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

JINWEI LIU, University of Central Florida
HAIYING SHEN, University of Virginia
HUSNU S. NARMAN, Marshall University
WINGYAN CHUNG, University of Central Florida
ZONGFANG LIN, Huawei US R&D Center

Mobile crowdsensing serves as a critical building block for 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 redundant data, data size and load can be significantly reduced, thereby reducing resource cost and facilitating the timely delivery of unique, probably critical information and enhancing QoS. This article presents a survey of existing works on mobile crowdsensing strategies with an emphasis on reducing 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 future research directions for 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 prior works, but also to discuss their applicability toward the IoT and provide guidance on future research directions for mobile crowdsensing.

18

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

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

ACM Reference format:
Jinwei Liu, Haiying Shen, Husnu S. Narman, Wingyan Chung, and Zongfang Lin. 2018. A Survey of Mobile Crowdsensing Techniques: A Critical Component for The Internet of Things. *ACM Trans. Cyber-Phys. Syst.* 2, 3, Article 18 (June 2018), 26 pages.
https://doi.org/10.1145/3185504

This research was supported in part by U.S. NSF grants OAC-1724845, ACI-1719397 and CNS-1733596, and Microsoft Research Faculty Fellowship 8300751.
Authors' addresses: J. Liu, Institute for Simulation and Training, University of Central Florida, 3100 Technology Pkwy, Orlando, FL 32826; email: jliu@ist.ucf.edu; H. Shen, Department of Computer Science, University of Virginia, 85 Engineer's Way, P.O. Box 400740, Charlottesville, VA 22904; email: hs6ms@virginia.edu; H. S. Narman, Department of Computer Science, Marshall University, 1676 3rd Ave, Arthur Weisberg Family Applied Engineering Complex, WAEC 3107, Huntington, WV 25755; email: narman@marshall.edu; W. Chung, Institute for Simulation and Training, University of Central Florida, 3100 Technology Pkwy, Orlando, FL 32826; email: wchung@ucf.edu; Z. Lin, Huawei US R&D Center, 2330 Central Expy, Santa Clara, CA 95050; email: Zongfang.Lin@huawei.com.

https://doi.org/10.1145/3185504

1 INTRODUCTION

In recent years, an increasing number of sensing devices and wireless networks emerged in our living environments to create the Internet of Things (IoT) integrating 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 Gartner, IoT will have the fast-growing, largest market potential 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 by which 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 with things, things with people, and people with people. Usually, 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 a 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 devices with built-in sensors. It is well-known that in such a large-scale system, sensing devices continuously generate huge amounts of data (raw sensor data), which consumes large amounts of resources (e.g., bandwidth, energy, etc.) (Liu et al. 2015). However, the sensing devices have limited resources. Due to these limited resources, the quality of the data collected can be sacrificed in bandwidth-constrained networks because of the heavy traffic load and high power consumptions (Hua et al. 2015; Dao et al. 2014). Therefore, 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 a disaster area play an important role in disaster relief, but the images collected may not be able to be uploaded in time due to limited bandwidth, which can incur a huge cost. Therefore, resource limitations always hinder the necessary participation and wide-scale adaptation of the targeting applications (Xu et al. 2015a).

Although MCS is a newly 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 an MCS approach in the business context requires a guarantee for Quality-of-Service (QoS). Hence, QoS is one of the most important emerging issues. QoS-driven policies are needed to deal with an application's nonfunctional issues to guarantee QoS.

In this article, we review MCS techniques and challenges. Different aspects of MCS are also reviewed. Ganti et al. (2011) introduced MCS, briefly overviewed existing MCS applications with their unique characteristics, and discussed several research challenges with possible solutions.

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

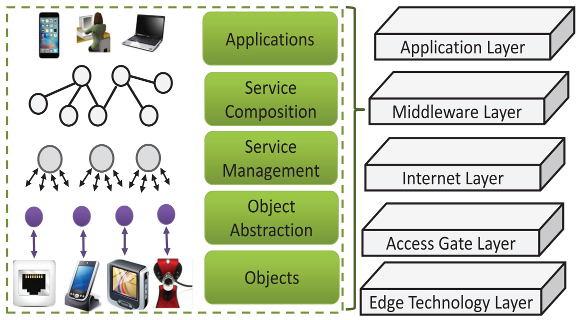


Fig. 1. Architecture of the Internet of Things (IoT) (right) and its five-layer middleware (left).

Vergara-Laurens et al. (2017) surveyed privacy issues and privacy-preserving mechanisms for crowdsensing systems. Zhang et al. reviewed the literature for incentives that encourage users to participate in MCS applications under entertainment, service, and money categories (Zhang et al. 2016). Wang et al. (2016) introduced sparse MCS, discussed sparse MCS challenges. and developed a framework with potential solutions to thosee challenges. Khan et al. (2013) comprehensively explained mobile sensing systems according to personal, social, and public sensings. By contrast, our focus is to discuss the resource limitations and QoS (e.g., data quality) issues and solutions in MCS. A better understanding of resource management and QoS estimation in MCS can help us design a cost-effective crowdsensing system that can reduce costs by fully utilizing resources and improve QoS for users, thus summing up the significance of our survey.

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

The remainder of this article is organized as follows. Section 2 introduces the concepts of IoT and MCS. Section 3 describes the strategies of MCS. Section 4 describes crowdsensing strategies for different application domains. Section 5 describes the challenges of MCS and future research directions. Section 7 concludes this article 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 from both academia and business communities, spurred by the potential capabilities of IoT (Carnot Institutes 2011; Atzori et al. 2010). 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 first 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 a broader version of IoT:

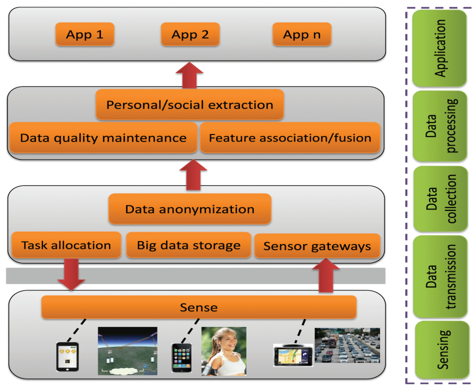


Fig. 2. Architecture of MCS system. Raw sensor data are collected via different mobile sensing devices (GPS, etc.) in the sensing layer. In order to preserve privacy, the data will be sent to the data collection layer and be modified using methods such as data anonymization.

• "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 newly emerging paradigm, and it is a very broad vision. The research into the IoT is still under way, but its potentialities 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): the transportation and logistics domain, the healthcare domain, the smart environment (e.g., home, plant) domain, and the personal and social domain.

2.2 Mobile Crowdsensing

MCS uses 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 tasks to reliable participants (those participants who are expected to complete the assigned sensing tasks), effectively gather the required data from participants, process and manage the data according to the purpose, and dynamically improve itself for the next crowdsending events by self-learning mechanism (Bellavista et al. 2015). MCS usually requires a large quantity of participants to sense the environment using 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, MCS applications can be divided into three categories: (a) environmental applications, (b) infrastructure applications, and (c) social applications (Ganti et al. 2011). Figure 2 shows the architecture of an 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 are manually collected by a user controlling the sensing device, such as a smartphone 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, as 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 other work (Pietschmann et al. 2008), context-aware data sensing can be accomplished using two methods: push and pull:

—**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. Periodic or instant pushing is able to facilitate a publish and subscribe model.
—**Pull:** The software component in charge of acquiring data (e.g., sensor data) from sensing devices (e.g., sensors) makes a request, like a query, of the hardware of the sensing devices, periodically or instantly, to acquire data.

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 compression on the raw data (e.g., images) collected by sensing devices. The deduplication is conducted with the constraint that the significance of the data is maintained. Deduplication of crowdsensing data makes full use of limited sensor storage and reduces the bandwidth consumption caused by data transfer. In deduplication, data are 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 bandwidth consumption will be reduced.

As the size of the data to be processed increases rapidly, deduplication technology is improving rapidly to meet the requirements of global industry and research. From the perspective of the phase at which deduplication occurs, data deduplication approaches can be categorized as real-time deduplication and post-process deduplication. At the layer where deduplication happens, either local or server deduplication can occur.

Real-time deduplication refers to hashing and compressing the data when it is acquired. Duplicated data acquired by the sensing device will be detected based on the stored data chunk. If the new data are judged duplicates, they will not be stored in the sensing device, nor be uploaded to the data center. The advantage of this strategy is that it lowers 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, 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 are stored first and then processed for deduplication. The tradeoff 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 other work (Dimov 2014) and taking into account the degree of stability of crowdsensing platform mobility and context, we can categorize crowdsensing applications in Table 1.

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

Table 1. Summary of Mobile Crowdsensing Applications

Degree of human participation	**Example**
Completely controlled by user	Application monitoring 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

3 EXISTING MOBILE CROWDSENSING STRATEGIES

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

Previous works demonstrate that significant redundancy exists in the content of 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 redundant data, which on the one hand can reduce 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 traffic load. One of the key challenges here, however, is detecting similar data. Another key challenge is how to eliminate similar data while ensuring high QoS (e.g., without compromising the quality of the data or timely delivery of valuable data). To handle the problem caused by limited available resources, many methods have been proposed. here, we present a review of previously proposed strategies.

3.1 Different Mobile Crowdsensing Strategies for Reducing Resource Cost

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 data only from a minimal set of sensors. Hua et al. (2015) presented a near-real-time and cost-effective solution for a cloud-assisted disaster environment. SmartEye (Hua et al. 2015) leverages two main methods, semantic hashing and space-efficient filters, to aggregate flows with similar features and provide communication services for the aggregated flow.

In bandwidth constrained network, Dao et al. (2014) introduced a method focusing on recognizing similar content in images and videos by leveraging the metadata uploaded first to distinguish data similarity. 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 data redundancy reduction by detecting similarities among data, such as images or videos. For example, Weinsberg et al. (2012) proposed a framework called CARE, which eliminates the redundancy of images to 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 a disaster happens, and makes use of a peer-to-peer strategy to eliminate redundant data. In mobile platform real-time crowdsensing, Sherchan et al. (Sherchan et al. 2012) designed a system for collecting data via instantaneous data analysis and processing. To reduce bandwidth consumption and save energy for mobile devices, their CAROMM is able to acquire various streamed data by mobile devices and process them based on the context (e.g., the location and time mark on photos), thus contributing to relevant data retrieval from the dataset.

Roemer et al. (Roemer et al. 2014) adopted a data deduplication strategy to reduce the storage and bandwidth consumption for 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, Lane et al. (Lane et al. 2013) proposed an energy effective crowdsensing strategy that takes advantage of an opportunistic application run by 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 to lower overhead (e.g., taking into account location information). Data (i.e., images) from exactly the same location tend to contain the same information. In addition, their analysis on the application specifics can also contribute to overall cost-reduction. 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 more than 40 participants, and an energy allocation algorithm with accessible IoT node solution design has been developed and evaluated based on the collected measurements.

3.2 Different Crowdsensing Strategies for Achieving Good QoS

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

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

Yan et al. (2010) proposed CrowdSearch for searching images using mobile phones. CrowdSearch integrates the strategy of automated image search into real-time human validation. 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 tradeoff 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 determines when and how to validate these results.

Due to limited resources, it is a challenge to transfer a huge amount of crowdsensed data. To address this challenge, 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 information on the device's orientation, location, and all related parameters of the built-in camera. With metadata, it can be inferred where and how the photo was taken. Also, SmartPhoto only transmits the most useful photos. They also studied three optimization problems on the tradeoffs between photo utility and resource constraints. Moreover, they designed efficient algorithms with theoretical proofs of their performance. Finally, 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. (2015a) studied compressive sensing under scenarios in which different samples have different costs. This work tries to balance the minimization of total sample cost and recovery accuracy and designs Cost-aware Compressive Sensing (CACS) for incorporating the samples' diversity on cost into the compressive sensing framework. CACS has been applied to networked sensing systems.

To maximize aggregate data utility, Li et al. (2015) studied how to 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 bid determination and a critical payment scheme.

Table 2 lists different techniques in MCS with an emphasis on redundancy reduction. Table 3 gives a comparison of common MCS strategies in recent years and an example work on each strategy. We cite the most representative case for each strategy. As we can see from the table, most of the data sensed by mobile devices are images that contains rich information and consume a small amount of storage due to data deduplication.

3.3 Dataset

In the preceding section, we discussed different techniques for handling the resource limitation issue to achieve low cost while achieving good QoS. Here, 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 a 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. (Chen 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 are available now.[5] CiteULike,[6] developed by Oversity Ltd., is a free website that allows users to save and share citations to academic papers and is used to help academics keep a record of articles they are reading. CiteULike encourages users to share their libraries on the website so that others can benefit from resource sharing to discover articles useful to them. To better facilitate research, Körner and Strohmaier (2010) released a list of social tagging datasets.[7]

4 CROWDSENSING STRATEGIES FOR DIFFERENT APPLICATION DOMAINS

In addition to the preceding strategies focusing on the data processing phase, we can also divide the categories of these strategies based on their application domains (Guo et al. 2015). Accordingly, common domains for crowdsensing are described next.

[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/.

Table 2. A List of Different Mobile Crowdsensing Strategies

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

4.1 Natural Environment Monitoring

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

As for environment protection-oriented crowdsensing strategies, academy and industry both currently are likely to take advantage of a vast number of smart devices owned by the public to do research or make a profit with a relatively low investment. Mun et al. (Mun et al. 2009) made use of a 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 processes the data in the data center. Two features of the

Table 3. Comparison of Different Types of Techniques in Mobile Crowdsensing

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	Requires 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 left 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)

strategy are context-triggered feedback and data visualization, which make the software interact effectively with the smart device owner to perform a specific task. For example, when a person arrives at a particular place, the context-aware device will automatically notify the person to take a photo of a targeted item or scene to upload to the cloud for further analysis. In terms of pollution-targeted crowdsensing, relevant strategies are designed to focus on noise and air pollution aspects.

Crowdsensing plays an important role in measuring and reducing noise pollution. Maisonneuve et al. (Maisonneuve et al. 2010) presented a participatory noise pollution detection method called NoiseTube based on a mobile phone platform, with the aim of acquiring first-hand noise data suffered by individuals. NoiseTube records the magnitude of the noise combined with position and time information for further statistical analysis. For similar purpose, Rana et al. (2010) established a noise map to enhance efficienct noise monitoring in cities. Crowdsensing techniques were adopted by them to avoid the high cost of building a noise map using traditional infrastructure-based methods. They implemented the method on Nokia N95 and HP iPAQ platforms and addressed the problem of ensuring 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 human effects on the environment with the aid of crowdsensing research, is conducted by Mun et al. (2009). The sensing system consists of mobile handset GPS receivers for collecting position data, server data classification processors for detecting different modes of transportation, and a database for looking up 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. (2013) proposed a holistic method combining data sources from crowdsensing applications monitor air status and a 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 basis on an Artificial Neural Network (ANN) taking into account the spatial information of an area, and the other a linear-chain Conditional Random Field (CRF) that considers real-time dependency among factors such as temperature, humidity, and the like. The paper sheds light on the use of artificial intelligence to solve crowdsensing problems, which we believe is a trend for the near future.

4.2 Traffic Information Collection and Management

Crowdsensing strategies play an important role in collecting traffic information and helping both the public and government with related decision-making. Here, we introduce three aspects.

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

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

4.3 Urban Dynamics Sensing

Understanding urban dynamics is critical for urban development and quality of life improvements and presents a key challenge. Urban dynamics sensing has become possible and has attracted keen interest from both industry and academic research societies. For human urban mobility/behavior patterns, some works study how to reveal human mobility and behavior patterns in urban areas. 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 incentivize mobile phone users to participate. Pan et al. (2013) addressed the problem of detecting and describing traffic anomalies using crowdsensing with two forms of data: human mobility and social media. Phithakkitnukoon et al. (2011) used 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), and more.

4.5 Social Network-Based Applications

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

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 (2009) introduced a system combining human-involved crowdsensing, Web 2.0, and mobile computing to establish a platform for recognizing an emergency, analyzing the situation, and calling for 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 online. 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 responders. Even though situation-aware crowdsensing is an effective way to detect and convey emergency information, the privacy of involved individuals should be protected. Thus, Ballesteros et al. (2013) designed iSafe, a privacy-preserving method that analyzes crowdsensing data from individuals' phones to improve the safety of the city. By their approach, snapshots taken by both the user's phone and geosocial network users are used for analysis e.g., the level of safety of a place is determined).

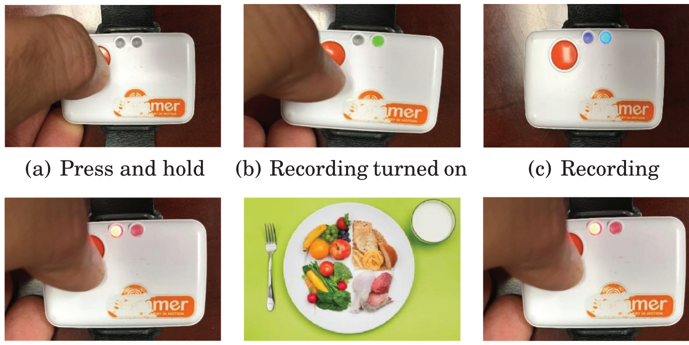
In disaster scenarios, smart wearable device-based crowdsensing plays an important role in guaranteeing communications with the outside world in a congested environment. Sakaki et al. (2010) presented an algorithm taking advantage of the real-time characteristics of Twitter to immediately detect earthquakes. The key idea is as simple as observing tweet activities: 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 methods to obtain an optimal estimation of an earthquake's center and the trace of a typhoon.

4.6 Healthcare

Health is becoming an increasingly important challenge. Wireless sensors are worn by people for heart rate and blood pressure monitoring, and they can communicate their information to users' equipment. MCS can utilize these existing data for large- scale healthcare studies. Based on the wealth of data collected from MCS systems, 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 enable quick responses to the disease outbreaks and can better control the spread of diseases. Also, 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. (2011) presented a mobile sensing system for measuring mental well-being from behavioral indicators in natural

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



(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 a watch-like device for detecting and recording an individual's eating activity.

Table 4. Classification of Applications for Crime Prevention

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)

everyday settings. To help with weight loss, other 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-electromechanical 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.

4.7 Public Safety

Public safety means detection or protection from social or natural events such as crimes and 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. (2012) presented iSafe, a privacy-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 social network activity and crime levels. Cvijikj et al. (2015) implemented a mobile application—a crowdsourcing approach—for crime prevention; it focuses on usage intention and motivations for content creation and consumption. Table 4 shows different application types for crime prevention.
- **Disaster management and relief:** Events such as the big flood in mid-Europe in 2013 and Typhoon Haiyan in the Philippines show that people become increasingly active in responding to disasters. MCS has been used for disaster management and relief. 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. (2015) proposed SmartEye, a near-real-time and cost-efficient mobile device-based crowdsourcing application, for rapid disaster relief in the cloud-assisted disaster environment. SmartEye utilizes an in-network deduplication strategy to obtain fast operation response and significant bandwidth savings so that it can efficiently support 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 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 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) (IEEE Instrumentation and Measurement Society 2007) and the Open Geospatial Consortium (OGC) Sensor Web Enablement-related standards like Sensor Markup Languages (SensorML) (Botts and Robin 2007) show future trends in research work that addresses the challenge of the 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 these limitations arise as a challenge for crowdsensing. Although more resources (e.g., computing, bandwidth) are provided for mobile phones compared to remote-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 of each other because of the multimodality sensing capabilities of sensing devices. In practical scenarios, different types of sensed data may be used for the same purpose. However, the diversity of quality and resource consumption of the sensed data poses 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 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 consume great resources. By intelligently reducing redundant data and the transfers of redundant content, the volume of data and the traffic load can be significantly reduced. Hence, it is important and necessary to eliminate redundant data, which can help reduce resource consumption (e.g., storage resource and bandwidth resource) and thereby reduce costs. A key challenge here, however, is detecting redundant data; that is, detecting "what content is similar" (Dao et al. 2014). For example, Dao's work (Dao et al. 2014) designs a framework for detecting similarity among data content and finding similar content. By restricting the transfer of similar content, this work (Dao et al. 2014) reduces resource consumption and thereby reduces cost and provides good QoS in bandwidth-constrained wireless networks.

However, another issue of 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 problems.

In addition, data derived via the crowdsensing process are often noisy and incomplete, which affects the quality of the data. Also, as redundant data is 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 volunteers. Prior literature demonstrated the role of motivation and incentives as a key factor. Due to privacy issues, many potential contributors are reluctant to carry out tasks. In many cases, they do not 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 active in carrying out small tasks. Therefore, providing effective motivation and incentives is still a challenge for MCS.

5.1.5 Privacy, Security, and Data Integrity. Sensing devices potentially collect sensitive data from individuals (Krontiris and Dimitriou 2013; Ballesteros et al. 2013; Zanella et al. 2014; Chen et al. 2016b; 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, GPS sensor readings usually record private information of individuals (e.g., the routes they take during their daily commutes and locations (Krumm 2009)). By sharing GPS sensor measurements, individuals' privacy can be compromised. Hence, it is important and necessary to preserve the security and privacy of an individual. Also, GPS records information from daily commutes shared within a larger community and can be used to learn about traffic congestion in a city (Hull et al. 2006). Thus, it is also necessary to enable crowdsensing applications so that individuals can better understand their surroundings and can ultimately benefit from their information sharing. To preserve the enormous amounts of individual private information, not only methodology efforts but also systematic studies are needed. The AnonySense architecture, proposed earlier (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 are not revealed to untrustworthy third parties. For example, malicious individuals often contribute erroneous sensor data and may intentionally pollute the sensing data for their own benefit. 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 obtained data.

The problem of ensuring data integrity from individuals' sensor data also needs to be 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 these methods prohibitive and unavailable at times. The reason behind this is that the approach relies on inputs from the installation of expensive infrastructure. Another approach for handling the data integrity problem is to sign the sensor data (e.g., typically, trusted hardware installed on mobile phones are used for this purpose); for example, a trusted platform module signs a SHA-1 digest of the sensor data. This approach is potentially problematic because 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 at 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 always available and are compatible with 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; it is even more critical than availability and it has stricter 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 services in IoT systems, reliability should be implemented in both software and hardware throughout all IoT layers.

5.2.3 Mobility. Since most of the services of IoT are expected to be delivered to mobile users and to connect 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. The future suggests a more ubiquitous and mobile Internet. As the number of smart devices increases sharply in IoT systems, mobility management becomes necessary. The Internet of Vehicles (IoV) is an emerging area of the IoT, and it needs careful attention paid to mobility issues. Some works (Zhu et al. 2011) study mobility in vehicle-to-vehicle networking, and it Zhu et al. discuss 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 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 solutions of remote network management for smart home and other control solutions.

5.2.5 Scalability. The scalability of IoT addresses 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 to support unique identification and transparency. A new network architecture is needed to overcome 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 embraces 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 and 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 emerging 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 developed AllJoyn, an open source project that provides a universal software frame and set of system services enabling interoperability.

5.3 Future Research Directions

Here, we present some future research directions for crowdsensing for IoT, and Figure 4 summarizes these trends.

5.3.1 Optimization of Multiple Factors Like Localization, Prediction, Energy Budget. The tradeoff 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. (2013) to lower the energy overhead based on context information such as position, real-world

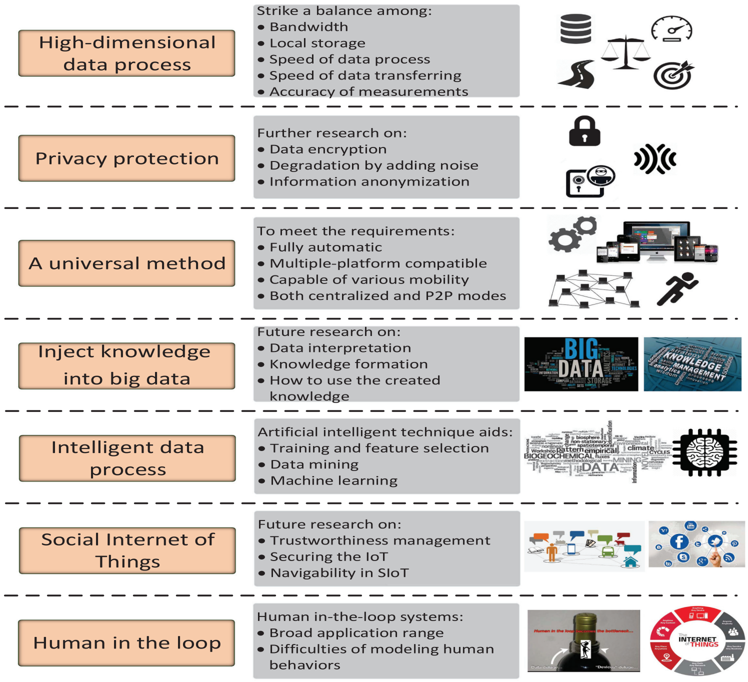


Fig. 4. Future research directions of mobile crowdsensing for IoT.

performance suffers from the inaccurate localization model. In addition, for MCS and especially for mobile device platforms, 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 context information may be increased, as in 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 the context attached (e.g., the location and time marks 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. 2015), 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 data from user's mobile devices, bears high risks of leakage of user privacy information since information like location and time must 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. (2013) conducted research on automatic data anonymization by masking particular information from the raw data sensed by local mobile devices. Also, the IoT additionally introduces unique challenges to privacy, and many of them go beyond current existing data privacy issues. This mainly stems from integrating devices into environments without users consciously using them. Moreover, many IoT scenarios involve device deployments and data collection activities with a multinational or global scope that crosses 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 platforms (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. 2013) can only be applied to those tasks that can be done without human participation and cannot be used in dynamic conditions (e.g., in a moving car). The data redundancy handling method proposed in Dao et al. (2014) is able to manage image data successfully, but can do nothing on videos although the image and video are both common information media. A universal strategy will be able to significantly reduce the cost of modifications to meet the requirements of various crowdsensing scenarios, especially applications (i.e., only one application is required to perform multiple tasks). Indeed this limitation impairs its communication efficiency in disasters. A crowdsensing strategy capable of centralized and distributed data collection is also a direction for future research, since researchers usually focus on one of these issues. CARE (Weinsberg et al. 2012) only uses a peer-to-peer structure to implement an information-aware redundant data reduction, while other authors (Dao et al. 2014) designed their redundancy elimination method in bandwidth-constrained wireless networks with the aid of infrastructure. Since in some particular scenarios, such as the disaster field, the feasibility and flexibility quality of the solution is nontrivial, we believe that a hybrid approach combining advantages from different system architectures is necessary.

5.3.4 Injecting Knowledge into Big Data. In an IoT world, a huge amount of raw data is continuously collected. It is worthwhile to develop techniques to convert the raw data into knowledge. Taking 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 need to know priori probabilities and the cost of computations. Although rule-based systems may be adopted, they are too 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 knowledge has been created, another challenge is how to better control or make good decisions in using that 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 to guarantee safety in making decisions, which is a nontrivial task.

5.3.5 Intelligent Data Processing. Current methodologies for data deduplication can be mainly categorized as real-time or post-process (Zhan et al. 2015; Kazmi et al. 2014). However, with an increasingly enlarging dataset and complex data types under limited time, bandwidth, and other resource budgets, machine learning techniques may play a non-negligible role. The data management method presented elsewhere (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 machine learning techniques, priority levels can be assigned to data automatically based on their relevance to the requirement (e.g., images of injured people with their 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 methods can improve system performance. Zheng et al. (2013) designed a semi-supervised learning approach consisting of a spatial classifier and a temporal classifier to learn the features of air quality in an entire city and then used it to classify the degree of air pollution. Two essential parts of their method are to select effective features of the air for machine learning and pollution classification and to ensure the size of the training set for the machine learning program. Machine learning-based crowdsensing approaches are also able to detect road surface problems automatically. Eriksson et al. (2008) proposed efficient methods to use taxis and on-board sensor sensing to contribute to road maintenance with the aid of machine learning. The establishment of 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. 2017). 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), and this attracts research interests. Social networks have the advantage of efficiently discovering and distributing services, and social networks are utilized by many systems, such as Yahoo! Answers, and Facebook, for sharing 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 trustworthiness is another challenge faced by the social IoT. Atzori et al. (2012) identified appropriate policies for the establishment and management of social relationships between objects in the way that the resulting social network is navigable. Nitti et al. (2014) defined the problem of trustworthiness management in the social IoT, and they presented two models: subjective and objective, for trustworthiness management starting from the solutions proposed for P2P and social networks. Table 5 summarizes the representative works on the social IoT. There is great potential and prospects for integrating social networking into the Internet of Things, and this will be an important research direction.

5.3.7 Humans in the Loop. Since many applications of IoT involve humans, humans and things will operate synergistically. Human in-the-loop systems bring opportunities to broaden a range of applications which include energy management (Lu et al. 2010), healthcare (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 humans in the loop has its advantage.

6 BIG DATA ANALYTICS AND CLOUD IN SUPPORT OF THE IOT

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

7 CONCLUSION

The IoT has attracted much attention over the past few years. Numerous sensing devices emerge in our living environments that create the IoT, integrating both cyber and physical objects. MCS

Table 5. A List of Representative Works on Social IoT (SIoT)

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 networking 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	Analyze 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	Analyze 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. 2016a)	Trust-based Service Management in SIoT	Adaptively control and manage trust

plays an important role in the IoT paradigm. Sensors continuously generate enormous amounts of data, which consumes much resources, such as storage for storing data and bandwidth for data transfer. Previous works demonstrate a significant amount of redundancy in sensor data. Thus, redundancy elimination of sensor data is important and worthwhile, and it can significantly reduce 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 article, we review mobile crowdsensing techniques and challenges. We focus on the discussion of resource limitations 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 cost by fully utilizing resources and improving the QoS for users. Here, we describe some challenges related to crowdsensing for IoT and discuss some of the trends in mobile crowdsensing for IoT. In the future, we will give an in-depth study of challenges, techniques, and solutions for addressing challenges in mobile crowdsensing for IoT, and we will also analyze production systems and provide case studies.

ACKNOWLEDGMENTS

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 ICCCN 2016 (Liu et al. 2016).

REFERENCES

IEEE Instrumentation and Measurement Society. 2007. IEEE standard for a smart transducer interface for sensors and actuators wireless communication protocols and transducer electronic data sheet (TDES) formats. *IEEE Std 1451.5-2007.* C1–236. 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 *Proceedings of the 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 (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 Communications Magazine* 52, 1 (2014), 97–105.

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. *Computer Networks* 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 *Proceedings of the LCN*.

P. Bellavista, A. Corradi, L. Foschini, and R. 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 Med* 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 *Proceedings of the CHI*.

M. Botts and A. Robin. 2007. *OpenGIS sensor model language (SensorML) implementation specification*. Technical Report. Open Geospatial Consortium Inc.

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 *Proceedings of the Urb-IoT*.

G. Burnham, J. Seo, and G. Bekey. 1974. Identification of human driver models in car following. *IEEE Transactions on Automated 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, 1 (2011), 141–151.

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

L.-J. Chen, B.-C. Wang, and W.-Y. Zhu. 2010. The design of puzzle selection strategies for ESP-like GWAP systems. In *IEEE Transactions on Computational Intelligence and AI in Games* 2, 2 (2010), 120–130.

C. Chen, D. Zhang, N. Li, and Z. Zhou. 2014. B-Planner: Planning bidirectional night bus routes using large-scale taxi GPS traces. *IEEE Transactions on Intelligent Transportation Systems* 15, 4 (2014), 1451–1465.

I. R. Chen, F. Bao, and J. Guo. 2016a. Trust-based service management for social internet of things systems. *IEEE Transactions on Dependable and Secure Computing (TDSC)* 13, 6 (2016), 684–696.

Y. Chen, J. Zhou, and M. Guo. 2016b. A context-aware search system for internet of things based on hierarchical context model. *Telecommunication Systems* 62, 1 (2016), 77–91. 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 *Proceedings of the UbiComp*.

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

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

I. Cvijik, Bogdan Ivan, Cristina Kada, and Y. Te. 2015. Towards a crowdsourcing approach for crime prevention. In *Proceedings of the ACM UbiComp/ISWC Adjunct*. 1367–1372.

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 *Proceedings of the CoNEXT*. 349–361.

D. Dimov. 2014. Crowdsensing: State of the Art and Privacy Aspects. Retrieved from 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 *Proceedings of the 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. *Applied Psychophysiology 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 *Proceedings of the ACM MobiSys*. Breckenridge.

S. Forsström and T. Kanter. 2014. Enabling ubiquitous sensor-assisted applications on the internet-of-things. *Personal and Ubiquitous Computing* 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 Communications Magazine* 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 *Proceedings of the HCI*.

A. Garyfalos and K. Almeroth. 2008. Coupons: A multilevel incentive scheme for information dissemination in mobile networks. *IEEE 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 *Proceedings of the 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 *Proceedings of the 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. 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 (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, 6 (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 *Proceedings of the INFOCOM*.

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 *Proceedings of the 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 *Proceedings of the ACM UbiComp*. 547–558.

A. Kamra, V. Misra, J. Feldman, and D. Rubenstein. 2006. Growth codes: Maximizing sensor network data persistence. In *Proceedings of the SIGCOMM*.

M. Kay, E. Choe, J. Shepherd, B. Greensten, N. Watson, S. Consolvo, and J. Kientz. 2012. Lullaby: A capture & access system for understanding the sleep environment. In *Proceedings of the 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 TOSN* 10, 4 (2014), 66:1–66:43.

W. Khan, Y. Xiang, M. Aalsalem, and Q. Arshad. 2013. Mobile phone sensing systems: A survey. *IEEE Communications Surveys Tutorials* 15, 1 (2013), 402–427.

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

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

O. Kounadi, K. Bowers, and M. Leitner. 2014. Crime mapping on-line: Public perception of privacy issues. *Europ Journal on Crime and Policing R* 21, 1 (2014), 1–24.

I. Krontiris and T. Dimitriou. 2013. Privacy-respecting discovery of data providers in crowd-sensing applications. In *Proceedings of the 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 *Proceedings of the ACM 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 *Proceedings of the SenSys*.

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

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

J. Li, X. Chen, M. Li, J. Li, P. 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 *Proceedings of the ICDCS*.

A. Liu and D. Salvucci. 2001. Modeling and prediction of human driver behavior. In *Proceedings of the HCI*.

J. Liu, H. Shen, and L. Yu. 2017. Question quality analysis and prediction in community question answering services with coupled mutual reinforcement. *IEEE TSC* 10, 2 (2017), 286–301.

J. Liu, H. Shen, and X. Zhang. 2016. A survey of mobile crowdsensing techniques: A critical component for the internet of things. In *Proceedings of the 6th International Workshop on Context-Aware Performance Engineering for the Internet of Things (ContextQoS) in conjunction with ICCCN*.

J. Liu, L. Yu, H. Shen, Y. He, and J. Hallstrom. 2015. Characterizing data deliverability of greedy routing in wireless sensor networks. In *Proceedings of the SECON*.

L. Liu, A. Biderman, and C. Ratti. 2009. Urban mobility landscape: Real time monitoring of urban mobility patterns. In *Proceedings of the 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 *Proceedings of the SenSys*.

N. Maalel, E. Natalizio, A. Bouabdallah, P. Roux, and M. Kellil. 2013. Reliability for emergency applications in internet of things. In *Proceedings of the DCOSS*.

N. Maisonneuve, M. Stevens, M. E. Niessen, and L. Steels. 2009. Noisetube: Measuring and mapping noise pollution with mobile phone. In *Information Technologies in Environmental Engineering*. Ioannis N. Athanasiadis, Andrea E. Rizzoli, Pericles A. Mitkas, and Jorge Marx Gómez (Eds.). Springer Berlin Heidelberg, 215–228.

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. Lane, K. Fodor, R. Peterson, H. Lu, M. Musolesi, S. Eisenman, X. Zheng, and A. Campbell. 2008. Sensing meets mobile social networks: The design, implementation and evaluation of the cenceme application. In *Proceedings of the SenSys*.

P. Mohan, V. Padmanabhan, and R. Ramjee. 2008. Nericell: Rich monitoring of road and traffic conditions using mobile smartphones. In *Proceedings of the ACM 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 *Proceedings of the ACM MobiSys*.

M. Nitti, L. Atzori, and I. Cvijikj. 2014. Network navigability in the social internet of things. In *Proceedings of the WF-IoT*.

M. Nitti, L. Atzori, and I. Cvijikj. 2015. Friendship selection in the social internet of things: Challenges and possible strategies. *IEEE Internet Things Journal* 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 *Proceedings of the SIGSPATIAL GIS*.

D. Peng, F. Wu, and G. Chen. 2015. Pay as how well you do: A quality based incentive mechanism for crowdsensing. In *Proceedings of the 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 *Proceedings of the ISSNIP*. Melbourne.

S. Phithakkitnukoon and P. Oliver. 2011. Sensing urban social geography using online social networking data. In *Proceedings of the 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. Notional Policing Improvement Agency.

M.-R. Ra, B. Liu, T. L. Porta, and R. Govindan. 2012. Medusa: A programming framework for crowd-sensing applications. In *Proceedings of the ACM 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 *Proceedings of the UbiComp*.

A. Rai, K. Chintalapudi, V. Padmanabhan, and R. Sen. 2012. Zee: Zero-effort crowdsourcing for indoor localization. In *Proceedings of the 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 *Proceedings of the ISPN*.

J. Roemer, M. Groman, Z. Yang, Y. Wang, C. C. Tan, and N. Mi. 2014. Improving virtual machine migration via deduplication. In *Proceedings of the IEEE MASS*.

J. Rogstadius, M. Vukovic, C. Teixeira, V. Kostakos, E. Karapanos, and J. Laredo. 2013. Crowdsourced social media curation for disaster awareness. *IBM Journal 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 *Proceedings of the URB-IOT*.

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

S. Saroiu and A. Wolman. 2009. Enabling new mobile applications with location proofs. In *Proceedings of the 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 *Proceedings of the 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 Journal* 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 *Proceedings of the SecurIT*.

K. Tangwongsan, H. Pucha, D. G. Andersen, and M. Kaminsky. 2010. Efficient similarity estimation for systems exploiting data redundancy. In *Proceedings of the 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 *Proceedings of the 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 *Proceedings of the SenSys*.

I. Vergara-Laurens, L. Jaimes, and M. Labrador. 2017. Privacy-preserving mechanisms for crowdsensing: survey and research challenges. *IEEE Internet of Things Journal* 4, 4 (2017), 855–869.

L. von Ahn and L. Dabbish. 2004. Labeling images with a computer game. In *Proceedings of the ACM CHI*.

L. Wang, D. Zhang, Y. Wang, C. Chen, X. Han, and A. 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 *Proceedings of the MOBIHOC*. Philadelphia, 113–122.

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 *Proceedings of the 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. Da Xu. 2015. The Internet of Things—a survey of topics and trends. *Information Systems Frontiers* 17, 2 (2015), 261–274.

W. Willett, S. Ginosar, A. Steinitz, B. Hartmann, and M. Agrawala. 2013. Identifying redundancy and exposing provenance in crowdsourced data analysis. *TVCG* 19, 12 (2013), 2198–2206.

L. Xu, X. Hao, N. Lane, X. Liu, and T. Moscibroda. 2015a. Cost-aware compressive sensing for networked sensing systems. In *Proceedings of the IPSN*.

L. Xu, X. Hao, N. Lane, X. Liu, and T. Moscibroda. 2015b. More with Less: Lowering user burden in mobile crowdsourcing through compressive sensing. In *Proceedings of the ACM UbiComp*. 659–670.

T. Yan, V. Kumar, and D. Ganesan. 2010. CrowdSearch: Exploiting crowds for accurate real-time image search on mobile phones. In *Proceedings of the 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 *Proceedings of the ACM CIKM*.

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

A. Zanella, N. Bui, A. Castellani, L. Vangelista, and M. Zorzi. 2014. Internet of things for smart cities. *IEEE Internet Things Journal* 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. *Computer 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 *Proceedings of the ACM UbiComp*. 703–714.

W. Zhang, J. Wang, and W. Feng. 2013. Combining latent factor model with location features for event-based group recommendation. In *Proceedings of the KDD*.

X. Zhang, Z. Yang, W. Sun, Y. Liu, S. Tang, K. Xing, and X. 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 *Proceedings of the 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. *TMC* 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 *Proceedings of the VLDB*. 358–369.

Z. Zhu, L. Zhang, and R. Wakikawa. 2011. Supporting mobility for internet cars. *IEEE Communications Magazine* 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 TOSN* 11, 1 (2014), 1–34.

Received August 2016; revised June 2017; accepted January 2018