A

# A Survey of Mobile Crowdsensing Techniques: A Critical Component for The Internet of Things

Jinwei Liu, Clemson University
Haiying Shen, University of Virginia
Husnu S. Narman, Marshall University

Mobile crowdsensing serves as a critical building block for the emerging Internet of Things (IoT) applications. However, the sensing devices continuously generate a large amount of data, which consumes much resources (e.g., bandwidth, energy and storage), and may sacrifice the quality-of-service (QoS) of applications. Prior work has demonstrated that there is significant redundancy in the content of the sensed data. By judiciously reducing the redundant data, the data size and the load can be significantly reduced, thereby reducing resource cost, facilitating the timely delivery of unique, probably critical information and enhancing QoS. This paper presents a survey of existing works for the mobile crowdsensing strategies with emphasis on reducing the resource cost and achieving high QoS. We start by introducing the motivation for this survey, and present the necessary background of crowdsensing and IoT. We then present various mobile crowdsensing strategies and discuss their strengths and limitations. Finally, we discuss the future research directions of mobile crowdsensing for IoT. The survey addresses a broad range of techniques, methods, models, systems, and applications related to mobile crowdsensing and IoT. Our goal is not only to analyze and compare the strategies proposed in the prior works but also to discuss their applicability towards the IoT, and provide the guidance on the future research direction of mobile crowdsensing.

CCS Concepts: •**Computer systems organization** → **Embedded systems;** *Redundancy;* Robotics; •**Networks** → Network reliability;

General Terms: Algorithms, Theory, Security, Performance

Additional Key Words and Phrases: Mobile crowdsensing, redundancy elimination, cost-effectiveness; quality of service; Internet of things

**ACM Reference Format:**
Jinwei Liu, Haiying Shen and Husnu S. Narman. Mobile crowdsensing, redundancy elimination, cost-effectiveness; quality of service; Internet of things. *ACM Trans. Cyber-Phys. Syst.* V, N, Article A (January YYYY), 25 pages.
DOI: http://dx.doi.org/10.1145/0000000.0000000

## 1. INTRODUCTION

In recent years, an increasing number of sensing devices and wireless networks emerge in our living environments, creating the Internet of Things (IoT) integrating the cyber and physical objects [Zordan et al. 2014; Shen et al. 2015b; Zhu and Shasha 2002; Zhang et al. 2013; Tangwongsan et al. 2010; Willett et al. 2013; Hasan and Curry 2014; Kirak et al. 2013; Kumbhare et al. 2013; Lane et al. 2013; Li et al. 2014]. As exposed in [Atzori et al. 2010], IoT will have a high impact on potential users' behavior because it integrates five layer middleware architecture (i.e., applications, service composition, service management, object abstraction, and objects) and identification, sensing and communication technologies. Figure 1 shows the architecture of IoT (right) and the architecture of its five layer middleware (left). According to the *Top 10 predictions of 2014* from the Gartner, IoT will be the fast-growing, largest market potential

Author's addresses: J. Liu, Department of Electrical and Computer Engineering, Clemson University, Clemson, SC 29634 USA; H. Shen, Department of Computer Science, University of Virginia, Charlottesville, VA 22904 USA; H. S. Narman, Computer Science, Marshall University, Huntington, WV 25755 USA.

DOI: http://dx.doi.org/10.1145/0000000.0000000

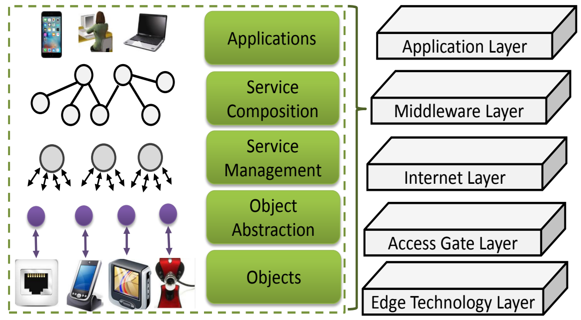



**Fig. 1:** Architecture of the Internet of Things (IoT) (right) and its five layer middleware (left).

and the most attractive emerging economy, thereby becoming the focus of attention in the field of networking.[1]

Mobile crowdsensing (MCS) refers to the wide variety of sensing models in which the individuals collectively share data and extract information to measure and map phenomena of common interest [Ganti et al. 2011; Peng et al. 2015]. MCS is emerging as a distributed paradigm, and it lies at the intersection between the IoT and the volunteer/crowd-based scheme. MCS creates a new way of perceiving the world to greatly extend the service of IoT and explore a new generation of intelligent networks, interconnecting things-things, things-people and people-people. Usually, the MCS applications are deployed on contributing nodes, such as mobile, personal devices that can be used to sense the physical environment and provide sensor data to mobile application server. Recently, various kinds of applications have been developed to realize the potential of MCS throughout daily life, such as environmental quality monitoring[2], noise pollution assessment [Maisonneuve et al. 2009; Rana et al. 2010], and traffic monitoring [Zhang et al. 2014].

MCS requires a large number of participants (individuals) to sense the surrounding environment using the sensing devices with built-in sensors. It is well-known that in such a large-scale system, the sensing devices continuously generate a huge amounts of data (raw sensor data), which consumes much resource (e.g., bandwidth, energy, etc.) [Liu et al. 2015]. However, the sensing devices have limited resources. Due to the limited resource, the quality of the data collected can be even sacrificed in the scenario of bandwidth constrained networks because of the heavy traffic load and high power consumptions [Hua et al. 2015; Dao et al. 2014]. Therefore, the resource limitation imposes a key challenge [Xu et al. 2015a; Dao et al. 2014; Hua et al. 2015; Gorlatova et al. 2014]. For example, images collected in the disaster area take an important role in disaster relief, the images collected may not be able to be uploaded in time due to the limited bandwidth, which can incur a huge cost. Therefore, resource limitation always hinders the necessary participation and wide-scale adaptation of the targeting applications [Xu et al. 2015a].

Although MCS is a new emerging paradigm, it has been applied in real applications [Chon et al. 2012; Mohan et al. 2008]. The application of MCS attracts great attention from both academic and business communities, which started investigating the commercial exploitation of MCS [Ra et al. 2012]. However, the adoption of MCS approach in business context requires the guarantee of the quality-of-service (QoS). Hence, QoS is one of the most important arising issues. Therefore, QoS-driven policies are needed to deal with the application non-functional issues to guarantee QoS.

[1]http://www.gartner.com.
[2]Creek watch: http://creekwatch.researchlabs.ibm.com/, 2010; Opensense: http://www.opensense.ethz.ch/trac/, 2010.

In this paper, we review the MCS techniques and challenges. Different aspects of the MCS are also reviewed by the researchers. In [Ganti et al. 2011], Ganti *et al.* introduced MCS, briefly overviewed of existing MCS applications with their unique characteristics, and discussed several research challenges with possible solutions. In [Vergara-Laurens et al. 2016], Vergara-Laurens *et al.* surveyed privacy issues and privacy-preserving mechanisms for crowdsensing systems. Zhang *et al.* reviewed the literature for the incentives that encourage users to participate in MCS applications under entertainment, service, and money categories [Zhang et al. 2016]. In [Wang et al. 2016], Wang *et al.* introduced sparse MCS, discussed sparse MCS challenges and developed a framework with potential solutions to the challenges. In [Wazir Zada Khan et al. 2013], Khan *et al.* comprehensively explained mobile sensing systems according to personal, social and public sensings. On the other hand, our focus is to discuss the resource limitation and QoS (e.g., data quality) issues and solutions in MCS. Apparently, a better understanding of resource management and QoS estimation in MCS can help us design a cost-effective crowdsensing system that can reduce the cost by fully utilizing the resource and improve the QoS for users, which manifests the significance of our survey.

Our objectives in reviewing the literature are threefold: 1) to learn what are the problems existing in MCS and how the proposed techniques have helped to develop solutions in the past; 2) to learn the strengths and limitations of different MCS techniques for smartly managing the resource to achieve low cost and good QoS, and how can we use those techniques to better solve similar problems in the future in different paradigms such as the IoT; 3) to provide guidance on the future research directions of MCS for IoT.

The remainder of this paper is organized as follows. Section 2 introduces the concepts of IoT and MCS. Section 3 describes the strategies of MCS. Section 4 describes the crowdsensing strategies for different application domains. Section 5 describes the challenges of MCS and the future research directions. Section 7 concludes this paper with remarks on our future work.

## 2. BACKGROUND

In this section, we introduce the main concepts of the IoT and MCS.

### 2.1. Internet of Things

During the past 10 years, the IoT has drawn great attention in both academia and business communities. The potential capabilities of IoT [Carnot Institutes 2011; Atzori et al. 2010] bring the interest of both academia and business communities. IoT is expected to create a world where all the objects around us are connected to the Internet, and eventually, it aims at creating ‘a better world for human beings’ [Dohr et al. 2010].

The term ‘Internet of Things’ was firstly coined by Kevin Ashton [Ashton 2009] in 1998. Later, the International Telecommunication Union (ITU) formally introduced the concept of IoT in 2005 [International Telecommunication Union 2005]. Currently, there is no standard definition for IoT. We use the definition of IoT from [Guillemin and Friess 2009] because it characterizes the broader version of IoT.

• Definition by the work [Guillemin and Friess 2009]: “The Internet of Things allows people and things to be connected Anytime, Anyplace, with Anything and Anyone, ideally using Any path/network and Any service.”

IoT is a new emerging paradigm, and it is a very broad version. The research into the IoT is still on the way. The potentialities of the IoT enable the development of a large number of applications in many domains. The application domains can be primarily divided into four categories [Atzori et al. 2010]: transportation and logistics domain, healthcare domain, smart environment (e.g., home, plant) domain, personal and social domain.

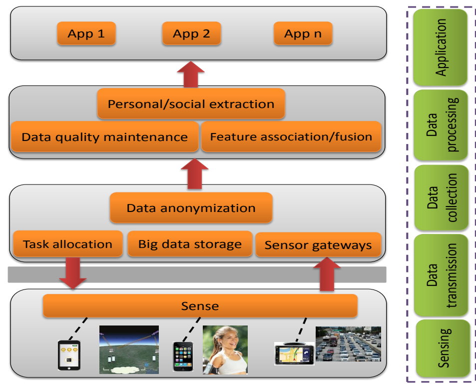



**Fig. 2:** Architecture of MCS system. Raw sensor data is collected via different mobile sensing devices (GPS, etc.) in sensing layer. In order to preserve privacy, the data will be sent to data collection layer and will be modified by using methods such as data anonymization.

### 2.2. Mobile Crowdsensing

Mobile Crowdsensing (MCS) uses the sensing devices equipped with sensors, to collect data (raw sensor data) from the surrounding environment. Therefore, the objectives of any MCS platform are to operate in a harmony with participants and individuals (e.g. load balancing), assign the tasks to the reliable participants (the participants which are expected to complete the assigned sensing tasks), effectively gather the required data from the participants, process and manage the data according to the purpose, and dynamically improve itself for next crowdsending events by self-learning mechanism [Bellavista et al. 2015]. MCS usually requires a large quantity of participants to sense the environment using the sensing devices. Based on the involvement of participants in sensing actions, MCS can be categorized as: participatory and opportunistic [Jami et al. 2015]. MCS has many applications. Based on the type of phenomenon being measured or mapped, the MCS applications can be divided into three categories: (a) Environmental application, (b) Infrastructure, and (c) Social application [Ganti et al. 2011]. Figure 2 shows the architecture of MCS system with five layers: sensing, data transmission, data collection, data processing, and application. In Figure 2, a certain number of challenges in MCS are indicated.

The basic MCS procedure includes three steps: data collection, data storage and data upload. Data collection is the first phase of MCS. The strategies for data collection usually can be divided into three categories [Lane et al. 2013]:

• All the data is manually collected by the user when controlling the sensing devices, such as smartphones with a specific application. This approach is attention-consuming and inefficient.

• Data collection is partially controlled by the user and by sampling, which is performed periodically. Sometimes, the data can be collected opportunistically, i.e., when the user opens some applications.

• Context-aware data sensing is triggered by predefined contexts, such as a particular location or time slot. This method releases the user from focusing on the crowdsensing tasks and makes it practical.

Based on the work [Pietschmann et al. 2008], context-aware data sensing can be almost accomplished by using the following two methods: push and pull. The definitions of push and pull are in below.

• Push: The physical or virtual sensing device (e.g., sensor) sends the data to the software component that is used to acquire data (e.g., sensor data) periodically or instantly. Periodical or instant pushing are able to help facilitate a publish and subscribe model.

• Pull: The software component which is in charge of acquiring data (e.g., sensor data) from sensing devices (e.g., sensors) makes a request like a query from the hardware of the sensing devices periodically or instantly to acquire data.

| **Degree of human participation** | **Example** |
|---|---|
| Completely controlled by user | Application monitoring the water levels, e.g., creek watch |
| Partially participation | Mobile energy efficient crowdsourcing, e.g., piggyback crowdsensing (PCS) |
| Control-free | Mobility information collection for pattern analysis on smartphone platform |
| **Platform mobility** | |
| Stationary | Roadside unit for traffic information collection |
| Mobile | Vehicular localization and vibration sensors, e.g., GPS, accelerometers |
| **Platform context** | |
| Social network | Location-based restaurant recommendation system, e.g., foursquare, whirl, etc. |
| Natural environment | Real-time forest soil temperature and moisture monitoring |
| Traffic management | Traffic congestion monitoring |

**Table I:** Summary of mobile crowdsensing applications.

**Deduplication.** Deduplication is a method for eliminating redundant data in the data collection phase to reduce resource cost and improve application QoS. Data deduplication is an essential part for reducing the cost of MCS implementation. As in most computation scenarios, data deduplication in crowdsensing performs filtering and compressing on the raw data (e.g., images) collected by sensing devices. The deduplication is conducted with the constraint that the significance of the data being kept. Deduplication of crowdsensing data makes full use of the limited sensor storage and reduces the bandwidth consumption caused by data transfer. In deduplication, data is usually partitioned into chunks, and unique chunks of data are identified and stored. Other chunks are compared to the stored chunks, and the redundant chunks are replaced with a small reference that points to the stored chunk. Only the unique chunks and the references are stored and uploaded. Thus the size of the uploaded data is reduced and the bandwidth consumption will be reduced.

As the size of the data to be processed increases rapidly, the deduplication technology is improving rapidly to meet the requirement of the industry and research all over the world. From the perspective of the phase at which the deduplication occurs, the data deduplication approaches can be categorized as real-time deduplication and post-process deduplication. On the other hand, as for the layer where the deduplication happens, there are local and server deduplication.

Real-time deduplication refers to hashing and compressing the data when acquiring the data. Duplicated data acquired by the sensing device will be detected based on the stored data chunk. If the new data is judged as duplicated, it will not be stored in the sensing device, neither be uploaded to the data center. The advantage of this strategy is to lower the required storage of local sensing devices. However, it shifts the computation burden from the data center to the terminals. For some commodity sensing devices like smart wearable gadgets, the real-time computation capacity is limited, so this strategy may not be practical. Hence, the post-process strategy can be adopted to relieve the real-time computation burden of local sensing devices. Specifically, the data acquired is stored first and then be processed for deduplication. The trade-off of this method is the relative high storage requirement and the storage overwriting risk when the storage margin is small.

Based on the criteria proposed in the work [Dimov 2014] and taking into account the stability degree of the crowdsensing platform mobility and context, we can categorize the crowdsensing applications in Table I.

Although crowdsensing has advantages of low cost, high flexibility and large data, the privacy of user requires to be protected when the crowdsensing process involves the human participation tightly, especially in the social network application. The users who contribute to the data collection are vulnerable to intended privacy attack [Krontiris and Dimitriou 2013]. Generally, the protection can be arranged at user end or performed by the cloud agent. Specifically, common strategies that help the users avoid privacy leakage are data anonymization, encryption and degradation.

## 3. EXISTING MOBILE CROWDSENSING STRATEGIES

In this section, we describe different MCS strategies aiming to reduce the resource consumption in order to reduce the resource cost and improve QoS.

Previous works demonstrate that significant redundancy exists in the content of the data [Dao et al. 2014; Aggarwal et al. 2011]. In many cases, sensors are likely to collect similar kinds of data from related sensors [Aggarwal et al. 2011]. Thus, it is important and necessary to eliminate the redundant data, which on the one hand can reduce the resource consumption and thus reduce the cost (e.g., bandwidth cost, energy cost, etc.), and on the other hand can improve the QoS of timely information delivery by reducing the traffic load. One of the key challenges here however, is detecting 'what data is similar'. Another key challenge is how to eliminate the similar data while ensuring high QoS (e.g., without compromising the quality of the data, timely delivery of valuable data). To handle the problem caused by limited available resources, many methods have been proposed. Below, we present a review of previously proposed strategies.

### 3.1. Different Mobile Crowdsensing Strategies for Reducing Resource Cost

Aggarwal *et al.* [Aggarwal et al. 2011] discussed real-time algorithms for reducing the volume of the data collected in sensor networks by determining the functional dependencies between sensor streams efficiently in real time, and actively collecting the data only from a minimal set of sensors. Hua *et al.* [Hua et al. 2015] presented a near-real-time and cost-effective solution under cloud assisted disaster environment. SmartEye [Hua et al. 2015] leverages two main methods, semantic hashing and space-efficient filters to aggregate the flows with similar features and provide communication services for the aggregated flow.

In bandwidth constrained network, Dao *et al.* [Dao et al. 2014] introduced a method focusing on recognizing the similar contents in images and videos, by leveraging the metadata uploaded first to distinguish the similarity of the data. According to their experimental results on a testbed and the simulation results using NS3, the rate of successful similarity detection is up to 70%. A number of researchers also dealt with the data redundancy reduction by detecting the similarity among the data, such as images or videos. For example, Weinsberg *et al.* [Weinsberg et al. 2012] proposed a framework called CARE, which eliminates the redundancy of the image for transferring data with constrained bandwidth while maintaining the quality of the service. In comparison with the former method in [Dao et al. 2014], CARE assumes that the infrastructure is unavailable, which is reasonable when the disaster happens, and makes use of peer-to-peer strategy to eliminate redundant data. In mobile platform real-time crowdsensing, Wanita *et al.* [Sherchan et al. 2012] designed a system for collecting data via instantaneously data analysis and process. To reduce bandwidth consumption and save the energy for mobile devices, their CAROMM is able to acquire various stream data by mobile devices and process them based on context attached, e.g., the location and time mark on photos, finally contributing to the relevant data retrieval from the dataset.

Riteau *et al.* [Roemer et al. 2014] adopted a data deduplication strategy to reduce the storage and bandwidth consumption for the applications which require a great deal of data to be kept and conveyed. Based on WANs, a distributed data deduplication method and a message-delivery model were provided. However, the semantics of the content was not considered to further improve the performance of the approach.

To address the high energy consumption problems involved in smartphone based crowdsensing applications, Nicholas *et al.* [Lane et al. 2013] proposed an energy effective crowdsensing strategy by taking advantage of opportunistic application run by the users. The solution is called Piggyback CrowdSensing (PCS), and it depends on a predictive model to find the optimal time slot to perform the sensing task. Prediction is an effective way to avoid meaningless cost and lower the overhead, e.g., taking into account location information. The data (i.e., images) from exactly the same location tend to contain the same information. Besides, their analysis on the application specifics can also contribute to the overall cost-reduction. Gorlatova *et al.* [Gorlatova et al. 2014] presented solutions on estimating harvested energy from acceleration records. In order to characterize the energy availability related to particular human behaviors, the work [Gorlatova et al. 2014] analyzes a motion dataset with over 40 participants, and an energy allocation algorithm with accessible IoT node solution designing has been developed and evaluated based on the collected measurements.

### 3.2. Different Crowdsensing Strategies for Achieving Good QoS

Below, we introduce a list of methods for achieving good QoS in MCS.

Xu *et al.* [Xu et al. 2015b] proposed Compressive CrowdSensing (CCS) which is a framework for applying compressive sensing techniques to mobile crowdsourcing scenarios. CCS enables compressive sensing techniques to be applied to MCS by providing significantly reduced amounts of manually collected data and maintaining acceptable levels of overall accuracy at the same time.

Yan *et al.* [Yan et al. 2010] proposed CrowdSearch for searching images using mobile phones. CrowdSearch integrates the strategy of automated image search into the real-time validation of human. They combined local processing on mobile phones and backend processing on remote servers to implement the process of image search. By balancing accuracy and monetary cost, CrowdSearch finds a trade-off between accuracy and monetary cost and ensures user-specified deadlines for responses to search queries simultaneously. To improve the quality of images, CrowdSearch presents a new prediction algorithm to determine the results needed to be validated, and determine when and how to validate these results.

Due to the limited resource, it is a challenge to transfer a huge amount of crowdsensed data. To address this challenge, Wang *et al.* [Wang et al. 2014] proposed a framework called SmartPhoto, to quantify the quality (utility) of crowdsensed photos based on the accessible geographical and geometrical information (referred to as metadata), which contains the information of the device's orientation, location and all related parameters of the built-in camera. With the metadata, it can be inferred where and how the photo is taken. Also, SmartPhoto only transmits the most useful photos. They also studied three optimization problems on the trade-offs between photo utility and resource constraints. Moreover, they designed efficient algorithms with theoretical proofs of the performance of the algorithms. Finally, by using Android based mobile phones, they implemented SmartPhoto in a testbed with techniques designed to improve the accuracy of the collected metadata by reducing sensor reading errors.

Xu *et al.* [Xu et al. 2015a] studied compressive sensing under the scenarios in which different samples have different costs. This work tries to balance the minimization of the total sample cost and the recovery accuracy, and designs Cost-aware Compressive Sensing (CACS) for incorporating the samples' diversity on cost into the compressive sensing framework. The CACS has been applied to networked sensing systems.

To maximize the aggregate data utility, Li *et al.* [Li et al. 2015] studied how to the aggregate data utility under the constraint on budget in MCS. They presented a combinatorial auction mechanism that utilizes a redundancy-aware reverse auction framework. The auction mechanism is mainly composed of two parts: an approximation algorithm used for winning bids determination and a critical payment scheme.

Table II lists different techniques in MCS with an emphasis on redundancy reduction. Table III gives a comparison among common strategies of MCS in recent years and an example work of each strategy. We cite the most representative case for each strategy. As we can see from the table, most of the data sensed by the mobile device is images which contains rich information and consumes a small amount of storage due to data deduplication.

### 3.3. Dataset

In the above, we discuss different techniques for handling the resource limitation issue to achieve low cost while achieving good QoS. Below, we provide some datasets for research in MCS.

There are several typical datasets available for crowdsensing research. The dataset contributed by von Ahn *et al.* [von Ahn and Dabbish 2004] consists of 100,000 images with English labels which are from their ESP Game[3]. TagATune[4] is a research dataset for human computation game, and it was published by Law *et al.* [Law and von Ahn 2009]. TagATune contains human annotations. Another dataset is the ESP Lite game developed by Chen *et al.* [Yuen et al. 2009; Huang et al. 2010]. The ESP Lite game is similar to the ESP game introduced by von Ahn *et al.* [von Ahn and Dabbish 2004]. The statistics for players playing the game is available now.[5] CiteULike[6] developed by Oversity Ltd is a free website, which allows users to save and share citations to academic papers and is used to help academics record of the articles they are reading on. CiteULike encourages users to share their libraries on the website so that others can benefit from the resource sharing for discovering articles that are useful to them. To better facilitate research, Körner and Strohmaier [Körner and Strohmaier 2010] released a list of social tagging datasets.[7]

## 4. CROWDSENSING STRATEGIES FOR DIFFERENT APPLICATION DOMAINS

Apart from the above strategies focusing on the data processing phase, we can divide the categories of these strategies based on their application domains [Guo et al. 2015]. Accordingly, common domains of the crowdsensing are as follows.

### 4.1. Natural Environment Monitoring

The main purpose of the crowdsensing strategies for natural environment monitoring is to keep track of the status of the natural environment in order to prevent avoidable disaster and human pollution. For example, tracking the real-time temperature in a particular area of a forest can monitor the sign of the fire and signal the warning to prevent from the disaster.

As for environment protection oriented crowdsensing strategies, academy and industry currently both are likely to take advantage of a vast number of smart devices owned by the public to do the research or make a profit at a relative low investment. Shilton *et al.* [Mun et al. 2009] made use of participatory sensing strategy to measure the impacts of climate changes and pollution sources. Participatory, a process of collecting and analyzing data, leverages the individuals' smart devices, conveys data by wireless network and process the data in the data center. Two features of the strategy are context-triggered feedback and data visualization, which make the software interact with smart device owner effectively to perform the specific task. For example, when

[3]ESP Game dataset: http://server251.theory.cs.cmu.edu/ESPGame100k.tar.gz.
[4]Tagatune Dataset: http://tagatune.org/Magnatagatune.html.
[5]The website of IIS-NRL Games With A Purpose - ESP Lite. http://hcomp.iis.sinica.edu.tw/dataset/dataset_esplite20100101.php.
[6]CiteULike website: http://www.citeulike.org and the dataset website: http://svn.citeulike.org/svn/plugins/HOWTO.txt.
[7]A List of Social Tagging Datasets Made Available for Research: http://kmi.tugraz.at/staff/markus/datasets/.

| Reference | Sensing Task | Technology | Collected Data |
|---|---|---|---|
| Hua *et al.* [Hua et al. 2015] | Real-time image sharing in disaster situation | QoS-sensible redundancy reduction in the software-defined networks | Image |
| Dao *et al.* [Dao et al. 2014] | Image/video uploading in disaster environment | Comparing the metadata of images to eliminate the redundancy | Image |
| Xu *et al.* [Xu et al. 2015b] | Data compression aided crowdsensing | Indirectly reducing the signal dimension | Responses to questionnaire |
| Gorlatova *et al.* [Gorlatova et al. 2014] | Kinetic energy sensing and analyzing | Energy allocation algorithm based on accelerometer acquisition | Kinetic energy |
| Yan *et al.* [Yan et al. 2010] | Smartphone based crowdsensing management | Participation pattern recognition, incentive modeling and cost reduction | Mobile sensing data |
| Aggarwal *et al.* [Aggarwal et al. 2011] | Sensor stream selection | Real-time data redundancy-reduction algorithm for data collection | Intel-humidity and intel-temperature |
| Willett *et al.* [Willett et al. 2013] | Redundancy recognition and provenance detection | Copying and paraphrasing determination | Crowdsourcing data like text |
| Wang *et al.* [Wang et al. 2014] | Smartphone based image crowdsensing | Metadata aided selective image sharing | Image |
| Xu *et al.* [Xu et al. 2015a] | Cost-sensible crowdsensing | Optimization algorithm for balancing the recovery accuracy and sample quantity | Air pollution data |
| Li *et al.* [Li et al. 2015] | City mobility pattern monitoring | Spatio-temporal analysis on vehicle traces | GPS data |
| Chon *et al.* [Chon et al. 2013] | Place-centered crowdsensing coverage and scalability analysis | Crowdsensing property modeling | Smartphone sensing data |

**Table II:** A list of different mobile crowdsensing strategies.

a person arrives at a particular place, the context-aware device will automatically notify the owner to take a photo with the natural target to upload to the cloud for further analysis. In terms of nature pollution targeted crowdsensing, relevant strategies are designed on noise and air pollution aspects.

Crowdsensing plays an important role in measuring and reducing the noise pollution. Nicolas *et al.* [Maisonneuve et al. 2010] presented a participatory noise pollution detection method, called NoiseTube, based on a mobile phone platform, aiming to acquire the first-handed noise data suffered by individuals. NoiseTube records the magnitude of the noise combined with the position and time information for further statistical analysis. For similar purpose, Rana *et al.* [Rana et al. 2010] established a noise map to enhance the efficiency for monitoring noise in cities. Crowdsensing techniques were adopted by them to avoid the high cost to build the noise map by traditionally infrastructure based methods. They implemented the method on Nokia N95 and HP iPAQ platforms and addressed the problem for ensuring the noise measurement accuracy. To reduce the computation overhead for mobile phones, the data analysis is conducted in the data center.

Crowdsensing is also used for air pollution measurement. PEIR, a project for monitoring the human effect on the environment with the aid of crowdsensing research, is conducted by Mun *et al.* [Mun et al. 2009]. The sensing system consists of mobile handset GPS receivers for collecting the position data, server data classification processors for detecting different modes of the transportation and a database for looking up the weather and road condition data. The main contribution of the work lies in two aspects: an innovative map-matching and pattern recognizing algorithm and a mechanism for protecting the user's privacy from leakage. Zheng *et al.* [Zheng et al. 2013] proposed a holistic method, combining data sources from crowdsensing application, monitor station and historical air pollution database, to actually report the real-time air quality all over the city. The core methods involved in their solution are two classifiers, one is on the basis of artificial neural network (ANN) taking into account the spatial information of an area; the other is called linear-chain conditional random field (CRF), considering the real-time dependency among factors, such as temperature, humidity and etc. The paper sheds light on making use of artificial intelligence to solve the crowdsensing problems, which we believe is a trend in the near future.

### 4.2. Traffic Information Collection and Management

Crowdsensing strategies play an important role in collecting traffic information, and Crowdsensing strategies help the public and the government on related decision-making. Below, we introduce three aspects.

*4.2.1. Traffic flow information collection.* The real-time road condition crowdsensing and monitoring has drawn much attention. Calabrese *et al.* [Calabrese et al. 2011] proposed a real-time road condition monitoring system with the aid of the LocHNESs platform performing the data collection and uploading task via a cellular network. The mobility information, position, speed and time of buses, taxis and pedestrians of the entire Roma city are collected to analyze the instantaneous traffic status in the city. The system focuses on the unexpected traffic trend comparing with predicted traffic information to further improve the city traffic management for a higher transportation efficiency. Specific to resident transport behavior, Liu *et al.* [Liu et al. 2009] designed a method to monitor the mobility patterns of citizens and visualize the data to show the trend of the development of city economy and infrastructure. They used the smart card including information like date, time, and taxi GPS records containing vehicle ID, company, longitude and latitude to extract the features to obtain travel distances, durations and zones. The experiments conducted in Shenzhen city show that the data analysis result can improve the citizens' quality of life by improving the traffic management efficiency.

*4.2.2. Transport service improvement.* The crowdsensing data also can be used to improve public transport quality. The application includes the optimization of the bus routes and schedules and modification of taxis zone allocation. By using the Taxi GPS traces, Chen *et al.* [Chen et al. 2014] presented a method to recognize the resident mobility pattern to contribute to the night-bus route modification. The solution comprises two phases. In the first phase, the high density spots of pick-up/drop-off are detected and an optimal bus stop to split the amount of the flow is applied. In the second phase, the constraints, bus route origin, destination and time, on the allocation of bus station are taken into account to obtain a global optimal arrangement of the bus stations. From the individual's perspective, Zhou *et al.* [Zhou et al. 2014] proposed an approach to predict the waiting time for the next bus with the aid of the crowdsensing techniques. On the basis of the commodity cell phones, the ambient of the bus passengers is detected and used to estimate the arrival time of buses. The highlight of the paper is that, instead of GPS-only localization method, the authors combined various context factors, cell tower positioning information, inertial measurements, voice records, etc. to obtain an energy-efficient and highly robust scheme. As proved by their experiments, during 7-week period with a variety of Android based cell phones, the crowdsensing improves the passengers' experience when they are waiting for the buses. Just like the information of the weather, real-time road condition is tightly related to the individual's daily activities, commuting, traveling, etc. Many researchers and companies (e.g., Google Map is able to visualize the traffic condition near the driver for choosing an easy path efficiently) are interested in keeping track of up-to-dated traffic conditions, such as congestion, accident and severe weather. The Pothole Patrol, a crowdsensing application researched and developed by Eriksson *et al.* [Eriksson et al. 2008] aims to test the road surface condition using GPS and vibration sensors equipped on the moving vehicles. The data collection is triggered opportunistically and road problems like potholes can be detected by using a fundamental machine learning approach. The advantages of their approach include low cost, due to accessible on-board positioning and inertial sensors, and high rate of successful road problem detection, e.g., more than 90% detected road anomalies require to be fixed.

### 4.3. Urban Dynamics Sensing

The understanding of urban dynamics is critical for urban development and quality improvement of citizen life. Understanding urban dynamics is a key challenge. Urban dynamics sensing has become possible and has attracted many interests from both industry and academic research societies. Human urban mobility/behavior patterns: Some works study how to reveal human mobility and behavior patterns in urban areas. Adeel *et al.* [Adeel et al. 2014] studied how to provide a cost-effective networking service for real-time and delay tolerant applications in Mobile Urban Sensing System (MUSS). They proposed a novel networking scheme that supports both real-time and delay-tolerant urban sensing applications. The core of the scheme is the trading of mobile sensor data in a virtual market where the scheme was demonstrated to be able to incentivize mobile phone users to participate. Pan *et al.* [Pan et al. 2013] addressed the problem of detecting and describing traffic anomalies using crowd sensing with two forms of data, human mobility and social media. Phithakkitnukoon *et al.* [Phithakkitnukoon and Oliver 2011] used a location-based online social networking data to sense geo-social activity and analyze the underlying social activity distribution of three different cities.

### 4.4. Location Services

With the development of sensing devices equipped with sensors, MCS has been widely used in location services. The benefits of location awareness promote many popular mobile applications, such as location search, location-based advertising (e.g., disseminating electronic coupons in a market [Garyfalos and Almeroth 2008]), indoor localization (using WiFi signal strength to locate people/objects) [Rai et al. 2012; Kumar et al. 2014], etc.

### 4.5. Social Network Based Applications

Since individuals are highly involved in the crowdsensing activities, there are a large number of social network applications developed from the crowdsensing data. Zheng *et al.* [Zheng and Xie 2011] proposed an adaptive travel recommendation system resulting from travelers' historical GPS position records. By collecting and analyzing the GPS trace of individuals, two types of recommendation approach are given. For the first approach, the recommendation system generates a general list of hot places of interest for users. The second approach is to provide a customized option based on particular needs of users. A tree-based structure is designed and combined with Hypertext Induced Topic Search aided model to estimate the attraction of a place and user's inclination. Similarly, position information sensed from the crowdsensing devices, like mobile phones, is analyzed to create a recommendation according to individual's interest [Ye et al. 2011]. The recommendation is derived from the location information among user's social network, and it takes into account the social influence. For example, close friends are likely to share the similar interests, and geographical influence, i.e., people tend to visit the nearest place where their requirement can be satisfied. Furthermore, random walk technique is used to compensate the bias of the basic approach when friends sometimes hold different preferences.

| Techniques | Pros | Cons | Applicability | Ref |
|---|---|---|---|---|
| Deduplication | Using metadata for reducing redundancy | Additional information of data is required | Redundant content management | [Dao et al. 2014] |
| Compression | Low bandwidth and storage requirement | Existing data accuracy loss | Bandwidth-constrained data transferring | [Zordan et al. 2014] |
| Machine learning | Fully automatic information classification | Requiring large training dataset | Data-driven city security maintenance | [Ballesteros et al. 2013] |
| Context-aware | Monitor and visualize service of a virtual world | High bandwidth requirement | 3-D Web-based interface | [Yao et al. 2014] |
| Peer-to-peer | Independent to centralized infrastructure | Low reliability | Android based distributed crowdsensing | [Rothenpieler et al. 2014] |
| Opportunistic sensing | Energy-efficient | Poor real-time performance | Collecting mobile sensor data | [Lane et al. 2013] |
| Optimal estimation | Low storage requirement | High computation workload | Earthquake center estimation | [Sakaki et al. 2010] |
| Data filtering | Increase the accuracy of information prediction | Priori knowledge and accurate model is necessary | Recommendation system for taxi service | [Yuan et al. 2013] |
| Content-aware | Content similarity detection contributes to redundancy reduction | High energy overhead for the similarity detection | Image-transferring in disaster area | [Weinsberg et al. 2012] |

**Table III:** Comparison of different types of techniques in mobile crowdsensing.

Crowdsensing applications always involve a huge amount of data related to social activities. Thus, many researchers focus on public security and communication enhancement in disasters. As for the public security issues, Sheth *et al.* [Sheth 2009] introduced a system combining human-involved crowdsensing, Web 2.0 and mobile computing, to establish a platform for recognizing the emergency, analyzing the situation of accidents and calling for the help automatically. The holistic situation awareness model proposed includes three pivotal phases, namely observation, perception and communication. For instance, when a traffic accident happens, individuals around the scene may take images to share on-line. Then, the metadata, longitude and latitude of the image are extracted and the emergency situation is recognized while the signal can be immediately broadcast to related people. Even the situation-aware crowdsensing is an effective way to detect and convey the emergency information, the privacy of involved individuals should be protected. Thus, Ballesteros *et al.* [Ballesteros et al. 2013] designed iSafe, a privacy preserving method, analyzing the crowdsensing data from individuals' phones, to improve the safety of the city. By their approach, snapshots of taken by both the user's phone and the geosocial network users are used for analysis, e.g., the level of the safety of a place is determined.

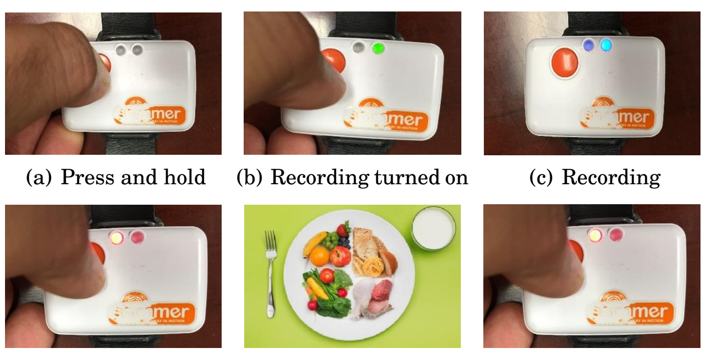
In disaster scenarios, smart wearable device based crowdsensing plays an important role in guaranteeing the communication with the outside in congested environment. Sakaki *et al.* [Sakaki et al. 2010] presented an algorithm taking advantage of the real-time characteristic of Twitter to immediately detect the earthquakes. The key idea is as simple as observing the tweet activities, e.g., whenever an earthquake happens, tons of Twitter posts relevant to the earthquake will be created during a short period of time. Similarly, large social events are also able to be captured in the same way. During the capturing process, a well-deigned event classifier is necessary. They used Kalman and particle filtering method to obtain an optimal estimation of the earthquake center and the trace of typhoon.

### 4.6. Healthcare

Health is becoming an increasingly important challenge. Wireless sensors are worn by people for heart rate monitoring and blood pressure monitoring, and they can communicate their information to users' equipment. MCS can utilize these existing data for large scale healthcare study. Based on the wealth of the data collected from MCS systems, the health monitoring and management services can be roughly categorized as: public health monitoring and personal well-being management [Guo et al. 2015].

• Public health monitoring: MCS can facilitate the monitoring of disease outbreaks and crisis management, which potential brings economic benefits. For example, the Ministry of Health in Cambodia uses GeoChat[8], a crowdsensing interactive mapping application, for disease reporting and staff alerts which enables quick responses to the disease outbreaks and can better control the spread of diseases. Also, Wesolowski *et al.* [Wesolowski et al. 2012] use large scale spatially mobile phone data and malaria prevalence information from Kenya to identify the dynamics of human carriers that drive parasite importation between regions.

• Personal well-being management: MCS can also facilitate personal well-being management by monitoring users' daily activities. For example, Rabbi *et al.* [Rabbi et al. 2011] presented a mobile sensing system for measuring mental well-being from behavioral indicators in natural everyday settings. In order to lose weight, the work [Dong et al. 2012] presents a method for measuring intake via automated tracking of wrist motion. The method uses a watch-like device embedded with a micro-electro-mechanical gyroscope to detect and record an individual's eating activity. Figure 3 illustrates the steps for using the watch-like device to detect and record an individual's eating activity.



(a) Press and hold (b) Recording turned on (c) Recording

(d) Pre-meal marker (e) Eat and drink (f) Post-meal marker

**Fig. 3:** Process of using watch-like device for detecting and recording an individual's eating activity.

### 4.7. Public Safety

Public safety means the detection or protection from social or natural events such as crimes, disasters that could endanger the safety of average citizens.

• Crime prevention and investigation: Crime is becoming one of the key problems in modern society. Ballesteros *et al.* [Ballesteros et al. 2012] presented iSafe, a privacy

[8]InSTEDD. (2006). GeoChat, Sunnyvale, CA, US: http://instedd.org/technologies/geocha.

| Application type | Typical Applications |
|---|---|
| Combining crime mapping and crowdsourcing | [Blom et al. 2010; Garbett et al. 2015] |
| Perception of crime mapping | [Kounadi et al. 2014; Quinton 2011] |

**Table IV:** Classification of applications for crime prevention.

preserving algorithm for computing safety snapshots of co-located mobile device users and integrated their approach into an Android application for visualizing safety level. They also investigated relationships between location dependent soia network activity and crime levels. Cvijikj *et al.* [Cvijik et al. 2015] implemented a mobile application, a crowdsourcing approach, for crime prevention, which focuses on the usage intention and motivations for content creation and consumption. Table IV shows different application types of crime prevention.

• Disaster management and relief: Events such as the big flood in mid-Europe 2013 and the typhoon Haiyan in Philippines show that people become increasingly active in responding to disasters. MCS has been used for disaster management and relief. Rogstadius [Rogstadius et al. 2013] presented CrisisTracker, a crowdsourced social media curation system, for disaster awareness. CrisisTracker collects data from Twitter based on predefined filters (i.e., keywords), and groups these tweets into stories for analysis. Hua *et al.* [Hua et al. 2015] proposed SmartEye, a near-real-time and cost-efficient mobile device based crowdsourcing, for rapid disaster relief in the cloud-assisted disaster environment. SmartEye utilizes the in-network deduplication strategy to obtain fast operation response and significant bandwidth savings so that it can efficiently support the image retrieval in the context of disaster relief.

## 5. CHALLENGES AND FUTURE RESEARCH DIRECTIONS

In this section, we first discuss further challenges related to crowdsensing for IoT, and then we provide guidance on the future research trends of crowdsensing for IoT.

### 5.1. Challenges in Mobile Crowdsensing

*5.1.1. Automated configuration of sensors.* In traditional pervasive/ubiquitous computing, only a limited number of sensing devices (e.g., sensors) are connected to the applications (e.g., smart farm, smart river). However, in IoT, a large number of sensing devices are expected to be connected together over the Internet. Therefore, the connection and configuration of sensing devices to applications become a key challenge. It is infeasible to connect all sensing devices manually to an application or to a middleware [Perera et al. 2013]. An automated or at least semi-automated process should be available to connect sensing devices to applications. To accomplish the tasks of connecting sensing devices to applications, applications should be capable of understanding the sensing devices (e.g., capabilities). Several recent developments such as Transducer Electronic Data Sheet (TEDS) [IEE 2007], Open Geospatial Consortium (OGC) Sensor Web Enablement related standards like Sensor Markup Languages (SensorML) [Botts and Robin 2007] show the future trends of carrying out research work for addressing the challenge of connection and configuration of sensors to applications.

*5.1.2. Resource limitations.* Sensing devices (e.g., sensors and mobile phones) usually have limited resources, and the resource limitations arise as a challenge for crowdsensing. Although more resources (e.g., computing, bandwidth) are provided for mobile phones compared to mote-class sensors, mobile phones still face the problem of resource limitations [Miluzzo et al. 2008; Guo et al. 2015].

Different types of sensed data may be independent with each other because of the multi-modality sensing capabilities of sensing devices. In practical scenarios, different types of sensed data may be used for the same purpose. However, the diversities on the quality and resource consumption of the sensed data pose an obstacle for improving the quality of data with low resource consumption. Therefore, it is still a challenge to improve the quality of data and minimize the resource consumption.

*5.1.3. Data redundancy, quality, and inconsistency.* Multiple participants involved in the same sensing activity usually incur data redundancy. Enormous amounts of data consumes much resource. By intelligently reducing the redundant data and the transfers of redundant content, the volume of the data and the traffic load can be significantly reduced. Hence, it is important and necessary to eliminate the redundant data, which can help reduce the resource consumption (e.g., storage resource and bandwidth resource) and thereby reduce the cost. A key challenge here however, is detecting redundant data. That is, detecting 'what content is similar' [Dao et al. 2014]. For example, the work [Dao et al. 2014] designs a framework for detecting similarity among data contents and finding similar content. By restricting the transfer of similar content, the work [Dao et al. 2014] reduces the resource consumption and thereby reducing the cost, and provides good QoS in bandwidth constrained wireless networks.

However, another issue data inconsistency arises, which poses another challenge. For example, due to the different capabilities of sensing and computing, a set of mobile devices that run the same algorithm and sense the same event may obtain different inference results, which results in data inconsistency problem.

In addition, the data derived via the crowdsensing process is often noisy and incomplete, which affects the quality of the data. Also, as the redundant data has been reduced, another potential issue is 'how to ensure the quality of data', which also poses a challenge, and needs to be handled.

*5.1.4. Motivation and incentives.* Motivation and incentives are an important part of MCS because they encourage users to participate in a crowdsensing application and the success strongly depends on the contribution of the volunteers. Prior literature demonstrated the role of motivation and incentives as a key factor. Due to the privacy issues, contributors are reluctant to carry out tasks. In many cases, they do not get benefit from their work (participation). Although, some strategies have been carried out to motivate users to participate in crowdsensing tasks, users are still not very actively in carrying out small tasks. Therefore, providing effective motivation and incentives still is a challenge for MCS.

*5.1.5. Privacy, security, and data integrity.* The sensing devices potentially collect sensitive data of individuals [Krontiris and Dimitriou 2013; Ballesteros et al. 2013; Zanella et al. 2014; Chen et al. 2015; Li et al. 2014; Stansfeld 2003; Rothenpieler et al. 2014; Teixeira et al. 2015; Sherchan et al. 2012], thus privacy arises as a key problem. For example, the GPS sensor readings usually record the private information of individuals (e.g., the routes they take during their daily commutes, and locations [Krumm 2009]). By sharing the GPS sensor measurements, individuals' privacy can be revealed. Hence, it is important and necessary to preserve the security and privacy of an individual. Also, the GPS records the information which is from daily commutes shared within a larger community and can be used to learn the information of traffic congestion in a city [Hull et al. 2006]. Thus, it is also necessary to enable the crowdsensing applications so that individuals can better understand their surroundings and can ultimately benefit from the information sharing. To well preserve the enormous amounts of private information of individuals, not only methodology efforts but also systematic studies are needed. The AnonySense architecture, proposed in [Cornelius et al. 2008], can support the development of privacy-aware applications based on crowdsensing. Also, it is important to guarantee that an individual's data is not revealed to untrustworthy third parties. For example, malicious individuals usually contribute erroneous sensor data. Meanwhile, for their own benefit, malicious individuals may intentionally pollute the sensing data. The lack of control mechanisms to guarantee source validity and data accuracy can result in information credibility issues. Therefore, it is necessary to develop trust preservation and abnormal detection technologies to ensure the quality of the obtained data.

The problem of data integrity that ensures the integrity of individuals' sensor data, also needs to be well addressed. In the existing literature [Lenders et al. 2008; Saroiu and Wolman 2009], although some methods have been proposed, they typically rely on co-located infrastructure that may not be installed as a witness and have limited scalability, which makes such kind of methods prohibitive and unavailable at times. The reason behind this is that the approach relies on the inputs which is from the installation of expensive infrastructure. Another approach for handling data integrity problem is to sign the sensor data (e.g., typically, trusted hardware installed on mobile phones are used for this purpose), i.e., a trusted platform module signs a SHA-1 digest of the sensor data. This approach is potentially problematic due to the reason that the verification process has to be done even in the software.

### 5.2. Challenges of IoT

*5.2.1. Availability.* Availability is an important challenge in IoT systems. Availability of IoT can be realized in the hardware and software levels to provide anywhere and anytime services for customers. Software availability is the ability of IoT applications to provide services for everyone at different places simultaneously, and hardware availability means that the devices are available all the time that are compatible with the IoT functionalities and protocols. Replicating critical devices and services is a common solution for achieving high availability of IoT services.

*5.2.2. Reliability.* Reliability refers to the power working of the system based on the system's specification. Reliability reflects the success rate of the delivery of IoT service, and it is even more critical than availability and it has more strict requirements when it involves the field of emergency response applications [Maalel et al. 2013]. The critical part must be resilient to failures so that the IoT system can provide reliable information distribution. To ensure the quality of the services in IoT systems, the reliability should be implemented in both software and hardware throughout all the IoT layers.

*5.2.3. Mobility.* Since most of the services of IoT are expected to be delivered to mobile users and connecting users with their desired services continuously, mobility is also a challenge. Service interruption for mobile devices can occur when the devices transfer from one gateway to another. Future Internet presents a more ubiquitous and mobile Internet. As the number of smart devices increases sharply in the IoT systems, the mobility management becomes necessary. Internet of Vehicles (IoV) becomes an emerging area of the IoT, and it needs a precise attention to the mobility issues. The work [Zhu et al. 2011] studies the mobility in vehicle-to-vehicle networking, and it discusses various solutions for handling the mobility issue in vehicle-to-vehicle networking.

*5.2.4. Management.* The connection of billions or trillions of devices poses another challenge for managing the Fault, Configuration, Accounting, Performance and Security (FCAPS). To address the challenge of device management, a number of companies have proposed unique solutions to the market. For example, UpdateLogic proposed a device management solution called NetReady, and it has found a market in supporting smart TVs and other connected consumer electronics. Ihiji provides the solutions of remote network management for smart home and other control solutions.

*5.2.5. Scalability.* The scalability of IoT indicates the ability to add new devices, services and functions for customers without compromising the quality of existing services. Adding new operations and supporting new devices is a nontrivial task, and the diversity of hardware platforms and communication protocols makes it more difficult. The Internet lacks the capability of the support for unique identification and transparency. A new network architecture is needed to overcome the current Internet limitations. The expected ubiquity of computing and communication resources should be considered to improve the connectivity and robustness of wireless sensor and actual sensor networks.

*5.2.6. Interoperability.* IoT is the connectivity between people, processes and things. A large number of heterogeneous things belonging to different platforms need to be handled in the IoT, thus end-to-end interoperability becomes another challenge. Addressing this challenge is essential to unlock the full potential of the IoT. To ensue the quality of the deliverability of services for customers, interoperability should be considered by both application developers and IoT device manufacturers. To handle the issue of interoperability, an increasing number companies and products are beginning to emerge that enable interoperability through open-source development. For example, third party associations such as the IEEE are working with global engineering communities to standardize and facilitate collaboration. Qualcomm develops AllJoyn, which is an open source project that provides a universal software frame and set of system services enabling interoperability.

### 5.3. Future Research Directions

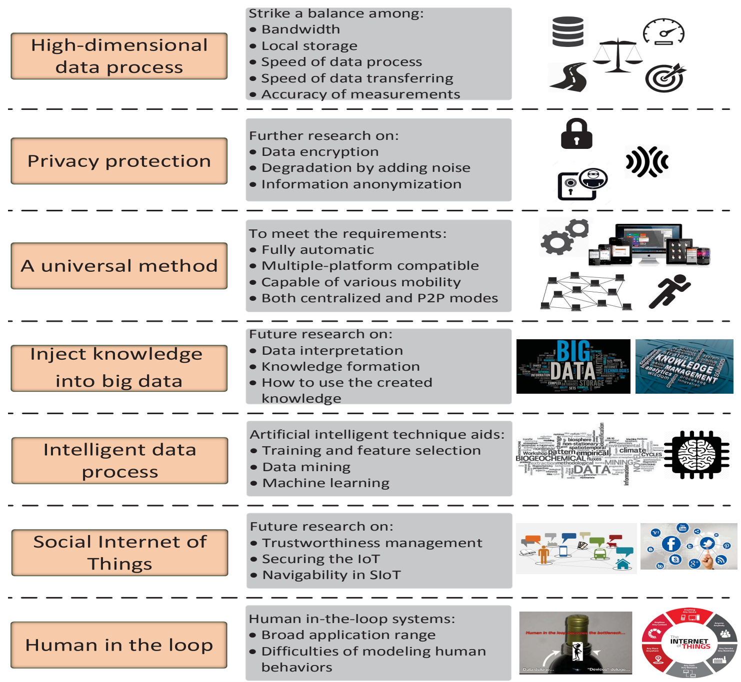
Below we present some future research directions of crowdsensing for IoT. Figure 4 summarizes the future research directions of crowdsensing for IoT.

*5.3.1. Optimization of multiple factors like localization, prediction, energy budget.* The trade-off between higher location accuracy and lower energy consumption for MCS devices is critical to successfully implement various algorithms [Howe 2006; Susmita and Anjali 2012; Kirak et al. 2013; Hasan and Curry 2014; Vasilescu et al. 2005]. For example, in the solution proposed by Lane *et al.* [Lane et al. 2013], to lower the energy overhead based on the context information, such as position, its real-world performance suffers from the inaccurate localization model. Besides, for MCS, especially mobile device platform, more than one sensor can be used to collect data and sense the context, such as dynamic status, localization, and noise magnitude. Thus, the reliability and the amount of information of context may be increased as in the work [Sherchan et al. 2012] in which the proposed CAROMM is able to acquire various stream data from mobile devices and process them based on context attached, e.g., the location and time mark on photos. This further contributes to the performance of crowdsensing.

*5.3.2. Privacy protection.* Privacy protection is a principal issue that has not yet been well addressed [Stankovic 2014; Whitmore et al. 2014], especially in the crowdsensing area. There is a large body of work focusing on privacy protection [Lane et al. 2013; Krontiris and Dimitriou 2013; Sherchan et al. 2012]. The CAROMM framework, making use of the context of the data from user's mobile devices, bears high risks to leak the privacy information of users since the information like location and time, which are required to be protected. Obviously, the privacy risk must be reduced to an acceptable level before any crowdsensing activity is conducted. Otherwise, the user's privacy may be exposed to the public. Lane *et al.* [Lane et al. 2013] conducted research on the automatic data anonymization by masking particular information from the raw data sensed by the local mobile devices. Also, the IoT additionally introduces unique challenges to privacy, and many of them go beyond the current existing data privacy issues. This mainly stems from integrating devices into the environments without users consciously using them. Moreover, many IoT scenarios involve device deployments and data collection activities with a multinational or global scope that cross social and cultural boundaries, which poses a new challenge for developing a broadly applicable privacy protection model for the IoT.

*5.3.3. A universal method.* To the best of our knowledge, current crowdsensing strategies can only be applied to limited contexts, i.e., either mobile or stationary plat-

**Fig. 4:** Future research directions of mobile crowdsensing for IoT.




form [Lane et al. 2013; Forsström and Kanter 2014; Kamra et al. 2006; Distefano et al. 2015; Brambilla et al. 2014; Bisdikian et al. 2013; Bengtsson et al. 2011]. PCS (Lane *et al.* [Lane et al. 2013]) can only be applied to the tasks that can be done without human participation and cannot be used in dynamic condition, e.g., in a driving car. The data redundancy handling method proposed in [Dao et al. 2014] is able to manage the image data successfully while it can do nothing on videos although the image and video are both the common information medium. Since a universal strategy is able to significantly reduce the cost for modification to meet the requirements of various crowdsensing scenarios, especially for the applications, i.e., only one application is required for performing multiple tasks. Indeed, the limitation impairs its communication efficiency in disasters. A crowdsensing strategy being capable of centralized and distributed data collection is also a direction for future research, since researchers usually focus on one of them. CARE [Weinsberg et al. 2012] only uses a peer-to-peer structure to implement an information-aware redundant data reduction while authors in [Dao et al. 2014] designed their redundancy elimination method in bandwidth constrained wireless networks with the aid of the infrastructure. Since in some particular scenarios, such as disaster field, the feasibility and flexibility quality of the solution is non-trial, we believe that a hybrid approach combining advantages from different system architectures would be necessary.

*5.3.4. Injecting Knowledge into Big Data.* In an IoT world, a huge amount of raw data are continuously collected. It is worthwhile developing techniques to convert the raw data into knowledge. Take raw data in medical area as an example, raw streams of sensor values should be converted into semantically meaningful activities performed by or about a person, i.e., eating, respiration, or exhibiting signs of depression. The main challenges are how to interpret data and how to format the knowledge. Specifically, the challenges are mainly reflected in how to address noisy, physical world data and how to develop new techniques without suffering the limitations of Bayesian or Dempster-Shafer schemes, which needs to know priori probabilities and the cost of computations. Although rule based systems may be adopted, they are two ad hoc for some applications.

The amount of sensor data collected is enormous. A huge amount of real-time sensor data streams will exist, thus it is not rare that a given sensor data stream will be used in many different ways for different inference purposes. Therefore, enabling data streams to act as primitives for unexpected future inferences will be an interesting research problem.

After the knowledge has been created, another challenge is how to better control or make good decisions in using the created knowledge. However, to ensure the reliability of the system, it is important and necessary to minimize the number of false negatives and false positives and guarantee safety in making decisions, which is a non-trivial task.

*5.3.5. Intelligent data processing.* Current methodology for data deduplication can be mainly categorized as: real-time process and post-process [Zhan et al. 2015; Kazmi et al. 2014]. However, as the increasingly enlarging dataset and complex data type, under the limited time, bandwidth and other resource budget, machine learning techniques may play a non-negligible role. The data management method presented in [Dao et al. 2014] can be improved by training the algorithm in advance and then using the trained parameters to improve the efficiency of redundancy detection. On the other hand, a portion of the sensing device's storage can be used to store the metadata, which can be used to set different priorities for different data types. Combined with the machine learning techniques, the priority can be arranged to the data automatically based on their relevance to the requirement, e.g., the image of injured people with exact location in a disaster scenario may be uploaded with high priority.

Since data classification plays a pivotal role in crowdsensing technology and machine learning is the current focus in that field, machine learning based crowdsensing method can improve the performance of the system. Zheng *et al.* [Zheng et al. 2013] designed a semi-supervised learning approach consisting of a spatial classifier and a temporal classifier to learn the features of the quality of the air in the entire city and then used it to classify the degree of the air pollution. Two essential parts of their method are to select effective features of the air for machine learning and pollution classification and ensure the size of the training set for the machine learning program. Besides, machine learning based crowdsensing approaches are also able to detect the road surface problems automatically. Eriksson *et al.* [Eriksson et al. 2008] proposed efficient methods to use the taxis and on-board sensor sensing to contribute to the road maintenance with the aid of machine learning. The establishment of the training data and the design of the features are their future research work.

*5.3.6. Social Internet of Things.* Real humans are believed to understand and answer better than a machine, and they are the most "intelligent machines" [Shen et al. 2015a; Liu et al. 2015]. A large number of individuals tied in a social network can provide better answers to complicated problems than a single individual (or even a knowledgeable individual) [Atzori et al. 2012]. The collective intelligence emerging in social networks can help users find information (e.g., answers to their problems), which attracts many interests. Social networks have the advantage of efficiently discovering and distributing services, and social networks are utilized by many systems, such as Yahoo! Answers, Facebook, for sharing the information (e.g., knowledge). Although many techniques have been proposed for social networks and IoT, the integration of social networks and IoT still faces some challenges. For example, the scalability problem will emerge as the number of embedded computing and communication devices will soon become too large. Also the trustworthiness is another challenge faced by the SIoT. Atzori *et al.* [Atzori et al. 2012] identified appropriate policies for the establishment and

| Reference | Addressing Problem | Technology |
|---|---|---|
| [Nitti et al. 2014] | Trustworthiness management | Subjective & objective model Real-time image sharing in disaster situation |
| [Atzori et al. 2012] | Integration of social netwo-rking concepts into the IoT | Design an architecture for the IoT; analyze the characteristics of the SIoT network structure using simulation |
| [Nitti et al. 2015] | Friendship selection | Analyzing possible strategies for selection of appropriate links for the benefit of overall network navigability |
| [Teixeira et al. 2015] | Secure the IoT | Treat the distributed system as a single body; crosscheck information inferred from different nodes |
| [Atzori et al. 2014] | Increasing levels of social involvement of the objects | Analyze the major opportunities arising from the integration of social networking concepts into the IoT |
| [Nitti et al. 2014] | Network Navigability in SIoT | Analyzing possible strategies for selection of appropriate links for the benefit of overall network navigability |
| [Girau et al. 2013] | Implementation of the SIoT Platform | Use RESTful approach |
| [Chen et al. 2015] | Trust-based Service Management in SIoT | Adaptively control and manage trust |

**Table V:** A list of representative works on SIoT.

the management of social relationships between objects in the way that the resulting social network is navigable. Nitti *et al.* [Nitti et al. 2014] defined the problem of trustworthiness management in the social IoT, and they presented two models: subjective model and objective model, for trustworthiness management starting from the solutions proposed for P2P and social networks. Table V summarizes the representative works on SIoT. Therefore, there is a great potential and prospect for integrating social networking into Internet of Things, which will be an important research direction.

*5.3.7. Humans in the Loop.* Since many applications of IoT involve humans, that is, humans and things will operate synergistically. Human in-the-loop systems bring opportunities to broad the range of applications which include energy management [Lu et al. 2010], health care [Kay et al. 2012], and automobile systems [Burnham et al. 1974; Liu and Salvucci 2001]. Modeling human behaviors is still a long way to go though having human in the loop has its advantage.

## 6. BIG DATA ANALYTICS AND CLOUD IN SUPPORT OF THE IOT

As the development of IoT system, the demand on the storage for the IoT system for big data analytics increases. Although, some platforms for big data analytics like Apache Hadoop have been developed, the systems are not strong enough for big data needs of IoT. Cloud computing offers a new management scheme for big data, however it involves many challenges for employing cloud computing for IoT. For example, securing the IoT cloud-based service poses a challenge. Therefore, big data analytics and cloud in support of the IoT will be a research direction.

## 7. CONCLUSIONS

The IoT has attracted much attention over the past few years. Numerous sensing devices emerge in our living environments, which creates the IoT integrating the cyber and physical objects. MCS plays an important role in the IoT paradigm. Sensors continuously generate enormous amounts of data, which consumes much resource, such as storage resource for storing data and bandwidth resource for data transfer. Previous works demonstrate that there is significant amount of redundancy in sensor data. Thus, redundancy elimination of sensor data is important and worthwhile, which can significantly reduce the cost (e.g., bandwidth cost for data transfer) and facilitate the timely delivery of critical information by reducing the traffic load, and thereby help achieving good QoS. In this paper, we review the mobile crowdsensing techniques and challenges. We focus on the discussion of the resource limitation and QoS (e.g., data quality) issues and solutions in mobile crowdsensing. A better understanding of resource management and QoS estimation in mobile crowdsensing can help us design a cost-effective crowdsensing system that can reduce the cost by fully utilizing the resource and improve the QoS for users. In the end of the paper, we describe some challenges related to crowdsensing for IoT, and discuss some of the trends in the mobile crowdsensing for IoT. In the future, we will give an in-depth study of challenges and techniques, solutions for addressing challenges in mobile crowdsensing for IoT, and we will also analyze the production systems and provide case studies.

**Acknowledgements**

This research was supported in part by U.S. NSF grants OAC-1724845, ACI-1719397 and CNS-1733596, and Microsoft Research Faculty Fellowship 8300751. We would like to thank Mr. Xiang Zhang for his valuable discussions. An early version of this work was presented in the Proceedings of the Workshop on Context-aware Performance Engineering for the Internet of Things (ContextQoS) in conjunction with ICCN 2016 [Liu et al. 2016].

## REFERENCES

2007. IEEE standard for a smart transducer interface for sensors and actuators wireless communication protocols and transducer electronic data sheet (TDES) formats. (2007). http://dx.doi.org/10.1109/IEEESTD.2007.4346346

U. Adeel, S. Yang, and J. A. McCann. 2014. Self-Optimizing Citizen-centric Mobile Urban Sensing Systems. In *Proc. of USENIX ICAC*.

C. C. Aggarwal, Y. Xie, and P. S. Yu. 2011. On Dynamic Data-Driven Selection of Sensor Streams. In *KDD*.

K. Ashton. 2009. That 'internet of things' thing in the real world, things matter more than ideas. *RFID Journal* (June 2009).

L. Atzori, A. Iera, and G. Morabito. 2010. The internet of things: A survey. *Computer Networks* 54, 15 (October 2010), 2787–2850.

L. Atzori, A. Iera, and G. Morabito. 2014. From "smart objects" to "social objects": The next evolutionary step of the Internet of Things. *IEEE Commun. Mag.* 52, 97-105 (2014), 1.

L. Atzori, A. Iera, G. Morabito, and M. Nitti. 2012. The Social Internet of Things (SIoT) – When social networks meet the Internet of Things: Concept, architecture and network characterization. *Comput. Netw.* 56 (2012), 3594–3608.

J. Ballesteros, B. Carbunar, M. Rahman, N. Rishe, and S. Iyengar. 2013. Towards Safe Cities: A Mobile and Social Networking Approach. *TPDS* PP, 99 (2013), 1. DOI:http://dx.doi.org/10.1109/TPDS.2013.190

J. Ballesteros, M. Rahman, B. Carbunar, and N. Rishe. 2012. Safe Cities. A Participatory Sensing Approach. In *Proc. of LCN*.

Paolo Bellavista, Antonio Corradi, Luca Foschini, and Raffaele Ianniello. 2015. Scalable and cost-effective assignment of mobile crowdsensing tasks based on profiling trends and prediction: The participact living lab experience. *Sensors* 15, 8 (2015), 18613–18640.

L. Bengtsson, X. Lu, A. Thorson, R. Garfield, and J. Schreeb. 2011. Improved response to disasters and outbreaks by tracking population movements with mobile phone network data: A post-earthquake geospatial study in haiti. *PLoS Medicine* 8, 8 (2011), e1001083 1–9. DOI:http://dx.doi.org/10.1371/journal.pmed.1001083

C. Bisdikian, L. Kaplan, and M. Srivastava . 2013. On the Quality and Value of Information in Sensor Networks. *ACM TOSN* 9, 4 (2013), 48. DOI:http://dx.doi.org/10.1145/2489253.2489265

J. Blom, D. Viswanathan, M. Spasojevic, J. Go, K. Acharya, and R. Ahoniu. 2010. Fear and the city: role of mobile services in harnessing safety and security in urban use contexts. In *Proc. of CHI*.

M. Botts and A. Robin. 2007. *Opengis sensor model language (sensorml) implementation specification*. Technical Report. Open Geospatial Consortium Inc, Tech. Rep.

G. Brambilla, M. Picone, S. Cirani, M. Amoretti, and F. Zanichelli. 2014. A Simulation Platform for Large-Scale Internet of Things Scenarios in Urban Environments. In *Proc. of Urb-IoT*.

G. Burnham, J. Seo, and G. Bekey. 1974. Identification of Human Driver Models in Car Following. *IEEE Trans. Automat. Control* 19, 6 (1974), 911–915.

F. Calabrese, M. Colonna, P. Lovisolo, D. Parata, and C. Ratti. 2011. Real-time urban monitoring using cell phones: A case study in Rome. *IEEE Transactions on Intelligent Transportation Systems* 12 (2011), 141–151. DOI:http://dx.doi.org/10.1109/TITS.2010.2074196

Carnot Institutes. 2011. Smart networked objects and internet of things. Carnot Institutes Information Communication Technologies and Micro Nano Technologies alliance, White Paper.

C. Chen, D. Zhang, N. Li, and Z. Zhou. 2014. B-planner: Planning bidirectional night bus routes using large-scale taxi GPS traces. *IEEE Trans. Intell. Transp. Syst.* 15, 4 (2014), 1451–1465. DOI:http://dx.doi.org/10.1109/TITS.2014.2298892

I. Chen, F. Bao, and J. Guo. 2015. Trust-based Service Management for Social Internet of Things Systems. *IEEE Trans. on Dependable and Secure Computing (TDSC)* PP, 99 (2015), 1–1.

Y. Chen, J. Zhou, and M. Guo. 2015. A context-aware search system for Internet of Things based on hierarchical context model. *Telecommunication Systems* (2015). DOI:http://dx.doi.org/10.1007/s11235-015-9984-x

Y. Chon, N. Lane, Y. Kim, F. Zhao, and H. Cha. 2013. Understanding the Coverage and Scalability of Place-centric CrowdSensing. In *UbiComp*.

Y. Chon, N. D. Lane, F. Li, H. Cha, and F. Zhao. 2012. Automatically characterizing places with opportunistic crowdsensing using smartphones. In *Proc. of ACM UbiComp*. New York, NY, USA, 481–490.

C. Cornelius, A. Kapadia, D. Kotz, D. Peebles, M. Shin, and N. Triandopoulos. 2008. Anonysense: Privacy-aware people-centric sensing. In *Proc. of ACM MobiSys*. New York, NY, 211–224.

I. Cvijik, Bogdan Ivan, Cristina Kada, and Y. Te. 2015. Towards a Crowdsourcing Approach for Crime Prevention. In *Proc.of ACM UbiComp / ISWC Adjunct*.

T. Dao, A. K. Roy-Chowdhury, H. V. Madhyastha, S. V. Krishnamurthy, and T. L. Porta. 2014. Managing Redundant Content in Bandwidth Constrained Wireless Networks. In *Proc. of CoNEXT*. 349–361.

D. Dimov. 2014. Crowdsensing: State of the Art and Privacy Aspects. (2014). http://resources.infosecinstitute.com/crowdsensing-state-art-privacy-aspects/.

S. Distefano, F. Longo, and M. Scarpa. 2015. QoS Assessment of Mobile Crowdsensing Services. *Journal of Grid Computing* (2015). DOI:http://dx.doi.org/10.1007/s10723-015-9338-7

A. Dohr, R. Modre-Opsrian, M. Drobics, D. Hayn, and G. Schreier. 2010. The internet of things for ambient assisted living. In *ITNG*.

Y. Dong, A. Hoover, J. Scisco, and E. Muth. 2012. A New Method for Measuring Meal Intake in Humans via Automated Wrist Motion Tracking. *Appl Psychophysiol Biofeedback* 37, 3 (2012), 205–215.

J. Eriksson, L. Girod, B. Hull, R. Newton, S. Madden, and H. Balakrishnan. 2008. The pothole patrol: using a mobile sensor network for road surface monitoring. In *Proc. of ACM MobiSys*. Breckenridge.

S. Forsström and T. Kanter. 2014. Enabling ubiquitous sensor-assisted applications on the internet-of-things. *Pers. Ubiquit. Comput.* 18, 4 (2014), 977–986. DOI:http://dx.doi.org/10.1007/s00779-013-0712-9

R.K. Ganti, F. Ye, and H. Lei. 2011. Mobile Crowdsensing: Current state and future challenges. *IEEE Comm. Mag.* 49, 11 (November 2011), 32–39.

A. Garbett, J. Wardman, B. Kirman, C. Linehan, and S. Lawso. 2015. Anti-social media: communicating risk through open data, crime maps and locative me. In *Proc. of HCI*.

A. Garyfalos and K. Almeroth. 2008. Coupons: A Multilevel Incentive Scheme for Information Dissemination in Mobile Networks. *TMC* 7, 6 (2008), 792–804.

R. Girau, M. Nitti, and L. Atzori. 2013. Implementation of an Experimental Platform for the Social Internet of Things. In *IMIS*.

M. Gorlatova, J. Sarik, G. Grebla, M. Cong, I. Kymissis, and G. Zussman. 2014. Movers and Shakers: Kinetic Energy Harvesting for the Internet of Things. In *Proc. of SIGMETRICS*. Austin, 407–419.

P. Guillemin and P. Friess. 2009. Internet of things strategic research roadmap. The Cluster of European Research Projects, Tech. Rep. DOI:http://dx.doi.org/pdf/IoT_Cluster_Strategic_Research_Agenda_2009.pdf

B. Guo, Z. Wang, Z. Yu, Y. Wang, N. Y. Yen, R. Huang, and X. Zhou. 2015. Mobile Crowd Sensing and Computing: The Review of an Emerging Human-Powered Sensing Paradigm. *ACM CSUR* 48, 1 (Aug. 2015), 1–31.

S. Hasan and E. Curry. 2014. Approximate semantic matching of events for the Internet of Things. *ACM TOIT* 14, 1 (2014), Article 2. DOI:http://dx.doi.org/10.1145/2633684

J. Howe. 2006. The Rise of Crowdsourcing. *Wired Magazine* 14, 06 (2006), 1–5. DOI:http://dx.doi.org/10.1086/599595

Y. Hua, W. He, X. Liu, and D. Feng. 2015. SmartEye: Real-time and Efficient Cloud Image Sharing for Disaster Environments. In *INFOCOM*.

E. Huang, H. Zhang, D. C. Parkes, K. Z. Gajos, and Y. Chen. 2010. Toward automatic task design: a progress report. In *Proc. of SIGKDD*.

B. Hull, V. Bychkovsky, Y. Zhang, K. Chen, M. Goraczko, A. Miu, E. Shih, H. Balakrishnan, and S. Madden. 2006. CarTel: A Distributed Mobile Sensor Computing System. In *Proc. of ACM SenSys*. Boulder.

International Telecommunication Union. 2005. Itu internet reports 2005: The internet of things. International Telecommunication Union, Workshop Report.

S. Jami, A. Basalamah, A. Lbath, and M. Youssef. 2015. Hybrid Participatory Sensing for Analyzing Group Dynamics in the Largest Annual Religious Gathering. In *Proc. of ACM UbiComp*. Osaka, 547–558.

A. Kamra, V. Misra, J. Feldman, and D. Rubenstein. 2006. Growth Codes: Maximizing Sensor Network Data Persistence. In *Proc. of SIGCOMM*.

M. Kay, E. Choe, J. Shepherd, B. Greenstein, N. Watson, S. Consolvo, and J. Kientz. 2012. Lullaby: a Capture & Access System for Understanding the Sleep Environment. In *Proc. of UbiComp*.

A. Kazmi, M. O'grady, D. Delaney, A. Ruzzelli, and G. O'hare. 2014. A Review of Wireless-Sensor-Network-Enabled Building Energy Management Systems. *ACM Trans. Sen. Netw.* 10, 4 (2014), 66:1–66:43. DOI:http://dx.doi.org/10.1145/2532644

H. Kirak, L. David, R. Umakishore, O. Beate, and K. Boris. 2013. Opportunistic spatio-temporal event processing for mobile situation awareness. In *Proc. of ACM DEBS*. Arlington, 195–206.

C. Körner and M. Strohmaier. 2010. A call for social tagging datasets. SIGWEB Newsl., 2:1C2:6.

Ourania Kounadi, Kate Bowers, and Michael Leitner. 2014. Crime Mapping On-line: Public Perception of Privacy Issues. *Europ J Crim Pol R* 21, 1 (2014), 1–24.

I. Krontiris and T. Dimitriou. 2013. Privacy-respecting discovery of data providers in crowd-sensing applications. In *Proc. of DCoSS*.

J. Krumm. 2009. A survey of computational location privacy. *Personal and Ubiquitous Computing* 13, 6 (2009), 391–399.

S. Kumar, S. Gil, D. Katabi, and D. Rus. 2014. Accurate Indoor Localization With Zero Start-up Cost. In *Proc. of MOBICOM*.

A. Kumbhare, Y. Simmhan, and V. K. Prasanna. 2013. Exploiting application dynamism and cloud elasticity for continuous dataflows. In *SC*.

N. D. Lane, Y. Chon, L. Zhou, Y. Zhang, F. Li, D. Kim, G. Ding, F. Zhao, and H. Cha. 2013. Piggyback CrowdSensing (PCS): Energy Efficient Crowdsourcing of Mobile Sensor Data by Exploiting Smartphone App Opportunities. In *Proc. of SenSys*. Rome.

E. Law and L. von Ahn. 2009. Input-agreement: A new mechanism for data collection using human computation games. In *Proc. of CHI*.

V. Lenders, E. Koukoumidis, P. Zhang, and M. Martonosi. 2008. Location-based trust for mobile user-generated content: Applications, challenges, and implementations. In *Proc. of HotMobile*. 60–64.

J. Li, X. Chen, M. Li, P. P. C. Lee, and W. Lou. 2014. Secure Deduplication with Efficient and Reliable Convergent Key Management. *TPDS* 25, 6 (2014), 1615–1625. DOI:http://dx.doi.org/10.1109/TPDS.2013.284

J. Li, Y. Zhu, J. Yu, Q. Zhang, and L. Ni. 2015. Towards Redundancy-Aware Data Utility Maximization in Crowdsourced Sensing with Smartphones. In *Proc. of ICDCS*.

A. Liu and D. Salvucci. 2001. Modeling and Prediction of Human Driver Behavior. In *Proc. of HCI*.

J. Liu, H. Shen, and L. Yu. 2015. Question Quality Analysis and Prediction in Community Question Answering Services with Coupled Mutual Reinforcement. *IEEE Trans. on Services Computing (TSC)* PP, 99 (2015), 1–14.

Jinwei Liu, Haiying Shen, and Xiang Zhang. 2016. A survey of mobile crowdsensing techniques: A critical component for the internet of things. In *Computer Communication and Networks (ICCCN), 2016 25th International Conference on*. IEEE, 1–6.

J. Liu, L. Yu, H. Shen, Y. He, and J. Hallstrom. 2015. Characterizing Data Deliverability of Greedy Routing in Wireless Sensor Networks. In *Proc. of SECON*. Seattle.

L. Liu, A. Biderman, and C. Ratti. 2009. Urban Mobility Landscape: Real Time Monitoring of Urban Mobility Patterns. In *Proc. of CUPUM*.

J. Lu, T. Sookoor, V. Srinivasan, G. Gao, B. Holben, J. Stankovic, E. Field, and K. Whitehouse. 2010. The Smart Thermostat: Using Occupancy Sensors to Save Energy in Homes. In *Proc. of SenSys*.

N. Maalel, E. Natalizio, A. Bouabdallah, P. Roux, and M. Kellil. 2013. Reliability for emergency applications in Internet of Things. In *Proc. of DCOSS*.

N. Maisonneuve, M. Stevens, M. Niessen, and L. Steels. 2009. Noisetube: Measuring and mapping noise pollution with mobile phone. In *Information Technologies in Environmental Engineering*.

N. Maisonneuve, M. Stevens, and B. Ochab. 2010. Participatory noise pollution monitoring using mobile phones. *Information Polity* 15 (2010), 51–71. DOI:http://dx.doi.org/10.3233/IP-2010-0200

E. Miluzzo, N. D. Lane, K. Fodor, R. Peterson, H. Lu, M. Musolesi, S. B. Eisenman, X. Zheng, and A. T. Campbell. 2008. Sensing Meets Mobile Social Networks: The Design, Implementation and Evaluation of the CenceMe Application. In *Proc. of SenSys*. Raleigh.

P. Mohan, V. Padmanabhan, and R. Ramjee. 2008. Nericell: rich monitoring of road and traffic conditions using mobile smartphones. In *SenSys*.

M. Mun, S. Reddy, K. Shilton, N. Yau, J. Burke, D. Estrin, M. Hansen, E. Howard, R. West, and P. Boda. 2009. PEIR, the personal environmental impact report, as a platform for participatory sensing systems research. In *MobiSys*.

M. Nitti, L. Atzori, and I.P. Cvijikj. 2014. Network navigability in the social Internet of Things. In *WF-IoT*.

Michele Nitti, Luigi Atzori, and Irena Pletikosa Cvijikj. 2015. Friendship selection in the Social Internet of Things: challenges and possible strategies. *IEEE Internet Things J.* 2, 3 (June 2015), 240–247.

M. Nitti, R. Girau, and L. Atzori. 2014. Trustworthiness Management in the Social Internet of Things. *TKDE* 26, 5 (May 2014), 1253–1266.

B. Pan, Y. Zheng, D. Wilkie, and C. Shahabi. 2013. Crowd Sensing of Traffic Anomalies based on Human Mobility and Social Media. In *Proc. of SIGSPATIAL GIS*.

D. Peng, F. Wu, and G. Chen. 2015. Pay as How Well You Do: A Quality Based Incentive Mechanism for Crowdsensing. In *Proc. of MOBIHOC*.

C. Perera, P. Jayaraman, A. Zaslavsky, P. Christen, and D. Georgakopoulos. 2013. Dynamic configuration of sensors using mobile sensor hub in internet of things paradigm. In *Proc. of ISSNIP*. Melbourne.

S. Phithakkitnukoon and P. Oliver. 2011. Sensing urban social geography using online social networking data. In *Proc. of ICWSM*.

S. Pietschmann, A. Mitschick, R. Winkler, and K. Meissner. 2008. Croco: Ontology-based, cross-application context management. In *SMAP*. http://dx.doi.org/10.1109/SMAP.2008.10

P. Quinton. 2011. The impact of information about crime and policing on public perceptions: The results of a randomised controlled trial. *London: NPIA* (2011).

M.-R. Ra, B. Liu, T. L. Porta, and R. Govindan. 2012. Medusa: A Programming Framework for Crowd-Sensing Applications. In *MobiSys*.

M. Rabbi, S. Ali, T. Choudhury, and E. Berke. 2011. Passive and In-situ Assessment of Mental and Physical Well-being using Mobile Sensors. In *Proc. of UbiComp*.

A. Rai, K. Chintalapudi, V. Padmanabhan, and R. Sen. 2012. Zee: Zero-Effort Crowdsourcing for Indoor Localization. In *Proc. of MOBICOM*.

R. Rana, C. Chou, S. Kanhere, N. Bulusu, and W. Hu. 2010. Ear-phone: An end-to-end participatory urban noise mapping system. In *ISPN*.

J. Roemer, M. Groman, Z. Yang, Y. Wang, C. C. Tan, and N. Mi. 2014. Improving Virtual Machine Migration via Deduplication. In *MASS*.

J. Rogstadius, M. Vukovic, C. Teixeira, V. Kostakos, E. Karapanos, and J. Laredo. 2013. Crowdsourced social media curation for disaster awareness. *IBM J. of Research and Development* 57, 5 (2013), 4:1–4:13.

P. Rothenpieler, B. Altakrouri, O. Kleine, and L. Ruge. 2014. Distributed Crowd-Sensing Infrastructure for Personalized Dynamic IoT spaces. In *Proc. of URB-IOT*. Rome.

T. Sakaki, M. Okazaki, and Y. Matsuo. 2010. Earthquake shakes Twitter users: real-time event detection by social sensors. In *Proc. of WWW*.

S. Saroiu and A. Wolman. 2009. Enabling new mobile applications with location proofs. In *Proc. of HotMobile*. 1–6.

H. Shen, Z. Li, J. Liu, and J. E. Grant. 2015a. Knowledge Sharing in the Online Social Network of Yahoo! Answers and Its Implications. *IEEE Transactions on Computers (TC)* 64, 6 (June 2015), 1715C1728.

H. Shen, J. Liu, K. Chen, J. Liu, and S. Moyer. 2015b. SCPS: A Social-Aware Distributed Cyber-Physical Human-Centric Search Engine. *IEEE Transactions on Computers (TC)* 64 (2015), 518–532.

W. Sherchan, P. P. Jayaraman, S. Krishnaswamy, A. Zaslavsky, S. Loke, and A. Sinha. 2012. Using on-the-move mining for mobile crowdsensing. In *Proc. of MDM*. 115–124.

A. Sheth. 2009. Citizen sensing, social signals, and enriching human experience. *IEEE Internet Computing* 13, 4 (2009), 87–92. DOI:http://dx.doi.org/10.1109/MIC.2009.77

J. Stankovic. 2014. Research directions for the Internet of Things. *IEEE Internet Things J.* 1, 1 (2014), 3–9.

S. A. Stansfeld. 2003. Noise pollution: non-auditory effects on health. *British Medical Bulletin* 68, 1 (2003), 243–257. DOI:http://dx.doi.org/10.1093/bmb/ldg033

H. Susmita and S. Anjali. 2012. Identity management framework for cloud based internet of things. In *Proc. of SecurIT*. Kollam.

K. Tangwongsan, H. Pucha, D. G. Andersen, and M. Kaminsky. 2010. Efficient Similarity Estimation for Systems Exploiting Data Redundancy. In *Proc. of INFOCOM*.

F. A. Teixeira, G. V. Machado, F. M. Q. Pereira, and L. B. Oliveira. 2015. SIoT: Securing the Internet of Things through Distributed System Analysis. In *Proc. of IPSN*. 310–321.

I. Vasilescu, K. Kotay, D. Rus, M. Dunbabin, and P. Corke. 2005. Data Collection, Storage, and Retrieval with an Underwater Sensor Network. In *Proc. of SenSys*.

Idalides J Vergara-Laurens, Luis G Jaimes, and Miguel A Labrador. 2016. Privacy-preserving mechanisms for crowdsensing: Survey and research challenges. *IEEE Internet of Things Journal* (2016).

L. von Ahn and L. Dabbish. 2004. Labeling Images with a Computer Game. In *Proc. of ACM CHI*.

Leye Wang, Daqing Zhang, Yasha Wang, Chao Chen, Xiao Han, and Abdallah M'hamed. 2016. Sparse mobile crowdsensing: challenges and opportunities. *IEEE Communications Magazine* 54, 7 (2016), 161–167.

Y. Wang, W. Hu, Y. Wu, and G. Cao. 2014. SmartPhoto: A Resource-Aware Crowdsourcing Approach for Image Sensing with Smartphones. In *Proc. of MOBIHOC*. Philadelphia, 113–122.

Yang Xiang Wazir Zada Khan, Mohammed Y Aalsalem, and Quratulain Arshad. 2013. Mobile phone sensing systems: A survey. *IEEE Communications Surveys Tutorials* 15, 1 (2013), 402–427.

U. Weinsberg, Q. Li, N. Taft, A. Balachandran, V. Sekar, G. Iannaccone, and S. Seshan. 2012. CARE: Content Aware Redundancy Elimination for Challenged Networks. In *Proc. of Hotnets*. Seattle.

A. Wesolowski, N. Eagle, A. Tatem, D. Smith, A. Noor, R. Snow, and C. Buckee. 2012. Quantifying the impact of human mobility on malaria. *Science* 338, 6104 (2012), 267–270.

A. Whitmore, A. Agarwal, and L. Xu. 2014. The Internet of Things—A survey of topics and trends. *Information Systems Frontiers* (Mar. 2014).

W. Willett, S. Ginosar, A. Steinitz, B. Hartmann, and M. Agrawala. 2013. Identifying Redundancy and Exposing Provenance in Crowdsourced Data Analysis. *IEEE Trans. Vis. Comput. Graphics* 19, 12 (2013), 2198–2206.

L. Xu, X. Hao, N. D. Lane, X. Liu, and T. Moscibroda. 2015a. Cost-Aware Compressive Sensing for Networked Sensing Systems. In *IPSN*.

L. Xu, X. Hao, N. D. Lane, X. Liu, and T. Moscibroda. 2015b. More with Less: Lowering User Burden in Mobile Crowdsourcing through Compressive Sensing. In *Proc. of ACM UbiComp*. Osaka, 659–670.

T. Yan, V. Kumar, and D. Ganesan. 2010. CrowdSearch: Exploiting Crowds for Accurate Real-time Image Search on Mobile Phones. In *MobiSys*.

L. Yao, Q. Sheng, A. Ngu, and B. Gao. 2014. Keeping You in the Loop : Enabling Web-based Things Management in the Internet of Things. In *Proc. of ACM CIKM*.

M. Ye, P. Yin, W. Lee, and D. Lee. 2011. Exploiting geographical influence for collaborative point-of-interest recommendation. In *SIGIR*.

N. Yuan, Y. Zheng, L. Zhang, and X. Xie. 2013. T-finder: A recommender system for finding passengers and vacant taxis. *TKDE* 25, 10 (2013), 2390–2403. DOI:http://dx.doi.org/10.1109/TKDE.2012.153

M.-C. Yuen, L.-J. Chen, and I. King. 2009. A survey of human computation systems. In *Proc. of CSE*. 72–728.

A. Zanella, N. Bui, A. Castellani, L. Vangelista, and M. Zorzi. 2014. Internet of Things for Smart Cities. *IEEE Internet Things J.* 1, 1 (2014), 22–32. DOI:http://dx.doi.org/10.1109/JIOT.2014.2306328

Z. Zhan, X. Liu, Y. Gong, and J. Zhang. 2015. Cloud Computing Resource Scheduling and a Survey of Its Evolutionary Approaches. *Comput. Surveys* 47, 4 (2015).

D. Zhang, H. Xiong, L. Wang, and G. Chen. 2014. CrowdRecruiter: Selecting Participants for Piggyback Crowdsensing under Probabilistic Coverage Constraint. In *ACM UbiComp*. Seattle, 703–714.

W. Zhang, J. Wang, and W. Feng. 2013. Combining Latent Factor Model with Location Features for Event-based Group Recommendation. In *KDD*.

Xinglin Zhang, Zheng Yang, Wei Sun, Yunhao Liu, Shaohua Tang, Kai Xing, and Xufei Mao. 2016. Incentives for mobile crowd sensing: A survey. *IEEE Communications Surveys & Tutorials* 18, 1 (2016), 54–67.

Y. Zheng, F. Liu, and H. Hsieh. 2013. U-Air: when urban air quality inference meets big data. In *Proc. of KDD*. 1436–1444.

Y. Zheng and X. Xie. 2011. Learning travel recommendations from user-generated GPS traces. *ACM TIST* 2, 1 (2011), 1–29. DOI:http://dx.doi.org/10.1145/1889681.1889683

P. Zhou, Y. Zheng, and M. Li. 2014. How long to wait? Predicting bus arrival time with mobile phone based participatory sensing. *IEEE Trans. Mobile Comput.* 13, 6 (2014), 1228–1241. DOI:http://dx.doi.org/10.1109/TMC.2013.136

Y. Zhu and D. Shasha. 2002. StatStream: Statistical Monitoring of Thousands of Data Streams in Real Time. In *Proc. of VLDB*. 358–369.

Z. Zhu, L. Zhang, and R. Wakikawa. 2011. Supporting mobility for Internet cars. *IEEE Commun. Mag.* 49, 5 (2011), 180–186.

D. Zordan, B. Martinez, I. Vilajosana, and M. Rossi. 2014. On the Performance of Lossy Compression Schemes for Energy Constrained Sensor Networking. *ACM Transactions on Sensor Networks* 11, 1 (2014), 1–34. DOI:http://dx.doi.org/10.1145/2629660