

Qualitative Collaborative Sensing in Smart Phone Based Wireless Sensor Networks

Wilson Thomas, E. Madhusudhana Reddy

Abstract: Collaborative sensing has become a novel approach for smart phone based data collection. In this process individuals contribute to the participatory data collection by sharing the data collected using their smart phone sensors. Since the data is gathered by human participants it is difficult to guarantee the Quality of the data received. Mobility of the participant and accuracy of the sensor also matters for the quality of data shared in such environment. If the data shared by such participants are of low quality the purpose of collaborative sensing fails. So there must be approach to gather good quality of data from participants. In this paper we propose a Truth Estimation Algorithm (TEA) to identify the truth value of the data received and filter out anomalous data items to improve the quality of data. To encourage the participants to share quality information we also propose an Incentive Allocation Algorithm (IAA) for qualitative data collection.

Keywords—collaborative sensing; Truth Value; Smart Phones; Truth Discovery, Incentive based approach.

I. INTRODUCTION

Smart phones are unprecedented tool for collecting empirical data. Modern smart phones are equipped with accelerometer, gyroscope, GPS, and high resolution camera. They are capable of sensing, mobile computing, and data sharing information with the support of wide mobile networks and social networking sites. Since these devices travel with people they can be used to collect more precise and real-time data in space and time. Collaborative sensing [1] is the process where individuals and communities use their smart phones and internet to collect and analyze real-time data on the fly. This technique is useful for developing applications for social well being like Pollution detection, health care applications, security and surveillance etc. The major component of this architecture is the participants with smart phones who collect and share sensor data such as images, voice records, accelerometer data etc. The collaborative sensing application can be distributed among participants through application store such as Google play or App store. In collaborative sensing using smart phones, data gathering can be triggered by participants themselves or through campaigning [2] in which individuals shares sensor data in response to queries of the organizer.

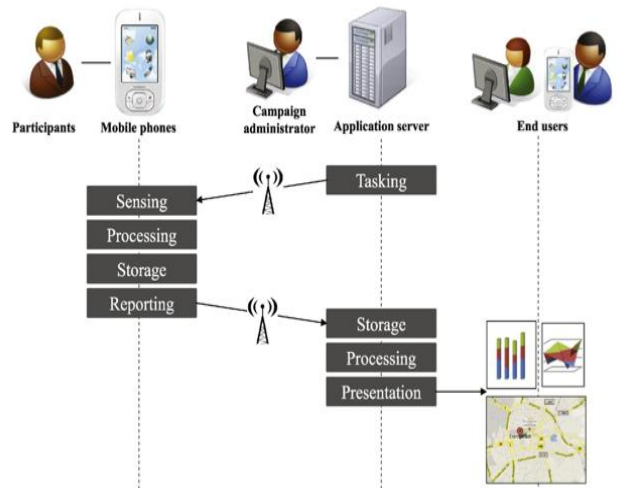


Figure 1: Collaborative Sensing & data Analysis

Data collected through the campaign need to be analyzed by application servers to extract useful information. However it is difficult to ensure quality of the data collected by these smart phones as the participant behave unreliable and maliciously. Hence the major challenge in the use of social sensing application is the noisy nature of data received. In this context the algorithms that can estimate the data quality of the data and underlying truth have drawn significant attention. There many Truth estimation Algorithms available in different research areas. Truth estimation approaches are generally not supervised, hence the originality of the source can be trusted based on the quality of data. High trust values will be assigned to the sources which give reliable data. These data collected from such reliable sources will be considered as truths. Teseus is a payment based approach that deals with workers' deliberate behavior, and reward good quality data sharing from participants. In Bayesian Nash Equilibrium of the non-cooperative game induced by Teseus, all participants will best work on sensing, which improves their data quality. Hence the final results calculated afterward by truth discovery algorithms based on participant's data will be highly accurate. Incentives based approach in collaborative sensing focus on inspiring the participants to share useful data by giving monetary reimbursement. Mobility based incentivizing the participants are an approach for quality data collection in collaborative sensing. In this approach people are encouraged to move to the unknown regions for data collection. With this approach the participant gets incentivized and organizer gain more accurate data.

Revised Manuscript Received on October 22, 2019.

Wilson Thomas, Research Scholar, Research & Development Center
Bharathiar University, Coimbatore, India

E. Madhusudhana Reddy, Professor, Department of CSE, Guru Nanak
Institutions Technical Campus, Telangana, India
E.Mail id: e_mreddy@yahoo.com

II. RELATED WORK

There has been many research conducted on addressing the problem of data accuracy through data aggregation, trust value calculation, malicious behavior of participants. These people can disrupt the sensing purpose by supplying low quality information. They can be spamming or misleading other participants. [3] Propose an approach for preserving privacy and quality of data in collaborative data sharing with Quality-aware Attribute-based Filtering technique. Anonymization Techniques for Preserving Data Quality in Participatory Sensing [4] discuss optimization techniques to preserve data quality. [5] Explains data quality estimation in crowd sensing platforms based on context. [6] Proposes techniques for achieves higher decision accuracy of sensor nodes. [7] Suggests quality estimation and monetary incentive, for participatory sensing.

III. QUALITATIVE DATA COLLECTION

Truth Discovery Techniques

We have proposed a series of approaches for Truth estimation of the source of the data.

a. Truth estimation Algorithm (TEA)

Research on trust estimation [8] considers the fact that data gathered about the same event by two or more participants are likely to conflict. This can be due to many reasons like imperfect data collection, network delays, bad weather etc. The following notations are used in the proposed TEA algorithm.

- Object: An object $i \in G$ is the object to be monitored for data gathering.
- Smart Phone: A Smart Phone $s \in S$ is the individual participant in the event. $\{X_{i,s}\}$ denote object i is monitored by smart phone s .
- Source-claim matrix: This matrix is of the order $S \times G$. Each entry $S_i G_j$ says whether a smart phone s has claimed the ownership of entry j .
- Quality: Q_s express the quality metric for the probability that smart phone s provide reliable data.
- Truth: This value for an object i , considered as T_j , is the most reliable data collected from a set of observations.

Problem statement: Let S be a set of smart phones and G be the set of events being monitored. For each smart phone $s \in S$, estimate an optimal truth value $\{T_s\}$ where $s \in S$, from a set of observations $\{X_{i,s}\}$ where $i \in G, s \in S$. received from different sources.

Algorithm1: Framework for Truth value calculation

Input: Observations from sources $\{X_{i,s}\}$ where $i \in G, s \in S$

Output: Optimal truths $\{T_j\}_{j \in G}$ and the estimated source probability $\{P_s\}_{s \in S}$

Initialize: Quality estimation $\{Q_i\}_{i \in S}$

while optimization not achieved **do**

for each $i \in G$ **do**

Truth value finding: infer this value for object

i

Using the recent finding of quality estimation Q_i

end

Quality estimation: update quality value $\{Q_i\}_{i \in S}$ using the latest calculated truth value.

end

Return: $\{T_i\}_{i \in G}, \{Q_i\}_{i \in S}$

One of the pioneer approach in truth finding is proposed in [9] where TruthFinder decides reputation of each source based on probability that it sends the correct data. It then averages the reputation value to calculate the trustworthiness of the source.

$$Q_i^j = \frac{\sum_{s \in S} x_{i,j}^s}{|\{S_i G_j\}|} \quad (1)$$

In [10] it estimates the truthiness of a claim by Truth calculates the "probability" of a claim by considering that reliability of data source is the probability that it sends correct data. It takes the average of reliability values to calculate truthiness for every claim from participants.

Incentive based Approach

In this approach we encourage the participants to contribute valuable data by providing incentives. This algorithm takes set of target events E and participants P and sensor data s , and $\{a_p, b_p\}_{p \in P}$ where a_p and b_p are quality factors related to the participants. The calculation of incentive to any participant is based on peer prediction method [11] which decides the incentives based on the data collected by a participant and the difference of it with a reference participant data.

b. Incentive Allocation Algorithm (IAA)

Input: $T, P, P', s, \{a_p, b_p\}_{p \in P}$;

Output: $\{Q_p \mid p \in P\}$;

foreach participant $p \in P$ **do**

if $p \in P'$ **then**

Pickup another participant $w \in P'$ in random;

$$Q_s \leftarrow b_p - a_p \frac{1}{T} \sum_{t=1}^T (s_t^p - s_t^w);$$

else

$$Q_s \leftarrow 0;$$

return $\{Q_p \mid p \in P\}$;

As per the above IAA algorithm if the participant p take part in data collection algorithm picks another random participant. The payment Q_s for the participant is calculated as

$$Q_s \leftarrow b_p - a_p \frac{1}{T} \sum_{t=1}^T (s_t^p - s_t^w)$$

If the participant's data shows more similarity to the reference data higher the payment Q_s will be. If the participant moves out from the task the payment Q_s is set to 0 for him. At the end the algorithm return incentives

allocated for all the participants.

VI. PERFORMANCE EVALUATIONS

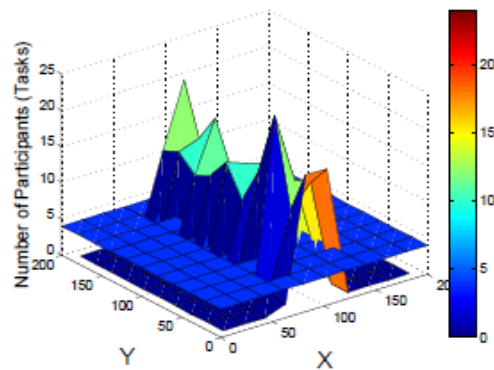
Here, we analyze the performance of our approaches for qualitative data collection, Truth Estimation Algorithm (TEA) and Incentive Allocation Algorithm (IAA). We compare the performance of our IAA algorithm with the well known research work in this area called MSensing [15] auction technique.

A. Simulation Setup

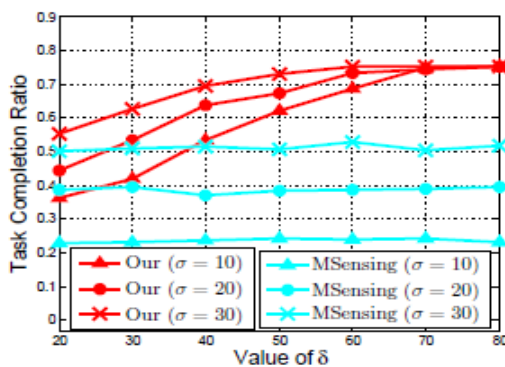
For simulation set up we deploy the smart-sense application in an area of $200m \times 200m$. This square region is again partitioned into 100 small squares of $20m \times 20m$ area. We assume that 400 sensing tasks and 300 participants are distributed randomly the area.

We consider below parameter for analysis.

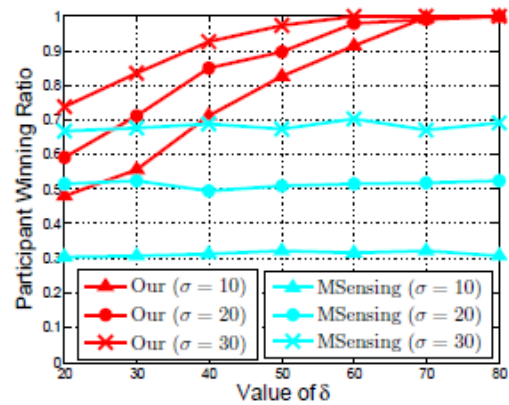
- **Completion Ratio** is the ratio of data collection task finished by the individuals.
- **Allocation ratio** is the ration of participants winning the auction being allocated incentives.
- **Social Welfare** is completed data collection jobs minus total data collection cost.



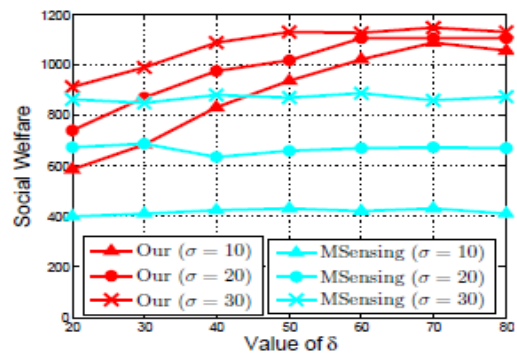
(a) Simulation Model



(a) Completion Ratio



(b) Allocation Ratio



(c) Social Welfare

B. Performance comparison

Simulation results shows that our approach outperforms MSensing with respect to the performance parameters (a) task completion ratio, (b) allocation ratio and (c) social welfare. It performs better when the number of participants are high. When $n = 150$ and $\sigma = 10$, with respect to MSensing, our approach improves parameter (a) by 100%, parameter (b) by 66.6%, and parameter (c) by 63%. Now consider $\sigma = 10$ and $n=300$, with respect to MSensing, our approach improves the parameter (a) by 145.4%, parameter(b) by 143.3%, parameter (c) by 146.7%. The reason for this is, in MSensing mechanism the participants are grouped in popular regions and there is a chance of losing the auction. But incentive based approach encourage the individual to change location to unknown regions and finish data collection.

VII. CONCLUSION

The TEA algorithm helps to calculate truthiness of the collaborative data received. In IAA approach we have proposed a technique to incentivize participants of collaborative data collection based on mobility where participants are motivated for data collection from unpopular areas. Based on theoretical analysis it is clear that the proposed approach holds true for truthfulness and efficiency in qualitative data collection. Hence the incentive based mechanism outperforms the traditional campaigning approach for collaborative data collection.

REFERENCES

1. Xiang Sheng, J. Tang and W. Zhang, "Energy-efficient collaborative sensing with mobile phones," *2012 Proceedings IEEE INFOCOM*, Orlando, FL, 2012, pp. 1916-1924.
2. F. Qiu, F. Wu and G. Chen, "Privacy and Quality Preserving Multimedia Data Aggregation for Participatory Sensing Systems," in *IEEE Transactions on Mobile Computing*, vol. 14, no. 6, pp. 1287-1300, June 1 2015.
3. T. Sabrina, M. Murshed and A. Iqbal, "Anonymization Techniques for Preserving Data Quality in Participatory Sensing," *2016 IEEE 41st Conference on Local Computer Networks (LCN)*, Dubai, 2016, pp. 607-610.
4. F. Qiu, F. Wu and G. Chen, "Privacy and Quality Preserving Multimedia Data Aggregation for Participatory Sensing Systems," in *IEEE Transactions on Mobile Computing*, vol. 14, no. 6, pp. 1287-1300, June 1 2015.
5. S. Liu, Z. Zheng, F. Wu, S. Tang and G. Chen, "Context-aware data quality estimation in mobile crowdsensing," *IEEE INFOCOM 2017 - IEEE Conference on Computer Communications*, Atlanta, GA, 2017, pp. 1-9.
6. S. Yang, F. Wu, S. Tang, X. Gao, B. Yang and G. Chen, "Good Work Deserves Good Pay: A Quality-Based Surplus Sharing Method for Participatory Sensing," *2015 44th International Conference on Parallel Processing*, Beijing, 2015, pp. 380-389.
7. Yaliang Li, Jing Gao, Chuishi Meng, Qi Li, Lu Su, Bo Zhao, Wei Fan, and Jiawei Han. 2016b. A Survey on Truth Discovery. SIGKDD Exploration Newsletter 17, 2 (Feb. 2016), 1-16.
8. X. Yin, J. Han and P. S. Yu, "Truth Discovery with Multiple Conflicting Information Providers on the Web," in *IEEE Transactions on Knowledge and Data Engineering*, vol. 20, no. 6, pp. 796-808, June 2008. doi: 10.1109/TKDE.2007.190745
9. R. W. Ouyang, M. Srivastava, A. Toniolo and T. J. Norman, "Truth Discovery in Crowdsourced Detection of Spatial Events," in *IEEE Transactions on Knowledge and Data Engineering*, vol. 28, no. 4, pp. 1047-1060, April 1 2016.
10. N. Miller, P. Resnick, and R. Zeckhauser, "Eliciting informative feedback: The peer-prediction method," in *Management Science*, 2005.
11. Yan Sun, Hong Luo, and Sajal K. Das. 2012. A Trust-Based Framework for Fault-Tolerant Data Aggregation in Wireless Multimedia Sensor Networks. *IEEE Transactions on Dependable and Secure Computing* 9, 6 (Nov 2012), 785-797. DOI: <http://dx.doi.org/10.1109/TDSC.2012.68>
12. Xiaoyong Li, Feng Zhou, and Junping Du. 2013. LDTS: A Lightweight and Dependable Trust System for Clustered Wireless Sensor Networks. *IEEE Transactions on Information Forensics and Security* 8, 6 (June 2013), 924-935. DOI: <http://dx.doi.org/10.1109/TIFS.2013.2240299>
13. Gang Wang, Sarita Schoenebeck, Haitao Zheng, and Ben Zhao. 2016b. "Will Check-in for Badges?": Understanding Bias and Misbehavior on Location-Based Social Networks. (2016). <http://www.aaii.org/ocs/index.php/ICWSM/ICWSM16/paper/view/13027>
14. M. Talasila, R. Curtmola, and C. Borcea, "Alien vs. mobile user game: Fast and efficient area coverage in crowdsensing," in *IEEE MobiCASE*, 2014.
15. G. Xue, X. Fang, and J. Tang, "Crowdsourcing to smartphones: incentive mechanism design for mobile phone sensing," in *ACM Mobi-Com*, 2012.