. In [9], the authors have developed an innovative approach to deal with this issue using the technique of compressive sensing, in the context of a noise monitoring application.

B. Inferring User Context and Activities

Inferring the surrounding context (e.g., is the user in a party or a quiet room) and activities (is the user walking, running, traveling by car, etc) undertaken by the phone carrier is of interest in various participatory sensing applications (see for example Section II-A4). This is usually achieved using data collected by one or more embedded sensors in the phone which include accelerometers, microphone, GPS, etc. At the high-level, this is essentially a machine learning problem, since we want to the phones to be embedded with the smarts to recognise human activity. The typical approach adopted is to make use of supervised learning [6]. This involves the following three steps. The first step is to collect properly labelled sensor data for the various categories of interest (i.e. training). The second step involves identifying important features from this training data, which can uniquely identify each category (i.e. fingerprints). The final step is to choose an appropriate classification algorithm (e.g., support vector machines, neural networks, etc), which can be used to achieve accurate classification.

C. Preserving User Privacy

Current participatory sensing applications are primarily focused on the collection of data on a large scale. Without any suitable protection mechanism however, the mobile phones are transformed into miniature spies, possibly revealing private information about their owners. Possible intrusions into a user's privacy include the recording of intimate discussions, taking photographs of private scenes, or tracing a user's path and monitoring the locations he has visited. Many users are aware of the possible consequences, and may therefore be reluctant to contribute to the sensing campaigns. Since participatory sensing exclusively depends on user-provided data, a high number of participants is required. The users' reluctance to contribute would diminish the impact and relevance of sensing campaigns deployed at large scale, as well as limiting the benefits to the users. To encounter the risk that a user's privacy might be compromised, mechanisms to preserve user privacy are mandatory. The authors in [11] present a comprehensive application independent architecture for anonymous tasking and reporting. The infrastructure enables applications to task a mobile device using a new tasking language, anonymously distribute tasks to mobile devices and collect anonymous yet verifiable reports from the devices.

D. Evaluating Trustworthiness of Data

The success of participatory sensing applications hinges on high level of participation from voluntary users. Unfortunately, the very openness which allows anyone to contribute data, also exposes the applications to erroneous and malicious contributions. For instance, users may inadvertently position their devices such that incorrect measurements are recorded, e.g., storing the phone in a bag while being tasked to acquire urban noise information. Malicious users may deliberately pollute sensor data for their own benefits, e.g., a leasing agent may intentionally contribute fabricated low noise readings to promote the properties in a particular suburb. Without confidence in the contributions uploaded by volunteers, the resulting summary statistics will be of little use to the user community. Thus, it is imperative that the application server can evaluate the trustworthiness of contributing devices so that corrupted and malicious contributions are identified. Recent work [12] proposes a novel reputation system that employs the Gompertz function for computing device reputation score as a reflection of the trustworthiness of the contributed data.

E. Conserving Energy

Note that, the primary usage of mobile phones should be reserved for the users's regular activities such as making calls, Internet access, etc. Users will only volunteer to contribute data if this process does not use up significant battery so as to prevent them from accessing their usual services. Even though, most users charge their phones on a daily basis, it is thus important that participatory sensing applications do not introduce significant energy costs for users. Energy is consumed in all aspects of participatory applications ranging from sensing, processing and data transmission. In particular, some sensors such as GPS consume significantly more energy that others. As such, it is important for participatory applications to make use of these sensors in a conservative manner. In [13], the authors present an adaptive scheme for obtaining phone location by switching between the accurate but energy-expensive GPS probing to energy-efficient but less accurate WiFi/cellular localization. Similar approaches can be employed for duty-cycling other energy-hungry sensors.

IV. Conclusion

In this seminar, we introduce a novel paradigm called participatory sensing for achieving large-scale urban sensing by crowdsourcing sensing data from the mobile phones of ordinary citizens. We provide an overview of some of the exciting applications that have been proposed by the research community using this new idea. These range from collecting and sharing people-centric data to monitoring environmental parameters. We presented an overview of the key research challenges posed in implementing real-world participatory sensing systems and briefly discussed some of the existing solutions.

References

[1] J. Burke, D. Estrin, M. Hansen, A. Parker, N. Ramanathan, S. Reddy, M. B. Srivastava, "Participatory Sensing", in Proceedings of the World Sensor Web Workshop, in conjunction with ACM SenSys'06, November 2006.

[2] A. Campbell, S. Eisenman, N. Lane, E. Miluzzo and R. Peterson, "People-centric Urban Sensing", in Proceedings of Second Annual International Wireless Internet Conference (WICON), pp. 2-5, August 2006.

[3] S. Reddy, A. Parker, J. Hyman, J. Burke, D. Estin and M. Hansen, "Image Browsing, Processing and Clustering for Participatory Sensing: Lessons from a DietSense Prototype" in Proceedings of the Workshop on Embedded Networked Sensors (EmNetS), Cork, Ireland, June 2007.

[4] M. Mun, S. Reddy, et al. "PEIR, the Personal Environmental Impact Report, as a Platform for Participatory Sensing Systems Research", in Proceedings of ACM MobiSys, Krakow, Poland, June 2009.

[5] S. Eisenman, E. Miluzzo, N. Lane, R. Peterson, G. Ahn and A. Campbell, "The Bikenet Mobile Sensing System for Cyclist Experience Mapping", in Proceedings of ACM SenSys, Sydney, Australia November 2007.

[6] E. Miluzzo, N. Lane, K. Fodor, R. Peterson, S. Eisenman, H. Lu, M. Musolesi, X. Zheng, A. Campbell, "Sensing Meets Mobile Social Networks: The Design, Implementation and Evaluation of the CenceMe Application", in Proceedings of ACM SenSys, Raleigh, NC, USA, November 2008.

[7] Y. Dong, S. S. Kanhere, C. T. Chou and N. Bulusu, "Automatic Collection of Fuel Prices from a Network of Mobile Cameras", in Proceedings of IEEE DCOSS 2008, Santorini, Greece, June 2008.

[8] J. Carrapetta, N. Youdale, A. Chow and V. Sivaraman, "Haze Watch Project", Online: http://www.pollution.ee.unsw.edu.au.

[9] R. Rana, C.T. Chou, S. Kanhere, N. Bulusu and W. Hu, "Ear-Phone: An End-to-End Participatory Urban Noise Mapping System", in Proceedings of ACM/IEEE IPSN, Stockholm, Sweden, April 2010.

[10] P. Mohan, V. Padmanabhan, R. Ramjee, "Nericell: Rich Monitoring of Road and Traffic Conditions using Mobile Smartphones", in Proceedings of ACM SenSys, Raleigh, NC, USA, November 2008.

[11] A. Kapadia, N. Triandopoulos, C. Cornelius, D. Peebles and D. Kotz, "AnonySense: Opportunistic and Privacy Preserving Context Collection", in, Proceedings of Sixth International Conference on Pervasive Computing (Pervasive), May 2007.

[12] K. Huang, S. S. Kanhere and W. Hu, "Are You Contributing Trustworthy Data? The Case for A Reputation Framework in Participatory Sensing", in Proceedings of ACM MSWiM, Bodrum, Turkey, October 2010.

[13] S. Gaonkar, J. Li, R.R. Choudhury,"Micro-Blog: Sharing and Querying Content Through Mobile Phones and Social Participation", in Proceedings of ACM MobiSys, Breckenridge, CO, USA, June 2008.