Mobile Crowd Sensing and Computing: The Review of an Emerging Human-Powered Sensing Paradigm

BIN GUO, ZHU WANG, and ZHIWEN YU, Northwestern Polytechnical University
YU WANG, University of North Carolina at Charlotte
NEIL Y. YEN, University of Aizu
RUNHE HUANG, Hosei University
XINGSHE ZHOU, Northwestern Polytechnical University

With the surging of smartphone sensing, wireless networking, and mobile social networking techniques, Mobile Crowd Sensing and Computing (MCSC) has become a promising paradigm for cross-space and large-scale sensing. MCSC extends the vision of participatory sensing by leveraging both participatory sensory data from mobile devices (offline) and user-contributed data from mobile social networking services (online). Further, it explores the complementary roles and presents the fusion/collaboration of machine and human intelligence in the crowd sensing and computing processes. This article characterizes the unique features and novel application areas of MCSC and proposes a reference framework for building human-in-the-loop MCSC systems. We further clarify the complementary nature of human and machine intelligence and envision the potential of deep-fused human–machine systems. We conclude by discussing the limitations, open issues, and research opportunities of MCSC.

Categories and Subject Descriptors: J.9.d [**Mobile Applications**]: Pervasive Computing; K.4.m [**Computers and Society**]: Miscellaneous

General Terms: Algorithms, Theory

Additional Key Words and Phrases: Mobile phone sensing, cross-space sensing and mining, human-machine systems, urban/community dynamics, crowd intelligence

ACM Reference Format:
Bin Guo, Zhu Wang, Zhiwen Yu, Yu Wang, Neil Y. Yen, Runhe Huang, and Xingshe Zhou. 2015. Mobile crowd sensing and computing: The review of an emerging human-powered sensing paradigm. ACM Comput. Surv. 48, 1, Article 7 (August 2015), 31 pages.
DOI: http://dx.doi.org/10.1145/2794400

1. INTRODUCTION

Successful society and city management relies on efficient monitoring of urban and community dynamics for decision and policy making. To achieve this, traditional

This work was partially supported by the National Basic Research Program of China 973 (No. 2015CB352400), the National Natural Science Foundation of China (No. 61332005, 61373119, 61222209, 61402369), the Program for New Century Excellent Talents in University (No. NCET-12-0466), and the Scientific and Technology New Star of Shaanxi Province (2014KJXX-39).
Authors' addresses: B. Guo, Z. Wang, Z. Yu, and X. Zhou, School of Computer Science, Northwestern Polytechnical University, 127 West Youyi Road, Xi'an 710072, China; emails: {guob, wangzhu, zhiwenyu, zhouxs}@nwpu.edu.cn; Y. Wang, Department of Computer Science, University of North Carolina at Charlotte, 9201 University City Blvd., Charlotte, NC 28223; email: yu.wang@uncc.edu; N. Y. Yen, School of Computer Science and Engineering, University of Aizu, Tsuruga, Ikki-machi Aizuwakamatsu 965-8580 Fukushima, Japan; email: neilyyen@u-aizu.ac.jp; R. Huang, Department of Computer & Information Sciences, Hosei University, Kajino-cho, Koganei-shi, Tokyo 184-8584, Japan; email: rhuang@hosei.ac.jp.

DOI: http://dx.doi.org/10.1145/2794400

sensing techniques (e.g., sensor networks) usually leverage distributed sensors to acquire real-world conditions [Lewis 2005; Stankovic 2008]. However, though there has been a growing body of studies on sensor networks, commercial sensor network techniques have never been successfully deployed in the real world for several reasons, such as high installation cost, insufficient spatial coverage, and so on [Liu et al. 2013; Burke et al. 2006].

Mobile Crowd Sensing and Computing (MCSC) is a large-scale sensing paradigm based on the power of user-companioned devices, including mobile phones, smart vehicles, wearable devices, and so on [Guo et al. 2014]. MCSC allows the increasing number of mobile phone users to share local knowledge (e.g., local information, ambient context, noise level, and traffic conditions) acquired by their sensor-enhanced devices, and the information can be further aggregated in the cloud for large-scale sensing and community intelligence mining [Zhang et al. 2011]. The mobility of large-scale mobile users makes MCSC a versatile platform that can often replace static sensing infrastructures. A broad range of applications are thus enabled, including traffic planning, environment monitoring, mobile social recommendation, public safety, and so on.

A formal definition of MCSC is as follows: *a new sensing paradigm that empowers ordinary citizens to contribute data sensed or generated from their mobile devices and aggregates and fuses the data in the cloud for crowd intelligence extraction and human-centric service delivery*. From the AI perspective, MCSC is founded on a distributed problem-solving model [Brabham 2008]. In the literature history, the concept of crowd-powered problem solving has been explored in several research areas. One decade ago, Surowiecki wrote a book titled *The Wisdom of Crowds* (or crowd wisdom) [Surowiecki 2005], where a general phenomenon—that the aggregation of data or information from a group of people often results in better decisions than those made by a single person from the group—is revealed. It identifies four key qualities that make a crowd smart: *diversity in opinion, independence of thinking, decentralization, and opinion aggregation*. A similar concept to crowd wisdom is collective intelligence [Malone et al. 2009]. Different from the two concepts that focus on the advantages of group decision making, MCSC is mainly about crowd-powered data collection and processing. In 2005, two senior editors from *Wired Magazine,* Jeff Howeand and Mark Robinson, coined the term "crowdsourcing" [Howe 2006]. According to the Merriam-Webster Dictionary,[1] crowdsourcing is defined as the practice of obtaining needed services or content by soliciting contributions from a large group of people, and especially from an online community. A typical example is Wikipedia, where tens of thousands of contributors collaboratively create the largest encyclopedia of the world. However, compared to MCSC, crowdsourcing focuses on the participation of online crowds. The closest concept to MCSC is participatory sensing, proposed in Burke et al. [2006]. It tasks average citizens and companioned mobile devices to form participatory sensor networks for local knowledge gathering and sharing. The definition of participatory sensing emphasized explicit user participation when it was proposed. In recent years, with the rapid development of smartphone sensing and mobile Internet techniques, the scope of crowd problem-solving systems using mobile devices has been broadened. To this end, we extend the definition of participatory sensing from two aspects and term the new concept MCSC [Guo et al. 2014]. First, MCSC leverages both sensed data from mobile devices (from the physical community) and user-contributed data from mobile social network services (from the online community). In other words, MCSC counts both explicit and implicit user participation for data collection. Second, MCSC presents the fusion of machine and human intelligence in both the sensing and computing process. A summary of the literature history is given in Table I.

[1]http://www.merriam-webster.com/dictionary/crowdsourcing.

Table I. The Literature History of Crowd-Powered Problem Solving

Concept	Definition and Relationship
Crowd wisdom/collective intelligence	The aggregation of data or information from a group of people often results in better decisions than those made by a single person from the group.
Crowdsourcing	The practice of obtaining needed services or content by soliciting contributions from a crowd of people, especially from an online community.
Participatory sensing	It tasks average citizens and companioned mobile devices to form participatory sensor networks for local knowledge gathering and sharing.
Mobile crowd sensing and computing (MCSC)	An extension to the participatory sensing concept to have user participation in the whole computing lifecycle: (1) leveraging both offline mobile sensing and online social media data; (2) addressing the fusion of human and machine intelligence in both the sensing and computing process.

MCSC benefits a number of application areas regarding urban/community dynamics monitoring and beyond. At the same time, numerous and unique research challenges and open issues arise from the MCSC paradigm. Specifically, the main contributions of our article are as follows:

—Characterizing the unique features of MCSC, such as grassroots-powered sensing, human-centric computing, transient network connection, and cross-space crowdsourced data processing.
—Reviewing existing applications and techniques on community/urban sensing, including environment monitoring, traffic planning, mobile social recommendation, healthcare, and public safety. Based on these studies, a layered reference framework to build MCSC systems is proposed.
—Investigating the complementary features of human capabilities and machine intelligence and exploring the collaboration patterns of them in the crowd sensing and computing process.
—Identifying challenges and future research directions of MCSC. We examine the challenges such as sensing with human participation, incentive mechanisms, data selection in transient networks, data quality and data selection, and heterogeneous crowdsourced data mining. The future research trends and emerging techniques to address these challenges are also discussed.

The remainder of the article is organized as follows. In Section 2, we characterize the unique features of MCSC. Section 3 presents various novel applications enabled by MCSC, followed by a conceptual framework of MCSC systems in Section 4. Section 5 investigates the deep fusion of human and machine intelligence and potential collaboration patterns. The limitations, challenging issues, and research opportunities to MCSC are discussed in Section 6. We conclude the article in Section 7.

2. CHARACTERIZING MCSC

In this section, we first characterize the unique features of MCSC and then present its taxonomy.

2.1. Grassroots-Powered Sensing

Compared to traditional sensor networks, the key difference of MCSC is the involvement of grassroots for large-scale sensing. Specifically, grassroots participation offers MCSC a number of advantages: (1) MCSC leverages existing sensing and communication infrastructure, and therefore, its deployment costs are quite low; (2) the inherent mobility of mobile device users provides unprecedented spatiotemporal coverage compared to static sensor network deployments. A summarization of the differences between MCSC and traditional sensor networks is given in Table II.

Table II. Differences Between MCSC and Traditional Sensor Networks

	Operators	Deployment Cost	Coverage	Data Quality
Wireless sensor network (WSN)	Government agencies, public institutions	High: Expensive sensors and infrastructure to deploy the network	Limited coverage with static sensor nodes	High, sound level sensors
Mobile crowd sensing and computing (MCSC)	Potentially everyone	Low: Leveraging existing infrastructure, i.e., broad proliferation of cellular network and mobile device usage	The inherent mobility of the phone carriers provides unprecedented spatiotemporal coverage	Low, suffering from issues such as built-in sensor performance and the trustworthiness of user-contributed data

The involvement of citizens in the sensing loop is the chief feature of participatory sensing. The same is true for MCSC, but it moves a step further than participatory sensing and introduces several new features.

—*The data generation mode*. Compared with participatory sensing, which mainly collects data from the physical space using mobile devices, there are two different data generation modes in MCSC. (1) *Mobile sensing*. It typically functions at a context-based and individual manner, leveraging the rich sensing capabilities from individual devices [Lane et al. 2010]. It is the data collection method used by participatory sensing. (2) *Mobile social networking* (MSN). With the rapid development of mobile Internet, MSN services [Zhang et al. 2014d] that bridge the gap between online interactions and physical elements (e.g., check-in places, activities [Liu et al. 2012], and objects [Guo et al. 2013]) are fast growing. The large-scale user-contributed data opens a new window to understand the dynamics of the city and society, which is counted as the other data source for MCSC. The combination of participatory mobile sensing and participatory MSN data is a unique feature of MCSC.

—*The sensing style*. According to Ganti et al. [2011], traditional crowd-powered sensing spans a wide spectrum regarding the degree of human involvement, where participatory sensing and opportunistic sensing are placing at the two ends. By having two distinct participatory data contribution modes, the sensing style of MCSC changes and we can categorize it from a new dimension: *user's awareness to the sensing task*. On one hand, for participatory/opportunistic sensing, data collection is the primary purpose of the application. As a consequence, the sensing task is *explicit* to the user (as he or she is informed). On the other hand, the primary usage of MSN services is for social interaction, while the data generated along with online interactions is used *implicitly* for collective intelligence extraction. In other wards, we can characterize the sensing style of MCSC at two levels: *explicit* and *implicit*.

—*Volunteer organization*. The participants can be self-organized citizens with varying levels of organizational involvement, ranging from total strangers to loosely organized groups of neighbors facing a shared problem to well-organized previously existing activism groups [Maisonneuve et al. 2010]. Specially, we categorize volunteer organizations into the following three modes, which work at distinct scales.

(1) *Group*. This refers to a loosely or opportunistically organized group of neighbors (i.e., spatially nearby phones) that collaboratively address a shared problem. For example, in SignalGuru [Koukoumidis et al. 2011], vehicles that are passing through an intersection can sense and share the traffic signal information and adjust their driving speed. Movi [Bao and Choudhury 2010] allows the collection

Table III. Characterizing the Scales of MCSC

Scale	Data Quality	Collaboration	Social Tie Strength	Key Techniques	Major App Areas
Group	Medium	Close	Weak or strong	Opportunistic connection and group formation	Local area sensing, local event replay
Community	High	Close	Medium	Community creation/detection	Scientific research, community-specific services
Urban	Low	Loose	Weak	Data quality, trust maintenance	Urban dynamics sensing, environment monitoring

and replay of a local social event leveraging the local-captured information. In this mode, people usually form the so-called opportunistic or ad hoc groups [Guo et al. 2013], and the relation among group members can be weak [Koukoumidis et al. 2011] or strong [Bao and Choudhury 2010]. The key techniques include group formation/identification and management. For instance, we need to partition the set of mobile devices based on the associated physical or social contexts [Bao and Choudhury 2010].

(2) *Community*. According to the Cambridge Advanced Learner's Dictionary,[2] community is defined as "the people living in one particular area or people who are considered as a unit because of their common interests, social group or nationality." Community-based MCSC is the crowdsourcing where task takers come from an existing community or could easily form a new community. The members trust each other and are more likely to interact with each other during the task execution process, resulting in high-quality data. For instance, the MIT Owl project[3] takes advantage of the sensor-enhanced smartphones (with GPS, directional microphones, etc.) to assess owl populations in a huge region. Hundreds of amateur botanists have been reported to use their mobile phones to gather pictures of plants to study the link between climate change and our ecosystems [Shilton et al. 2009].

(3) *Urban*. It targets the applications with a broad appeal, such as urban traffic dynamics [Chen et al. 2013] and air/noise pollution monitoring [Zheng et al. 2013]. Any citizens (mostly strangers) can participate in the urban-scale sensing activity. However, compared with the previous two modes, the data contributed by ordinary users often has low quality (e.g., people may contribute fake data).

We further summarize the characteristics of MCSC at the previous three different scales in Table III.

2.2. Hybrid Human–Machine Systems

To motivate full-scaled user participation and enhance user experience, MCSC systems should be developed in a human-centric manner. Particularly, it should address the following key problems:

—*Motivation of human participation*. Mobile devices usually have limited resources (e.g., energy, bandwidth, etc.), and the sensory information from them is often highly sensitive. To facilitate data sharing among peers, the development of an incentive model is significant [Guo et al. 2013]. The deployment of a large-scale MCSC system

[2] http://dictionary.cambridge.org/dictionary/british/.
[3] http://owlproject.media.mit.edu.

usually requires a considerable amount of participants, and some of them may drop out of the collecting loop unless the return on investment is greater than their expectations. Questions about human motivation have been central in philosophy and economics. For example, the promise of *financial* or *monetary* gain has been an important incentive method for most actors in markets and traditional organizations. *Interest* and *entertainment* are also important motivators in many situations, even when there is no prospect of monetary gain [Malone et al. 2009]. In addition, people could be motivated to participate in an activity by *social* and *ethical* reasons, such as socializing with others, reputation, or recognition by others. Some crowd-sensing services ask users to contribute their data for service usage. There are also indirect ways to enhance user participation, such as *energy conservation* and *privacy protection* mechanisms in data contribution.

—*Fusion of human–machine intelligence*. The involvement of human participation in the sensing and computing process will lead to a mixture of human and machine intelligence in MCSC. On one hand, when the crowd participates in the loop of data sensing, human intelligence is embedded in the obtained data (e.g., contributors can decide what to be captured). On the other hand, based on the wealth of data that is collected in MCSC systems, both machine intelligence (e.g., machine learning and data mining techniques) and human intelligence (e.g., cognition, reasoning) can be used for data processing (e.g., classification, decision making). By integrating the intelligence of both crowds and machineries, MCSC allows the creation of hybrid human–machine systems. However, such human intelligence and machine intelligence could either be complementary or conflicting, and they show strength and weakness to distinct computing problems. Therefore, novel approaches and techniques should be studied to have "optimal" fusion of human and machine intelligence in MCSC systems. More details will be presented in Section 5.

—*User security and privacy*. The sharing of personal data (e.g., user location, ambient sound) in MCSC applications can raise significant concerns about security and user privacy. To motivate user participation, we should explore new techniques to protect information security in data contribution. Particularly, the definition of security and privacy might continue to evolve as MCSC systems mature, which is highlighted by the fact that even if personal information is not directly obtained by an unwanted party, much of the information can be inferred from aggregated data. For example, the fact that an object with an RFID tag can be uniquely identified and tracked back to its user might bring up many privacy issues [Acampora et al. 2013]. Therefore, new privacy-preserving approaches need to be developed to ensure that important information cannot be gleaned from mined patterns [Matatov et al. 2010; Fung et al. 2010; Christin et al. 2011].

2.3. Transient Networking

The success of MCSC relies on leveraging ubiquitous and heterogeneous communication capabilities to provide transient network connection and effective collection of mobile sensing data. While different MCSC applications or systems may have various connection architectures and communication requirements, most of them share the following characteristics:

—*Heterogeneous network connection*. Current mobile devices are usually equipped with multiple wireless communication interfaces and supported via different wireless technologies. For example, a smartphone can at least have GSM, WiFi, and Bluetooth interfaces. While GSM and WiFi interfaces can provide network connectivity with preexisting communication infrastructure (e.g., via cellular base stations in an urban area or WiFi access points in a campus building) in relevantly large regions, Bluetooth or WiFi can also provide short-range connection among mobile devices and form self-organized opportunistic networks for data sharing [Conti and Kumar 2010; Guo et al. 2013]. Heterogeneous networks raise new research challenges and enriched opportunities for MCSC applications by supporting transient networking services, such as connectivity, collaborative sensing, and data routing/transmission, for the participants crossing these multiple wireless networks.

—*Time-evolving network topology and human mobility*. The mobility of mobile devices and their carriers not only provides a nice coverage for sensing tasks in MCSC but also brings challenges to communications. Network topology evolves over time, which makes it hard to find stable routes among mobile devices. Traditional routing protocols designed for static wireless networks or wired networks cannot handle highly dynamic topology to fulfill basic communication tasks in MCSC (especially for pure ad hoc deployment). Fortunately, recent advances in mobile ad hoc networks and wireless sensor networks provide many low-cost ad hoc routing solutions [Chen et al. 2011]. In addition, since human mobility plays a major role in governing the network dynamic and behavior in MCSC, recent research efforts have been devoted to mobility modeling and new routing schemes based on human mobility [Karamshuk et al. 2011; Pasquale and Silvestri 2013].

—*Disruption tolerance service*. In some MCSC applications, the sensing data from each individual mobile device does not need to be transmitted in real time or guaranteed to be complete and accurate. Therefore, such MCSC systems can take advantage of disruption-tolerant networks [Gao and Cao 2011], which only rely on intermittent network connectivity and have a much lower deployment cost. Note that many mobile sensing devices cannot be guaranteed to be connected all the time due to poor network coverage (e.g., poor signal strength due to interference or no signal in a rural area), energy constraint (e.g., low battery), or user preference (e.g., phone shut off in a meeting) [Ma et al. 2014]. Thereby, it is possible that the network in MCSC has a very sparse connection, and disruption-tolerant services should be used. In such situations, the sensing data is usually stored at local storage and waits for future transmission opportunities.

—*High scalability requirement*. Since MCSC relies on sensing data from a large volume of mobile users, the scalability is clearly a basic requirement and key challenge for underlying communication systems. To achieve sufficient scalability, the MCSC communication protocols and network architectures are usually highly distributed and decentralized. Such solutions can also improve the robustness of the overall MCSC system. In addition, energy-efficient design has to be considered for MCSC due to the limited power resources of each individual device and the large number of devices in the system.

2.4. Crowd Data Processing and Intelligence Extraction

The aim of MCSC is to extract high-level intelligence from a large volume of user-contributed data. Regarding the value of intelligence and its beneficiary, we can classify it into three distinct dimensions [Guo et al. 2013], namely, *user*, *ambient*, and *social awareness*.

—*User awareness* refers to the extraction of personal contexts (e.g., location, activity) and behavioral patterns (e.g., mobility patterns).

—*Ambient awareness* is to learn the status (e.g., noise level of a bus stop, traffic dynamics of a street) or the semantics (the logical type) of a particular space. The space can be small (e.g., a restaurant) or large (e.g., a city-wide area).

—*Social awareness* is about the contexts of a group or a community, such as social activity type, interpersonal relations, and so on.

However, data collected from MCSC systems has brought forth numerous new issues to the effective extraction of such intelligence, for example, its low quality and cross-space characters.

—*Low-quality data*. Data quality is usually defined as the degree of how fit the data is for its intended uses in operations, decision making, and planning. In other words, low-quality data is user-related data collected regardless of fitness for use. Particularly, the involvement of human participation brings forth critical challenges to the data quality of MCSC systems. For example, MCSC participants may send incorrect or low-quality data to the backend server [Zhang et al. 2014c], data contributed by different people may be redundant or inconsistent [Uddin et al. 2011], and the same device might be used to record the same activity under distinct sensing conditions (e.g., sensing ambient noise when placing the mobile phone in a pocket or at hand). Therefore, *data selection* is often needed to improve data quality, and we should explore methods on fault filtering, quality estimation, expert contributor encouragement, and so forth.
—*Heterogeneous, cross-space data mining*. The mobile crowd data can be collected from both offline/physical and online/virtual communities. Different communities represent distinct interaction manners (e.g., comments/transfers online, colocation offline) and contain different knowledge (e.g., friendship in online communities, movement patterns in offline communities). Therefore, how to effectively associate and fuse cross-space data (e.g., how physical events are mirrored in online social networks) and how to integrate their complementary features (e.g., from pieces of data collected from different sources to a comprehensive picture of a sensing object [Wang et al. 2014]) become important yet challenging research issues for MCSC.

2.5. The Taxonomy of MCSC

According to the main characters of MCSC, as summarized in Figure 1, a taxonomy is proposed in this section, which can steer the discussions in the following sections.

Mobile sensing. Both *hardware* sensors and *software* sensors are involved in MCSC. The prior one refers to embedded sensors (e.g., accelerometer, GPS, camera) in mobile devices, while the latter one counts user-generated data from MSN portals. In terms of the degree of user involvement in MCSC, we can group the sensing styles as *participatory* sensing, *opportunistic* sensing, and *hybrid* sensing (a combination of machine computation and user control). On the other hand, in terms of user awareness to the sensing task, we can categorize them as *implicit* sensing and *explicit* sensing. Sensing tasks can work continuously (*continuous sensing*) or in the *event-triggered sensing* manner. For the prior case, the relevant sensor works continuously under the parameter settings (e.g., sampling rate, duration) of a sensing task. However, continuous sensing on mobile devices may be too energy expensive [Zhang et al. 2014a], and some sensing tasks are only meaningful under certain sampling contexts (e.g., certain time periods or places). In such cases, triggers should be defined to capture data in a context-aware manner [Bao and Choudhury 2010].

Crowd data collection. Regarding the scale of crowd data contribution, it can be small or large, ranging from a *group* to a *community* to the *urban* scale. The MCSC network connection can be broadly categorized into three types: *infrastructure based*, *ad hoc*, and *hybrid*. The first one leverages existing infrastructure (e.g., cellular, 3G, access points), the second one usually forms opportunistic networks [Conti and Kumar 2010], while the last one mixes the features of the prior two types. Furthermore, as discussed in Section 2.2, incentives for user participation used for the crowd sensing can include *monetary*, *ethical*, *entertainment*, *service provision*, *privacy protection*, and so on.

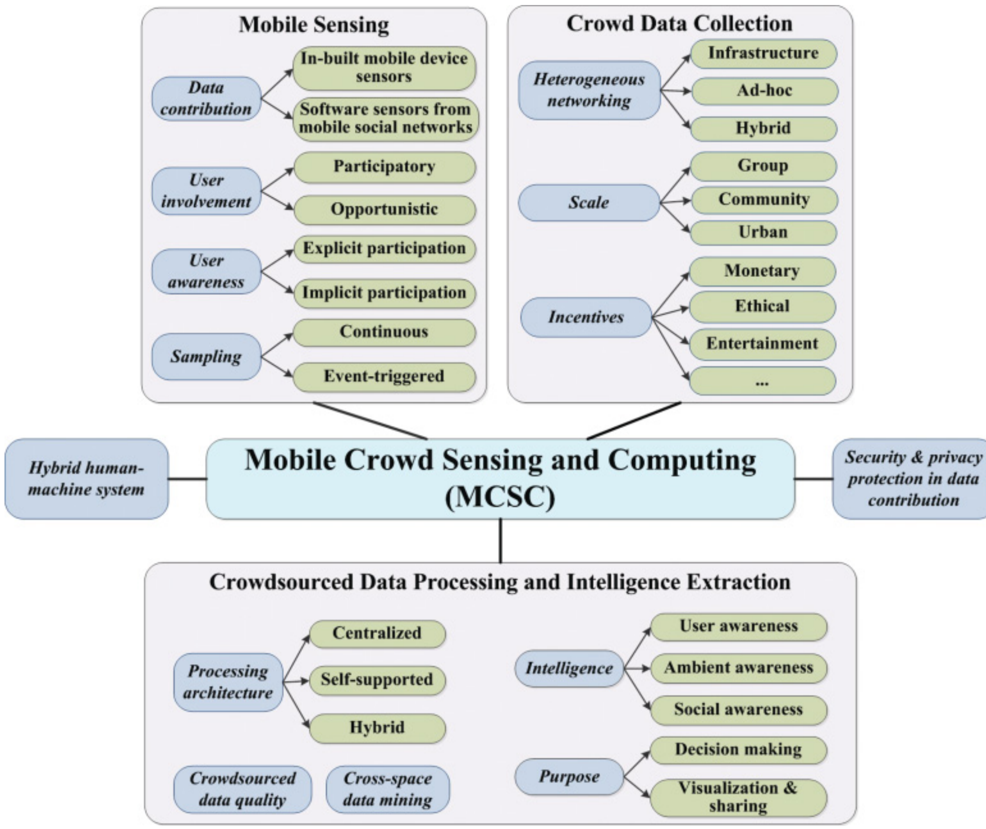

Fig. 1. A taxonomy graph of MCSC.

Crowdsourced data processing. The involvement of human participation in crowd sensing brings forth redundant, low-quality, or even fake data to MCSC systems, and thus *data selection* is often needed to improve data quality. Meanwhile, MCSC systems usually use heterogeneous devices, and some of these devices may have limited computing resources while others have more. It thus results in two different data processing methods in MCSC: the *centralized* method transmits all gathered data to a backend server for processing, whereas the *self-supported* method endows the device itself with the ability for data processing [Guo et al. 2013]. Recently, some studies have been trying to balance between them based on *hybrid* methods [Guo et al. 2013; Miluzzo et al. 2010].

Crowd intelligence extraction and usage. As presented in Section 2.4, there are three main types of crowd intelligence: *user*, *ambient*, and *social awareness*. The learned crowd intelligence can be used by authorities, public institutions, and ordinary citizens in different application domains, for example, public health, urban planning, and environmental protection. We further derive the two major purposes of MCSC in application usage: *decision making* and *visualization and sharing*. For decision making, it refers to making decisions (e.g., object classification/recognition, event prediction/replay, policy making) or cueing recommendations according to the learned knowledge. For visualization and sharing, the collected information is visualized and shared among citizens.

3. MCSC-EMPOWERED APPLICATIONS

The development of MCSC has resulted in various novel sensing applications. A selection of representative applications will be presented and discussed in this section, as shown in Table IV, V, VI, and VII, respectively.

3.1. Crowd-Powered Environment Monitoring

The participatory and mobile nature of MCSC provides a novel way for environment monitoring, such as nature preservation and pollution measurement.

Nature preservation. Many scientific studies are based on the large-scale data collected from the real world. During the past few years, people's mobile devices have been explored to contribute data for various scientific studies. For example, with the help of human volunteers, the Great Backyard Bird Count project[4] continuously reported the count of watched birds in the United States. The MIT Owl project, on the other hand, leveraged the network of sensor-equipped smartphones to study owl populations. Scientists also took advantage of citizens to collect data to study the impacts of climate change, such as the link between increasing temperatures and the timing of specific events of plants (e.g., emergence of the first leaf, fruiting) [Shilton et al. 2009]. We are sure that the prospect of working with large-scale citizen-generated information for scientific research and nature preservation is becoming a reality.

Noise pollution measurement. Noise pollution has become a worldwide problem. Studies have linked noise exposure to a decline in health and in quality of life, citing raised blood pressure, hearing loss, annoyance, and so on [Stansfeld and Matheson 2003]. In light of its risks, the European Commission mandates the creation of noise contour maps to gather information about the exposure.[5] However, government efforts are limited because in most cases the deployed sensing nodes are not able to cover all areas of the city. This issue is remedied with the emergence of MCSC, where microphones in smartphones can be used to measure the ambient noise level and the data aggregated from the volunteers at the city scale can be used to generate a fine-grained noise map. NoiseTube [Maisonneuve et al. 2010] was a system that could measure personal exposure to environmental noise in users' daily lives. Shared measurements were used to create city noise maps. EarPhone [Rana et al. 2010] was also a participatory noise mapping system that also used mobile phones to determine environmental noise levels (particularly roadside ambient noise). SoundOfTheCity [Ruge et al. 2013] allowed users to link their feelings and experiences with the measured noise level, for example, is it in a party or in a crowded street.

Air pollution measurement. The PEIR project [Mun et al. 2009] linked users' transportation mode (e.g., walking, driving) and its environmental impact (e.g., carbon exposure during daily activities). Mobile phones were also used to gather the information of on-the-road diesel trucks to study community exposure to urban air pollution [Shilton et al. 2009]. ExposureSense [Predic et al. 2013] explored the integration of WSN and participatory sensing paradigms for personal air quality exposure measurement. U-Air [Zheng et al. 2013] inferred fine-grained air quality information (e.g., PM2.5) of a city using heterogeneous crowdsourced data, including the air quality data reported by existing sensing stations, points of interest (POIs), traffic information, and so on. The BikeNet application [Eisenman et al. 2010] could measure and report the CO_2 level along the path of a cyclist's activity.

[4]http://www.birdsource.org/gbbc/.
[5]http://ec.europa.eu/environment/noise/directive.htm.

Table IV. A Summary of Main MCSC Applications and Sensing Modalities

Application Type	Function	Typical Applications (Sensors Used)
Environment monitoring	Nature preservation	MIT Owl (GPS, compass, directional microphone)
	Noise pollution measurement	NoiseTube [Maisonneuve et al. 2010], EarPhone [Rana et al. 2010], SoundOfTheCity [Ruge et al. 2013] (GPS, microphone)
	Air pollution measurement	PEIR [Mun et al. 2009] (GPS) BikeNet [Eisenman et al. 2010] (GPS, accelerometer, CO_2 meter, microphone)
Transportation and traffic planning	Traffic dynamics	Real Time Rome [Calabrese et al. 2011] (GPS from citizens, buses, taxies) Liu et al. [2009], Wen et al. [2008] (GPS from taxis)
	Public transportation	B-Planner [Chen et al. 2013] (GPS, pick-up/drop-off records from taxis) Morency et al. [2007] (bus boarding records) Zhou et al. [2012] (cell tower signals, audio recordings, etc.)
	Individual travel planning	VTrack [Thiagarajan et al. 2009], T-Share [Wolfson et al. 2013], SmartTrace+ [Costa et al. 2011] (GPS) SingalGuru [Koukoumidis et al. 2011] (camera)
	Road condition	Nericell [Mohan et al. 2008] (GPS, accelerometer, microphone) PotHole [Eriksson et al. 2008] (GPS, accelerometer)
Urban dynamics sensing	Human urban mobility/behavior patterns	Noulas et al. [2012] (check-ins, GPS) UBhave [Lathia et al. 2013] (microphone, Bluetooth)
	Urban social events/structure	Crooks et al. [2013] (tweets) mHealth [Consulting 2009] (cell tower) Tang and Liu [2010], Wang et al. [2014] (blogs, check-ins) FlierMeet [Guo et al. 2015] (images, GPS)
	Social functions of urban regions	Pan et al. [2013] (GPS from taxis) Karamshuk et al. [2013] (check-ins, tips)
Location services	Logical localization	CrowdSense@Place [Chon et al. 2012] (images, audio clips, GPS)
	Interesting location discovery	GeoLife [Zheng et al. 2011] (GPS trajectories)
Mobile social recommenda-tion	Place recommendation	GeoLife [Zheng et al. 2012] (GPS trajectories) Ye et al. [2011], Yang et al. [2013] (check-ins)
	Itinerary suggestion	Zheng and Xie [2011] (GPS) Yu et al. [2014] (check-ins)
	Service/activity recommendation	T-Finder [Yuan et al. 2013] (GPS from taxis) Zhang et al. [2013] (location, activity)
Healthcare	Public health	Google [Ferguson et al. 2006] (search query, location) Wesolowski et al. [2012] (location)
	Personal well-being	Neat-o-Games [Fujiki et al. 2008] (accelerometer) DietSense [Reddy et al. 2007] (camera, microphone, GPS)
Public safety	Crime prevention	Lee et al. [2011] (geo-tagged Tweets) Boston Marathon explosion [Fowler and Schectman 2013] (images, videos, texts) iSafe [Ballesteros et al. 2014] (location, reviews)
	Disaster relief	Haiti [Bengtsson et al. 2011] (GPS) Sakaki et al. [2010] (geo-tagged Tweets)

Table V. A Taxonomy of Major MCSC Applications on User Awareness

Application Name	Sensing Style and Scale	Sampling	Incentives	Network Architecture	Data Processing	Usage
CrowdSense @Place [Chon et al. 2012]	Participatory, urban	Triggered	Service	Infrastructure	Centralized	Decision making (classification)
GeoLife [Zheng et al. 2011]	Opportunistic, urban	Continuous	——	Infrastructure	Centralized	Decision making (recommendation)
Neat-o-Games [Fujiki et al. 2008]	Opportunistic, community	Continuous	Entertainment	Infrastructure	Centralized	Visualization and sharing
DietSense [Reddy et al. 2007]	Participatory, community	Triggered	Service	Infrastructure	Centralized	Decision making (recommendation)

Table VI. A Taxonomy of Major MCSC Applications on Ambient Awareness

Application Name	Sensing Style and Scale	Sampling	Incentives	Network Architecture	Data Processing	Usage
MIT Owl	Opportunistic, community	Continuous	Ethical	Infrastructure	Centralized	Decision making (recognition)
NoiseTube [Maisonneuve et al. 2010]	Hybrid, urban	Continuous	Ethical	Infrastructure	Centralized	Visualization and sharing
EarPhone [Rana et al. 2010]	Opportunistic, urban	Continuous	Energy saving	Infrastructure	Hybrid	Visualization and sharing
PEIR [Mun et al. 2009]	Hybrid, community	Continuous	Ethical & service	Infrastructure	Centralized	Visualization and sharing
BikeNet [Eisenman et al. 2010]	Hybrid, community	Continuous & triggered	Energy saving, privacy protection	Infrastructure	Centralized	Decision making (recommendation)
Real Time Rome [Calabrese et al. 2011]	Opportunistic, urban	Continuous	——	Infrastructure	Centralized	Visualization and sharing
B-Planner [Chen et al. 2013]	Opportunistic, urban	Continuous	——	Infrastructure	Centralized	Decision making (policy making)
Nericell [Mohan et al. 2008]	Opportunistic, urban	Triggered	Energy saving	Infrastructure	Centralized	Visualization and sharing
T-Share [Wolfson et al. 2013]	Hybrid, urban	Continuous	Service	Infrastructure	Centralized	Decision making (recommendation)
SingalGuru [Koukoumidis et al. 2011]	Opportunistic, group	Continuous	Service	Ad hoc	Self-supported	Decision making (prediction)

Table VII. A Taxonomy of Major MCSC Applications on Social Awareness

Application Name	Sensing Style and Scale	Sampling	Incentives	Network Architecture	Data Processing	Usage
TagSense [Qin et al. 2014]	Participatory, group	Triggered	Privacy protection, energy saving	Ad hoc	Self-supported	Visualization and sharing
MoVi [Bao and Choudhury 2010]	Participatory, group	Triggered	——	Ad hoc	Centralized	Visualization and sharing
Sakaki et al. [2010]	Participatory, community	Continuous	——	Infrastructure	Centralized	Decision making (prediction)
Haiti [Bengtsson et al. 2011]	Opportunistic, community	Continuous	——	Infrastructure	Centralized	Decision making (prediction & policy making)
iSafe [Ballesteros et al. 2014]	Participatory, group	Continuous	——	Infrastructure	Centralized	Decision making (prediction)

3.2. Transportation and Traffic Planning

Crowd data in urban areas can be used for traffic forecasting, public transportation system design, travel planning, and so on.

Traffic dynamics. A number of studies have investigated traffic dynamics using large-scale data from GPS-equipped vehicles and mobile phones. By using data from buses, taxies, and mobile phones, the Real Time Rome project [Calabrese et al. 2011] could report real-time urban dynamics. Similarly, a combination of the data from GPS-equipped taxis and smart card records from buses was used by Liu et al. [2009] for traffic dynamics understanding (e.g., hotspot detection). Wen et al. [2008] used GPS-equipped taxis to analyze traffic congestion changes around the Olympic games in Beijing.

Public transportation. As a shared passenger transport service, it is important to improve its design (e.g., planning of routes, efficiency in different weather conditions) and offer real-time information about the transport entities (e.g., buses, trains) to the general public. B-Planner [Chen et al. 2013] used crowdsourced GPS data and pick-up/drop-off records from taxis to suggest the planning of night-bus routes. Morency et al. [2007] investigated the spatiotemporal dynamics (e.g., examining the effects of weather on transit demand) of public transportation networks based on bus boarding records in Canada. Considering that the bus arrival time was the primary information to most travelers, Zhou et al. [2012] utilized the bus passengers' surrounding environmental context to estimate the bus arrival time of different bus stations.

Individual travel planning. In addition to the support for public transportation, it is also important to support individual travelers, such as suggesting driving routes with low traffic delays. VTrack [Thiagarajan et al. 2009] was a system that used mobile phones to accurately estimate the traffic time between different venues. SmartTrace [Costa et al. 2011] helped identify mobility patterns or a given trajectory's popularity for transit planning. T-Share [Wolfson et al. 2013] was a taxi ridesharing service that could generate optimized ridesharing schedules based on crowd-powered data. Traffic signals inevitably enforce a stop-and-go movement pattern that increases fuel consumption. SignalGuru [Koukoumidis et al. 2011], a novel software service that relied solely on a collection of mobile phones to detect and predict the traffic signal schedule, was developed to enable drivers to adjust speed, avoiding coming to complete halts.

Road condition. The condition of roads (e.g., volume and speed of traffic flowing, icy or bumpy roads) is significant to vehicular traffic in our daily lives. Compared to developed countries, road conditions in the developing world tend to be more varied, with bumpy roads and potholes being commonplace even in the heart of cities. Nericell [Mohan et al. 2008] could detect and report road conditions (e.g., potholes) using the built-in sensors (e.g., accelerometer, GPS) in mobile phones. The information was further integrated into traffic maps to be shared by the public. A similar application is PotHole [Eriksson et al. 2008], which can identify holes in streets using the crowdsourced vibration and position data collected from smartphones.

3.3. Urban Dynamics Sensing

Awareness and understanding of urban dynamics are critical for sustainable urban development and improving the quality of citizen life in terms of convenience, comfort, safety, and security. However, understanding urban dynamics is an increasingly important challenge. With the recent surge of MCSC, urban dynamics sensing has become possible and has been receiving more attention from many major well-established companies and academic research societies.

Human urban mobility/behavior patterns. A number of research works have focused on revealing human mobility and behavior patterns in urban areas. For instance, Noulas et al. [2012] studied human mobility patterns by analyzing the check-in histories of a large set of LBSN (location-based social network) users. They found that the rank distance rather than the pure physical distance played a dominant role in human movement prediction. The UBhave project [Lathia et al. 2013] investigated how to leverage the mobile crowd sensing data to study, understand, and positively affect human behaviors. In particular, an application named EmotionSense was developed under the project to autonomously capture the user's emotive, behavioral, and social signals, based on which one's real-time mood could be identified.

Urban social events/structure. Another research trend is to detect social events or structures by exploring MCSC data. Crooks et al. [2013] studied the spatial and temporal characteristics of the Twitter feed activity responding to a 5.8-magnitude earthquake, which could be used to identify and localize the impact area of the earthquake event. Several studies have investigated urban social structures based on crowdsourced data. For example, Tang and Liu [2010] studied the community detection problem of online social media. Wang et al. [2014] discussed the community discovery and profiling issues in LBSNs.

Social functions of urban regions. The detailed land use is an integral part of urban planning, which is normally difficult to obtain. Fortunately, the large-scale data collected from MCSC systems can be used to unveil the social functions of urban regions. For example, Pan et al. [2013] proposed to classify the social functions of urban land by exploring the GPS log of taxi cabs. The Geo-Spotting system [Karamshuk et al. 2013] focused on identifying the optimal placement for new retail stores. Experiments based on various supervised learning algorithms showed that both geographic and human mobility features were important to the retail success of a new business.

3.4. Crowd-Sensing-Enabled Location Services

Location remains the most successful and widely used context in everyday usage. Awareness of user location underpins many popular and emerging mobile applications, including location search, location-based advertising (disseminating electronic coupons in a market [Garyfalos and Almeroth 2008]), indoor positioning (using WiFi signal strength to locate people [Yu et al. 2011]), and so on. However, these studies mainly rely on individual data, whereas crowd-powered data can spark a variety of new location-based services. For example, CrowdSense@Place [Chon et al. 2012] was a framework that exploited opportunistically captured images and audio clips crowdsourced from smartphones to link place visits with place categories (e.g., store, restaurant). GeoLife [Zheng et al. 2011] was able to suggest friends based on human location history. Crowdsourced user-location histories were also used for discovering interesting locations, which could help users understand an unfamiliar city within a short period [Zheng et al. 2012].

3.5. Mobile Social Recommendation

Based on the wealth of data collected from MCSC systems, a number of mobile social recommendation techniques and services could be enabled, including place/friend recommendation, itinerary planning, and service/activity suggestion.

Place recommendation. A number of research works have focused on providing personalized place recommendations by exploring mobile crowd sensing data. One approach is to leverage the historical location trajectories recorded by mobile devices for recommendation. For instance, by understanding the location history of different individuals, GeoLife [Zheng et al. 2011] measured the similarity among users, based on which a personalized place recommendation service was provided. Another approach is to exploit user-generated data in LBSNs to recommend venues and friends. For example, Ye et al. [2011] developed a place recommendation service based on the data from FourSquare. They found that the geographical influence among POIs played an important role in user check-in behaviors, and thereby developed a collaborative recommendation algorithm that took into account a combination of user preference, social influence, and geographical influence. Yang et al. [2013] also developed a personalized place recommendation system by leveraging both user check-ins and user comments from FourSquare, where they proposed a hybrid preference model that used both place preferences and user sentiments (learned from user comments).

Itinerary planning. Different from place recommendation, itinerary planning can recommend travel routes to tourists under given constraints (e.g., time budget, user preference). For instance, Zheng and Xie [2011] developed a travel recommendation service that could suggest travel sequences under given spatial constraints. Yu et al. [2014] developed a travel package recommendation system that helped users make travel plans by leveraging the data from LBSNs. It could generate personalized travel packages by considering user preferences, POI characteristics, and temporal-spatial constraints (such as travel time and start location).

Service/activity suggestion. Another research trend is to provide adaptive service and activity suggestions to mobile users. For example, T-Finder [Yuan et al. 2013] was a recommending system that could guide taxi drivers to the places where passengers could more easily be picked up. Zhang et al. [2013] focused on effectively recommending groups to users for participating in offline social events and proposed a unified recommendation model that considered location features, social features, and implicit patterns simultaneously. GroupMe [Guo et al. 2014c] was a mobile social activity aid system that facilitated group suggestion in social activity organization (e.g., having a dinner party, sporting together), leveraging a combination of mobile sensing and social graph mining techniques.

3.6. Healthcare

With the advent of the era of an aging society, healthcare is becoming a more and more important challenge. Based on the wealth of data collected from MCSC systems, a number of health monitoring and management services could be enabled.

Public health monitoring. MCSC can facilitate the monitoring of disease outbreaks. For instance, by mining health-related search queries (which can be localized by their IP address), Google researchers can estimate the level of influenza-like illnesses in America [Ferguson et al. 2006]. Similar to influenza, malaria is another public health concern, and human movements may impact its transmission. Wesolowski et al. [2012]

reported the use of large-scale mobile phone location data and malaria prevalence information from Kenya to study their implicit links.

Personal well-being management. MCSC also facilitates personal well-being management by logging user daily activity trajectories. For example, Neat-o-Games detected human movement information (e.g., how much a user walked in the past hour) and shared it in an online group race game to motivate more exercises [Fujiki et al. 2008]. Dietary patterns have been recognized as contributing factors to many chronic diseases. The DietSense system [Reddy et al. 2007] allowed people to photograph and share their dietary choices and get suggestions from online experts for weight loss.

3.7. Public Safety

Public safety refers to the detection of or protection from social or natural events (e.g., crimes, disasters) that could endanger the safety of average citizens.

Crime prevention and investigation. Recently, user-contributed data has been used for crime prevention [Sheth 2009]. For instance, Ballesteros et al. [2014] showed that data collected by geo-social networks bears relations with crimes and proposed iSafe, which is able to evaluate the safety of users based on their spatial and temporal dimensions. Similarly, by analyzing the large number of geo-tagged Twitter messages posted from mobile devices, Lee et al. [2011] proposed a method to detect unusually crowded places (e.g., a terrorist activity). The spread of technology from security cameras to smartphones in every pocket has proved helpful to criminal investigations. It has been reported that, during the Boston Marathon explosion event in April 2013, photos and videos shot by onlookers after the explosion were used as evidence in the investigation by the FBI [Fowler and Schectman 2013].

Disaster management and relief. Mobile-phone-based participatory sensing systems can also be used for assisting in disaster relief [Consulting 2009]. A powerful demonstration of MCSC to disasters came after the January 2010 earthquake in Haiti [Bengtsson et al. 2011]. By analyzing large-scale mobile phone user data on pre- and postearthquake movement behaviors, they built a model that could predict community responses to future disasters. Similarly, it was reported in Sakaki et al. [2010] that Twitter could give a near-real-time report of earthquakes of a region by analyzing geo-tagged user posts.

4. CONCEPTUAL FRAMEWORK FOR MCSC

Based on the elaboration of MCSC characters and applications, we propose a reference architecture to illustrate the key functional blocks and explain the key techniques of MCSC. It is intended to be the starting point that advances this new research area. Figure 2 shows the proposed architecture, which consists of five layers: *crowd sensing*, *data transmission*, *data collection*, *crowd data processing*, and *applications*.

(1) *Crowd sensing*. This layer involves heterogeneous data sources for mobile crowd sensing, including the sensory information from mobile devices and user-contributed data from mobile Internet applications. Considering the significance of security and privacy issues in MCSC, it is important that participants be able to determine what kind of information to publish and whom to share with. To this end, *access control* becomes an important function of the participatory sensing clients.

(2) *Data transmission*. In MCSC, sensed data from mobile devices should be shipped to the backend server for further processing. MCSC applications should make data uploading tolerant of transient networking environments and inevitable network interruptions. Therefore, data forwarding and routing protocols become significant for data transmission. If the network infrastructure is not available, the data should be transmitted to the destination in a pure opportunistic networking manner [Conti

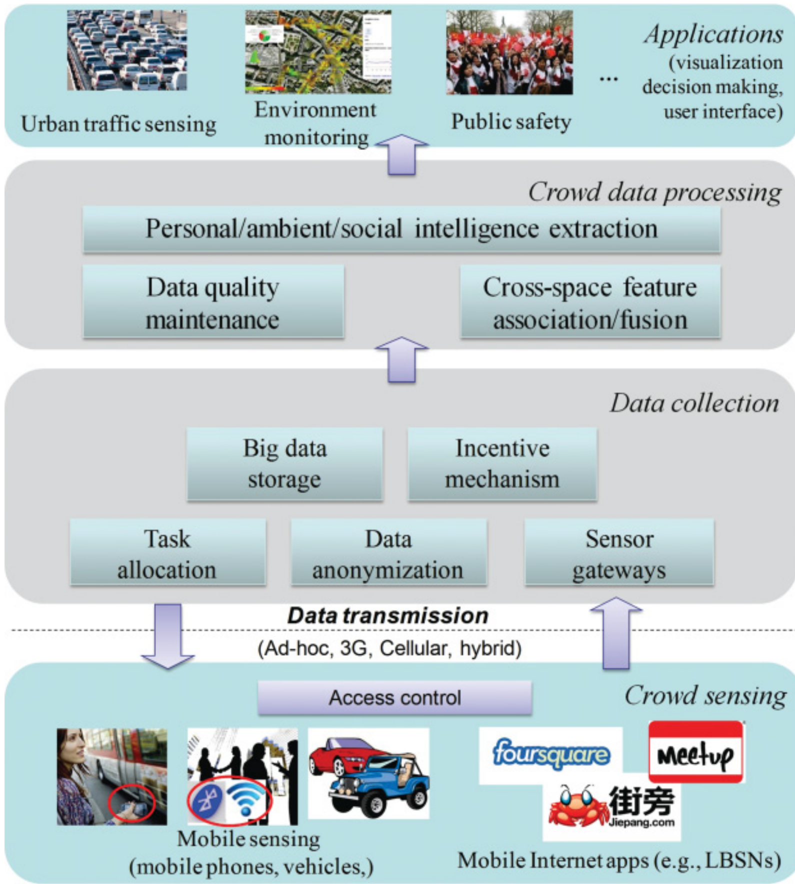

Fig. 2. A reference architecture for MCSC.

and Kumar 2010]. The cooperation of heterogeneous networking nodes to enhance the performance of data transmission is also important to this layer [Guo et al. 2013].

(3) *Data collection infrastructure*. This layer gathers data from selected sensor nodes and provides privacy-preserving mechanisms for data contributors. The following components are involved:
    (a) *Task allocation*. It can analyze a sensing task from an application requester and assign it to a selected number of human nodes in terms of specified requirements, such as sampling contexts (e.g., time, location), device capability, user willingness, and the given budget (e.g., monetary cost, time limit).
    (b) *Sensor gateways*. It provides a standard approach (e.g., the web service techniques used in the SenseWeb [Kansal et al. 2007]) to facilitate data collection from various crowdsensing sources; that is, sensor gateways serve all the top-level components (e.g., data processing and applications) by supplying a unified interface.
    (c) *Data anonymization*. One major concern when sharing personal data is privacy. In addition to the access control function at local stakeholders, this component also supports privacy protection by providing anonymization mechanisms before data is published.

(d) *Incentive mechanism*. This component provides strategies for incentives and reputation to data contributors.
(e) *Big data storage*. The collected data in MCSC systems has two characters: large scale and multimodality. First, the volume of data to be stored and managed is so large and complex (e.g., collecting pollution information at the city scale) that it becomes more and more inefficient to process using existing data management and processing approaches and tools. Second, the characteristics and attributes of different types of sensors usually vary a lot, leading to big differences in the accuracy of crowd sensing. Therefore, in order to boost further processing (e.g., learning and reasoning), the raw data collected from different sensors must be first transformed and represented in a unified manner, for example, based on the same vocabulary or ontology [Zhang et al. 2011].

(4) *Crowd data processing*. It aims to extract high-level intelligence from the raw sensory data by leveraging a wide variety of machine-learning and logic-based inference techniques. In other words, the focus of crowd data processing is to discover frequent data patterns to obtain the three dimensions of crowd intelligence at an integrated level.
(a) *Data processing architecture*. Instead of a purely *centralized* or *self-supported* method (see Section 2.5), we advocate the *hybrid* data processing architecture and propose the Hybrid Data Processing (HDP) solution. In HDP, some of the data processing tasks are allowed to execute on mobile devices to fulfill local perception (e.g., analyzing individual behavior on a smartphone), and after that, local results will be transmitted to servers for further processing. By adopting such a hybrid data processing approach, the communication cost between clients and servers can be notably cut down, and the resilience of the entire network will be enhanced.
(b) *Data quality maintenance*. The data from different contributors has distinct quality and creditability and is often redundant. Therefore, quality measurement and data selection metrics are needed to preprocess the data and eliminate the data with low quality.
(c) *Cross-space feature association/fusion*. The mobile crowd data can be collected from both offline and online spaces. This component studies approaches for cross-space data association and complementary feature fusion [Guo et al. 2014b].
(d) *Crowd intelligence extraction*. It aims to extract the three types of crowd intelligence, that is, user awareness, ambient awareness, and social awareness, by applying various data processing techniques.

(5) *Applications*. This layer consists of different types of applications and services that could be enabled by MCSC. Associated key functions include data visualization, user interface, and so on. Visualization (e.g., mapping, animation, graphing) displays the crowd computing results in a legible format to the users. User interface is designed for the interaction between humans and machines. Both of them facilitate decision making and knowledge sharing to users (e.g., government officials, citizens, service providers).

5. TOWARDS HYBRID HUMAN–MACHINE SYSTEMS

Research on the combination of human and machine intelligence has a long history. In 1950, Alan Turing proposed that "The idea behind digital computers may be explained by saying that these machines are intended to carry out any operations which could be done by a human computer" [Turing 1950]. He also put forward the so-called Turing test to examine the intelligence of an agent. It represents that human intelligence and machine intelligence have been interlinked since the very beginning of AI research.

Similarly, Licklider [1960] also presented the idea that humans and computers can work together in complementary roles. Recently, there have been studies about having humans in the loop in cyber-physical systems [Schirner et al. 2013], which can measure the cognitive activity of a human through body and brain sensors and then use the "inferred" intent for physical environment controlling. Without using any intrusive sensors, MCSC employs the sensing capabilities from the companioned devices (e.g., smartphones, wearable devices) of humans as well as the computing ability of humans. It aims to solve the large-scale sensing and computing problems by having humans in the loop. The reason is that human and machine intelligence often show distinct strengths and weaknesses in MCSC systems.

—*Human intelligence (HI)*. Knowledge, cognition, perception, and social interaction are general abilities of human beings. With these, humans can have deep context and understanding of the sensing tasks. However, they are limited in memory and speed. Also, people vary in quality, and they often introduce errors or low-quality data in MCSC.
—*Machine intelligence (MI)*. Machines are powerful in large-scale storage and computing. Advanced data mining and machine-learning algorithms also enable automatic knowledge discovery and event/society understanding. However, there are still numerous problems that cannot be well addressed by machines.

By combining the intelligence of crowds and computing machinery, MCSC allows the creation of hybrid human–machine systems. These hybrid systems enable applications and experiences that neither crowds nor machines could support alone. As far back as Ivan Sutherland's [1964] Sketchpad, human–computer interaction has been structured around a tradeoff between user control and system automation. The same is true for MCSC, where we should investigate how to design the MCSC system by mixing human and machine intelligence—a question that has not yet been solved. More specifically, with advanced digital technologies, the Internet of Things, and mobile social networks, we are living in a merging world consisting of cyber, physical, and social spaces [Wang 2010]. The problems to solve and the tasks to complete in such a hyperspace are much more complex and may require both human and machine intelligence for achieving better efficiency and effectiveness, which raises new research challenges and opportunities.

Figure 3 gives a description of our vision on the potential combination of human and machine intelligence in MCSC systems. Specifically, according to the MCSC framework shown in Figure 2, there are four functional layers: *crowd sensing*, *data transmission*, *data collection*, and *crowd data processing*. While the *data collection* layer mainly depends on machine intelligence, the combination of human and machine intelligence can take complementary roles in terms of their distinct abilities over the other three layers. For example, in the crowd sensing layer, people can understand and execute the tasks using their knowledge and cognition abilities. Machines, nevertheless, can decompose complex tasks and allocate them to proper human nodes and further provide platforms for information sharing (e.g., user-contributed data in MSNs). In the data transmission layer, user movement and interaction facilitate the development of hybrid networking protocols (e.g., a combination of opportunistic networking and infrastructure-based networking) in MCSC. In the data processing layer, the introduction of human cognition and expert knowledge can attain higher efficiency and accuracy than pure machine processing (e.g., classification, reasoning), especially for the situations that are difficult to tackle by existing machine intelligence techniques.

There have recently been a few studies that try to leverage the complementary strengths of human and machine intelligence in MCSC. On one hand, crowdsourcing applications derive values by letting participants perform tasks that request a certain

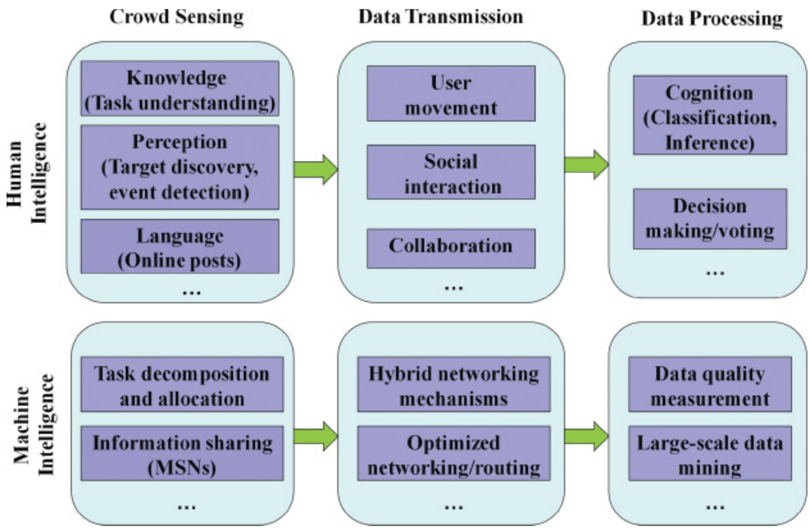

Fig. 3. MCSC: Hybrid human–machine systems.

level of human cognition and cannot be mimicked by computers [Marcus et al. 2011]. For example, DietSense [Reddy et al. 2007] used both automatic image processing techniques and manual image review, due to the complexity or ambiguity of the recognition tasks. Similarly, to obtain the minimum number of verification tasks, CrowdER [Wang et al. 2012] developed a two-tiered hybrid human–machine approach for creating batched tasks, where machines are used to do an initial, coarse pass over all the data, and people are used to verify only the most likely matching pairs. On the other hand, machine intelligence is important and often employed for assisting humans in their decision making and future prediction. For instance, Kamar et al. [2012] investigated how machine-learning techniques can be utilized to explore the complementary capabilities of humans and machineries to fulfill crowdsourcing tasks, for example, providing guidance on recruiting participants and assigning tasks based on certain metrics so that the utilities of large-scale crowdsourcing tasks can be optimized.

Although there have been pioneering studies and applications on combining human and machine intelligence, it is believed that research in this field is still in its infant stage. There are many challenging issues such as how to take the best of machine intelligence, the best of human intelligence, and the best of human–computer interfaces to create “blended” human–computer systems with unprecedented capabilities. Specifically, human and machine intelligence can be combined within or across the three layers of MCSC systems in different patterns. We present a vision on the potential combination patterns of human intelligence and machine intelligence in MCSC systems in Figure 4. Three potential patterns, *sequential*, *parallel*, and *iterative*, are identified, but there should be more to be explored.

Given a complex crowd sensing task, for example, machines can decompose it into a set of subtasks and allocate them to proper human nodes in the data collection layer. Afterward, people can understand and execute the subtasks using their knowledge and cognition abilities in the crowd sensing layer. While this example belongs to *sequential* combination, human and machine intelligence can also be combined in a *parallel* manner. Still taking the accomplishment of a complex crowd sensing task as an example, human nodes and machines (e.g., static sensing nodes) might have complementary sensing abilities and need to work in a parallel way to fully capture the required

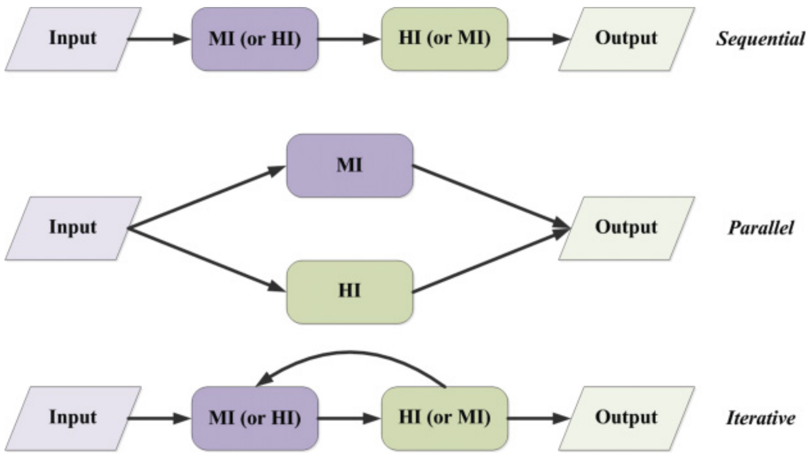

Fig. 4. Potential combination patterns of human and machine intelligence in MCSC systems.

information. Similarly, the aforementioned DietSense system [Reddy et al. 2007] also adopted the *parallel* combination approach, which uses both automatic image processing techniques and manual image review in recognition tasks. Furthermore, human and machine intelligence can be combined *iteratively* as well. For instance, in the data processing layer, humans and machines can work iteratively to achieve higher efficiency and accuracy, especially in case of situations that are difficult to be tackled by existing machine intelligence techniques.

In addition to the vision in Figure 4 and the previous studies, we have two more suggestions for the design of MCSC systems. First, the combination of human and machine intelligence should be "application centric," which means that we should create systems that dynamically trade off human and machine intelligence in terms of application needs. Second, an investigation of formal models and design patterns for crowd computing systems should be studied, which may make use of multidisciplinary knowledge, including social science, management, computer science, and so on.

6. LIMITATIONS, CHALLENGES, AND OPPORTUNITIES

We have presented the unique characters and potential of MCSC. However, translating the MCSC vision into a practical system entails a range of challenges. In this section, we will give a summary of the key technical challenges and research opportunities of this emerging research area.

6.1. Sensing with Human Participation

Compared with traditional static, centrally controlled sensor networks, the involvement of mobile human volunteers in gathering, analyzing, and sharing local knowledge in an interactive sensing infrastructure leads to a number of new challenges.

(1) *Task allocation and data sampling*. In MCSC, a swarm of highly volatile mobile sensors can potentially provide coverage where no static sensing infrastructure is available. Nevertheless, because of a potentially large population of mobile nodes, a sensing task must identify which node(s) may accept a task. A set of criteria should be considered in filtering irrelevant nodes, such as the specification of a required region (e.g., a particular street) and time window, acceptance conditions (e.g., for a traffic-condition capture task, only the phones out of users' pockets and with good illumination conditions can satisfy requirements), device capabilities,

and termination conditions (e.g., sampling period). Some preliminary studies on these issues have already been initiated. For example, in Cornelius et al. [2008], a task description language called AnonyTL was proposed to specify the sample context for a sensing task. Reddy et al. [2010] proposed a recruitment framework to identify well-suited participants for data collections based on spatial-temporal availability and participation habits. Similarly, Zhang et al. [2014b] proposed a participant selection framework named CrowdRecruiter, which operates on top of the energy-efficient Piggyback Crowdsensing (PCS) task model and minimizes incentive payments by selecting a small number of participants while still satisfying probabilistic coverage constraints. However, improving the efficiency of the decision-making process in task assignment and data sampling necessitates further efforts. Successful approaches to task assignment in online crowdsourcing markets can be explored [Ho and Vaughan 2012].

(2) *Human grouping*. Interactions among volunteers are necessary during the sensing process but absent in most existing crowd sensing systems. For example, Vukovic [2009] claimed that one of the research challenges in crowd sensing is “designing a mechanism for virtual team formation, incorporating not only skill-set, but also discovered social networks.” Lane [2012] also believed that crowdsourcing misses automated methods to identify and characterize user communities. The interaction among users also enhances the data quality in MCSC. Therefore, grouping users and facilitating the interaction among them should be a challenge of MCSC. Key techniques to address this include community creation approaches, dynamic group formation metrics, social networking methods, and so on. For example, to identify the people who are involved in the same social event, MoVi [Bao and Choudhury 2010] proposed a multidimensional sensing approach, where a combination of visual and acoustic ambience of phones was used.

(3) *Coverage, reliability, and scalability*. MCSC is akin to the event coverage in conventional sensor networks. To design and deploy successful MCSC systems, the relation between the event coverage and the number of participants should be studied. Possible issues involve spatial-temporal coverage, the impact of user skills/preferences to task coverage, and so on. For example, Chon et al. [2013] found that the amount of people’s place visits follows a power-law distribution. A generative model of location coverage was developed based on user population and city characteristics, which can be used to predict how many participants are sufficient to achieve a certain level of coverage.

6.2. Incentive Mechanisms

Incentive is another challenge to the human involvement in MCSC. While sensing devices are usually possessed and administrated by a single authority in traditional sensor networks, they belong to different individuals with diverse interests in MCSC. In order to sense, process, and collect the desired data, participants have to make either implicit efforts (e.g., energy and monetary costs) or explicit efforts (e.g., give some input or assessments). Without strong incentives, individuals may not be willing to participate in the sensing task with the cost because of their own limited resources. Therefore, an efficient incentive mechanism is essential for the success of MCSC applications [Luo et al. 2014]. Similar to other crowdsourcing systems [Quinn and Bederson 2011; Yuen et al. 2011], broadly two types of incentives can be used in MCSC applications: intrinsic incentives or financial incentives.

Intrinsic incentives include interest (where the volunteers are willing to help when they think the task is interesting and important), enjoyment (making the MCSC task an entertainment activity, such as a game), and social/ethical (where the participators are attracted by the chance of receiving public recognition). While these intrinsic incentives work well in certain types of MCSC applications (e.g., environment monitoring, location services), financial incentives are probably the easiest way to motivate user participation in almost all types of MCSC applications. In MCSC systems with financial incentives, the financial rewards to their participates for performing sensing or communication tasks could be money, virtual cash, or credits (which are redeemable for later services or online goods). However, once money is involved, the participants are more likely to deceive the system to get more financial gains. Therefore, how to provide valuable incentive mechanisms that enable honest contributions in MCSC becomes a critical challenge. Recently, several game theory approaches [Yang et al. 2012; Lee and Hoh 2010; Huangfu et al. 2013] have been proposed for MCSC to motivate and reward truthful contributions. These game theory methods are typically based on auction mechanisms but are rather complex to implement in a fully distributed and time-evolving system. Therefore, for a highly dynamic MCSC system, there is still a need for new incentive and pricing mechanisms to attract, encourage, and reward truthful and high-quality sensing data contributors.

6.3. Data Delivery in Transient Networks

How to ship the sensed data from distributed participants to the backend server is another challenge due to a variety of MCSC characteristics, such as the low bandwidth of wireless communication, frequent network partitioning caused by human mobility, and large number of energy-constrained devices. While this is a well-known research challenge in both wireless sensor networks and general mobile systems, MCSC adds new requirements on scalability. In particular, the following four research issues need to be addressed for MCSC:

(1) *Robust data delivery among highly mobile devices*. Though many data forwarding and routing protocols have been developed for mobile ad hoc networks and mobile opportunistic networks in the past decade, it is still a hard problem to achieve robust and reliable data delivery with a large amount of mobile devices without infrastructure. Recently there have been emerging social-based approaches [Zhu et al. 2013], which attempt to exploit social behaviors of participants to make better forwarding and routing decisions, though the research results are mostly obtained from simulation environments and the large-scale, real-world deployment is a necessity to validate the usefulness of such approaches.
(2) *Tradeoffs between communication and processing via localized analytics*. In many MCSC applications, certain localized analytics are performed on raw sensing data at the individual device level. By doing so, MCSC systems can consume less network bandwidth than directly transmitting raw sensor readings. The challenge is how to make a balance between the energy cost on local computing and the data transmission cost [Xiong et al. 2014].
(3) *Distributed caching and replication schemes*. In order to make the storing and retrieving of MCSC data effective, efficient, and robust in a large-scale mobile system, it is essential to develop new distributed data caching and replication schemes. Such schemes should be integrated with the data processing and management issues we discuss next.
(4) *Hybrid networking protocols*. The coexistence of heterogeneous network connections is a basic feature of MCSC. Combining the complementary merits of heterogeneous networks provides opportunities to develop improved networking protocols for MCSC. Pioneering studies on this have been done in Guo et al. [2013] and Ding and Xiao [2010].

6.4. Data Redundancy, Quality, and Inconsistency

In MCSC, there can be multiple participants involved in the same sensing activity, for example, sensing the traffic information in an intersection. One of the issues caused by multiparticipant sensing is data redundancy. In other words, to accomplish a particular sensing task, it is important to smartly select data from multiple available contributors. Note that for the same task, nearby mobile sensing devices may have various sensing qualities, which could be caused by the mobility of devices and the differences in energy levels or communication channels. Certain quality estimation and prediction methods are thus necessary to evaluate the quality of sensing data, and statistical processing can be used to identify outliers. For example, in Movi [Bao and Choudhury 2010], a view selection module was developed to select videos of high quality, leveraging multidimensional sensing to obtain the "best view" of the recorded event, for example, accelerometer readings for selecting stable images and light intensity to deprioritize darker pictures.

Another potential issue caused by "redundant" sensing is data inconsistency. For example, due to the differences in sensing and computing capabilities, a set of collocated smartphones running the same algorithm (e.g., sound-based social context recognition) and sensing the same event can obtain different inference results (e.g., a party or a meeting), thereby causing the problem of inconsistency. To handle such problems, a collaborative method was put forward in Miluzzo et al. [2010]. Furthermore, a more complex issue is the inconsistence of semantics derived from multimodal sensory data (e.g., audio clips, images, videos, texts). For example, while both the audio and video data can be used to predict the social context of a user, the inference results can be different. Further studies should be done to address the inconsistency caused by the multimodel data contributed by crowds.

6.5. Cross-Space, Heterogeneous Crowdsourced Data Mining

The strength of MCSC relies on the usage of crowdsourced data from both physical and virtual societies. The development of the Internet of Things and mobile Internet bridges the gap between the physical space and the cyber space. For the same sensing object (e.g., a social gathering in a street corner), it will interact with both spaces and leave fragmented data in each space, making the information obtained from different communities (online or offline) differ. For instance, we can learn social relationships from online social networks and infer group activities and interaction behaviors using mobile phone sensing in the real world. Obviously, the complementary nature of heterogeneous communities will bring new opportunities to develop new human-centric services. Therefore, we should integrate the information from heterogeneous data sources to attain a comprehensive picture about the sensing object. A typical example is the Social Contact Manager, which is a system that leverages a combination of the data collected from the web and smartphones to enable intelligent social contact information management and associative search [Guo et al. 2014].

As people live and traverse among different communities, the properties (social, geographical, thematic) of distinct social networks are thus interweaved and highly associated. For instance, more and more evidence shows that when we want to model the behavior of a person, the best predictor is often not based on the person himself or herself, but rather on his or her social links [Cho et al. 2011]. Meanwhile, the correlation between human social ties and geographic coincidences has also been investigated in Eaglea et al. [2007] and Guo et al. [2014a]. The event-based social network [Liu et al. 2012] is an emerging mobile social network that connects online and offline social activities. By exploring a set of features that connect the physical and the cyber space, our recent work [Du et al. 2014] proposes an approach to predict the activity attendance behaviors of users in Douban.[6] To deeply understand the sensing object and predict its behaviors, it is crucial to study the correlation among the properties learned from different social communities. Further discussions about heterogeneous community sensing and association have been presented in Guo et al. [2014b].

In short, with the increase in the large-scale, interlinked data collected from heterogeneous communities, advanced techniques on complex network modeling, data mining, data association and aggregation, and semantic fusion will become more and more important.

6.6. Trust, Security, and Privacy

The involvement of human participation in crowd sensing also brings forth certain trust issues. MCSC participants are likely to provide incorrect or even fake data to the system. For example, incorrect recordings might be collected when mobile devices are improperly placed by the participants; for example, one may put his or her phone in the pocket when assigned with a noise sensing task. Meanwhile, for their own benefit, malicious users may intentionally pollute the sensing data. The lack of control mechanisms to guarantee source validity and data accuracy can lead to information credibility issues. Thereby, we have to develop trust preservation and abnormal detection technologies to ensure the quality of the obtained data. For example, Huang et al. [2010] proposed a reputation system based on the Gompertz function, which is able to estimate the trustworthiness of the collected data.

To motivate user participation, an MCSC system must be capable of providing effective privacy protection mechanisms so that participants can conveniently and safely share high-quality data using their devices. The following two ways are promising:

(1) *Local versus remote data processing*. One possible way to preserve privacy is to upload the processed data rather than the original raw data (as the HDP solution presented in Section 5). For example, in case an MCSC application needs to collect the environmental noise using microphones, by leveraging some phone-based noise identification methodologies, we might only need to upload the obtained results (i.e., noise level) rather than the raw audio files to the cloud. While such an approach avoids the disclosure of people's conversation, phone-based algorithms would incur severe energy consumption. Meanwhile, it's more efficient to process audio files in the cloud than locally on the phone, but this requires implementing privacy-preserving data mining techniques for remote processing.
(2) *Privacy-aware sensing model and architecture*. To effectively preserve the privacy of a huge amount of MCSC participants, not only methodology efforts but also systematic studies are needed. In other words, a privacy-aware architecture should be provided to support the development of MCSC applications. One such effort is the AnonySense architecture proposed in Cornelius et al. [2008], which supported the development of privacy-aware applications based on crowd sensing.

7. CONCLUSIONS

We have presented MCSC, a cross-space, heterogeneous, crowdsourced sensing paradigm for large-scale sensing and computing. Layered on participatory sensing, it presents two unique features. First, it leverages both sensed data from mobile devices and user-contributed data from mobile social networking services. Second, it propels the fusion of machine and human intelligence in both sensing and the computing process. We clarify the main characters of MCSC by having humans in the loop for large-scale sensing and computing, including human-powered sensing, human-centered

[6] http://www.douban.com/.

computing, transient networking, and crowd data processing and intelligence extraction. MCSC will enhance or nurture numerous application areas, such as environment monitoring, intelligent transportation, urban sensing, mobile social recommendation, and so on. We have made a summary and comparison of existing projects/studies over a derived taxonomy of MCSC. Based on the reviewing of existing systems and the identified characters, we have proposed a reference framework for developing MCSC systems. The success of MCSC relies on the deep fusion of human–machine intelligence, so we discuss the potential techniques and approaches that can be leveraged. We finally identify several key challenging areas and research opportunities of MCSC.

Over the longer term, MCSC is able to stimulate basic research across a number of areas, of which there are several we consider crucial. First, we are living in a world consisting of cyber, physical, and social spaces, and each sensing object leaves fragmented, incomplete data in the three spaces. It is thus important to explore approaches for aggregating and fusing the cross-space, complementary data for urban/society dynamics understanding. Second, we should study the fusion of human and machine intelligence in the whole crowd sensing life cycle, from pervasive sensing, human-enhanced data transmission to cross-space data processing. Third, some of the ethical factors, such as inventiveness and user privacy, should be the fundamental building blocks of MCSC. Finally, the success of MCSC relies on the usage of multidisciplinary knowledge, including social science, cognitive science, economics, computing science, and so on. This should be considered in the design of MCSC techniques and systems.

ACKNOWLEDGMENTS

The authors would like to thank all the colleagues for their discussion and suggestions.

REFERENCES

Giovanni Acampora, Diane J. Cook, Parisa Rashidi, and Athanasios V. Vasilakos. 2013. A survey on ambient intelligence in healthcare. *Proceedings of the IEEE* 101, 12, 2470–2494.

Jaime Ballesteros, Bogdan Carbunar, Mahmudur Rahman, Naphtali Rishe, and S. S. Iyengar. 2014. Towards safe cities: A mobile and social networking approach. *IEEE Transactions on Parallel and Distributed Systems* 25, 9, 2451–2462.

Xuan Bao and Romit Choudhury. 2010. MoVi: Mobile phone based video highlights via collaborative sensing. In *Proceedings of the 8th International Conference on Mobile Systems, Applications, and Services (MobiSys'10)*. ACM, New York, NY, 357–370.

Linus Bengtsson, Xin Lu, Anna Thorson, Richard Garfield, and Johan von Schreeb. 2011. Improved response to disasters and outbreaks by tracking population movements with mobile phone network data: A post-earthquake geospatial study in haiti. *PLoS Medicine* 8, 8, e1001083.

Daren C. Brabham. 2008. Crowdsourcing as a model for problem solving: An introduction and cases. *Convergence: The International Journal of Research into New Media Technologies* 14, 1, 75–90.

Jeff Burke, Deborah Estrin, Mark Hansen, Andrew Parker, Nithya Ramanathan, Sasank Reddy, and Mani B. Srivastava. 2006. Participatory sensing. In *Proceedings of the International Workshop on World-Sensor-Web*. 1–5.

Francesco Calabrese, Massimo Colonna, Piero Lovisolo, Dario Parata, and Carlo Ratti. 2011. Real-time urban monitoring using cell phones: A case study in Rome. *IEEE Transactions on Intelligent Transportation Systems* 12, 1, 141–151.

Chao Chen, Daqing Zhang, Zhi-Hua Zhou, Nan Li, Tulin Atmaca, and Shijian Li. 2013. B-Planner: Night bus route planning using large-scale taxi GPS traces. In *Proceedings of the IEEE International Conference on Pervasive Computing and Communications*. 225–233.

Wai Chen, Ratul K. Guha, Taek Jin Kwon, John Lee, and Yuan-Ying Hsu. 2011. A survey and challenges in routing and data dissemination in vehicular ad hoc networks. *Wireless Communications and Mobile Computing* 11, 7, 787–795.

Eunjoon Cho, Seth A. Myers, and Jure Leskovec. 2011. Friendship and mobility: User movement in location-based social networks. In *Proceedings of the 17th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining (KDD'11)*. ACM, New York, NY, 1082–1090.

Yohan Chon, Nicholas D. Lane, Yunjong Kim, Feng Zhao, and Hojung Cha. 2013. Understanding the coverage and scalability of place-centric crowdsensing. In *Proceedings of the ACM International Joint Conference on Pervasive and Ubiquitous Computing (UbiComp'13)*. ACM, New York, NY, 3–12.

Yohan Chon, Nicholas D. Lane, Fan Li, Hojung Cha, and Feng Zhao. 2012. Automatically characterizing places with opportunistic crowdsensing using smartphones. In *Proceedings of the ACM Conference on Ubiquitous Computing (UbiComp'12)*. ACM, New York, NY, 481–490.

Delphine Christin, Andreas Reinhardt, Salil S. Kanhere, and Matthias Hollick. 2011. A survey on privacy in mobile participatory sensing applications. *Journal of System Software* 84, 11, 1928–1946.

Vital Wave Consulting. 2009. *mHealth for Development: The Opportunity of Mobile Technology for Healthcare in the Developing World*. UN Foundation-Vodafone Foundation Partnership.

Marco Conti and Mohan Kumar. 2010. Opportunities in opportunistic computing. *Computer* 43, 1, 42–50.

Cory Cornelius, Apu Kapadia, David Kotz, Dan Peebles, Minho Shin, and Nikos Triandopoulos. 2008. Anonysense: Privacy-aware people-centric sensing. In *Proceeding of the 6th International Conference on Mobile Systems, Applications, and Services*. ACM, New York, NY, 211–224.

Costandinos Costa, Christos Laoudias, Demetrios Zeinalipour-Yazti, and Dimitrios Gunopulos. 2011. SmartTrace: Finding similar trajectories in smartphone networks without disclosing the traces. In *Proceedings of the 29th IEEE International Conference on Data Engineering*. IEEE Computer Society, Los Alamitos, CA, USA, 1288–1291.

Andrew Crooks, Arie Croitoru, Anthony Stefanidis, and Jacek Radzikowski. 2013. Earthquake: Twitter as a distributed sensor system. *Transactions in GIS* 17, 1, 124–147.

Yong Ding and Li Xiao. 2010. SADV: Static-node-assisted adaptive data dissemination in vehicular networks. *IEEE Transactions on Vehicular Technology* 59, 5, 2445–2455.

Rong Du, Zhiwen Yu, Tao Mei, Zhitao Wang, Zhu Wang, and Bin Guo. 2014. Predicting activity attendance in event-based social networks: Content, context and social influence. In *Proceedings of the 2014 ACM International Joint Conference on Pervasive and Ubiquitous Computing*. ACM, New York, NY, 425–434.

Nathan Eaglea, Alex Pentlandb, and David Lazer. 2007. Inferring friendship network structure by using mobile phone data. *PNAS* 106, 36, 15274–15278.

Shane B. Eisenman, Emiliano Miluzzo, Nicholas D. Lane, Ronald A. Peterson, Gahng-Seop Ahn, and Andrew T. Campbell. 2010. BikeNet: A mobile sensing system for cyclist experience mapping. *ACM Transactions on Sensor Networks* 6, 1, Article 6, 39 pages.

Jakob Eriksson, Lewis Girod, Bret Hull, Ryan Newton, Samuel Madden, and Hari Balakrishnan. 2008. The pothole patrol: Using a mobile sensor network for road surface monitoring. In *Proceedings of the 6th International Conference on Mobile Systems, Applications, and Services (MobiSys'08)*. ACM, New York, NY, 29–39.

Neil M. Ferguson, Derek A. T. Cummings, Christophe Fraser, James C. Cajka, Philip C. Cooley, and Donald S. Burke. 2006. Strategies for mitigating an influenza pandemic. *Nature* 7101, 10, 448–452.

Geoffrey Fowler and Joel Schectman. 2013. Citizen surveillance helps officials put pieces together. *The Wall Street Journal*.

Yuichi Fujiki, Konstantinos Kazakos, Colin Puri, Pradeep Buddharaju, Ioannis Pavlidis, and James Levine. 2008. NEAT-o-Games: Blending physical activity and fun in the daily routine. *Computer Entertainment* 6, 2, Article 21, 22 pages.

Benjamin C. M. Fung, Ke Wang, Rui Chen, and Philip S. Yu. 2010. Privacy-preserving data publishing: A survey of recent developments. *ACM Computing Surveys* 42, 4, Article 14, 53 pages.

Raghu K. Ganti, Fan Ye, and Hui Lei. 2011. Mobile crowdsensing: Current state and future challenges. *IEEE Communications Magazine* 49, 11, 32–39.

Wei Gao and Guohong Cao. 2011. User-centric data dissemination in disruption tolerant networks. In *Proceedings of the IEEE International Conference on Computer Communications*. 3119–3127.

Anargyros Garyfalos and Kevin C. Almeroth. 2008. Coupons: A multilevel incentive scheme for information dissemination in mobile networks. *IEEE Transactions on Mobile Computing* 7, 6, 792–804.

Bin Guo, Huihui Chen, Zhiwen Yu, Xing Xie, Shenlong Huangfu, and Daqing Zhang. 2015. FlierMeet: A mobile crowdsensing system for cross-space public information reposting, tagging, and sharing. *IEEE Transactions on Mobile Computing*. DOI:http://dx.doi.org/10.1109/TMC.2014.2385097

Bin Guo, Yunji Liang, Zhiwen Yu, Minshu Li, and Xingshe Zhou. 2014a. From mobile phone sensing to human geo-social behavior understanding. *Computational Intelligence*. DOI:http://dx.doi.org/10.1111/coin.12050

Bin Guo, Zhiwen Yu, Daqing Zhang, Huilei He, Jilei Tian, and Xingshe Zhou. 2014c. Toward a group-aware smartphone sensing system. *IEEE Pervasive Computing* 13, 4, 80–88.

Bin Guo, Zhiwen Yu, Daqing Zhang, and Xingshe Zhou. 2014b. Cross-community sensing and mining. *IEEE Communications Magazine* 52, 8, 144–152.

Bin Guo, Zhiwen Yu, Xingshe Zhou, and Daqing Zhang. 2013. MemPhone: From personal memory aid to community memory sharing using mobile tagging. In *Proceedings of PERCOM Workshops*. 332–335.

Bin Guo, Zhiwen Yu, Xingshe Zhou, and Daqing Zhang. 2014. From participatory sensing to mobile crowd sensing. In *Proceedings of PERCOM Workshops*. 593–598.

Bin Guo, Daqing Zhang, Zhu Wang, Zhiwen Yu, and Xingshe Zhou. 2013. Opportunistic IoT: Exploring the harmonious interaction between human and the internet of things. *Journal of Networking and Computer Applications* 36, 6, 1531–1539.

Bin Guo, Daqing Zhang, Dingqi Yang, Zhiwen Yu, and Xingshe Zhou. 2014. Enhancing memory recall via an intelligent social contact management system. *IEEE Transactions on Human-Machine Systems* 44, 1, 78–91.

Bin Guo, Daqing Zhang, Zhiwen Yu, Yunji Liang, Zhu Wang, and Xingshe Zhou. 2013. From the internet of things to embedded intelligence. *World Wide Web* 16, 4, 399–420.

Chien-Ju Ho and Jennifer W. Vaughan. 2012. Online task assignment in crowdsourcing markets. In *Proceedings of the 26th AAAI Conference on Artificial Intelligence*. AAAI, 45–51.

Jeff Howe. 2006. The rise of crowdsourcing. *Wired* 14, 6.

Kuan Lun Huang, Salil S. Kanhere, and Wen Hu. 2010. Are you contributing trustworthy data?: The case for a reputation system in participatory sensing. In *Proceedings of the 13th ACM International Conference on Modeling, Analysis, and Simulation of Wireless and Mobile Systems (MSWIM'10)*. ACM, New York, NY, 14–22.

Shenlong Huangfu, Bin Guo, Zhiwen Yu, and Dongsheng Li. 2013. Using the model of markets with intermediaries as an incentive scheme for opportunistic social networks. In *Proceedings of the 10th IEEE International Conference on Ubiquitous Intelligence and Computing*. 142–149.

Ece Kamar, Severin Hacker, and Eric Horvitz. 2012. Combining human and machine intelligence in large-scale crowdsourcing. In *Proceedings of the 11th International Conference on Autonomous Agents and Multiagent Systems (AAMAS'12)*. International Foundation for Autonomous Agents and Multiagent Systems, Richland, SC, 467–474.

Aman Kansal, Suman Nath, Jie Liu, and Feng Zhao. 2007. SenseWeb: An infrastructure for shared sensing. *IEEE MultiMedia* 14, 4, 8–13.

Dmytro Karamshuk, Chiara Boldrini, Marco Conti, and Andrea Passarella. 2011. Human mobility models for opportunistic networks. *IEEE Communications Magazine* 49, 12, 157–165.

Dmytro Karamshuk, Anastasios Noulas, Salvatore Scellato, Vincenzo Nicosia, and Cecilia Mascolo. 2013. Geo-spotting: Mining online location-based services for optimal retail store placement. In *Proceedings of the 19th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining (KDD'13)*. ACM, New York, NY, 793–801.

Emmanouil Koukoumidis, Li-Shiuan Peh, and Margaret Rose Martonosi. 2011. SignalGuru: Leveraging mobile phones for collaborative traffic signal schedule advisory. In *Proceedings of the 9th International Conference on Mobile Systems, Applications, and Services (MobiSys'11)*. ACM, New York, NY, 127–140.

Nicholas D. Lane. 2012. Community-aware smartphone sensing systems. *IEEE Internet Computing* 16, 3, 60–64.

Nicholas D. Lane, Emiliano Miluzzo, Hong Lu, Daniel Peebles, Tanzeem Choudhury, and Andrew T. Campbell. 2010. A survey of mobile phone sensing. *IEEE Communications Magazine* 48, 9, 140–150.

Neal Lathia, Veljko Pejovic, Kiran K. Rachuri, Cecilia Mascolo, Mirco Musolesi, and Peter J. Rentfrow. 2013. Smartphones for large-scale behavior change interventions. *IEEE Pervasive Computing* 12, 3, 66–73.

Juong-Sik Lee and Baik Hoh. 2010. Sell your experiences: A market mechanism based incentive for participatory sensing. In *Proceedings of the IEEE International Conference on Pervasive Computing and Communications*. 60–68.

Ryong Lee, Shoko Wakamiya, and Kazutoshi Sumiya. 2011. Discovery of unusual regional social activities using geo-tagged microblogs. *World Wide Web* 14, 4, 321–349.

Frank L. Lewis. 2005. *Wireless Sensor Networks*. John Wiley and Sons, 11–46.

Joseph C. R. Licklider. 1960. Man-computer symbiosis. *IRE Transactions on Human Factors in Electronics* HFE-1, 1, 4–11.

Liang Liu, Assaf Biderman, and Carlo Ratti. 2009. Urban mobility landscape: Real time monitoring of urban mobility patterns. In *Proceedings of the 11th International Conference on Computers in Urban Planning and Urban Management*. 1–16.

Xingjie Liu, Qi He, Yuanyuan Tian, Wang-Chien Lee, John McPherson, and Jiawei Han. 2012. Event-based social networks: Linking the online and offline social worlds. In *Proceedings of the 18th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining (KDD'12)*. ACM, New York, NY, 1032–1040.

Yunhao Liu, Yuan He, Mo Li, Jiliang Wang, Kebin Liu, and Xiangyang Li. 2013. Does wireless sensor network scale? A measurement study on GreenOrbs. *IEEE Transactions on Parallel and Distributed Systems* 24, 10, 1983–1993.

Tie Luo, S. S. Kanhere, and Hwee-Pink Tan. 2014. SEW-ing a Simple Endorsement Web to incentivize trustworthy participatory sensing. In *Proceedings of the 2014 11th Annual IEEE International Conference on Sensing, Communication, and Networking (SECON'14)*. 636–644.

Huadong Ma, Dong Zhao, and Peiyan Yuan. 2014. Opportunities in mobile crowd sensing. *IEEE Communications Magazine* 52, 8, 29–35.

Nicolas Maisonneuve, Matthias Stevens, and Bartek Ochab. 2010. Participatory noise pollution monitoring using mobile phones. *Information Polity* 15, 1, 51–71.

Thomas W. Malone, Robert Laubacher, and Chrysanthos Dellarocas. 2009. *Harnessing Crowds: Mapping the Genome of Collective Intelligence*. MIT Sloan Research Paper No. 4732-09. Massachusetts Institute of Technology.

Adam Marcus, Eugene Wu, David Karger, Samuel Madden, and Robert Miller. 2011. Human-powered sorts and joins. *Proceedings of the VLDB Endowment* 5, 1, 13–24.

Nissim Matatov, Lior Rokach, and Oded Maimon. 2010. Privacy-preserving data mining: A feature set partitioning approach. *Information Sciences* 180, 14, 2696–2720.

Emiliano Miluzzo, Cory T. Cornelius, Ashwin Ramaswamy, Tanzeem Choudhury, Zhigang Liu, and Andrew T. Campbell. 2010. Darwin phones: The evolution of sensing and inference on mobile phones. In *Proceedings of the 8th International Conference on Mobile Systems, Applications, and Services (MobiSys'10)*. ACM, New York, NY, 5–20.

Prashanth Mohan, Venkata N. Padmanabhan, and Ramachandran Ramjee. 2008. Nericell: Rich monitoring of road and traffic conditions using mobile smartphones. In *Proceedings of the 6th ACM Conference on Embedded Network Sensor Systems (SenSys'08)*. ACM, New York, NY, 323–336.

Catherine Morency, Martin Trpanier, and Bruno Agard. 2007. Measuring transit use variability with smartcard data. *Transport Policy* 14, 3, 193–203.

Min Mun, Sasank Reddy, Katie Shilton, Nathan Yau, Jeff Burke, Deborah Estrin, Mark Hansen, Eric Howard, Ruth West, and Péter Boda. 2009. PEIR, the personal environmental impact report, as a platform for participatory sensing systems research. In *Proceedings of the 7th International Conference on Mobile Systems, Applications, and Services (MobiSys'09)*. ACM, New York, NY, 55–68.

Anastasios Noulas, Salvatore Scellato, Renaud Lambiotte, Massimiliano Pontil, and Cecilia Mascolo. 2012. A tale of many cities: Universal patterns in human urban mobility. *PLoS ONE* 7, 5, e37027.

Gang Pan, Guande Qi, Zhaohui Wu, Daqing Zhang, and Shijian Li. 2013. Land-use classification using taxi GPS traces. *IEEE Transactions on Intelligent Transportation Systems* 14, 1, 113–123.

Andrea Clementi Francesco Pasquale and Riccardo Silvestri. 2013. Opportunistic MANETs: Mobility can make up for low transmission power. *IEEE/ACM Transactions on Networking* 21, 2, 610–620.

Bratislav Predic, Zhixian Yan, Julien Eberle, Dragan Stojanovic, and Karl Aberer. 2013. ExposureSense: Integrating daily activities with air quality using mobile participatory sensing. In *Proceedings of PERCOM Workshops*. 303–305.

Chuan Qin, Xuan Bao, Romit Roy Choudhury, and Srihari Nelakuditi. 2014. TagSense: Leveraging smartphones for automatic image tagging. *IEEE Transactions on Mobile Computing* 13, 1, 61–74.

Alexander J. Quinn and Benjamin B. Bederson. 2011. Human computation: A survey and taxonomy of a growing field. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI'11)*. ACM, New York, NY, 1403–1412.

Rajib Kumar Rana, Chun Tung Chou, Salil S. Kanhere, Nirupama Bulusu, and Wen Hu. 2010. Ear-phone: An end-to-end participatory Urban noise mapping system. In *Proceedings of the 9th ACM/IEEE International Conference on Information Processing in Sensor Networks (IPSN'10)*. ACM, New York, NY, 105–116.

Sasank Reddy, Deborah Estrin, and Mani Srivastava. 2010. Recruitment framework for participatory sensing data collections. In *Pervasive Computing*, Patrik Floren, Antonio Krger, and Mirjana Spasojevic (Eds.). Lecture Notes in Computer Science, Vol. 6030. Springer, Berlin, 138–155.

Sasank Reddy, Andrew Parker, Josh Hyman, Jeff Burke, Deborah Estrin, and Mark Hansen. 2007. Image browsing, processing, and clustering for participatory sensing: Lessons from a DietSense prototype. In *Proceedings of the 4th Workshop on Embedded Networked Sensors (EmNets'07)*. ACM, New York, NY, 13–17.

Lukas Ruge, Bashar Altakrouri, and Andreas Schrader. 2013. SoundOfTheCity: Continuous noise monitoring for a healthy city. In *Proceedings of PERCOM Workshops*. 670–675.

Takeshi Sakaki, Makoto Okazaki, and Yutaka Matsuo. 2010. Earthquake shakes twitter users - real-time event detection by social sensors. In *Proceedings of the 19th International Conference on World Wide Web (WWW'10)*. ACM, New York, NY, 851–860.

Gunar Schirner, Deniz Erdogmus, Kaushik Chowdhury, and Taskin Padir. 2013. The future of human-in-the-loop cyber-physical systems. *Computer* 46, 1, 36–45.

Amit Sheth. 2009. Citizen sensing, social signals, and enriching human experience. *IEEE Internet Computing* 13, 4, 87–92.

Katie Shilton, Jeff Burke, Deborah Estrin, Mark Hansen, Nithya Ramanathan, Sasank Reddy, Vidyut Samanta, Mani Srivastava, Ruth West, and Jeffrey Goldman. 2009. *Participatory Sensing: A Citizen-Powered Approach to Illuminating the Patterns That Shape Our World*. White paper. Woodrow Wilson International Center for Scholars, Washington, DC.

John A. Stankovic. 2008. Wireless sensor networks. *Computer* 41, 10, 92–95.

Stephen A. Stansfeld and Mark P. Matheson. 2003. Noise pollution: Non-auditory effects on health. *British Medical Bulletin* 68, 5, 243–257.

James Surowiecki. 2005. *The Wisdom of Crowds: Why the Many Are Smarter Than the Few and How Collective Wisdom Shapes Business, Economies, Societies and Nations*. Knopf Doubleday Publisher.

Ivan E. Sutherland. 1964. Sketch pad a man-machine graphical communication system. In *Proceedings of the SHARE Design Automation Workshop (DAC'64)*. ACM, New York, NY, 6.329–6.346.

Lei Tang and Huan Liu. 2010. *Community Detection and Mining in Social Media*. Morgan & Claypool Publishers.

Arvind Thiagarajan, Lenin Ravindranath, Katrina LaCurts, Samuel Madden, Hari Balakrishnan, Sivan Toledo, and Jakob Eriksson. 2009. VTrack: Accurate, energy-aware road traffic delay estimation using mobile phones. In *Proceedings of the 7th ACM Conference on Embedded Networked Sensor Systems (SenSys'09)*. ACM, New York, NY, 85–98.

Alan M. Turing. 1950. Computing machinery and intelligence. *Mind* 59, 433–460.

Md Y. S. Uddin, Hongyan Wang, Fatemeh Saremi, Guo-Jun Qi, Tarek Abdelzaher, and Thomas Huang. 2011. PhotoNet: A similarity-aware picture delivery service for situation awareness. In *Proceedings of the 32nd IEEE International Conference on Real-Time Systems Symposium*. 317–326.

Maja Vukovic. 2009. Crowdsourcing for enterprises. In *Proceedings of the 9th IEEE World Congress on Services*. IEEE Computer Society, Los Alamitos, CA, USA, 686–692.

Fei-Yue Wang. 2010. The emergence of intelligent enterprises: From CPS to CPSS. *IEEE Intelligent Systems* 25, 4, 85–88.

Jiannan Wang, Tim Kraska, Michael J. Franklin, and Jianhua Feng. 2012. CrowdER: Crowdsourcing entity resolution. *Proceedings of the VLDB Endowment* 5, 11, 1483–1494.

Yi Wang, Wenjie Hu, Yibo Wu, and Guohong Cao. 2014. SmartPhoto: A resource-aware crowdsourcing approach for image sensing with smartphones. In *Proceedings of the 15th ACM International Symposium on Mobile Ad Hoc Networking and Computing (MobiHoc'14)*. ACM, New York, NY, 113–122.

Zhu Wang, Daqing Zhang, Xingshe Zhou, Dingqi Yang, Zhiyong Yu, and Zhiwen Yu. 2014. Discovering and profiling overlapping communities in location-based social networks. *IEEE Transactions on Systems, Man, and Cybernetics: Systems* 44, 4, 499–509.

Huimin Wen, Zhongwei Hu, Jifu Guo, Liyun Zhu, and Jianping Sun. 2008. Operational analysis on Beijing road network during the Olympic Games. *Journal of Transportation Systems Engineering and Information Technology* 8, 6, 32–37.

Amy Wesolowski, Nathan Eagle, Andrew J. Tatem, David L. Smith, Abdisalan M. Noor, Robert W. Snow, and Caroline O. Buckee. 2012. Quantifying the impact of human mobility on Malaria. *Science* 338, 6104, 267–270.

Ouri Wolfson, Yu Zheng, and Shuo Ma. 2013. T-share: A large-scale dynamic taxi ridesharing service. In *Proceedings of the IEEE International Conference on Data Engineering (ICDE'13)*. IEEE Computer Society, Washington, DC, 410–421.

Haoyi Xiong, Daqing. Zhang, Leye Wang, and H. Chaouchi. 2014. EMC3: Energy-efficient data transfer in mobile crowdsensing under full coverage constraint. *IEEE Transactions on Mobile Computing*. DOI:http://dx.doi.org/10.1109/TMC.2014.2357791

Dejun Yang, Guoliang Xue, Xi Fang, and Jian Tang. 2012. Crowdsourcing to smartphones: Incentive mechanism design for mobile phone sensing. In *Proceedings of the 18th Annual International Conference on Mobile Computing and Networking (Mobicom'12)*. ACM, New York, NY, 173–184.

Dingqi Yang, Daqing Zhang, Zhiyong Yu, and Zhiwen Yu. 2013. Fine-grained preference-aware location search leveraging crowdsourced digital footprints from LBSNs. In *Proceedings of the ACM International Joint Conference on Pervasive and Ubiquitous Computing (UbiComp'13)*. ACM, New York, NY, 479–488.

Mao Ye, Peifeng Yin, Wang-Chien Lee, and Dik-Lun Lee. 2011. Exploiting geographical influence for collaborative point-of-interest recommendation. In *Proceedings of the 34th International ACM SIGIR Conference on Research and Development in Information Retrieval (SIGIR'11)*. ACM, New York, NY, 325–334.

Zhiwen Yu, Yun Feng, Huang Xu, and Xingshe Zhou. 2014. Recommending travel packages based on mobile crowdsourced data. *IEEE Communications Magazine* 52, 8, 56–62.

Zhiwen Yu, Yunji Liang, Bukan Xu, Yue Yang, and Bin Guo. 2011. Towards a smart campus with mobile social networking. In *Proceedings of the 4th International Conference on Cyber, Physical and Social Computing*. 162–169.

Nicholas J. Yuan, Yu Zheng, Liuhang Zhang, and Xing Xie. 2013. T-Finder: A recommender system for finding passengers and vacant taxis. *IEEE Transactions on Knowledge and Data Engineering* 25, 10, 2390–2403.

Man-Ching Yuen, Irwin King, and Kwong-Sak Leung. 2011. A survey of crowdsourcing systems. In *Proceedings of the IEEE International Conference on Social Computing*. 766–773.

Daqing Zhang, Bin Guo, and Zhiwen Yu. 2011. The emergence of social and community intelligence. *Computer* 44, 7, 21–28.

Daqing Zhang, Leye Wang, Haoyi Xiong, and Bin Guo. 2014a. 4W1H in mobile crowd sensing. *IEEE Communications Magazine* 52, 8, 42–48.

Daqing Zhang, Haoyi Xiong, Leye Wang, and Guanling Chen. 2014b. CrowdRecruiter: Selecting participants for piggyback crowdsensing under probabilistic coverage constraint. In *Proceedings of the 2014 ACM International Joint Conference on Pervasive and Ubiquitous Computing (UbiComp'14)*. ACM, New York, NY, 703–714.

Daqing Zhang, Zhiyong Yu, Bin Guo, and Zhu Wang. 2014d. Exploiting personal and community context in mobile social networks. In *Mobile Social Networking*, Alvin Chin and Daqing Zhang (Eds.). Springer, New York, 109–138.

Wei Zhang, Jianyong Wang, and Wei Feng. 2013. Combining latent factor model with location features for event-based group recommendation. In *Proceedings of the 19th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining (KDD'13)*. ACM, New York, NY, 910–918.

Xinglin Zhang, Zheng Yang, Chenshu Wu, Wei Sun, Yunhao Liu, and Kai Liu. 2014c. Robust trajectory estimation for crowdsourcing-based mobile applications. *IEEE Transactions on Parallel and Distributed Systems* 25, 7, 1876–1885.

Vincent W. Zheng, Yu Zheng, Xing Xie, and Qiang Yang. 2012. Towards mobile intelligence: Learning from GPS history data for collaborative recommendation. *Artificial Intelligence* 184–185, 17–37.

Yu Zheng, Furui Liu, and Hsun-Ping Hsieh. 2013. U-Air: When Urban air quality inference meets big data. In *Proceedings of the 19th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining (KDD'13)*. ACM, New York, NY, 1436–1444.

Yu Zheng and Xing Xie. 2011. Learning travel recommendations from user-generated GPS traces. *ACM Transactions on Intelligent System Technologies* 2, 1, Article 2, 29 pages.

Yu Zheng, Lizhu Zhang, Zhengxin Ma, Xing Xie, and Wei-Ying Ma. 2011. Recommending friends and locations based on individual location history. *ACM Transactions on the Web* 5, 1, Article 5, 44 pages.

Pengfei Zhou, Yuanqing Zheng, and Mo Li. 2012. How long to wait?: Predicting bus arrival time with mobile phone based participatory sensing. In *Proceedings of the 10th International Conference on Mobile Systems, Applications, and Services (MobiSys'12)*. ACM, New York, NY, 379–392.

Ying Zhu, Bin Xu, Xinghua Shi, and Yu Wang. 2013. A survey of social-based routing in delay tolerant networks: Positive and negative social effects. *IEEE Communications Surveys Tutorials* 15, 1, 387–401.

Received July 2014; revised May 2015; accepted June 2015