When Crowdsourcing Meets Mobile Sensing: A Social Network Perspective

Pin-Yu Chen, Shin-Ming Cheng, Pai-Shun Ting, Chia-Wei Lien, and Fu-Jen Chu

ABSTRACT

Mobile sensing is an emerging technology that utilizes agent-participatory data for decision making or state estimation, including multimedia applications. This article investigates the structure of mobile sensing schemes and introduces crowdsourcing methods for mobile sensing. Inspired by social networks, one can establish trust among participatory agents to leverage the wisdom of crowds for mobile sensing. A prototype of social-network-inspired mobile multimedia and sensing application is presented for illustrative purposes. Numerical experiments on real-world datasets show improved performance of mobile sensing via crowdsourcing. Challenges for mobile sensing with respect to Internet layers are discussed.

INTRODUCTION

Wireless sensor networks (WSNs) explore avenues to collect and use information from the physical world by deploying low-cost tiny sensor nodes on the ground, in the air, under water, on bodies, in vehicles, and inside buildings. With sensing, processing, and communication capabilities, networked sensor nodes cooperatively collect information on entities of interest, and WSNs have emerged as a promising technology with numerous and various applications. As shown in Fig. 1, sensor nodes locally collect information and then forward the sensed result over a wireless medium to a remote static sink, where it is fused and analyzed in order to determine the global status of the sensed area. In order to successfully gather sufficient information, a static sink could send a mobile agent to collect data from individual sensor nodes by following a trajectory spanning all the nodes (Fig. 1).

To accomplish large-scale sensing, the WSN evolves not only at the sink side (e.g., mobile agents), but also at the sensor node side. Mature mobile networks consisting of mobile devices with advanced processing and communication capabilities become a possible sensing infrastructure of WSNs. By exploiting the rich set of embedded sensors (e.g., camera, gyroscope, GPS, accelerator, light sensor, and digital compass) on mobile phones as sensor nodes, a new paradigm of WSNs is realized, which is known as *mobile sensing* [1–9]. As shown in Fig. 1, mobile sensing utilizes crowdsensed information for data analysis and decision making due to penetration of mobile devices as well as human mobility and ubiquity. It relies on the wisdom of crowds [10] to successfully infer the information of interest and accomplish its tasks. The data from mobile crowds (users, sensors, robots, etc.) can be either numerical or categorical, depending on applications. Examples of crowd-sensed data include numerical environmental measurements such as temperature and air conditions [3, 5], personal activities such as daily life patterns and events [2], interactions among people such as crowd density [7] and common interests [6], categorical recommendations such as ratings for nearby restaurants [1], and user experience/quality feedback on wireless mobile multimedia applications [1].

It is worth noting that many multimedia applications lie within the scope of mobile sensing, since extracting and analyzing the information sensed or generated from the crowds is one of the core goals for many multimedia applications in order to attract users' attention. Better prediction of users' interests leads to longer multimedia stickiness, and hence more revenues can be expected. Modern multimedia applications often pull user-centric information from the crowds and offer personalized contents (e.g., the next video to watch). Typically, location and social network information are widely used for targeted advertisement and recommendation. Therefore, the major challenge is to efficiently and accurately extract user-centric information from the crowds and identify users of high similarity for improved content delivery.

In recent years, many machine learning tasks and business models have leveraged the wisdom of crowds to acquire crowdsourced data for discriminating unknown objects. The website *Galaxy Zoo* asks visitors to help classify the shapes of galaxies, and the website *Stardust@home* asks visitors to help detect interstellar dust particles in astronomical images. Business models such as *Amazon Mechanical Turk* (*MTurk*) and *CrowdFlower* provide crowdsourcing services at low prices. For *MTurk*, a minimum of US$0.01[1] is paid to a labeler/worker when he/she makes a click (i.e., generates a label) for an item. Despite

Pin-Yu Chen and Pai-Shun Ting are with the University of Michigan

Shin-Ming Cheng is with National Taiwan University of Science and Technology.

Chia-Wei Lien is with Amazon Corporate LLC.

Fu-Jen Chu is with the Georgia Institute of Technology.

0163-6804/15/$25.00 © 2015 IEEE

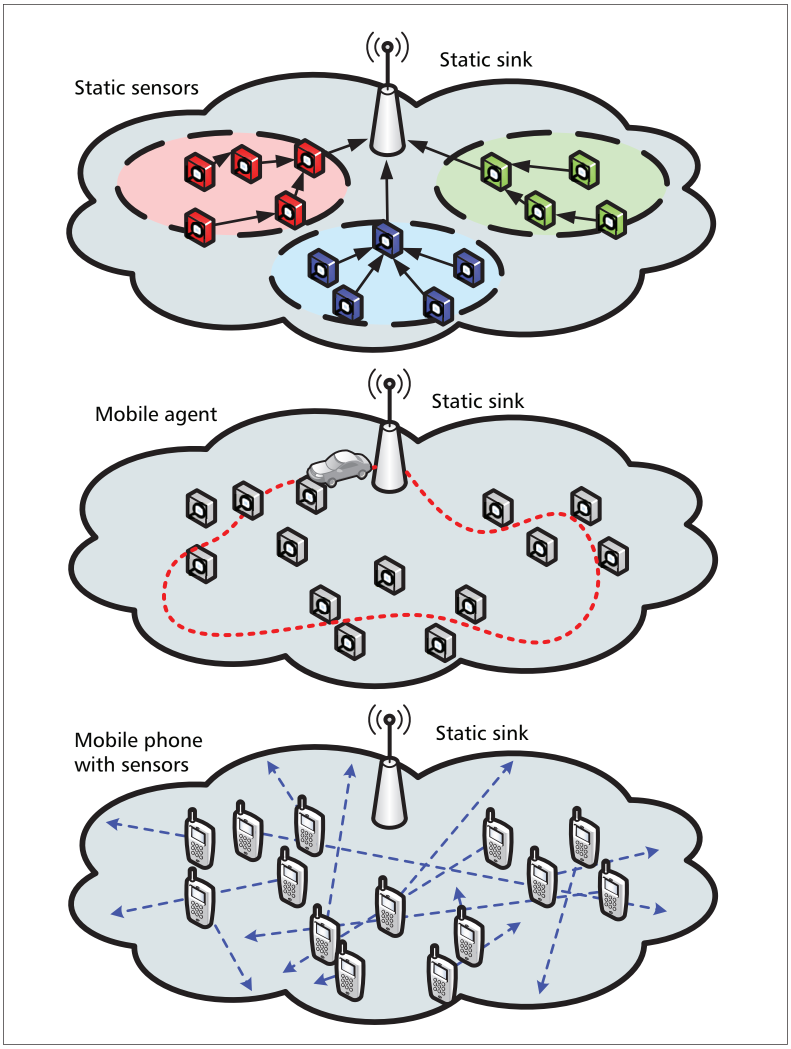

Figure 1. Evolution from wireless sensor networks to mobile sensing.

the low costs for acquiring crowdsourced data, one of the major challenges of mobile sensing is rooted in dealing with noisy and potentially erroneous data [4, 8, 11]. These undesired data can originate from environmental/object uncertainties (e.g., channel noise and difficulty of object discrimination) or user intentions (e.g., fraudulent recommendations and irresponsible user clicks). Consequently, identifying trustworthy data and reliable agents becomes an essential task in mobile sensing [4, 5, 8].

Utilizing the concept of trust in social networks, this article proposes trust-based data analysis approaches for crowdsourcing in mobile sensing schemes. In addition to scraping trustworthy data from the bulk, these approaches aim to identify reliable agents for performance enhancement. A weight of trust is built on the reliability of each agent for mobile sensing via a limited number of training queries. For spectrum sensing, these approaches can be implemented by broadcasting reference signals for reception power calibration. For annotation, these approaches can be implemented by uploading some items with known answers. For multimedia, these approaches can be associated with user behaviors based on the provided contents.

[1] https://requester.mturk.com/pricing

This article summarizes mobile sensing network paradigms and introduces several trust-based crowdsourcing methods for mobile sensing. We also illustrate a prototype application of a social-network-inspired mobile multimedia and sensing scheme. Numerical experiments on real-world datasets show that mobile sensing can benefit from crowdsourcing for improved performance. Potential challenges of social-network-based mobile sensing with respect to mobile multimedia Internet layers are discussed. Thus, this article sheds new light on integration of social networking and mobile sensing, and applications therein.

Mobile Sensing Paradigms

In mobile sensing, people share and distribute sensed information via physical proximity or social relations over portable sensors. As illustrated in Fig. 2, a mobile user plays the roles of both *querier* and *collector*, who request and provide information, respectively. A querier can simultaneously be a collector if he/she also participates in mobile sensing. Network structures of mobile sensing can be classified into two categories: the *direct* and *indirect* paradigms, which are described as follows.

Direct Mobile Sensing Paradigm

A direct mobile sensing paradigm involves direct communication between a querier and crowds (i.e., the collectors). Typically, it is achieved by adopting current device-to-device (D2D) communication technologies, such as WiFi Direct, ZigBee, Bluetooth Low Energy (BLE), and near field communication (NFC). In such a case, the *store-carry-forward* behavior facilitates information delivery in an ad hoc fashion. That is, sensed information may be stored in a sensor node in the absence of immediate connectivity to any other node, and relayed to other sensor nodes at encounters. Examples include the following.

Proximity sensing in mobile social networks (MSNs): It supports social platforms among physically proximate mobile users. For instance, one can simply scan the environment for discoverable Bluetooth devices to analyze crowd density and crowd flow direction [7]. By exploiting P2P communications, one can further make new social interactions with nearby devices. A popular example is sensing "potential friends with similar interests nearby." To enjoy such new activities, mobile users have to provide their own interests for profile matching by broadcasting their personal profiles to all nearby users, and then comparing their personal profiles and other users' profiles for friend matching [6].

Cooperative spectrum sensing in cognitive radio networks (CRNs): Unlicensed secondary users (SUs) sense the surrounding environment and exploit spectrum holes unoccupied by licensed primary users (PUs) for secondary transmission with minimal interference to PUs. To achieve better spectrum management and enhance radio resource utilization, a querier could exploit observations on local spectrum vacancy from surrounding SUs (i.e., crowds). The empowerment of cooperative spectrum sensing improves the throughput of wireless communications and reduces potential interference among heterogeneous systems.

Indirect Mobile Sensing Paradigm

In this paradigm a querier and crowds are indirectly connected through a communication system in a centralized or distributed fashion. Typically, access points in a WLAN and a base station in a cellular network or WiMAX are exploited as communication paths in the former case. In the latter case, a querier/collectors could download/upload data from/to nearby relays via localized communication technologies such as WiFi-direct, BLE, or NFC. Examples follow.

Environmental measurements: Collectors provide local measurements (e.g., temperatures, air pollution indexes) to a querier via an existing cellular infrastructure for event detection or state estimation. Consequently, the current environment can be understood and improved [3, 5]. For example, the PEIR project [3] exploits sensors in mobile phones to build a system that tracks the impact of individual actions on carbon emissions.

Personal activity sharing: A collector shares his/her daily life patterns, activities (e.g., sports), health (e.g., heart rate, blood pressure) with his/her friends using online social networks. For example, by automatically classifying events in people's lives via sensors on mobile phones, CenceMe [2] enables selective event sharing among friends using Facebook or Twitter.

Online recommendation: Crowds (e.g., data collected from proximal users or users of high similarity) provide recommendations to a user-centric query, such as the best seafood restaurant within two miles, or the next video to watch for multimedia applications. For example, MicroBlog [1] encourages users to record multimedia blogs manually or automatically (via sensors). Moreover, the blogs from collectors in the same area are integrated to enrich the contents. Consequently, a querier can browse multimedia blogs at a selected region for relevant information.

Annotation: Crowds (e.g., machines, people) annotate labels, such as scenery labels for a picture or comments and interactions for multimedia contents, for an item requested by a user. One typical example is the Amazon Mechanical Turk (MTurk) service.

LifeTie: A Prototype Social-Network-Inspired Mobile Multimedia and Sensing Application

For further illustration, in this section we introduce a social-network-inspired mobile multimedia and sensing application. In Taiwan, mountainous areas are hikers' paradise. Hikers are used to tie trail marking ribbons on trees for direction guidance. It is a matter of life and death to clearly know one's own location, especially at night. However, trail marking ribbons have several disadvantages such as misinterpretation, lack of detailed information, and environmental pollution caused by overuse. To ensure hikers' safety while overcoming the aforementioned drawbacks, we propose a mobile sensing system named LifeTie.

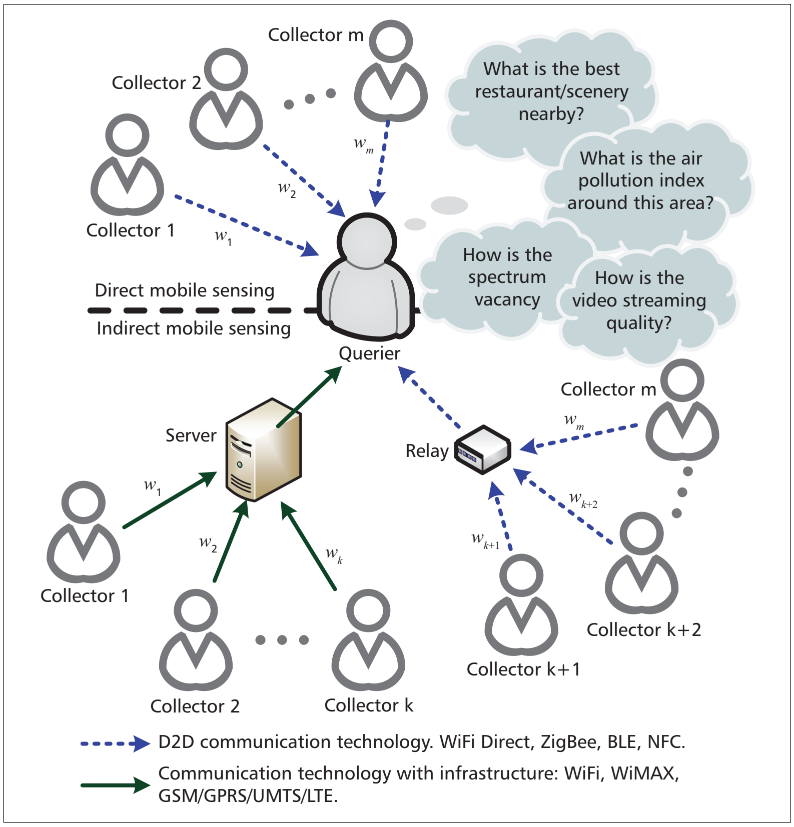

Figure 2. Network architecture of mobile sensing.

Unlike traditional WSNs deployed in mountains for wildlife tracking and ecological monitoring, LifeTie acts as an annotation platform where users can exchange their sensing results. The integration of tail marking ribbon and NFC technology replaces the traditional marking method with a smartphone app and achieves the purposes of information exchange, rescue, and search.

The main concept of LifeTie is to exploit NFC tags as the enabler for information exchange among people. To achieve that, NFC tags shall be attached around a mountain, thus creating an infrastructure. Hikers can trigger NFC tags nearby (typically in the range of only a few inches) using their NFC-enabled mobile phones. Then the NFC-enabled mobile phones can read/write data from/to the NFC tags. With embedded memory, LifeTie could handle a vast amount of data, which opens up the possibility of multimedia information and advances toward integration of social network and mobile sensing. The prototype of LifeTie is shown in Fig. 3. The features and functionalities of LifeTie are summarized as follows.

Features

Flexible: The shape of LifeTie was inspired by zap-straps. It can easily be tied to tree branches.

Recognizable: With the color of fluorescent orange and the addition of reflective stickers, it provides direction guidance even at night time.

Reliable: Polypropylene (PP) is used in LifeTie for its bendability and durability. An NFC tag is integrated in the internal parts of a strap to resist extreme weather conditions in mountains.

Affordable: The cost of an NFC tag is low.

Figure 3. LifeTie mobile sensing system — the physical device and its mobile multimedia interface.

Power saving: An NFC tag is not powered by electricity. As a result, no replacement for LifeTie is required, and therefore even lower deployment cost can be achieved.

Functionalities

Navigation and warning: Via the corresponding app, hikers can check out LifeTie's guestbook right after triggering NFC tags. Depending on the current environmental conditions, hikers can leave comments or draw a simple map to make a notification or assist navigation. Some useful warning icons (e.g., cliff, snake, wasps, slippery conditions) and guiding icons (e.g., cave, camp) are provided when drawing the map to enable diverse multimedia contents. The updated surrounding information facilitates following hikers.

Tracking and rescue: When a hiker is lost in a mountain and finds LifeTie, he/she can check regular comments to see if there is any shelter nearby. Moreover, he/she can leave urgent comments highlighted in red. If he/she can leave such information on several LifeTies, rescuers can easily identify a rough search area according to the positions of deployed LifeTie and timestamps of the comments. As a result, the rescuing operations become more effective and efficient.

Trust-Based Crowdsourcing Methods for Mobile Sensing

This section provides an overview of weight (trust) assignment methodologies on agents for crowd-sensed data in mobile sensing. The utility of these methods is investigated in the next section, and the challenges toward practice are addressed after that. For crowdsourcing-empowered mobile sensing, a user (or an intermediate system) evaluates a weight of trust w_i for the ith agent and fuse information from agents via weighted combination of agents' observations for the user's query. As shown in Fig. 4, the collected data from the agents can be viewed as a matrix with rows representing agents and columns representing observations associated with queries. The training queries refer to queries with known answers and are used for weight evaluation. The number of agents is denoted by m.

The final output for mobile sensing is the combination of each agent's observation multiplied by the associated weight. Here we discuss several crowdsourcing methods involving different weight evaluation approaches. These methods can be classified into two categories, unsupervised and supervised, separated by the need for training queries.

Unsupervised Crowdsourcing Methods

Majority votes: Majority votes adopts uniform trust among all agents (i.e., the weight $w_i = 1/m$ for all i) and selects the observation on which most agents agree as the final output. This method may lead to poor performance when the majority of agents have incorrect observations or some observations are maliciously manipulated.

Probabilistic inference: Probabilistic inference assumes that each observation made by an agent is statistically independent and imposes a statistical model to infer the weights from observations. One popular method is the weight evaluation method based on an expectation maximization (EM) algorithm [5, 12].

Supervised Crowdsourcing Methods

Supervised crowdsourcing methods aim to find the optimal weight of each agent by solving the optimization problem as

$$\text{minimize}_{\mathbf{w}} \ \text{cost(training-queries, final-output)} + \lambda \cdot R(\mathbf{w}),$$

where $\mathbf{w} = [w_1, w_2, \ldots, w_m]$ is the vector of weights, $R(\mathbf{w})$ is a regularization function for $\mathbf{w}$, and $\lambda \geq 0$ is the regularization parameter for $R(\mathbf{w})$. Here we introduce several supervised crowdsourcing methods.

Weighted averaging: Weighted averaging is a heuristic weight evaluation approach which assigns a weight that is proportional to the accuracy of each agent in the training queries. Let q_i be the fraction of correct queries responded by agent i. The weight of agent i is the normalized accuracy $w_i = q_i/\Sigma_{i=1}^{m} q_i$.

Exponential weighted algorithm: The exponential weighted algorithm adopts exponential cost function and zero regularization parameter ($\lambda = 0$) and sequentially adjusts the weight of each agent from the training queries. Interested readers can refer to [13, references therein] for more details.

Support vector machine: The support vector machine adopts the Hinge loss function as its cost function and assumes the regularization function $R(\mathbf{w}) = \Sigma_{i=1}^{m} w_i^2$ and positive regularization parameter (i.e., $\lambda > 0$). The support vector machine aims to find a separating hyperplane that best discriminates the training queries in the data sample space, and the weight of each agent can be determined by the resulting separating hyperplane. Interested readers can refer to [13, references therein] for more details.

Professional search: Inspired by social networks where problems are often resolved by professionals, professional search aims to assign weights to only a few agents that have outstanding accuracy in the training queries [14]. Professional search adopts the Hinge loss function as its cost function and assumes the regularization function $R(\mathbf{w}) = \Sigma_{i=1}^{m} |w_i|$. This regularization function is known as a surrogate function that promotes sparsity in $\mathbf{w}$ (i.e., most of the weights are zero), and hence the professionals hidden in the crowds can be selected for mobile sensing.

Numerical Experiments

In this section we use two crowd generated datasets to investigate the performance of the crowdsourcing methods in the previous section. For the first dataset, each agent only participates in some fraction of queries and hence resembles the dynamic participatory nature in mobile sensing. For the second dataset, almost every agent responds to each query, but none of the agents have correct answers to all queries, which resembles the imperfect sensing capability in mobile sensing. In both scenarios, crowdsourcing methods can improve the query classification performance by identifying trustworthy agents.

For crowdsourcing methods involving a regularization function R, we use a leave-one-out-cross-validation (LOOCV) approach [13] to determine the optimal regularization parameter λ, by swiping λ from 0 to 200 to select the optimal value that leads to minimum training error. A one-to-all classification approach is used for multiple (more than two) categorical datasets (e.g., the exam dataset).

Text Relevance Judgment

The text relevance judgment dataset is provided by the Text Retrieval Conference (TREC) crowdsourcing track in 2011,[2] where 689 agents (participants) are asked to judge the relevance of paragraphs excerpted from a subset of articles with given topics. Each agent then generates an observation, either "relevant" or "irrelevant," for an article. It is worth mentioning that this dataset is sparse in the sense that on average each agent only read roughly 26 out of 394 articles. For supervised crowdsourcing methods we use roughly 10 percent (40 training queries) of articles to evaluate each agent's weight. The rest of the data samples are used to test the accuracy of crowdsourcing algorithms, and the results are summarized in Table 1. It is observed that supervised methods can achieve higher accuracy than unsupervised methods via training queries. Also note that the support vector machine and professional search outperform other methods since their main objective is to assign more weights on the trustworthy agents/data samples possessing eminent discriminant capability.

Science Exam Dataset

The science exam dataset is collected by the authors[3] and contains 40 questions. Each question has four choices, and the correct answer is one of these four choices. There are 183 agents (students) taking the exam and producing observations (their answers). Unlike the TREC dataset, this exam dataset is dense in the sense

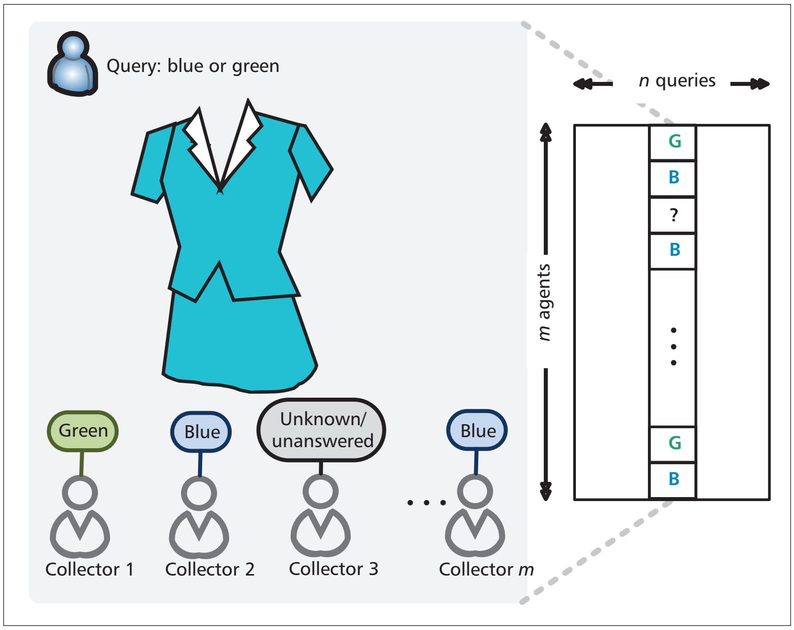

Figure 4. Illustration of crowdsourced data.

that almost every student has provided an answer for each question. We use 10 questions as training queries and the rest to test the accuracy. The results are summarized in Table 2. The baseline accuracy (random guess) is 25 percent. None of the methods can achieve accuracy as high as in the TREC dataset due to the following facts:

- The observations of the TREC dataset only have two categories, whereas the observations of the exam dataset have four categories, which renders the latter more difficult to discriminate.
- The exam is challenging since the majority votes method leads to low accuracy, and there is no student who answers all questions correctly. The best student only has 70 percent accuracy.

Nonetheless, crowdsourcing methods such as the support vector machine and professional search can still attain relatively good accuracy by identifying reliable agents.

Ongoing Challenges: Aspects from Mobile Multimedia Internet Layers

In this section we discuss some ongoing challenges toward integration of social networking and mobile sensing, particularly in aspects of the mobile multimedia Internet. Issues corresponding to each layer of the mobile multimedia Internet are specified as follows.

Application Layer

Conflict between privacy and trustworthiness: Although such integration of social networking and mobile sensing is exciting and promising, the collection and sharing of personal information related to human activity introduces the concern of privacy, where participants are reluctant to reveal any sensitive personal information (e.g.,

[2] https://sites.google.com/site/treccrowd/2011

[3] The science exam dataset can be downloaded from the first author's website, https://sites.google.com/site/pinyuchenpage/.

The amount of collected information and the number of participants are proportional to the degree of volunteering. When the amount of collected information is insufficient, the sensing results might not be precise. Thus, mobile sensing needs incentive and mechanism design to encourage people to participate.

	Unsupervised		Supervised			
Methods	Majority votes	Expectation maximization	Weighted averaging	Exponential weighted algorithm	Support vector machine	Professional search
Accuracy (%)	79.38	78.81	83.05	80.51	83.33	84.46

Table 1. The TREC 2011 dataset. Supervised methods attain higher accuracy than unsupervised methods via acquiring a few training queries for weight (trust) assignment.

	Unsupervised		Supervised			
Methods	Majority votes	Expectation maximization	Weighted averaging	Exponential weighted algorithm	Support vector machine	Professional search
Accuracy (%)	46.67	50	46.67	26.67	50	53.33

Table 2. The science exam dataset. Despite the fact that no students answer all questions correctly, professional search can still achieve more than twice the accuracy of random guess (25 percent accuracy).

time, location, pictures, sound, acceleration, and biometric data). As a result, it is crucial to design an approach to collecting sufficient information from participants without violating their privacy [8]. Specifically, authentication shall be supported to identify legal mobile users and adversaries. Moreover, anonymity shall be preserved to hide sensitive information by using technology such as k-anonymous or pseudonyms. These avenues prevent adversaries traversing a relationship between users' contributions and identities.

In particular, trust-based crowdsourcing methods aim to preserve trustworthiness in harsh environments where malicious participants may deliberately feedback fraudulent data. Obviously, to counteract this effect we need to observe contributions made by each user for a period of time and hence evaluate his/her trustworthiness. However, it may conflict with the privacy consideration where actual attribute values of a specific user are obscured, and the links between multiple contributions from the same user are broken. How to acquire linkability across multiple contributions from the same user while preserving privacy is a challenging issue [5].

Data integrity on complicated multimedia content: By manually recording via users or automatically collecting via sensors, a huge amount of information can be retrieved in mobile sensing. The multimedia contents generated from the retrieved materials via applications like MicroBlog [1] contain abundant but complicated information, which burdens data integrity. Unlike the example raised in Fig. 4 where we just need to identify the color of cloth, multiple parts in one clip or video might lead to a distinct conclusion that is highly dependent on a viewer's ideology. Extra meta information shall also be included to increase data integrity like user reviews (e.g., iMDB, Youtube) or user preference (e.g., Netflix).

Incentives for participation: Mobile sensing requires participants to spend their time, attention, and mobile phones' battery power for contributing data. Obviously, the amount of collected information and the number of participants are proportional to the degree of volunteering. When the amount of collected information is insufficient, the sensing results might not be precise. Thus, mobile sensing needs incentive and mechanism design to encourage people to participate [15].

Network Layer

Data retrieval in a distributed environment: In distributed scenarios (e.g., proximity sensing, spectrum sensing, and annotation), a querier can only retrieve information from localized collectors, which might lead to biased inference results. To overcome this issue, current researchers propose to enable information relay for each individual. Leveraging human mobility and store-and-forward features, the amount of data collected from crowds grows substantially via information exchange, thereby improving the accuracy of estimation. In addition, in the indirect mobile sensing paradigm involving data retrieval and analysis from distributed systems (e.g., data storage servers), distributed computation is known to be one of the big data challenges.

The limitations of D2D communications: In the direct sensing paradigm, a querier could communicate with collectors via D2D communications like WiFi Direct, ZigBee, BLE, and NFC. However, current D2D communication technologies typically require manual mutual authentication when making a connection, which is unfavorable for automatic data collection and device connectivity. Moreover, mobile sensing applications should be capable of integrating the features of different D2D communication technologies (e.g., transmission ranges, transmission rates, and power consumption) for ubiquitous multimedia content delivery.

Link Layer

Unreliable link due to mobility and interference: The mobile nature of users and agents may hinder the performance of mobile sensing due to change in location, environment, and participatory agents. The highly dynamic positions of agents incur ever changing received interference, thereby affecting link reliability. To accommodate this effect, a crowdsourcing-aided mobile sensing method should possess adaptivity and robustness in such a dynamic situation in order to identify inadequate agent participation and obsolete data collection.

Physical Layer

Trade-off between power consumption and sensing accuracy: Apparently, the higher the sensing accuracy attained, the greater the amount of energy is consumed in mobile devices with limited power due to the increase in data acquisition frequency, which might violate the design rationale of sensing paradigms. Consequently, power consumption fairness and energy-efficient scheduling for participatory devices should be considered jointly, and new throughput measures should be studied to balance the trade-off between power consumption and sensing accuracy.

Conclusion

This article proposes to incorporate crowdsourcing methods for mobile sensing and introduces several crowdsourcing methods for evaluating the weight of trust among agents. The direct and indirect mobile sensing network paradigms are discussed. A prototype of social-network-inspired mobile multimedia and sensing application is illustrated toward integration of social network and mobile sensing. Numerical experiments on real-world datasets show that mobile sensing can benefit from crowdsourcing methods for performance improvements. Ongoing challenges of integration of social networking and mobile sensing are addressed through the aspects of mobile multimedia Internet layers. This article therefore paves new avenues to various mobile applications and future mobile technology development.

Acknowledgment

This research is sponsored by the Ministry of Science and Technology, Taiwan, under the contracts of 103-2221-E-011-008-MY3 and 104-2221-E-011-051. We would also thank to Kai-Ching Wang, Wei-Ju Kao, and Ting-Wen Lin for their efforts on implementing the mobile sensing system LifeTie.

References

[1] E. Gaonkar *et al.*, "Macro-Blog: Sharing and Querying Content through Mobile Phones and Social Participation," *Proc. ACM MobiSys 2008*, June 2008, pp. 174–86.

[2] E. Miluzzo *et al.*, "Sensing Meets Mobile Social Networks: The Design, Implementation and Evaluation of the CenceMe Application," *Proc. ACM SenSys 2008*, Nov. 2008, pp. 337–50.

[3] M. Mun *et al.*, "PEIR, The Personal Environmental Impact Report, as a Platform for Participatory Sensing System Research," *Proc. ACM MobiSys 2009*, June 2009, pp. 55–68.

[4] R. Ganti, F. Ye, and H. Lei, "Mobile Crowdsensing: Current State and Future Challenges," *IEEE Commun. Mag.*, vol. 49, no. 11, Nov. 2011, pp. 32–39.

[5] C. Xiang *et al.*, "PassFit: Participatory Sensing and Filtering for Identifying Truthful Urban Pollution Sources," *IEEE Sensors J.*, vol. 13, no. 10, Oct. 2013, pp. 3721–32.

[6] R. Zhang *et al.*, "Privacy-Preserving Profile Matching for Proximity-Based Mobile Social Networking," *IEEE JSAC*, vol. 31, no. 9, Sept. 2013, pp. 656–68.

[7] J. Weppner *et al.*, "Participatory Bluetooth Scans Serving as Urban Crowd Probes," *IEEE Sensors J.*, vol. 14, no. 12, Dec. 2014, pp. 4196–4206.

[8] D. He, S. Chan, and M. Guizani, "User Privacy and Data Trustworthiness in Mobile Crowd Sensing," *IEEE Wireless Commun.*, vol. 22, no. 1, Feb. 2015, pp. 28–34.

[9] H. Liu, Y. Zhou, and Y. Zhang, "Estimating Users' Home and Work Locations Leveraging Large-Scale Crowd-Sourced Smartphone Data," *IEEE Commun. Mag.*, vol. 53, no. 3, Mar. 2015, pp. 71–79.

[10] J. Surowiecki, *The Wisdom of Crowds*, Anchor, 2005.

[11] R. Snow *et al.*, "Cheap and Fast — But Is It Good? Evaluating Non-Expert Annotations for Natural Language Tasks," *Proc. Conf. Empirical Methods in Natural Language Processing*, 2008, pp. 254–63.

[12] V. C. Raykar *et al.*, "Learning from Crowds," *J. Machine Learing. Research*, vol. 11, Aug. 2010, pp. 1297–1322.

[13] T. Hastie, R. Tibshirani, and J. Friedman, *The Elements of Statistical Learning*, Springer Series in Statistics, 2001.

[14] P.-Y. Chen *et al.*, "Supervised Collective Classification for Crowdsourcing," IEEE GLOBECOM Wksp. 2015, San Diego, CA; arXiv:1507.06682.

[15] D. Yang *et al.*, "Crowdsourcing to Smartphones: Incentive Mechanism Design for Mobile Phone Sensing," *Proc. ACM Mobicom 2012*, Aug. 2012, pp. 173–84.

Biographies

Pin-Yu Chen [S'10] received his B.S. degree in electrical engineering and computer science (undergraduate honors program) from National Chiao Tung University, Taiwan, in 2009, and his M.S. degree in communication engineering from National Taiwan University, Taipei, in 2011. He is currently working toward his Ph.D. degree in electrical engineering and computer science at the University of Michigan, Ann Arbor. His research interests include network science, interdisciplinary network analysis, graph clustering and community detection, statistical graph signal processing, cyber security, and their applications to data analysis and communication systems. He is a member of the Tau Beta Pi Honor Society and Phi Kappa Phi Honor Society, and was a recipient of the Chia-Lun Lo Fellowship. He was also the recipient of the IEEE GLOBECOM 2010 GOLD Best Paper Award, an IEEE ICASSP 2014 NSF travel grant, and an IEEE ICASSP 2015 SPS travel grant.

Shin-Ming Cheng received his B.S. and Ph.D. degrees in computer science and information engineering from National Taiwan University in 2000 and 2007, respectively. He was a postdoctoral research fellow at the Graduate Institute of Communication Engineering, National Taiwan University, from 2007 to 2012. Since 2012, he has been with the Department of Computer Science and Information Engineering, National Taiwan University of Science and Technology, Taipei, as an assistant professor. His current research interests include mobile networks, wireless communication, information security, and complex networks. He was a recipient of the IEEE PIMRC 2013 Best Paper Award and the 2014 ACM Taipei/Taiwan Chapter K. T. Li Young Researcher Award.

Pai-Shun Ting is a Ph.D. student in the Department of Electrical Engineering and Computer Science at the University of Michigan, Ann Arbor. His current research concerns design automation and mathematical analysis of stochastic logic networks. He earned his M.S. degree from the University of Michigan in 2013 and his B.S. degree from National Taiwan University in 2011, both in electrical engineering. Before joining the Ph.D. program in 2014, he worked on machine learning, image processing, and computer graphics as a research assistant at the National Institute of Informatics, Chiyoda-ku, Tokyo, Japan. He is currently a student member of the University of Michigan's Advanced Computer Architecture Laboratory.

Chia-Wei Lien received her M.S. degree in electrical engineering systems from the University of Michigan in 2014 and a B.S degree in electrical engineering from National Taiwan University. She joined Amazon.com, Inc. in 2014 and has been working on supply chain optimization modeling Amazon worldwide network planning. Her fields of interest include statistical signal processing, artificial intelligence, and machine learning.

Fu-Jen Chu received his B.S. degree in electrical engineering from National Taiwan University in 2011 and his M.S. degree in electrical engineering systems from the University of Michigan in 2014. He is currently pursuing his Ph.D degree in robotics and intelligent machines at the Georgia Institute of Technology, Atlanta. His research interests include computer vision, machine learning, and neural networks, and he is also interested in control systems and multi-agent problems.

The highly dynamic positions of agents incur ever-changing received interference, thereby affecting the link reliability. To accommodate this effect, a crowdsourcing aided mobile sensing method should possess adaptivity and robustness in such a dynamical situation in order to identity inadequate agent participation and obsolete data collection.