Towards Mobile Sensor-Aware Crowdsourcing: Architecture, Opportunities and Challenges[§]

Jiyin He[1], Kai Kunze[2], Christoph Lofi[3], Sanjay. K. Madria[4], and Stephan Sigg[5]

[1] Centrum Wiskunde & Informatica, Netherlands
[2] Osaka Prefecture University, Japan
[3] National Institue of Informatics, Japan
[4] Missouri University of Science and Technology, USA
[5] Georg-August University Goettingen, Germany
Emails: { jiyinhe@gmail.com, kai.kunze@gmail.com, lofi@nii.ac.jp, madrias@mst.edu, ssigg@gwdg.de }

Abstract. The recent success of general purpose crowdsourcing platforms like Amazon Mechanical Turk paved the way for a plethora of crowd-enabled applications and workflows. However, the variety of tasks which can be approached via such crowdsourcing platforms is limited by constraints of the web-based interface. In this paper, we propose mobile user interface clients. Switching to mobile clients has the potential to radically change the way crowdsourcing is performed, and allows for a new breed of crowdsourcing tasks. Here, especially the ability to tap into the wealth of precision sensors embedded in modern mobile hardware is a game changer. In this paper, we will discuss opportunities and challenges resulting from such a platform, and discuss a reference architecture.

Keywords: Mobile Platforms, Sensor-Enabled Crowdsourcing, Location-Aware Crowdsourcing

1 Introduction

Crowdsourcing has become a popular approach to many problems that cannot be easily addressed by automated methods and algorithms, or problems that explicitly require significant amount of human intelligence or human feedback. Crowdsourcing can often be found in knowledge processing tasks such as data or media classification [8], data acquisition tasks such as data completion [6] or information extraction [16], as well as in providing training data for machine-learning-based approaches [20]. Furthermore, crowdsourcing has proven to be useful to the research community for performing large-scale user studies for evaluating new prototype implementations [11], or performing surveys with a large and diverse number of participants for investigating general human behavior or preferences [1]. Instead of laboriously growing own custom crowdsourcing platforms, these tasks mostly rely on general purpose crowdsourcing platforms such as Amazon Mechanical Turk, CrowdFlower, or ClickWorker. These platforms allow a complex task to be executed by dividing it into many smaller and simpler sub-tasks, i.e.,

---

[§] Author names listed in alphabetical order

HITs (Human Intelligence Tasks) – the smallest unit of crowdsourceable work, which are the distributed to a human worker pool. Workers are recruited and retained with payment. Hence, in theory such platforms can be used to perform any dividable tasks that require human intelligence. However, most of these services only offer a web-based interface for workers, and therefore tasks are limited to those that can be displayed and solved within a web browser.

In this paper, we propose an alternative architecture for a general-purpose crowdsourcing platform based on mobile as well as PC devices to interact with the worker pool, referred to as "hybrid crowdsourcing platform". This will increase not only the ease of use and acceptance of workers in an ever more mobile society, but also the utility and the range of possible crowdsourcing tasks for research applications as well as practical application. In particular, the access to GPS locations and mobile sensors will allow novel crowd-based applications that have not been possible before. Our contributions in this paper are as follows:

- We motivate and discuss need and benefits of mobile sensor-enabled crowdsourcing platforms.
- We highlight use cases of our platform, especially in the area of locality-sensitive services and ubiquitous computing.
- We present the design space and the generic architecture of such a platform, and discuss the impact of certain decisions on the system features and usability.

2 Background

Crowdsourcing can lead to significant cost savings [9, 15, 25], product quality improvements [2] and acceleration of time to market [3, 4].

However, crowdsourcing also has the potential to mitigate regional differences in the distribution of labor and human resources. Therefore, most previous work on mobile crowdsourcing platforms focused on societal aspects of crowdsourcing [5, 7, 18]. These approaches have been tailored for developing countries as an alternative source of labor and income. In developing countries, the spread of personal computers and wired internet connectivity is low. However, still many may have access to mobile phones or even mobile internet service. Therefore, the core challenge discussed in these works is how crowdsourcing can be adapted to the low-end hardware commonly available in developing countries, and how gaps in internet connectivity could be covered using SMS or alternative messaging methods.

In contrast, mobile crowdsourcing as discussed in this paper especially focused on exploiting the capabilities of modern, powerful mobile hardware to offer new functionality to crowdsourcing services. Especially the ability to tap into the user's geo-location or access to high-quality sensor data allows for completely new applications.

3 General Design

We envision a crowdsourcing platform that can be used in a stationary as well as mobile setting. The various instances in the private devices of users are interconnected through

a server in the cloud that takes care of the aggregation of responses, ranking, evaluation and source-selection for a given request. For this purpose, the server stores locations, end user profiles and source profiles among other information (cf. figure 1).
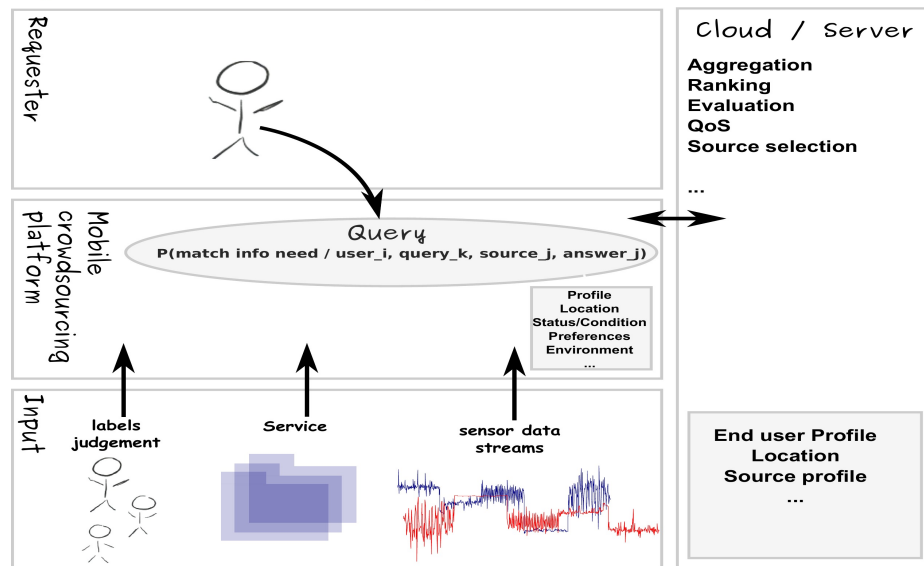
Hybrid crowdsourcing

Fig. 1. A proposal of a mobile crowdsourcing platform

Manual labels and judgements can be harvested as well as sensor data on mobile devices. The requester is likewise part of the crowd as any user or service may issue a query for input of users, services or sensors.

By the combination of mobile, pervasive and crowdsourcing concepts, we will be able to provide crowdsourcing for the masses: A more democratic crowdsourcing usage pattern in which everybody can be crowdsourced or equally state own queries. Mobile crowdsourcing will be seamlessly integrated into daily life with constantly up-to-date, personalised queries that can be completed anytime, anywhere. Instead of playing pointless mobile games to bridge waiting times, people can instead solve interesting queries and even earn money by completing these tasks. Through the integration of context sources in addition to human sources for content provisioning, queries can be highly personalized (e.g. location, environment, condition [19]) and in addition be automatically evaluated for their quality (e.g. fatigue [14]). Such a platform can be exploited to collect huge amounts of labelled sensor data (by asking users to perform certain tasks while being recorded by sensors on the mobile device [23]) from a tightly controlled target population. In addition, it might change the nature of crowdsourcing by empowering ordinary people to set up simple queries that might even reach into their real world (ask people to buy/bring something somewhere). Furthermore, a such a platform

might replace traditional data-bases in applications that rely on data which is changing at a high pace. For instance, imagine a dating service, in which a query for a potential partner is not stated to a database of registered users but instead towards the crowd.

4 Opportunities and challenges

State-of-the-art croudsourcing platforms are implemented through web-based services by international players such as Amazon. These platforms require explicit input and reach a maximally diverse population of possible content providers regardless of their location, gender, age, condition or further preferences. However, the result of a request is typically of medium or low quality and requires significant effort to filter out meaningful and quality responses [10]. The integration of crowdsourcing principles with mobile and Pervasive Computing has the potential to disruptively extend the possibilities underlying current crowdsourcing towards, among others, new applications, new classes of data and new possibilities to automatically evaluate quality of responses. We envision a platform with access to implicit information on, for instance, location, condition or further preferences that could restrict a given query to the most intended audience and also utilise sensor information (e.g. fatigue, crowd, loudness level) during the completion of a query in order to automatically estimate the quality of a response. In addition, with sufficient data at hand, prediction techniques might be applied in order to further boost the confidence on a result reached [22]. Figure 2 illustrates this concept.
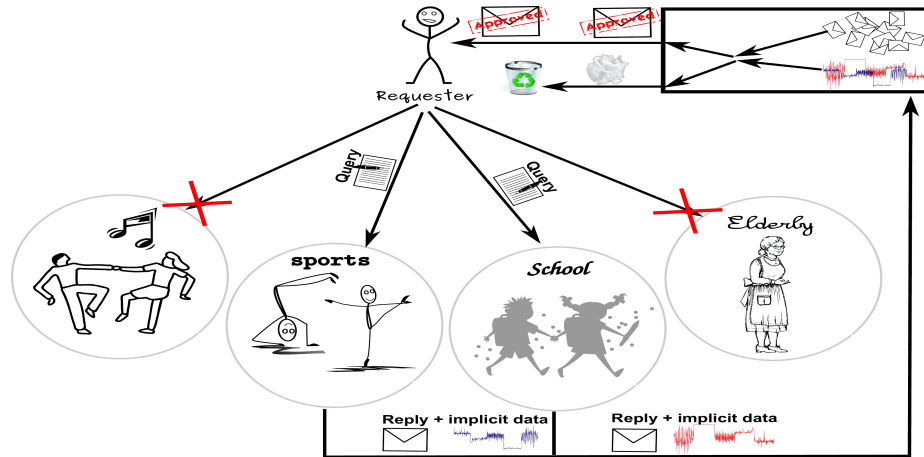

Fig. 2. Concept of mobile crowdsourcing

Expected advantages of a hybrid mobile sensor-aware crowdsourcing paradigm include: (1) improved task performance and efficiency; (2) enabling new crowdsourcing process; and (3) enabling new types of applications. Below, we discuss these aspects, as well as the challenges involved in realising the proposed platform.

4.1 Improved task performance and efficiency

Improved personalisation of request allocation and response aggregation Hybrid mobile sensor-aware crowdsourcing would enable personalised requests filtered by preferences set on mobile devices as well as by dynamic contextual parameters such as location, situation or condition.

Specifically, by maintaining worker profiles including a history of the tasks they have participated, their task performance, as well as the sensor inputs accompanied with this profile, e.g., their location and environment, the system can learn to predict the expertise of the worker, and under which conditions a task may be suitable for that particular worker. For instance, a worker may be able to accomplish a translation task with high quality in the morning when transiting from home to work, while this performance may decrease in the evening when he/she transits from work to home since she is exhausted after a long day already. Using such information, the system can (1) selectively recommend tasks to target workers, and (2) selectively return or aggregate worker response to the requester.

Crowdsourcing spontaneous feedback. With a mobile-based platform, both requesters and workers will have less constraints in terms of working locations. This may greatly reduce the time from stating a query to the reception of responses. As a result, responses can be very up-to-date and may include real-time assistance, for instance, in searching/recommendation for point-of-interest locations/navigation or spontaneous translation of foreign sentences (e.g. while ordering a menu at a restaurant).

New quality control mechanisms. With sensor data available alongside user input, this data may be utilised to estimate the quality of the provided input. For instance, by analysing the eye-gaze-movement, the platform can estimate fatigue or, reasoning from the loudness level [21] or amount of other people around [24], which can be used to judge whether the user is impaired in answering questions that require considerable concentration.

Information about situational impacts on cognitive performance. By utilising contextual information, a requester can gain knowledge about the performance of users in various environmental conditions. For instance, by stating a request to several groups of users in various contextual situations, the requester may learn about impacts on cognitive performance. Similarly, by controlling also the situational impacts for a series of queries to several sets of users, the requester can exclude side-effects on the result of a query.

4.2 New crowdsourcing processes

Crowdsourcing for the masses. A crowdsourcing platform on a mobile device, available anytime and anywhere at the convenience of users will change the principle nature of crowdsourcing. Constantly updated, up-to-date and personalised queries can be completed on-demand, interrupted and continued seamlessly. Another aspect is that mobile-based crowdsourcing mitigates hierarchies. Requester and source fall together to the same person as everybody is in the position to state a query. Consequently, quantity of queries will increase while their complexity will fall.

Weakening the strong correlation between labour and human resources. There is a strong relation between the physical location of labour and human resources. While crowdsourcing in general is capable of weakening this correlation, mobile crowdsourcing will further foster this development. In particular, since queries can be more personalised, companies are capable of stating more complex queries also for well-educated workers. This will open new possibilities for workers to offer their workforce without the necessity to relocate.

Participatory Sensing. The envisaged crowdsourcing platform provides access not only to manual input provided by users completing tasks, but also to sensors attached to the mobile platform (Gyroscope, Camera, GPS, Magnetometer, etc.) [17]. This might enable, for instance, quick requests for survey purposes even without manual user intervention. Devices and services might extend their contextual perception by harvesting (via automatically answered queries) for sensor information from devices in proximity. Similarly, a mobile crowdsourcing platform may be utilised to acquire labelled sensor data by requesting users to perform specific actions which are then recorded. For example, researchers can survey the relation between exercises and people's ability of solving math problems by requesting workers to solve a math problem before and after performing certain exercises. This process can be monitored and measures can be taken from both user input and sensor input from the mobile devices. This type of tasks are not feasible with traditional Web-based crowdsourcing platforms.

4.3 New applications

Mobile crowdsourcing enables new applications for crowdsourcing. For instance, crowdsourcing can replace a database when sensor-based or non-time critical manual feedback is required. There are new challenges introduced by this paradigm as data might then fluctuate in quality and quantity. In addition, crowdsourcing may partly leave the virtual space through a mobile platform. We envision, for instance, an event-hosting company that crowdsources actual manpower on demand. Also, crowdsourcing for educational purposes may serve the need of companies completing actual business-related tasks as well as the need of learners. For instance, a company active in language translations may provide users with text to be translated and later, after collection of all responses, with the corrected aggregated results for educational purposes.

Crowdsourcing as an anonymised customer information system. Mobile crowdsourcing can lower the burden and improve security and privacy in customer information systems. Instead of collecting and maintaining customer-related information for personalised interaction and product design, companies can reach a desired sub-set of customers on demand through mobile crowdsourcing platforms. This will significantly reduce cost and release companies from the burden to maintain huge databases of privacy-critical customer-related information.

Enabling technology for smart cities. A city is defined as smart when investments fuel sustainable economic growth in the respective aspects 'economy', 'mobility', 'environment', 'people', 'living' and 'governance' [12, 13]. A hyrid crowdsourcing platform connects people, government, industry and the environment as all can state queries or

provide input to requests stated. Mobile crowdsourcing can therefore serve as an interaction principle in such environments and constitute the backbone of a smart city, interconnecting all major entities.

Mobile crowdsourcing for energy management and smart buildings. Mobile crowdsourcing platform integrates environmental sensors and services. Humans and services acquire maintenance information from infrastructure and surrounding sensors via queries limited by proximity or belonging to a specific entity (building, room, etc). In addition, services can serve as actuators, completing queries designed to control smart buildings and automation. In particular, the controlled entity might change relative to the location of the requester.

Mobile crowdsourcing as a chance for ubiquitous computing research Currently, there is a trend to ubiquitous and smart computing, which especially includes wearable computing devices which transparently integrate into a user's life. Often, these devices need to detect the user's current activity and context in order to offer an appropriate service. However, detecting context and activity is an ongoing research challenge, and many approaches heavily rely on machine-learned models for this task. These, however, require exhaustive training data of users in clearly defined contexts performing a given activity in order to train those models. This training data is very hard to obtain efficiently and with the high diversity required for creating rich detection models (for example, sensor data of people riding a bicycle, or cooking in the kitchen, etc.) . Therefore, mobile crowdsourcing can provide a significant boost to this line of research by allowing to acquire such training data quickly and cheaply.

4.4 Challenges

High performance data processing and analysis mechanism With a mobile sensor-enabled crowdsourcing platform, we need to be able to process the vast amounts continuously generated explicit user inputs (requests and responses) as well as implicit sensor inputs in real time, e.g., in order to realise the above mentioned personalised request allocation and response aggregation. This requires high performance computational power as well as sophisticated data mining and machine learning algorithms that can scale to this type of data and give spontaneous responses. Further, sensor data as well as user inputs may be noisy. It is non-trivial to extract meaningful features from the raw sensor data as input for machine learning algorithms, or to derive human interpretable results.

Limitations of mobile devices While mobile devices provide great flexibility for people to perform tasks, there are also limitations. These include: the small screen, the limitation of battery life, and the limited types of interactions allowed. For instance, it is less convenient for people to type long sentences in a mobile device compared to that on a PC. With these limitations in mind, dedicated user experience studies need to be carried out while designing and implementing mobile based HITs.

Data security and privacy issues The proposed platform involves collecting data such as a user's location, activities, as well as other personal information measured by the sensors. A major concern is therefore data security and privacy issues. These personal

information and mobile users' activities may be disclosed or abused by malicious users, which will threaten all other users. This problem has to be approached by several angels, including well-designed policies restricting the flow of data and information, and also by carefully considering which data needs to be anonymised, and which can be encrypted.

5 Conclusion

In this vision paper, we have discussed a mobile crowdsourcing paradigm which can augment the current web-based crowdsourcing platforms to provide real-time location based query response using mobile devices. Towards this goal, we have provided a hybrid-crowdsourcing architecture and discussed several facets to realize this vision. This mobile crowdsourcing approaches can filter and target workers who more closely matches not only the queries, but also the location and context requirements. In addition, some of the processing can be done in the centralized web-based part of the proposed architecture, which reduces the burden of processing queries on the mobile devices. This allows several new application of crowdsourcing, especially in the areas of crowd-sensing, ubiquitous computing, and smart cities.

6 References

[1] T. S. Behrend, D. J. Sharek, A. W. Meade, and E. N. Wiebe. The viability of crowdsourcing for survey research. *Behavior research methods*, 43(3):800–813, 2011.

[2] E. Bonabeau. Decisions 2.0: the power of collective intelligence. *MIT Sloan Management Review*, pages 45–52, 2009.

[3] I. Borst. *Understanding Crowdsourcing – effects of motivation and rewards on participation and performance in vountary online activities*. PhD thesis, Erasmus research institute of Management, Rotterdam School of Management, Erasmus School of Economics, Erasmus University Rotterdam, 2010.

[4] J. V. den Ende, A. Villarroel, and C. Tucci. Strategic crowdsourcing, orchestrating innovation through the cream of the crowd, 2009. Panel symposium, Academy of Management Conference.

[5] N. Eagle. txteagle: Mobile crowdsourcing. In N. Aykin, editor, *Internationalization, Design and Global Development*, volume 5623 of *Lecture Notes in Computer Science*, pages 447–456. Springer Berlin Heidelberg, 2009.

[6] M. J. Franklin, D. Kossmann, T. Kraska, S. Ramesh, and R. Xin. Crowddb: answering queries with crowdsourcing. In *Proceedings of the 2011 ACM SIGMOD International Conference on Management of data*, pages 61–72. ACM, 2011.

[7] A. Gupta, W. Thies, E. Cutrell, and R. Balakrishnan. mclerk: Enabling mobile crowdsourcing in developing regions. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, CHI '12, pages 1843–1852. ACM, 2012.

[8] J. He, J. van Ossenbruggen, and A. P. de Vries. Do you need experts in the crowd?: a case study in image annotation for marine biology. In *Proceedings of the 10th Conference on Open Research Areas in Information Retrieval*, pages 57–60, 2013.

[9] J. Howe. The rise of crowdsourcing. 14(6), September 2009. http://www.wired.com/wired/archive/14.06/crowds.html.

[10] G. Jouret. Inside cisco's search for the next big idea. *Harvard Business Review*, pages 43–45, September 2009.

[11] A. Kittur, E. H. Chi, and B. Suh. Crowdsourcing user studies with mechanical turk. In *Proceedings of the SIGCHI conference on human factors in computing systems*, pages 453–456. ACM, 2008.

[12] N. Komninos. *Intelligent cities: innovation, knowledge systems and digital spaces*. Spon Press, 2002.

[13] N. Komninos. Intelligent cities: towards interactive and global innovation environments. *International Journal of Innovation and Regional Development*, 1(4): 337–355, September 2009.

[14] K. Kunze, H. Kawaichi, K. Yoshimura, and K. Kise. Towards inferring language expertise using eye tracking. In *CHI'13 Extended Abstracts on Human Factors in Computing Systems*, pages 217–222. ACM, 2013.

[15] J. Lampel and A. Bhalla. The role of status seeking in online communities: Giving the gift of experience. *Journal of Computer-Mediated Communication*, 122, 2007. http://jcmc.indiana.edu/vol12/issue2/lampel.html.

[16] C. Lofi, J. Selke, and W.-T. Balke. Information extraction meets crowdsourcing: A promising couple. *Datenbank-Spektrum*, 12(2):109–120, 2012.

[17] P. Lukowicz, A. Pentland, and A. Ferscha. From context awareness to socially aware computing. *IEEE Pervasive Computing*, 11(1):32–41, 2012.

[18] P. Narula, P. Gutheim, D. Rolnitzky, A. Kulkarni, and B. Hartmann. Mobileworks: A mobile crowdsourcing platform for workers at the bottom of the pyramid. In *Human Computation*, 2011.

[19] D. Schuermann and S. Sigg. Secure communication based on ambient audio. *IEEE Transactions on mobile computing*, 12(2), 2013.

[20] J. Selke, C. Lofi, and W.-T. Balke. Pushing the boundaries of crowd-enabled databases with query-driven schema expansion. *Proceedings of the VLDB Endowment*, 5(6):538–549, 2012.

[21] S. Sigg, D. Schuermann, and Y. Ji. Pintext: A framework for secure communication based on context. In *Proceedings of the Eighth Annual International ICST Conference on Mobile and Ubiquitous Systems:Computing, Networking and Services (MobiQuitous 2011)*, 2011.

[22] S. Sigg, D. Gordon, G. v. Zengen, M. Beigl, S. Haseloff, and K. David. Investigation of context prediction accuracy for different context abstraction levels. *IEEE Transactions on Mobile Computing*, 11(6):1047 –1059, june 2012.

[23] S. Sigg, M. Scholz, S. Shi, Y. Ji, and M. Beigl. Rf-sensing of activities from non-cooperative subjects in device-free recognition systems using ambient and local signals. *IEEE Transactions on Mobile Computing*, 13(4), 2013. doi: http://doi.ieeecomputersociety.org/10.1109/TMC.2013.28.

[24] S. Sigg, U. Blanke, and G. Troester. The telepathic phone: Frictionless activity recognition from wifi-rssi. In *IEEE International Conference on Pervasive Computing and Communications (PerCom)*, PerCom '14, 2014.

[25] C. Wu, J. Gerlach, and C. Young. An empirical analysis of open source software developers motivations and continuance intentions. *Information and Management*, 44:253–262, 2007.