A SURVEY ON DATA COLLECTION IN CROWDSENSING SYSTEMS

Xiaorong Yang

School of Computer Science and Technology, Tiangong University, Tianjin 300387, China.

ABSTRACT

Crowdsensing is considered as a promising data collection model in which platforms outsource sensing tasks to a large number of users. In mobile crowdsensing(MCS), citizens participate in the sensing process, contributing data with their mobile devices, such as smartphones, tablets and wearable devices. First of all, we discuss the basic concept, system architecture and application of mobile crowdsensing. Then, combined with the development status of data collection mechanism of mobile crowdsensing at home and abroad, three types of data collection mechanism are introduced in detail, including data collection mechanism based on data quality, data collection mechanism based on node coverage and data collection mechanism based on compressed sensing.Finally, the article is summarized.

KEYWORDS: Mobile crowdsensing, data collection mechanism.

1. INTRODUCTION:

In recent years, the use of smart phones has experienced rapid growth. According to Gartner (2016), more than 1.4 billion smartphones will be sold worldwide by 2015. Statista (2018) estimates that the number of smartphone users will reach 2.87 billion by 2020. Since the end of last century, mobile devices have been rapidly popularized and constantly updated in recent years. By 2020, with the rapid development of embedded technology, wireless sensor network and Internet of things technology, intelligent equipment has become a necessary item in people's daily life. Many powerful sensors are embedded in these intelligent devices, including accelerometers, cameras, GPS and so on, which make these intelligent devices have strong sensing ability, and can sense, collect, record and process various types of data. With the development of mobile devices in communication, computing, storage and multimodal sensing capabilities, the proliferation of mobile devices has gradually spawned a new sensing paradigm, the so-called mobile crowdsensing (MCS) [1].

Mobile crowdsensing refers to the interactive and participatory sensing network formed by people's existing mobile devices, and the sensing tasks are released to individuals or groups to complete in the network, so as to help professionals or the public to collect data, analyze information and share knowledge. In the early days, these social sensing tasks could be completed by using wireless sensor networks [2]. However, because different types of sensing tasks need to deploy sensors with different functions or even sensor networks with different architectures, with the increase and change of social sensing tasks, the deployment and maintenance of static sensor networks will cost a lot of human and material costs [3]. MCS takes advantage of human intelligence, universality and mobility. In this regard, people have the right to contribute data perceived or generated from their mobile devices, so as to be able to effectively (in terms of cost and time) monitor large-scale phenomena that professionals cannot easily measure or require expensive investment (hardware and software) and greatly reduce the waste of human and material resources. The concept of mobile crowdsensing is to cooperate unconsciously, let users complete the sensing task without knowing it, and break through the barriers of professional participation. Compared with wireless sensor networks, mobile crowdsensing has many advantages, such as flexible deployment, multi-source heterogeneous sensing data, wide and uniform coverage, high scalability and so on.

2. MOBILE CROWDSENSING:

2.1. Overview of mobile crowdsensing:

Mobile crowdsensing takes the mobile devices of ordinary users as the basic sensing unit, and carries out conscious or unconscious cooperation through the mobile Internet to realize the sensing task distribution and sensing data collection, and complete large-scale and complex social sensing tasks[4].

2.2. Architecture of mobile crowdsensing system:

Mobile crowdsensing is a new data acquisition mode combining crowdsourcing [5] idea and mobile device sensing ability, and is a manifestation of the Internet of things. A typical mobile crowdsensing architecture consists of three parts, including the mobile terminal composed of many sensing devices (such as smart phones), namely participants, the central server composed of many servers located in the cloud, and task publishers. As shown in Figure 1, the server in the cloud accepts the demand of a certain task from the task publisher, and assigns the sensing task to the mobile phone. After the mobile phone receives the sensing task, it perceives the required data, and then collects and returns the data to the server. The server processes the data and returns it to the task publisher. Through the whole process, the functions of data perception, data collection and information service provision are realized. Crowdsensing is a distributed, mobile, grass-roots and independent service mode. At the same time, mobile crowdsensing has benefits in many application areas such as environment sensing, traffic conditions, and social services [6] [7].

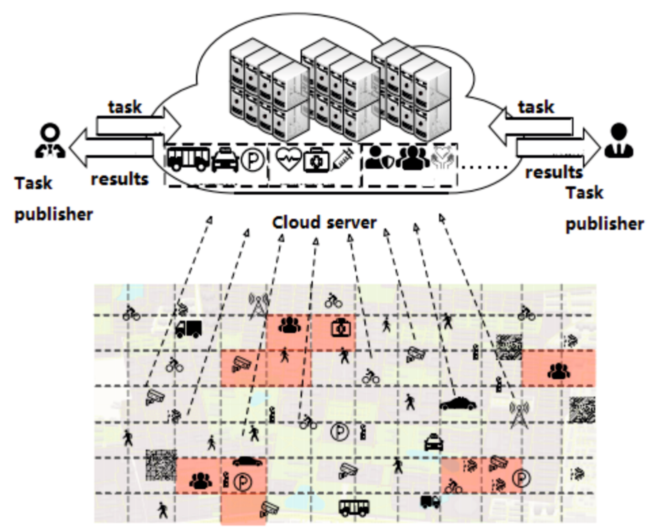

Figure 1: Typical architecture of mobile crowdsensing

2.3. Application of mobile crowdsensing application:

The application of mobile crowdsensing network is very extensive, and it has been applied in various fields, which can facilitate people's lives and also benefit the technological development of cities. The current crowdsensing applications can be roughly divided into:

(1) Driverless cars: Driverless cars have become more and more popular in recent years, and MCS can play an important role in this field. In fact, unmanned aerial vehicles (uav) and unmanned vehicles are equipped with different types of high-precision sensors, and often interact with mobile devices and users [8]. In [9], the author studied the energy-saving joint path planning and task assignment problems of the fixed-wing UAV-assisted MCS system. The corresponding joint optimization problem is developed into a two-stage matching problem. The first stage uses genetic algorithm to solve the path planning problem. In the second stage, the proposed Gale-Shapley algorithm solves the task assignment problem.

(2) Intelligent transportation system: MCS can also support ITS to provide innovative services and improve the cost-effectiveness and efficiency of transportation and traffic management systems [10]. It includes all applications related to traffic, vehicles, public transportation and road conditions. In [11], the author proposed a system that can predict the arrival time of a bus and rely on the participation and sensing of bus passengers. This model is only based on the collaborative efforts

Volume : 5 ¦ Issue : 9 ¦ Sep 2020 ¦ e-ISSN : 2456-5040

of participants and uses base station signals to locate users and prevent users from consuming batteries. In addition, it uses an accelerometer and microphone to detect when the user is on the bus. Nericell is a system for monitoring road and traffic conditions. It uses different sensors for rich sensing, detecting potholes, bumps, brakes and horns [12]. It uses the weight-bearing mechanism on the user's smartphone to conduct experiments on the roads of Bangalore.

(3) **Environmental detection:** Environmental perception is the basis for the sustainable development of urban space and improving the quality of life of residents. It aims to monitor the use of resources, the status quo of infrastructure and urban phenomena are well-known issues affecting citizens' daily lives. The author of the literature [13] proposed a platform to realize the environmental monitoring tasks of smart cities with fine-grained granularity, focusing on data heterogeneity. At present, air pollution is an important issue. Connecting sensors that use human mobility to monitor air quality to their devices may be a win-win solution with higher accuracy than fixed sensor networks. In [14], the author uses flash and smartphone cameras as light sources and receptors. In cooperation with a dedicated sensor, their goal is to measure the amount of dust in the air. Literature [15] In an indoor environment, users can use WiFi positioning to mark specific types of noise and locations, thereby creating a map of urban noise pollution.

3. DATA COLLECTION MECHANISM IN MOBILE CROWDSENSING NETWORK:

3.1. Data collection mechanism based on data quality:

Maintaining data quality in crowdsensing is an important and challenging problem. Due to the inherent open nature of the MCS application, inaccurate data provision may occur unintentionally and/or intentionally. In addition to noise and outdated data, malfunctions, low-quality wireless media, low-tech contribution users, selfish and/or malicious users deliberately providing low-quality and/or forged data should be considered.

The quality of the collected data in the MCS system depends on user contributions, which may be intentionally or unintentionally inaccurate. Unintentional inaccurate data configuration or low-quality data may be a wrong result, low-quality wireless connection, low resource level of mobile devices, limited sensors, few qualified workers, environmental uncertainty and outdated data Upload platform, which will lead to low quality of collected data. For example, when collecting sensor readings, workers may inadvertently place mobile devices in inappropriate locations, thereby reducing the quality of submitted data [16,17]. In addition, the limited availability of energy resources may result in the activation of different sensors to generate data for the same purpose, despite the lower quality of the data, and a lower level of energy consumption [18].

In crowdsensing, selecting appropriate participants to perform the required tasks at the minimum cost, while ensuring the maximum coverage and high perceptual quality is a complex problem in [19]. Although the data perceived by a small number of users is believed to lead to very inaccurate results, applying multiple perception processes to the same task makes the obtained data more effective. However, an increase in the number of participants does not necessarily mean that the overall quality of results will be improved. Therefore, predictive models have been proposed in many documents to estimate the number of personnel required for specific task assignments to achieve the required accuracy. In addition, the quality of the data may be affected by users who are not familiar with the target area, or require special effort because they do not know how to perform the task [20]. Literature [21] In order to improve the efficiency of data collection, it prompts participants to collect data at the location where the data has not been collected by displaying the location of the collected data on the map. This method can not only avoid the low-quality redundancy of the participant's collection Data can be checked for omissions.

Deliberately providing inaccurate information refers to situations where selfish and/or malicious entities deliberately provide low-quality and/or wrong information. The selfish party may provide low-quality data (for example, non-fresh or random sensor readings) in order to minimize the cost of their energy and resources, or maximize their utility, while being entitled to task execution Designated rewards. For example, in real-time traffic monitoring, selfish users can report false traffic jam warnings to prevent traffic on their own routes [22]. On the other hand, malicious parties try to destroy the extracted information and the usefulness of the MCS application, falsifying the provided data. For example, malicious participants can spoof GPS location [23] and start to provide fake location data. The presence of selfish and malicious users can lead to low-quality contributions that need to be identified and eliminated. Therefore, a mechanism to effectively verify and ensure the integrity of the collected data is necessary. As a simple first solution, the location verification mechanism can be enforced. For example, current solutions recommend checking whether the user is close to where they are shown in the answer.

3.2. Data collection mechanism based on node coverage:

Since traditional wireless sensor networks mainly include static sensors or mobile sensors with controlled motion trajectories, although static nodes can collect data in real time, the sensing range is fixed. At this stage, researchers only pay attention to the hardware system or algorithm design, but do not consider the coverage problem. However, the coverage quality is a key factor that affects the pros and cons of MCS applications. According to the mobile characteristics of nodes, the coverage can usually be divided into two types: random mobile sensor network coverage and controlled mobile sensor network coverage.

The coverage of random mobile sensor networks is mainly focused on target detection applications, that is, the continuous random movement of nodes to increase the coverage of the network so that each point in the monitoring area can be covered within a specific time delay. Literature [24] studies the performance of target detection in hybrid sensor networks, which includes fixed deployment sensor nodes and mobile sensor nodes.

The coverage of the controlled mobile sensor network is mainly for the monitoring of a limited number of discrete points of interest in the monitoring area, which can be regarded as point coverage. By controlling the moving speed and moving path of the sensor nodes, each point of interest is monitored within a specific time delay. Covered to. Literature [25] focuses on event detection applications, assuming that there are multiple discrete points of interest in the monitoring area, and random events occur at each point of interest with a certain probability and last for a period of time. This document defines the coverage quality as the probability that the event is not detected, and analyzes the relationship between the coverage quality and the node's moving speed, moving path, and the number of nodes. Literature [26] studied the saccade coverage (problem, that is, given a set of interest points and the time delay limit for each interest point to be covered, how to plan the moving path of each mobile node so as to use the least number of mobile nodes to satisfy each Coverage requirements for points of interest.

3.3. Data collection mechanism based on compressed sensing:

Compressed sensing [27], also known as compressed sampling or sparse sampling, is a technique for signal sampling. It uses some means to achieve "compressed sampling", which is exactly the process of data compression during the sampling process. This method takes advantage of the sparseness of the signal. Compared with the Nyquist theory, this method can restore the entire desired signal from fewer measured values. The data collection mechanism based on compressed sensing is often used in real-world applications of group intelligence perception. In the actual application of mobile crowdsensing technology for data collection, due to the limitations of objective conditions, there will be a small number of users participating in the perception. Compressed sensing (CS) algorithms are commonly used at this time. Only select cells in a small area for physical data sensing, and at the same time infer the data in the rest of the target area, and then restore all the data.

Assuming that the data collected by the sensing node is denoted as X, the data size is denoted as N, and the basis vector is denoted by $\Psi = [\Psi_1, \Psi_2, \Psi_3 \ldots \Psi_N]$, the coefficient transformation of the data is performed according to the following formula :

$$X=\Psi s \tag{1}$$

As shown in Equation (1), the data X and s collected by the sensing node are equivalent representations of the signal, where s is the sparse coefficient, which is the representation in the basis vector.

Design a M×N (M<<N)-dimensional measurement matrix to observe the signal change, and obtain the M×1 dimensional observation result, denoted as Y, then Y can be expressed as equation (2):

$$Y=\Phi X \tag{2}$$

In order to accurately reconstruct the coefficients in the relationship from the observations, and must meet the routing information protocol (RIP) conditions. In theory, RIP can guarantee that M measured values Y can reconstruct a k sparse signal X of length N, so choose a suitable algorithm to reconstruct X:

$$X=\min|(|X|)|; \quad \text{s.t. } \|\Phi X-Y\|_2 \leq \delta \tag{3}$$

When $\mathbf{M} \geq \mathbf{cKlog}(\frac{N}{K})$, the K sparse signal X can be reconstructed accurately.

In the past, selecting a small number of users to perceive people in crowdsensing would be considered as inaccurate as the result of data collection. However, a good compressed sensing bridge can solve this problem and at the same time solve the data redundancy caused by human uncontrollable mobility. Yu and the problem of poor data quality. Literature [28] uses mobile devices to collect human activity information, and successfully reduces the energy consumption of mobile devices based on CS technology. Literature [29] applies the CS algorithm to the urban road traffic monitoring system, and the designed method restores the high-precision urban traffic conditions without collecting most of the data.

4. SUMMARY:

At present, certain progress has been made in the research on the incentive mechanism of mobile crowdsensing. In the future, the research on the data collection mechanism of mobile crowdsensing, we can also conduct various researches from the aspects of privacy and security. There are few researches on existing data collection that threaten users' privacy (location information, etc.) exposure, especially for location-sensitive mobile cluster intelligence applications, research on how to reduce privacy risks caused by location information leakage is crucial important.

References:

I. Ganti, R. K. , Ye, F. , & Lei, H. . (2011). Mobile crowdsensing: current state and future challenges.IEEE Communications Magazine,49(11), 32-39.

II. Akyildiz, I. F. , Su, W. , & Sankarasubramaniam, Y. . (2002). Wireless sensor networks: a survey.Computer Networks. 38(4), 393-422.

III. Ren, Y. , Zhang, S., & Zhang, H. . (2006).Coverage control theory and algorithm in wireless sensor network. Journal of Software,422–433.

IV. Liu ,Y.. (2012).Crowd sensing computation.China Computer Federation, 8(10), 38-41.

V. Feng, H., Li, G. & Feng ,J.. (2014).Overview of crowdsourcing technology research.Journal of Computer.

VI. Froehlich, J., Dillahunt, T., & Klasnja, P..(2009). 27th International Conference Extended Abstracts on Human Factors in Computing Systems. Boston:Association for Computing Machinery.

VII. James, H., Alexei, E..(2008) 26th IEEE Conference on Computer Vision and Pattern Recognition.Anchorage:IEEE Computer Society.

VIII. Zhang, B. , Liu, C. H. , Tang, J. , Xu, Z. , & Wang, W. . (2018). Learning-based energy-efficient data collection by unmanned vehicles in smart cities. IEEE Transactions on Industrial Informatics,14(99), 1666-1676.

IX. Zhou, Z. , Feng, J. , Gu, B. , Ai, B. , Mumtaz, S. , & Rodriguez, J. , et al. (2018). When mobile crowd sensing meets uav: energy-efficient task assignment and route planning. IEEE Transactions on Communications.

X. Calabrese, F. , Colonna, M. , Lovisolo, P. , Parata, D. , & Ratti, C. . (2011). Real-time urban monitoring using cell phones: a case study in rome. IEEE Transactions on Intelligent Transportation Systems, 12(1), 141-151.

XI. Zhou, P., Zheng, Y., & Li, M. (2014). How Long to Wait? Predicting Bus Arrival Time With Mobile Phone Based Participatory Sensing.IEEE Transactions on Mobile Computing, 13(6), 1228-1241.

XII. Mohan, P., Padmanabhan, V. N., & Ramjee, R. (2008). Nericell: rich monitoring of road and traffic conditions using mobile smartphones.international conference on embedded networked sensor systems.

XIII. Montori, F. , Bedogni, L. , & Bononi, L. . (2017). A collaborative internet of things architecture for smart cities and environmental monitoring. IEEE Internet of Things Journal, 1-1.

XIV. Budde, M., Barbera, P., Masri, R. E., Riedel, T., & Beigl, M. (2013). Retrofitting smartphones to be used as particulate matter dosimeters.international symposium on wearable computers.

XV. Stevens, M., & D'Hondt, E. . (2010).Crowdsourcing of pollution data using smartphones. Workshop Ubiquitous Crowdsourcing, 1–4.

XVI. Guo,B., Yu, Z., Zhang, D., & Zhou, X. .(2014). From Participatory Sensing to Mobile Crowd Sensing. 2014 IEEE International Conference on Pervasive Computing and Communication Workshops, PERCOM WORKSHOPS 2014, Budapest, Hungary , March 24-28, 2014 .

XVII. Louta, M., Banti, K., Karetsos, G., & Lagkas, T. . (2015). Mobile crowd sensing architectural frameworks: a comprehensive survey. 2016 7th International Conference on Information, Intelligence, Systems & Applications ,IISA 2016, Chalkidiki, Greece, July 13-15 ,2016.

XVIII. Marjanovic, M. , Skorin-Kapov, L. , Pripuzic, K. , Antonic, A. , & Zarko, I. P. . (2015). Energy-aware and quality-driven sensor management for green mobile crowd sensing. Journal of Network & Computer Applications, 59(JAN.), 95-108.

XIX. Li, H. , Li, T. , Li, F. , Wang, W. , & Wang, Y. . (2016). Enhancing participant selection through caching in mobile crowd sensing.IEEE/ACM International Symposium on Quality of Service. ACM.

XX. Guo, B. , Chen, H. , Yu, Z. , Nan, W. , Xie, X. , & Zhang, D. , et al. (2016). Taskme: toward a dynamic and quality-enhanced incentive mechanism for mobile crowd sensing.International Journal of Human - Computer Studies, S107158191630101X.

XXI. Kathleen Tuite, Noah Snavely, Dunyu Hsiao, Nadine Tabing, & Zoran Popović. (2011). Photocity: training experts at large-scale image acquisition through a competitive game.

XXII. Cheng, L. , Niu, J. , Kong, L. , Luo, C. , Gu, Y. , & He, W. , et al. (2016). Compressive sensing based data quality improvement for crowd-sensing applications. Journal of Network & Computer Applications, S1084804516302338.

XXIII. Borcea, C. , Talasila, M. , & Curtmola, R. . (2015). Mobile Crowd Sensing.

XXIV. Wimalajeewa, T., Jayaweera, S. . (2010). Impact of mobile node density on detection performance measure in a hybrid sensor network. IEEE Transactions on Wireless Communications,9(5),1760-1769,2010.

XXV. Bisnik, N. , Abouzeid, A. A. , & Isler, V. . (2007). Stochastic event capture using mobile sensors subject to a quality metric. IEEE Transactions on Robotics, 23(4), 676-692.

XXVI. Li, M. , Cheng, W. , Liu, K. , He, Y. , Li, X. , & Liao, X. . (2011). Sweep coverage with mobile sensors. IEEE Transactions on Mobile Computing,10(11), 1534-1545.

XXVII. Donoho, D., L. . (2006). Compressed sensing. IEEE Transactions on Information Theory, 52(4), 1289-1306.

XXVIII. Akimura, D., Kawahara, Y. , & Asami, T. . (2012). Compressed sensing method for human activity sensing using mobile phone accelerometers. 2012 Ninth International Conference on Networked Sensing ,INSS 2012, Antwerp, Belgium, June 11-14,2012. IEEE.

XXIX. Li, Z. , Zhu, Y. , Zhu, H. , & Li, M. . (2011). Compressive Sensing Approach to Urban Traffic Sensing.2011 International Conference on Distributed Computing Systems, ICDCS 2011, Minneapolis, Minnesota, USA, June 20-24, 2011. IEEE.