Task Selective and Comfort-Aware User Recruitment with Incentives in Mobile Crowd-Sensing

by

Venkat Surya Dasari

Thesis Supervisor : Dr. Burak Kantarci

A thesis
presented to the University of Ottawa
in fulfillment of the
thesis requirement for the degree of
Master of Applied Sciences

School of Electrical and Computer Engineering
Faculty of Engineering
University of Ottawa

I hereby declare that I am the sole author of this thesis. This is a true copy of the thesis, including any required final revisions, as accepted by my examiners.

I understand that my thesis may be made electronically available to the public.

Abstract

With the significant improvement in IoT technology and smart devices, data collection and distributed computation have led a foundation for Mobile crowd-sensing (MCS). MCS utilizes the capabilities of embedded sensors in smart devices for gathering data. MCS benefits both data provider (participant/user), and data requester, i.e. data providers via incentives/rewards, data requesters by delivering required data.

Apart from the benefits gained through acquiring data, confronting challenges such as participant privacy, data trustworthiness, malicious attacks (from illegitimate users) need to be addressed to build robust and reliable data solicitation. In addition to that, it is necessary to consider user motivation and user preference, comfort during its engagement in crowd-sensing. User preferences/constraints can be due to privacy concerns in terms of location, the sensitivity of data or energy usage and many more. With this in mind, the main contributions of the thesis can be listed as follows.

1) We design user selective trustworthy data acquisition frameworks. We introduce a variety of user selection criteria to form participant communities based on participants reliability and income. To evaluate the trustworthiness of our selective reputation-based data acquisition, we consider malicious users in the environment and calculate the total rewards given to malicious users. Simulations results show that community formation based on the acquired income of participants ended up with a substantial loss to the cloud platform as well as participants. Contrary to that, reputation-based community formation has shown nearly equal platform utility (profit), negligible loss of user utility compared to benchmark Non-selective data acquisition with 7% malicious probability.

2) Moreover, we attempt to enable users to modify (allow/deny access to) their built-in sensor set according to their comfort levels. We formulate three comfort levels *high* (only allow access to sensors that would not directly reveal personal identity such as accelerometer, light sensor, etc.), *moderate* (obstruct access to sensitive data, e.g. camera), $zero\ comfort$ (allow access to all users). We introduce *Static* modification, where users pre-arrange their sensor set before the start of data collection. Our feasibility study shows that pre-arrangement of the sensor set favours user comfort, user utility at the cost of loss in platform utility and performs better than selective reputation-based recruitment for the considered settings.

3) We apply *Adaptive* sensor modification on top of pre-arrangement of sensor set through which participants are authorized to re-arrange their sensor availability based on reliability scores. Simulation results show that the *Adaptive* comfort-aware approach performed better than *static* in terms of platform utility and achieved comparatively better user comfort with reasonable loss in user utility.

Acknowledgements

I would like to express my sincere gratitude to Dr Burak Kantarci for his continuous support and valuable suggestions throughout my Masters. The constant feedback and motivation during the study helped me to publish an ample amount of research papers in esteemed venues. I wouldn't have chosen research-based Masters without meeting Dr. Kantarci in my first course of Masters in the University of Ottawa, which I always consider as one of my greatest achievements. I must thank Dr Maryam Pouryazdan and Dr Murat Simsek for their invaluable collaboration in the published articles.

I must thank my parents for their strong support and belief in me. I am glad to have them as my parents for everything they provided me. I am thankful to my Thesis Examiners, Prof. Hussein T. Mouftah, Prof. Ashraf Matrawy, for their substantial time and efforts spent on the thesis report. Last but not least, I would like to mention my research group for creating an admirable work environment.

Table of Contents

List of Tables

List of Figures

Glossaries

$CLAPS$:	Comfort Level-Aware Participant Selection scheme that considers only user comfort alongside the bids during the selection of a winner.
MCS :	Mobile Crowd-sensing is a technique that acquires data with the help of embedded sensors in smart devices carried by a large group of individual-sAdvanced High Strength Steel
NSR :	Non-Selective reputation-based recruitment strategy that recruits users based on their reputation without offering any choice to choose the community to join for sensing.
$RA-CLAPS$:	Reputation and Comfort Level-Aware Participant Selection scheme that considers both user comfort as well as user reputation alongside the bids during the selection of a winner.
SNR :	Selective data acquisition with no reputation-awareness is a user recruitment strategy that offers the user a choice to choose the community that has lower income than participants income. Winner selection is completely based on bids of participants.
$SRHI$:	Selective and Reputation-aware data acquisition with low-income service providers is a user recruitment strategy that recruits users based on their reputation while offering the choice to choose a community that has a comparatively higher income than the participant's income.
$SRLI$:	Selective and Reputation-aware data acquisition with low-income service providers is a user recruitment strategy that recruits users based on their reputation while offering the choice to choose a community that has lower income than the participants income.
SRR :	Selective and Reputation-aware Recruitment strategy that recruits users based on their reputation while offering the choice to choose a community

to participate in sensing on the basis of communities reputation.

WSN : Wireless Sensor Networks that is capable of collecting data through deployed sensor nodes.

List of Symbols

$A.E$ Average Energy consumption of a user

R_i Present Reputation of user/sensor i

$R_i(t-1)$ Previous reputation of user/sensor i

R_{total} Total reputation of users/sensors that can participate for task T

$T.E$ Total Energy consumption of all the selected users

T_U Set of tasks that can be sensed by the users/sensors in the set U in a participatory manner.

T'_W Set of tasks out of the set T_U that have been opted-in to be sensed by at least one user of the set W

$T_{\mathcal{C}}$ Sensor associated task T

Th_{High} Threshold value for high reputation level

$Th_{Moderate}$ Threshold value for high reputation level

Th_{Naive} Threshold value for high reputation level

n Factor of Index of campaign at which sensor set is reset

$U_{platform}$ Platform Utility

U_{user} User Utility

Γ_T List of users/sensors that can participate in the sensing of task T

Γ'_T List of users/sensors out of the set Γ_T that have opted-in to participate in the sensing of task T

δ Weight factor

κ_i^T Binary decision of user/sensor i to accept or decline task T

$\mathcal{C}_i$ Set of sensors corresponding to the comfort group of user/smartphone i

T Total crowd-sensing participants available in the complete terrain

T Index of a task

W Winner set

ϱ Closeness margin used in the decision of accepting/declining a sensing task

a_i Binary value that represents the availability of *Accelerometer* in sensor set of user i

c_i Sensing cost (i.e. auction bid) of user/sensor i

c_i Binary value that represents the availability of $Camera$ in sensor set of user i

l_i Binary value that represents the availability of *Lightsensor* in sensor set of user i

m_i Binary value that represents the availability of $Microphone$ in sensor set of user i

n_i Negative (outlier) readings of user/sensor i

p_i Positive readings of user/sensor i

$r_{max}^{\Gamma_T}$ Maximum total rewards received by the users/sensors in the set T

$r_{\Gamma_T}^{avg}$ Average total rewards received by the users/sensors in the set T

r_i Total rewards issued to participant i
r_i^t Total rewards issued to participant i in sensing campaign t

t_{total} Total number of sensing campaigns

$v^R(W)$ Reputation-based value of the set of participants W

$v^R(W^t)$ Total reputable value obtained from winner set W in the sensing campaign t

v_T Value of task T

$v_i^R(W)$ Reputation-based marginal value of user/sensor i over the set of participants W

w_a weight of *Accelerometer* or sensitivity of data that could be disclosed with use of it.

w_a weight of $Microphone$ or sensitivity of data that could be disclosed with use of it.

w_a weight of *Lightsensor* or sensitivity of data that could be disclosed with use of it.

w_c weight of $Camera$ or sensitivity of data that could be disclosed with use of it.

x_i^{Tt} Binary value that shows 1 when user i bid for the task T in campaign t.

Publications of the Candidate During MCS Studies

- **V.S. Dasari**, M. Pouryazdan , B. Kantarci, "On the impact of selective data acquisition in mobile crowd-sensing performance," in *IEEE Canadian Conference on Electrical and Computer Engineering*, (Montreal, QC, Canada), May 2018

- **V.S. Dasari**, M. Pouryazdan , B. Kantarci, "Selective Versus Non-Selective Acquisition of Crowd-Solicited IoT Data and its Dependability," in *IEEE International Conference on Communications*, (Kansas City, MO, USA), May 2018

- **V.S. Dasari**, B. Kantarci and M. Simsek "Trustworthiness and Comfort-Aware Participant Recruitment for Mobile Crowd-Sensing in Smart Environments," in *IEEE Symposium on Computers and Communications (ISCC)*, (Barcelona,Spain), June 2019
  (Accepted)

- **V.S. Dasari**, B. Kantarci and M. Simsek, Adaptively Managed Participant Comfort in the Acquisition of Trustworthy Mobile Crowdsensing Data, Elsevier Internet of Things, 2019 (Submitted).

Chapter 1

Introduction

Distributed sensors are used for acquiring data in traditional sensing techniques, such as wireless sensor networks (WSN) [39]. Capitalizing on the advancements in mobile technology would be a smart approach to overcome the challenges faced using WSN for data acquisition such as low coverage, high cost and low reusability. Today, many applications extract information from smartphones to provide real-time data. One of the popular examples is Google traffic update. In this context, Mobile crowd-sensing (MCS) appeared as a feasible tool for data collection using non-dedicated sensors such as smart embedded sensors[6]. MCS is a technique that collects the required sensing information through smart devices carried by people. It is also expressed as smartphone sensing, participatory sensing [10], mobiscopes [2], opportunistic sensing [12], people-centric sensing [13]. The Internet of Things (IoT) has transformed society with its wide variety of applications that require a considerable amount of data. Cisco states that 30 billion IoT connected devices are expected to generate 2.5 quintillions of data per day by 2020 [1]. Scalability, cost-effectiveness and mobility of participants make MCS as a data fueling agent to accomplish the needs of data craving applications [70]. Furthermore, MCS facilitates the participants to get profited through different types of incentive mechanisms alongside the cloud platform in the form of obtaining the requested data.

Advancement in MCS has led to the development of various applications by researchers in multiple fields such as sustainability [84, 60], health monitoring [85, 75, 54], traffic management, monitoring [4, 101, 16, 92], public safety such as emergency situations [88], and even psychological survey questionnaires [89] thus improving quality of services. Fig 1.1 showcase the various fields that make use of crowd-sensing.

Data collection in MCS can be achieved either through participatory sensing or oppor-

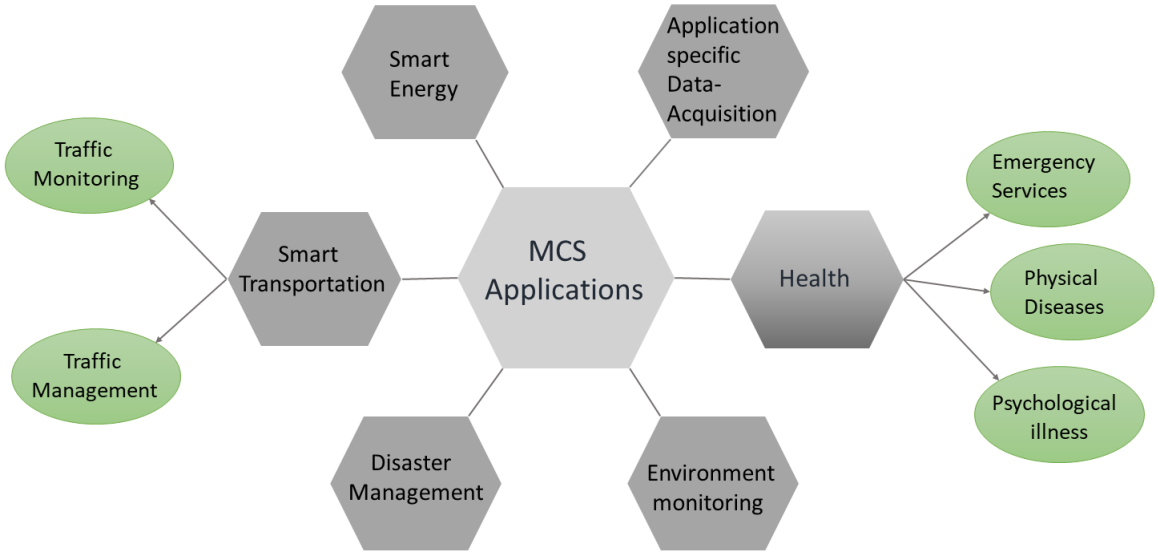


Figure 1.1: Applications of MCS

tunistic sensing. Participatory sensing requires human support for data collection such as capturing a scene, reporting an accident etc. On the other hand, in opportunistic sensing approach users' intervention is very minimal or negligible. Examples include health, traffic monitoring applications. Driving entities of MCS are

- Data requester: Entity that request for data acquisition. Some MCS systems consider data requester has its cloud platform.
- User/Participants: Also referred to as data providers, users sense tasks and transfer data to the cloud platform.
- Cloud Platform: Computationally efficient and data storage unit where collected data is computed and analyzed and transferred to data requester. Moreover, it plays a mediator role in assigning incentives to the participants.

A natural process for data acquisition through non-dedicated sensors include incentivization of data providers for sensing the data because sensors consume battery power, time, transmission cost, computational power during data collection. Incentives play a crucial role in motivating participants. Indeed, effective incentive mechanisms help in fetching qualitative and quantitative information according to the requirement of data requester.

Figure 1.2: Built-in sensors of Smart devices

The list of mostly observed built-in sensors in a regular smartphone is presented along with their main functionalities in Fig 1.2.

Despite MCS having fascinating advantages, it has considerable limitations such as

data quality which is indeed guaranteed to be high in WSN. Out of all, openness creates a significant problem concerned that people not necessarily be truthful and may deliberately send false data to cause harm to the reliability of the collected data. Obtaining untrustworthy data would cause substantial loss than fetching no data. Thus, ensuring the trustworthiness of data should be a key objective of MCS. Furthermore, participants use diverse types of networks while transmitting data. Besides this, data collection involves various kinds of mobile devices, responsible for diversity in the obtained sensor accuracy, transmission time, computation power based on the capabilities of an individual device. With the benefits and limitations mentioned, MCS could complement and strengthen WSN but need not be a complete replacement.

Hewlett Packard report [71] shows that there exist an average of 25 ways to attack an IoT device. Almost 9 of 10 devices tested had collected a minimum of some of part private information such as location, credit details, residential address, personal email etc. In addition to that, they also mentioned the collection of private data through mobile applications and suggested that end to end solutions such as HP Fortify on Demand would efficiently avoid software vulnerabilities. The privacy concerns of the participants exist in mobile crowd sensing since data collection is executed through non-dedicated sensors such as smart devices. MCS framework should also be conscious of energy constraints, privacy and security challenges due to the usage of personal devices for data acquisition and reporting. Many researchers contributed their work towards securing the data [24, 35, 90] as well as to protect the privacy of participants [83, 34, 79].

We contemplate that users have various concerns/constraints that depend on location, residual battery, duration of a task and many more. In this thesis, we address the user concerns by 1) Providing an opportunity for users to indirectly select the task to contribute to choosing a group to join. We denote it by user selectiveness throughout the thesis. 2) Secondly, one step further, we introduce another metric called user comfort that is a measure of how convenient the users are taking part in crowd-sensing. User comfort is gauged by weight of built-in sensors that can be switched off during the data acquisition. In other words, it allows users to announce their unwillingness to perform specific tasks deliberately.

1.1 Motivation

To utilize the embedded sensors of smart devices that are widely available, MCS appeared as a new paradigm for data sensing. In addition to the aforementioned challenges faced

by MCS, user recruitment policy plays a crucial role in the success of the data acquisition through non-dedicated sensors. The platform always aims at acquiring appropriate data/requested data for the lowest possible price. On the other hand, users aim at maximizing their income from data requester for contributing data to the crowd sensing system holding to their preferences/ constraints Preference may be due to concern in privacy or energy consumption or convenience in sensing a particular task and also time, maximum tasks that a participant wishes to sense etc. Incentive mechanisms deal with user costs to encourage users to continue to participate in data sensing. User costs include energy usage, transmission costs (mobile data, WiFi), user time and effort etc. Preferences, constraints of recruiter and data provider can be addressed through user recruitment schemes. To this end, we propose user selectiveness through which participants could opt to join a community based on different criteria (see chapter 3). Furthermore, we present a term called user comfort-awareness (refer chapter 4) that emphasizes the user preference to switch-off some built-in sensors that raise any concerns as mentioned earlier. Along with this, we simultaneously aim to collect reliable data in a vulnerable environment that contain some portion of malicious/ illegitimate users.

1.2 Objectives

Mobile crowd sensing can be efficient when a large number of participants join in sensing campaigns. One of the major challenges that MCS need to be aware of is to maintain substantial number of participants. There is a reasonable chance that participants would not contribute data to the platform because of privacy and security concerns [31]. Apart from these, incentives also play a crucial role in motivating participants to take part in crowd sensing. An efficient user preference/constraint aware recruitment scheme and an user-centric incentive mechanism could be beneficial to address this problem.

The main objectives of this thesis is

1) To address participants' preferences, comfort while ensuring data trustworthiness in presence of malicious attacks.

2) To achieve overall efficient system by maintaining balance among user comfort, platform and user utilities.

1.3 Contributions

Our contributions in this thesis can be summarized as follows.

- We investigate the impact of community selectiveness of user in the data acquisition process alongside the presence of malicious users. The selectiveness of a user can be summarized as the opportunity for a user to select a community/group that sense the same task based on the average reputation of the group (SRR) or average income of group (SRLI, SRHI). The selectiveness is introduced with obtaining acceptable platform utility.
- A feasibility study of user comfort in mobile crowd sensing. We propose a user comfort aware participant selection scheme (RA-CLAPS) to take into account user preferences while participating in data acquisition. Users modify sensor configuration based on comfort level before registering for crowd-sensing in RA-CLAPS.
- We further improve the proposed comfort aware model to investigate the impact of adaptive modification of access permission to the built-in sensor set. In such a process, the cloud platform permits users to choose to dynamically modify their sensor configuration based on the reliability level of participants for a specified period.

1.3.1 Structure of the Thesis

The remaining part of the thesis is organized as follows:

Chapter 2 represents the background and literature study: Section 2.1 gives overview of the background of MCS, in Section 2.2 we discuss the challenges and issues in MCS such that each Sub-Section concentrate on separate issue/chalenge. In Section 2.2.3 we present the works that presented user community formations in MCS Section 2.4 investigates the related research works that consider user preferences/comforts.

In chapter 3, we present a selective data acquisition model in MCS with Section 3.1presents introduction followed by the problem statement, Section 3.2 describes the overview of system with the detailed discussion presented in Sub-Section 3.2.1, 3.2.2. In Section 3.3, we explain the formulation for the performance metrics followed by simulation results in Sub-Section 3.3.1, 3.3.2. We then summarize the chapter in Section 3.4.

In chapter 4, we present the proposed comfort-aware participant recruitment scheme: Section 4.1 provides overview of the problem statement, Section 4.2 elaborates the proposed recruitment scheme and Section 4.3 represents the performance metrics, Section 4.4

provides the detailed discussion on obtained simulation results followed by a summary of chapter in Section 4.5.

Chapter 5 represents the third contribution of this thesis. We propose an adaptive version of reputation and comfort-aware participant recruitment strategy. Section 5.1 introduces the problem statement, Section 5.2 provides system overview followed by Sub-Section 5.2.1 presents the preliminary study, Sub-Section 5.2.2 explains the proposed algorithm and Section 5.3 represents the detailed discussion of obtained results and the chapter ends with summary in 5.4.

Chapter 2

Background and Literature study

2.1 Mobile Crowd-Sensing Background

The ubiquity of smart devices such as mobile phones, tablets that are equipped with various sensing and computing capabilities has given potential support for the rise of Mobile crowd sensing(MCS) in data acquisition domain. MCS posses distinct characteristics that wireless sensor networks do not own such as flexibility in the coverage area, low-cost data sensing, mobility of nodes etc. In the first place, the term 'Mobile crowd-sensing' was introduced by Ganti et al. [31] to describe community sensing. Community sense can be either *participatory* [10] where user involvement is comparatively more or *opportunistic* [64] in which user involvement and effort is minimal during the data collection process. Since MCS collects data through people carrying smartphones, it is essential to ensure the truthfulness of participants as well as trustworthiness of data. Moreover, participants need to be motivated by incentives for their contribution, indeed maintaining platform utility. Besides this, MCS experience various challenges and issues that are listed in the following section 2.2.

2.2 Issues and Challenges in MCS

MCS confront many challenges, such as resource limitations of smart devices such as network connectivity, energy, computational capabilities [76, 108]. In addition to that, participant's concern about their privacy and data security is another matter of concern. In

this thesis, we consider user privacy, constraints together under user preference or willingness and propose user recruitment models to suffice user requirements. In this section, we survey research works that address different challenges and issues of MCS. Figure 2.1 shows the research challenges that are being addressed by efficient frameworks proposed by various researchers in the MCS domain.

Figure 2.1: Challenges in MCS

2.2.1 Trustworthiness of Data

The trustworthiness of data has a vital role in the success of data acquisition. It describes the truthfulness/ reliability of the obtained data. In this section, we present various techniques used by researchers to maintain the truthfulness of data. Reasons for the generation of untrustworthy data are presented below.

- One of the primary reasons for generation of incorrect data is participants may be unaware of effects due to misplacing the sensors' while collecting the data [45]. However, the efficiency of the sensor plays a substantial role in providing valid data.
- Recruitment of malicious users' (who intentionally send false data) likely to cause degradation in the performance of the crowd-sensing system.
- Sybil attacks could take place while transmitting the data.

- Lack of experience in sensing a task may lead to misplacing of mobile phones or paying less attention to data collection process leads to significantly low-quality data that cannot be reliable [87]

One of the main objectives of crowd-sensing is to extract accurate, appropriate data at the location of interest while ensuring data trustworthiness. Availability of only a single user data can be seen often, rendering it challenging to determine the reliability of data. Thus, the trustworthiness of data could be calibrated by the reliability score of the participant. Choosing a credible user reduces the risk of obtaining misleading data. To address this issue, an additional attribute is annexed to all participants known as reputation score. It represents the reliability of the participant that continously change on based on the degree of accurate data submitted to the platform.

Besides the strategy to acquire trustworthy data, it is equally important to consider the cost of sensing/bid of a user for providing the sensed data. Authors of [86] considered the above fact and selected appropriate participants based on the relation between bid cost and truthfulness and latency of report. Moreover, tasks associated with users interest can motivate users to provide quality data. Conjunction with that, reputation score and social interests were also calculated that sequentially improves platform utility by eliminating malicious users.

Authors of [100] went a step further and evaluated the degree of willingness alongside QoI of received data, performance ratio. Willingness denotes users response time, and high performance can be obtained if the user expects lower bids for his quality submissions, regarding these three factors reputation of a user, is calculated. Zhou et al. [126] proposed a framework that considers both spatial connections with data as well as the participants' previously submitted data quality. Thus, effectively avoided the intriguing users' from being selected with the help of clustering algorithms. The objective of the authors of [46] is to prevent misinterpretation of the participant's qualitative effort. They succeeded in distinguishing the cause of erroneous either by bias or due to lack of ability to perform a task in Amazon mechanical trunk.

From an application point of view, there can be data requester, recruiter/task publisher and sensing participants, for instance, crowd-sourcing applications like amazon trunk. Task publisher/recruiter takes care of deployment of tasks as well as recruiting phase and forwards the received data. Since the recruiter and requester are from different organizations/entities, an alternative approach of evaluating reputation was mentioned in [122]. Accordingly, with the satisfaction of data requester, the reputation of a participant is updated by task publisher. That being said, data requester might pretend to be unsatisfied even for a quality submission.

Authors of [23] approached data trustworthiness by ensuring data integrity with the help of external micro-controller, which is carried by every data provider that is connected via Bluetooth. Indeed, claiming that such procedure of ensuring data integrity is effective since it is uncomplicated for verifying location, time and other required attributes that provide data reliability. With the vast number of crowd-sensing users, it may not be possible to supply each participant with a trusted platform module (micro-controller). This problem can be addressed by analyzing/comparing trustworthy data results.

Another compelling reliability assurance model [82] considered the presence of entirely reliable users named as anchors. Users' who share the sensing task including anchors', vote each other on comparing their sensed data within themselves. Each user is associated with a reputation score that is updated based on positive and negative votes it obtains from neighbours. The level of impact for the posted vote is calculated based on the number of proper/trustful votes. Evaluation of reliability of participants may be complicated with a different ground truth value. Under these circumstances, authors of [99] proposed a system model to determine the truthfulness of data with the help of maximum likelihood estimation(MLE) model. Another relevant work [44] that utilizes MLE with consideration of spatio-temporal information in a truth disclosure process to improve the precision.

The study in [26] proposes a privacy-aware trustworthy data aggregation to mitigate the malicious attacks and also to protect user privacy since users may not trust the data collector. Enabling users to choose a suitable value from the predefined data value array narrows down the options making the system vulnerable compared to a scenario in which users' may report any value. The proposed algorithm implements the data encryption by which platform concludes that on varying the data integrity without knowing the data submitted, thus ensuring user privacy. Furthermore, they claim that the proposed model detects the false data from aggregated data even though the range of data vector is broad and reduces the malicious attack up to a certain extent.

2.2.2 Reputation of Participants

The primary difference between trustworthiness and reputation is a measure of participants trutfulness and trustworthiness is a measure for the reliability of data. The authors of [30] proposed a reputation based framework that is built on the basis of the beta probability density function. Huang et al. [45] used trustworthiness of data for evaluation of participants reputation in a noise monitoring application. They formulated it based on Gompertz function for computing reputation. Many researchers focused on privacy-preserving reputation mechanisms to protect the information leakage that approximates user identity

through their reputation [47, 61, 93]. Amintoosi et al. [5] calculated the reputation of a participant based on two factors 1) Rating given by the data requester and 2) Trustworthiness of submitted data. Moreover, the reputation update is managed with the help of the page rank algorithm. Another research work formulated the participants' reputation based on the quality of contributed data and their willingness [100]. Moreover, Dellarocas et al. [22] stated that evaluation of reputation only on the basis of the rating/feedback given by the data requester might lead to misjudgement. Authors considered three levels for Willingness of participant such as willing, neutral, unwilling with willingness been issued with a higher value. Maryam et al. [81] proposed a collaborative reputation score that is computed with the ratio of positive readings and with the votes allotted by the neighbours performing the same task. Besides the huge prominence given to the evaluation of reputation, it is a fact that reputation scores may not show the impact in the case of the naive participant as historical information is limited. That being said, reputation score eventually shows critical impact over a longer period. On the other hand, illegitimate users/malicious users may intermittently attack the crowd-sensing system by playing with reputation scores.

2.2.3 Incentives

Incentives are the rewards received by the participants for reporting the data. Each user intends to amplify their rewards by contributing to the cloud platform/data requester. Impact of incentives in MCS can be explained as follows:

- First and foremost, the purpose of allotting incentives is to compensate for the sensing cost. It is unequivocal that sensing and reporting the data utilizes the resources of mobile devices such as the battery, computational power alongside users time and energy that collectively include in sensing cost.
- In accordance with the above mentioned impact, it is evident that profits encourage users to take part in crowd-sensing. Profits are directly related to the rewards obtained. To collect a large amount of data or to cover the wide area, it is essential to motivate users to get involved in the data acquisition process.
- Furthermore, it is worth mentioning that it is the only source available to control the users' trustworthiness, efforts to achieve the required information.

There may be different modes of incentives,such as monetary funds, entertainment, services [69], virtual rewards. In addition to that, incentives can be of two types that are built on the basis of the aforementioned key aspects.

- **Platform-centric Incentives**

  The title itself suggests that incentive mechanism intend to favor the platform, simultaneously satisfying users requirements for their contribution. Alternately explained as, users rewards entirely depend on their performance, i.e. effort, QoI submitted to cloud. Users compete with each other to gain high rewards while cloud ensures minimum payment that equals to sensing cost [112].

  Authors of [68] constructed a Bayesian game between platform and users' to maximize the profit gained by the platform. They formulated an all-pay incentive scheme that rewards all the participants, unlike traditional auction. Comparatively the uppermost participant in tuple gets a prize. Thus authors were successful in enabling users' to improve their contribution. Another interesting work [115] that proposed an incentive mechanism for a practical scenario, where the budget for crowd-sensing is generated from users contributions. Based on QoI of submitted data, user reward is calculated with the help of Shapley value. Author et al [117] proposed a platform-centric incentive mechanism by formulating bargaining game between platform and users while simultaneously ensuring the social welfare (i.e. satisfaction of participants).

  Another game theoretic incentive mechanism has used the Stackelberg game where the platform is the leader and participants are followers. The unique equilibrium point ensures the maximum profit for the platform [112]. Duan et al [25] proposed incentive mechanisms for data acquisition based on Stackelberg game. They considered users are identical in data reporting but heterogeneous in capabilities of computing power. Nash equilibrium shows lowest possible incentive for the user that benefits the data requester.

- **User-centric Incentives**

  User-centric incentive mechanism aims to benefit users' by means of directing reverse auction among them. This allows users' to bid their required reward and fulfill platforms' prerequisite of possible low-cost data acquisition. Authors of [48] considered the drop-out of users from sensing campaign due to unsatisfied bids set by the platform. As a solution to this, a dynamic bid mechanism was proposed where, participants have the liberty to declare the desired reward amount for their contribution.

  Authors of [52] were eager to investigate the scenario in which users' play with their efforts rather than demanding the bids they require to participate for data requester.

To tackle this situation, researchers formulated a Bayesian game between users' and data requester. With the help truth discovery algorithm, each users' QoI is calculated that determines the amount of reward it posses. Indeed, the Bayesian game ensures that at the equilibrium state, all the users' strategy to get selected should be performing the task with high effort.

Another similar research [66] considered a use case scenario of cognitive radio to utilize the spectrum effectively. The secondary user aims to utilize the un-used spectrum when the primary user is at rest. To sense the availability of spectrum, secondary user appoints the sensing helpers and incentivize them for their contribution with the assumption of zero malicious behaviours. They formulated a coalitional game in which participants strategy is the sensing time. A social incentive model that facilitates sensing helpers with community membership and also improves reputation.

Authors in [51] introduced a reverse combinatorial auction among users. Indeed, proposed an effective incentive mechanism that depends not only on user bids but also considers the QoI of previously reported data. Based on the historical QoI values, platform shortlists the users who can provide the required/sufficient qualitative data for a lower price.

Crowd-sensing systems may have budget constraints which are also an essential metric to consider. Under certain conditions where the availability of users around task location are less, either quality or budget need to be compromised [119]. However, an efficient incentive mechanism can be a viable tool for high-quality data acquisition.

2.2.4 Task Allocation

Task allocation is one of the crucial steps to obtain the desired objective of data requester. The process of choosing the appropriate task for the user is known as an optimized task allocation. The main objective of optimal task allocation is to allot a task to a user considering various factors that distract the accomplishment of a task. Task allocation also used to achieve data requesters' requirements,such as minimal cost data, high-quality acquisition. Reasons for the need of optimal task allocation can be described as follows.

- One of the main factors that invoke the need for optimal task allocation is the location of the user or mobility of users.
- Furthermore, participants' may not be willing to contribute to some tasks due to some privacy leakage.

- User may not be capable of fulfilling certain tasks while flawless in executing other tasks assigned.
- There might be a case where users may feel overwhelmed with the assigned task which,leads to unfinished tasks or lack of proper quality.

Authors of [96] stated that considering the privacy concerns of non-selected participants who are requested to share their location along with other selected participants need to be emphasized. They dealt with the task allocation problem based on the location-based tasks can be either in the standard route of the participant or away from it. The proposed approach uses location obfuscation to ensure the privacy of the user while the objective is to reduce the travel distance for sensing a task. Another research work [36] assigns the time-sensitive tasks for users to travel in standard route, whereas, delay tolerant tasks are assigned to the users who are willing to travel outside their standard route. While incentives assigned to contributors were based on the distance, they travelled. Unlike the work in [96], this framework needs the historical data of travel paths of participants.

Xiong et al. [109] focus on low-cost task allocation while ensuring acceptable coverage area under given budget constraints. The proposed iCrowd framework also ensures minimal energy consumption with the help of piggyback crowd-sensing. They claim it as near optimal framework since it outperforms baseline approaches using such as the greedy and partial enumeration procedures to achieve required spatio-temporal coverage within budget.

Another study proposed a task assignment algorithm simultaneously considering both participatory and opportunistic users for a budget constrained the system. The proposed Hytasker model achieved higher task completions with the awareness of user mobility and density of users assuming the equal quality of data [94]. Authors in [97] proposed an online task allocation method for air pollution monitoring, temperature monitoring for a specific area and divided it into smaller cells. They attempted to minimize the number of cells required to be sensed and applied bayesian interface based methods to estimate the data from the remaining cells while satisfying quality requirements.

Shibo et al. [41] handled task allocation problem by consedreing user spatial movements and also formulated a bargaining game between users and platform by varying price of tasks. They assumed users are individuals, do not form coalitions and so each user has separate agreement with the platform. Local ratio based algorithm (LRBA) which is an orienteering algorithm performed better than the popular greedy alogrithm in terms efficient incentives.

One of the primary details required for data acquisition through non dedicated sensor nodes is the location. Studies show that 85% of users in the top three most visited locations could be successfully predicted with the help of call data [116]. Jianbing et al. [72] proposed a privacy-preserving task allocation strategy that considers geographical data and reliability of participants by means of blind signatures. While authors in [125] utilized advancements in mobile edge computing and proposed a decentralized task allocation scheme through which latency due to simultaneous communications with a large number of users is minimized. Moreover, privacy leakage of participants, as well as the context of data requester through third-party entities, is addressed alongside recruiting reliable participant with the help of historical data.

Paolo et al. [8] took advantage of human-driven edge computing (HEC) over conventional mobile edge computing (MEC) in order to broadening the coverage area. To do so, they selected some of the users as edge nodes based on their history of behaviour. Results demonstrate that selecting a central user has benefited compared to co-operative behviour based selection. Dimitri et al. [9] selected users as mobile edges based on the number of connections to avoid the question of where to install edge nodes through traditional MEC. Results has shown that combination of both fixed and mobile edges satisfies both latency and satisfied requests.

2.2.5 Privacy and Security

Since MCS deals with sensory data, it is instinctive to expect the privacy of participants and security issues. Adversaries could attack either by submitting error data or by generating the illegitimate tasks in the environment that creates Denial of service attacks (DoS) [123]. Zhang et al. [123] have considered illegitimate tasks alongside the legal tasks that cause heavy energy consumption. Through machine learning they successfully reduced energy consumption by differentiating illegitimate tasks.

Authors of [105] proposed a secured transmission of data from participants to cloud without the use of pre defined keys by slicing the data into pieces. They also considered to protect the data security and participant information through blind identities. Haiqin et al. [107] considered privacy for both task allocator and participants by making participants unaware of details about task allocator simultaneously by making selected tasks of participants to task allocator. This objective was acheived with the help of encryption models and fog nodes that do not have information about either task allocator, participants.

Authors in [118] proposed a Privacy preserving QoI-aware Participant co-ordination in MCS (PPCM) through which users co-ordinate among themselves to achieve required QoI

of task by completing the pieces of task according to their abilities. Cloud server ensures to select users based on their previous data submissions. On the other hand the primary idea behind the privacy preservation method is to distribute the task among more number of users. Accordingly, though attackers find location tagged information it would be difficult because of large number of users which are to be matched at equal probability.

Jia Hu et al. [43] considered internal attackers with two different behviours where some attack with faking a part of information whereas other attackers cause user security threats by gaining control over users personal information. They proposed a trustworthy framework that make use of mutual trust between person to person, person to group, group to group in order to get rid of vulnerable attackers. Results has shown that the proposed framework has outperformed TSCM [57], PPCM [118] in defending users against internal security attacks.

Owoh et al. [74] mentioned the location information access by most of the smart applications in google store. To secure the sensitive information of participants they used AES-256 GCM encryption methods and results has shown the high execution time with low memory usage compared to benchmark scheme. Layla et al. [77] surveyed the types of privacy attacks as well as the proposed counter measures for them during task management. Survey also discusses about the limitation during participant privacy preservation such as efficiency, trust, contribution based rewards etc.

Many works considered user behavioural attributes during authentication process for end user to identify authorized user. Authors in [32] used patterns of each user during the keyboard usage to identify the person using SVM. Authentication methods are not only confined to smart devices but also can be effectively used in smart homes, smart health, traffic and many more [127]. Fig 2.2 summarizes the related work discussed in this chapter. Unlike the aforementioned research works, we consider to provide user to authorize sensor access modification which can be seen from privacy point of view to protect the private information leakage through sensing a task.

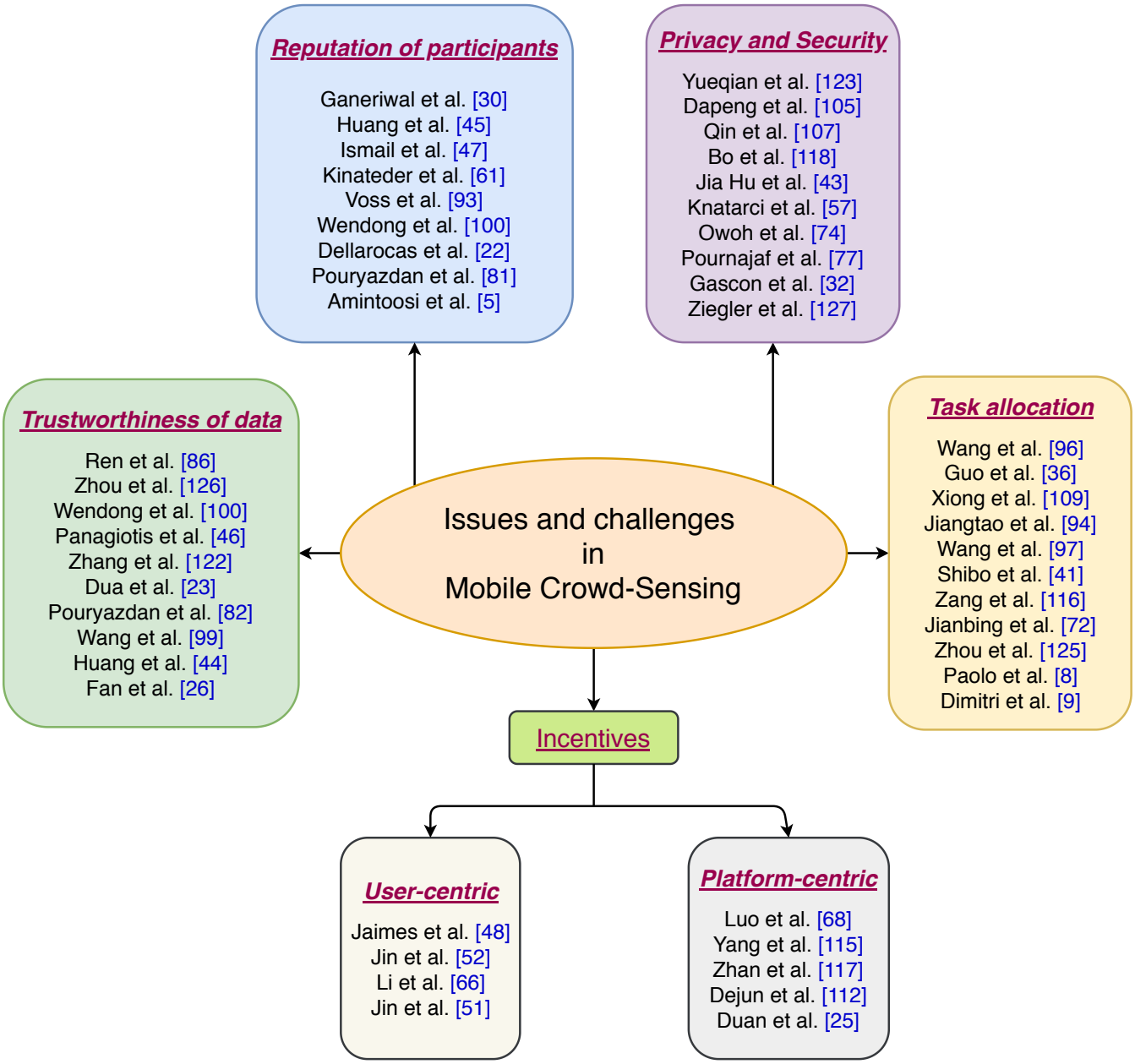


Figure 2.2: Classification of research works that addressed the challenges faced in MCS

2.3 Investigation of Related Work for User Selectiveness and Community Formations in MCS

Pouryazdan et al. [81] formulated a collaborative trust score which is determined by the votes of neighbours as well as instantaneous reliability score. They consider that participants group into communities to contribute for a task. Based on votes received by participants system was able to detect the malicious contributors among the others. While in their previous study [55] only vote based community sensing was considered alongside focusing on Sybil detection techniques to mitigate the effect of malicious users.

An application called MobiGroup [37] was proposed to suggest the private group based on the context awareness. They associated cross-community sensing [38] and mobile sound sensing and were able to determine the appropriate running groups. Moreover, they defined three types of groups such as 1) context based groups that depend on location, user preferences, 2) activity based groups that form communities for specific activity and 3) third party groups that are formed based on social relationships and interactions. Their objective is to support intelligent advertising for public activities and also suggest related tasks based on physical and social contexts.

Nicholas [63] states the importance of finding the reason, motive/criteria for community formation which can be based on social relationships, inter-dependencies among participants. Determination of criteria behind community formation and attributes of users such as skills, location, knowledge etc makes simple to cloud server to gain QoI at low cost.

Authors in [17] proposed a task selection algorithm that allows users to choose time sensitive tasks based on their mobility pattern. They formulated a non-cooperative game among the users in choosing their mobility path as well as task selection. Results had shown that the proposed task selection algorithm performed better in terms of Jian's fairness index than greedy task allocation algorithms. They also proved that it is NP hard to maximize the social surplus.

Another relevant work [59] provided users with an option to select one out of given two tasks based on distance and payments for task completion. Jin et al. [50] formulated user task selection behaviour where number of task requesters approach user for data collection. The probability of user opting a task was calculated based on their previous choices and also other factors such as age, gender. They also considered that users attract to comparatively high payment requesters and proposed Markov correlated equilibrium as solution for requesters dynamic pricing. Yan et al. [110] considered various user task selection criteria based on destination, to maximize the profit. They proposed a privacy preserving task selection algorithm to avoid any point of interest leakages.

Maryam et al. [80] considered coalition formations in MCS and showed that participants forming communities could increase their payoff by mitigating malicious payments. Moreover, community formations led to increase in participation as well as sensed tasks in the crowd sensing environment. Wang et al. [103] considered the problem of unbalanced task completion on basis of travel distance, since users opt tasks that are nearby. That being said user tend to select the nearest task and avoid contributing to the farther tasks which may end up with no data collector. To address this problem, they group the less popular tasks with high attract tasks forming task bundles. Through the application, layer user is modelled to choose a preferable set of tasks from the mixed task bundle. Each task requires multiple readings to get accomplished, thus eliminates the effect of malicious users in the environment.

Most of the aforementioned research works select tasks based on the location, distance between task and user and some works considered community formations to determine rewards based on votes. In this thesis, we consider user selection to join a community based on its income, reputation. With this we show the users intention to increase the income and also to increase the chance of getting selected. On top of this we implement this in malicious environment with the aim to maintain better user, platform utilities at the same time ensuring the data trustworthiness (see chap 3).

2.4 Investigation of Research Works That Consider User Comfort/Preferences

Considering the requirement of a large number of data collectors in mobile crowd-sensing, many research works have found solutions for various elements that bother user to perform data sensing. Some of them addressed the energy consumption problem faced by users. Authors of [98] solved the problem of energy consumption concerns while data uploading with efficient algorithms. Anjomshoa et al. [7] considered the low residual energy as the reason for users not being part of crowd-sensing in the long run. They formulated a strategy that assigns tasks to the users based on the residual energy. Another major part of research works considered privacy-related problems such as [67], [65].

Apart from the works that consider the specific problem of users the following works concentrate on user preferences. One of them is [67] where users were given an option to exhibit their preferences through tuples that consist details of their path, magnitude of the private place along the path, the maximum number of tasks that a user like to accept and privacy threshold. Based on the preference tuples given by the participant, recruiter

assigns the tasks that are to be accomplished.

Authors of [3] mentioned that users frequently report preference that they do not follow in reality. In addition to that, another work [62] has stated that user are likely to submit their preference based on the risk of information leakage vs benefit trade-off as well as trust in the entity to which they report. Moreover, data quality, number of participants will be effected due to many factors such as mobility, user preferences, energy levels [31].

Considering the fact of a wide variety of tasks and various factors that influence participants in crowd-sensing, authors et al. [95] proposed a framework to handle the task allocation problem. The factors include the maximum task limit of a user, participant willingness, available sensor configuration that can be modified by the user. In addition to the factors favoring user, they introduced an attribute for determining the percentage of accomplished tasks by each user. One of the major factors that contradict each other in implementation are privacy preservation and trustworthiness of participants. To tackle this issue, authors divided participants based on their level of privacy concerns and then implemented two layer neural network to identify the user ID [73].

Galinina et al. [29] considered the concept to provide energy as incentive for the collection of data. They also assumed users may turn of some of the available sensors according to their energy levels. That being said, cloud server can access the sensors of wearable devices only when user grants permission. Power transmission is done through the emitted radio waves from stationary wireless power beacons placed in the crowd-sensing environment. This idea could transform the crowd-sensing to a new level without the concern about the energy constraints that each wearables possess.

Zhang et.al [120] considered user preferences which can be either location and travel distance, type of tasks while allotting the tasks to users. They do not provide any choice to choose the task but rather consider their preference before allocation. Moreover, they considered users to bid untruthfully to increase their utilities. Results show that, users who submit their true departure times have acheived optimal utilities.

Contrary to the discussed related work, we consider user preferences can be influenced by various factors such as location privacy, energy constraints, past experience etc. We formulate user preference aware participant selection through authorizing users to switch off the some built-in sensors through which they can gain comfort. We implement this in a vulnerable environment where malicious users attack with maximum possibilities to damage platform utility.

Chapter 3

Selective Data Acquisition in Mobile Crowd-Sensing

3.1 Introduction

The advent of mobile technology has facilitated the data acquisition process with the help of embedded sensors of smart devices (mobiles, tablets, smartwatches, etc) [6]. Selection of appropriate participant is a significant challenge since it depends on many aspects such as reliability of user, location, cost of sensing [49]. Incentives provided to the users make a direct impact on the profit obtained by the platform (platform utility), user profit (user utility) and indirect influence on motivating users to participate, obtaining the quality of data and compensating the cost of sensing. Considering the above fact, most of MCS systems' rely on incentive dependent frameworks. Incentives can be through monetary rewards, entertainment, complementary services, virtual reward points.

In order to build a robust MCS system, it is important to consider various factors along with the trustworthiness of data such as users willingness, user utility. Thus consequently, encourages more participants to involve in the data gathering process. Contrarily, most of crowd sensing platform doesn't encourage users to exhibit their choice of interest to sense a task and push users' to collect data for a required task. However, users' try to guide themselves to maximize their payoff. From this perspective,users' may not be interested in contributing data for the task that is assigned to him. Thus, it is equally important to emphasize the assignment of task based on users' interest. Moreover, ensuring the trustworthiness of data in the presence of malicious users play a key role in data acquisition through non-dedicated sensors [40].

We introduce a user recruitment scheme with the primary objective to give prominence to users' selectiveness in choosing the task to contribute and improve user utility. In addition to this, it ensures the recruitment of trustworthy participants with the help of reputation score. Reputation score provides information about the degree of reliability of a participant up till then.

3.2 System Overview

The main components of the considered crowd sensing system are 1) End-user/data requester 2) Cloud platform 3) Participants/users. Data acquisition through non-dedicated sensors includes user recruitment, data sensing and processing, data reporting followed by data analysis and reward presentation [28], However, the fundamental consideration of the system model (both selective, non-selective) is to follow a non-aggressive payment process through which users' are guaranteed to pay the bid price, irrespective of the accuracy of submitted data [57]. This indeed alters the traditional procedure of awarding rewards after the analysis of data. Moreover, we do not stress on data processing and reporting since these phases relate more to data quality. In this thesis, we consider untrustworthy data is generated due to untruthful/malicious participants but not due to Sybil attacks, data manipulation issues. Malicious users are who voluntarily send false data to misinform the system. Accordingly, updates the database and publish the tasks, then the user recruitment process is initiated.

Table 3.1: Notation used in this chapter

Notation	Description
κ_i^T	Binary decision of user/sensor i to accept or decline task T
ϱ	Closeness margin used in the decision of accepting/declining a sensing task
T	Index of a task
U	Total crowd-sensing participants available in the complete terrain
$\|W\|$	Number of participants in winner set W
δ	Weight factor
$r_{max}^{\Gamma_T}$	Maximum total rewards received by the users/sensors in the set T
Γ_T	List of users/sensors that can participate in the sensing of task T
Γ'_T	List of users/sensors out of the set Γ_T that have opted-in to participate in the sensing of task T
$r_{\Gamma_T}^{avg}$	Average total rewards received by the users/sensors in the set T
T_U	Set of tasks that can be sensed by the users/sensors in the set U in a participatory manner.
T'_W	Set of tasks out of the set T_U that have been opted-in to be sensed by at least one user of the set W
$v_i^R(W)$	Reputation-based marginal value of user/sensor i over the set of participants W
$v^R(W)$	Reputation-based value of the set of participants W
v_T	Value of task T
R_i	Present Reputation of user/sensor i
$R_i(t-1)$	Previous reputation of user/sensor i
R_{total}	Total reputation of users/sensors that can participate for task T
c_i	Sensing cost (i.e. auction bid) of user/sensor i
p_i	Positive readings of user/sensor i
n_i	Negative (outlier) readings of user/sensor i
r_i	Total rewards issued to participant i
r_i^t	Total rewards issued to participant i in sensing campaign t
$v^R(W^t)$	Total reputable value obtained from winner set W in the sensing campaign t
$x_i^{T^t}$	Binary value that shows 1 when user i bid for the task T in campaign t.
$\|W^t\|$	Total number of selected participants in sensing campaign t
t_{total}	Total number of sensing campaigns

3.2.1 Centralized Data Acquisition

User Recruitment Phase

For the reputation-based reverse auction, we adopt Trustworthy Sensing for Crowd Management(TSCM) presented in [57] which is built on top of reputation unaware incentive mechanism referred to MSening [112]. In this thesis, we term TSCM as Non-selective Reputation-based Recruitment (NSR), also be stated as a centralized data acquisition model where users do not have the choice to select a community to perform data sensing. According to NSR, users start bidding to the tasks that are in the vicinity of 30m. After users bid their price, cloud server evaluates the total reputable value that can be described as the total expected benefit that could be obtained by recruiting the participant. It is worthwhile mentioning that users also share their location co-ordinates along with the bids. Platform checks, whether the users are bidding from the task location,are not. Then reverse auction is performed on the basis of bids and the users' respective reputation.

The fundamental consideration in TSCM is that every task has its value that indicates its importance, represented by v_T. Considering U as a set of total users in the sensing environment, $v(U)$ denotes the maximum value that can be obtained from the tasks surrounding the set U (see Eq 3.1).

$$v(U) = \sum_{\mathrm{T} \in T_U}^{M} v_{\mathrm{T}} \tag{3.1}$$

Γ_T represents the set of participants that are in the vicinity of task T and could contribute to data acquisition on winning the reverse auction. Furthermore, Γ'_T denotes the set of participants that bid for task T. Since the centralized data acquisition is a Non-selective process, we assume each and every user around the task to participate in the reverse auction. It is clear that under NSR satisfies the condition of $\Gamma_T = \Gamma'_T$.

$$\Gamma'_{\mathrm{T}} = \bigcup_{i=0}^{|\Gamma_{\mathrm{T}}|} \{i | i \in \Gamma_{\mathrm{T}} \wedge \kappa_i^T = 1\} \tag{3.2}$$

Besides users choice, there is a fair chance of tasks that aren't surrounded by any user. Eq. 3.3 T'_U denotes the set of tasks that are opted by user set U.

$$T'_U = \bigcup_{T} \{T | T \in T_U \wedge \exists i \subset \Gamma'_{\mathrm{T}}\} \tag{3.3}$$

The significant component of the recruitment process relies on this reputation factor which is evaluated as shown in Eq 3.4. Here $R_i(t)$ represents the cumulative evaluation of participant $i's$ reputation based on current positive and negative readings as well as previous reputation $R(t-1)_i$. Besides, $epsilon_1$ and $epsilon_2$ are non zero positive real numbers, were used to avoid indeterminate form 0/0. The reputation of each participant is calculated after the completion of one campaign and/or after determining the false readings of each user.

$$R_i \leftarrow \delta R_i(t-1) + (1-\delta)\frac{p_i + \epsilon_1}{p_i + n_i + \epsilon_2} \tag{3.4}$$

Now the system has the previous aggregated reputation of participants (since this step is before sensing the present tasks), location information as well as their bids to respective tasks. In most of the cases, it is expected that more than one user participates in the auction, and each of them has a good opportunity to cover more than one task simultaneously. Thus insisting on the comparison of profit that can be obtained by selecting a user with each other, as shown in Eq: 3.5. It is the difference of reputable value ($v^R(W \cup \{i\})$ obtained with the addition of user i to the winner set W and the reputable value of set W itself ($v^R(W)$), that gives users individual significance.

$$v_i^R(W) = v^R(W \cup \{i\}) - v^R(W) \tag{3.5}$$

Where $v^R(W)$ can be calculated by the average for the product of value obtained and reputation of set W. By this, we consider the value of data achieved from a participant is always directly proportional to its reputation (see Eq 3.6).

$$v^R(W) = \sum_{\mathrm{T} \in T'_W} \sum_{i \in \Gamma'_\mathrm{T}} \left(v_T . R_i / |\Gamma'_\mathrm{T}| \right) \tag{3.6}$$

The aforementioned assessment shortlists beneficial users based on the only reputation. Now, auction bids influence the winner selection process, as shown in Eq 3.7. Bid/cost of sensing of user i is divided by its reputation in order to eliminate low reputation users when it is subtracted from evaluated $v_i^R(W)$ in Eq: 3.5. This difference values are sorted in descending order that provides the opportunity to the highest contributor for a comparatively low price.

$$v_i^R(W) - b_i/R_i > ... > v_m^R(W) - b_m/R_m \tag{3.7}$$

3.2.2 Decentralized Data Acquisition

User Recruitment Phase

In the centralized data acquisition data analysis, incentivization, winner selection are handled at the central cloud platform. Authors of [106] state that the storage and processing capabilities of cloud computing can be improved with the help of edge devices in the network. Thus, Mobile edge computing(MEC) can be beneficial to distribute the workload among other edge nodes besides reducing network latency. Moreover, Network function virtualization technology helps edge devices to utilize virtual machines for improving computational and storage performance. We take advantage of edge computing and present the hybrid architecture consists of multiple MEC nodes and a cloud platform. Every edge node performs the following activities with the help of Data Service Function (DSF).

- Sensor selection
- Winner selection
- Payment determination.

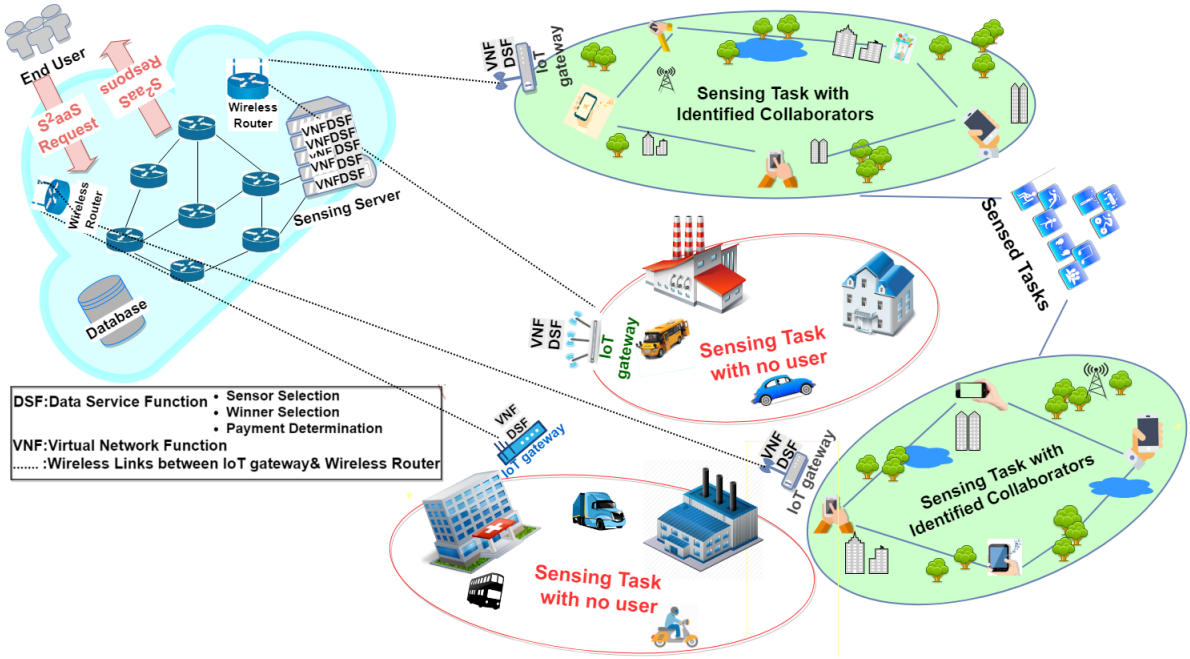


Figure 3.1: Decentralized System Model

Fig. 3.1 represents the minimalist overview of decentralized MCS model. Along with decentralized data acquisition, we introduce user selectiveness through which participants' choose a group/community to join with based on certain selection criteria. Community corresponds to the group of participants that are interested/opted to contribute to the same task T. We introduce two types of criteria,

1) *Income-based selection:* As we present in [91] Users compare their total income with the average income of the community. We further divide it by participants willing to join a comparatively lower average income community as Selective and reputation-aware data acquisition with low-income service providers (SRLI), whereas with Selective and reputation-aware data acquisition with high-income service providers (SRHI) participants chooses to join in comparatively higher income communities as formulated in Eq 3.8. L.H.S of the equation refers to the ratio of the difference between total rewards of user i and maximum total rewards among the participants opted to the same task $T(\Gamma_T)$ to the maximum total rewards. Besides ϱ denotes the closeness margin, more the ϱ higher the selectiveness.

$$\kappa_i^T = \begin{cases} 1, & \varrho[1 - \frac{r_i}{r_{max}^{\Gamma_T}}] \leq [1 - \frac{r_{\Gamma_T}^{avg}}{r_{max}^{\Gamma_T}}] \wedge SRLI\ mode \\ 1, & \varrho[1 - \frac{r_i}{r_{max}^{\Gamma_T}}] > [1 - \frac{r_{\Gamma_T}^{avg}}{r_{max}^{\Gamma_T}}] \wedge SRHI\ mode \\ 0, & Otherwise \end{cases} \tag{3.8}$$

While R.H.S denotes the ratio of the difference between maximum total rewards obtained by set Γ_T and average rewards obtained by set Γ_T to maximum rewards obtained by set Γ_T. Participants, when opted to selectiveness, is shown with the indication of $\kappa_i^T = 1$. Users join the community by activating the respective sensor.

2) *Reputation-based selection:* Similar to income based-selection as the name suggest users compare their latest reputation score with the communities average reputation score. This process is termed as Selective and Reputation-aware Recruitment (SRR) as presented in [21].

$$\kappa_i^T = \begin{cases} 1, & \varrho[R_i] \leq [\frac{R_{total}}{\Gamma_T}] \wedge SRR\ mode \\ 0, & Otherwise \end{cases} \tag{3.9}$$

Where R_{total} denotes the total reputation of users surrounds to task T. We denote consider the scenario participants joining in higher reputation groups since it is logically true that the participants with comparatively higher reputation score get selected by the platform.

Threat Model

We implement attack model of malicious users where malicious users attack with a strategy to sustain with the reputation-aware user recruitment process. The main objective of malicious users is to effect the platform utility by sending wrong data. To do so, malicious user need to get selected in the reverse auction for which reputation plays a crucial role. Malicious users attack the platform by considering upper (R_{UTA}) and lower thresholds (R_{LNTA}) comparing their previous reputation with the present reputation as illustrated in Fig 3.2.

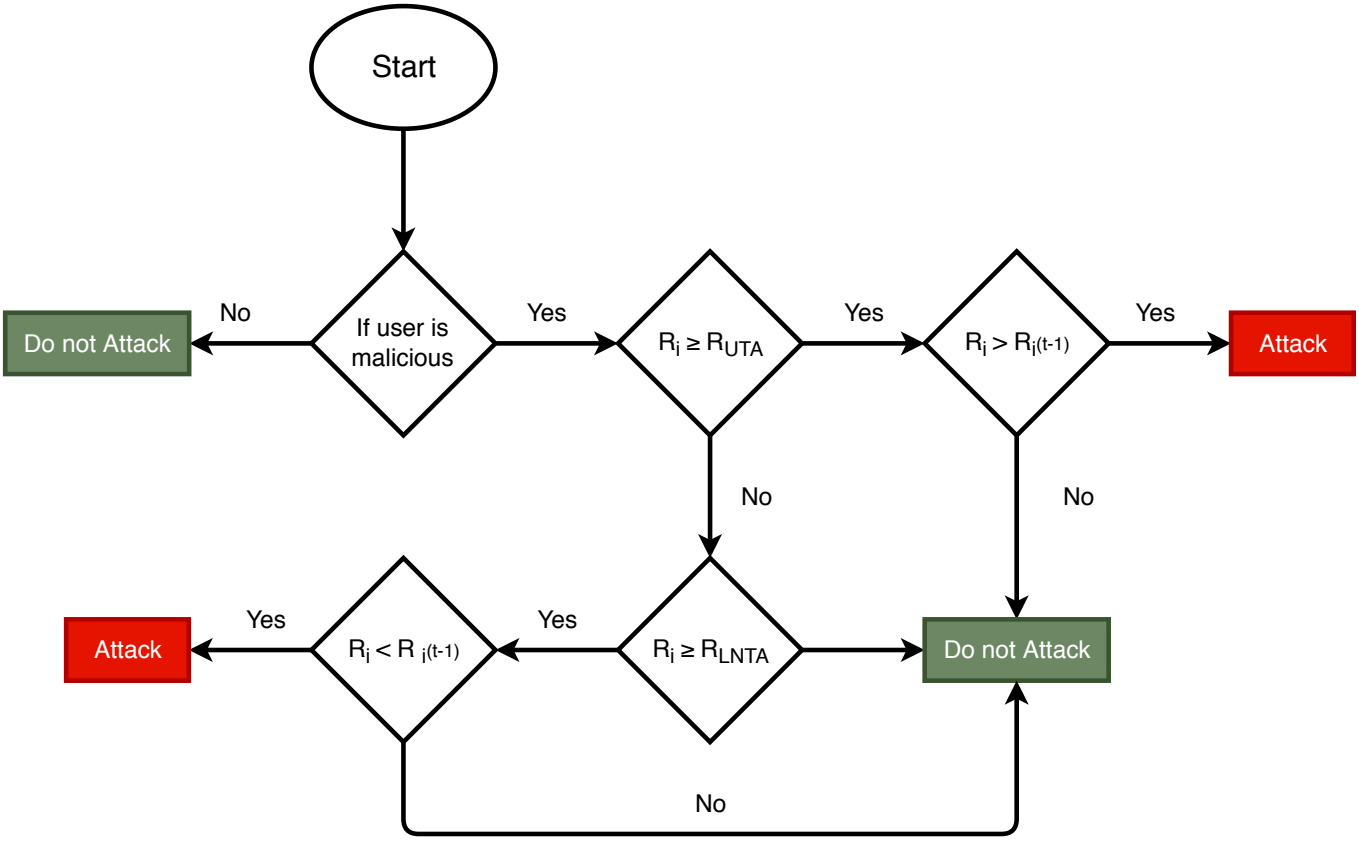


Figure 3.2: Threat model

Mobility of User

We formulate the mobility of user with Random waypoint mobility model adopted in [56] which was built based on the mobility model presented in [11]. The fundamental difference between random mobility and the random waypoint mobility is that participants pause for a while and continue moving to the next location in a random waypoint model [11]. In this thesis, we consider a terrain of 1000m*1000m that further divided into smaller grids. At the start participants are randomly placed in the terrain, then each participant selects their arbitrary destination and moves towards it with average velocity $\vec{V}$. On reaching the destination user pause for a certain period calculated by the probability of $P(pause)$ presented in Eq 3.10 and again set the next destination and reach it. It is essential to mention that we consider platform is unaware of the mobility of the user in the sensing environment (i.e. terrain). The aforementioned assumption could be more realistic with no history of past mobility as well as unpredictable movement or unexpected movements made by the participants while involved in crowd-sensing.

$$P(pause) = \frac{totalwalktime}{totalwalktime + totalpausetime} \tag{3.10}$$

Rewarding Phase

The reward/ incentive mechanism is adopted from [57] that is built upon principles of Msensing [112]. The cardinal rule of the presented incentive mechanism is to reward no less than the bid price for all the selected participants. Since the considered system is user-centric, each user is ensured to pay the maximal price for their contribution, as shown in equation 3.11. That being said, platform checks for the highest bid that a selected user can quote that he/she remain to be a winner out of other participants.

The first step is to prepare a set s' such that user w is not part of it. The system constructs a temporary winner set Δ and computes the maximum bid that an excluded user could propose that makes him get selected among other participants (see Eq. 3.11). Where ρ_w indicates payments to user w and w_n indicates the new user added to set Δ.

$$\rho_w = max(\rho_w, min(v_w(\Delta) - (v(\Delta)_{w_n} - b_{w_n}), v_w(\Delta))) \tag{3.11}$$

Since we adopt non-aggressive payment mechanism [57] (which was built with the reference of work presented in [111]), the quantified bids (b_i/R_i) are considered only during the recruitment phase and while arranging participants according to the difference between

their contribution value to requested bid (see Eq 3.12). With that said, the only actual bid of a participant is considered while rewarding.

$$w_n = argmax_{n \in s'/\Delta}\{v_n^R(W) - b_n/R_n\} \tag{3.12}$$

3.3 Performance Evaluation Metrics

We considered a java-based simulation environment to model the intended recruitment schemes. We considered 1000m x 1000m terrain in which users and tasks are randomly placed over the terrain. To avoid the impact of randomness we generate random number using 10 different seeds and average the output. Total simulation time is 30min out of which, each campaign lasts for 1min. One complete campaign includes, completion of user recruitment, rewards to winners and data collection for the tasks published/requested in that period of time. Further details of simulation settings are represented in Table 3.2.

To evaluate our impact of system model we consider four performance metrics such as

- **Platform Utility:** It is the difference of total reputable value obtained for the crowd-sensed tasks and incentives allotted to the winners. Where t denotes the index of campaign and r_i^t denotes rewards for user i in campaign t.

$$U_{platform} = \sum_t \left(v^R(W^t) - \sum_i {\mathrm{r}_i}^t \right), \tag{3.13}$$

  Given that, $v^R(W^t)$ denotes the total reputable value obtained from the winner set W at t time slot. Cloud platform aim to increase its utility by choosing reliable participants.

- **User Utility:** *User Utility* indicates the average profit gained by a user per campaign. It denotes the average difference between total received rewards and sensing costs of all selected participants normalized by the total number of sensing campaigns (t_{total}) as formulated in Eq. 3.14.

$$U_{user} = \frac{\sum_t \left(\left(\sum_i {\mathrm{r}_i}^t - \sum_i {b_i}^t \right) / |W^t| \right)}{t_{total}}. \tag{3.14}$$

Table 3.2: Simulation settings for implementing user selectiveness

Parameter	Value
Size of terrain	$1000\,\mathrm{m} \times 1000\,\mathrm{m}$
Total number of Users	1000
Range of task	30 m
Task Arrival Rates	$20; 40; 60; 80; 100/\mathrm{min}$
Initial reputation of a participant	0.7
Closeness margin (ϱ)	$\{0.6, 0.7, 0.8\}$
Probability of malicious users	$0.05, 0.07$
Probability of invalid reading submitted by an user	$0.02, 0.03$
Weight factor δ	0.6
Task Value	$\{1, 2, 3, 4, 5\}$
Bid value / Sensing costs	$\{1, 2, ..., 10\}$
Total number of sensing campaigns	30
Simulation Duration	30 min
Average walking speed($\vec{V}$)	$2m/s$
Mobility awareness	False
Upper reputation threshold for malicious users to turn on attacking (R_{UTA})	0.8
Lower reputation threshold for malicious users not to attack (R_{LTNA})	0.5

Where r_i^t denotes the rewards obtained and b_i^t denotes the sensing cost for user i at time t and/or campaign t.

- **Dis-information ratio:** The ratio of tasks affected by at least one malicious user to the total number of tasks. Affected tasks can be stated as the tasks for which malicious users' are recruited.
- **Rewards to malicious:** It is the summation of total payments made to the malicious users in all the sensing campaigns. It indicates the system robustness to adversaries.

  We consider disinformation ratio and rewards to malicious to quantify the impact of adversaries and also to take care of trustworthiness of data. Throughout the thesis we try the mitigate the malicious effect which directly lead to loss in platform utility and average user utility.

3.3.1 Simulation Results for Income-Based Selectiveness

Platform Utility:

In order to evaluate the selective low-income reputation-based recruitment scheme (SRLI), we compare SRLI under different selective coefficient(ϱ) with the Non-selective reputation based recruitment strategy (NSR). In addition to this, we compare Selective and no reputation aware user recruitment for the illustration of the impact of reputation awareness. Fig 3.4 describes the performance of SRLI, SNR and NSR for the utility of the platform. The equations 3.15, 3.16 are reflected in the Fig 3.4 i.e. in case of SRLI as we increase the ϱ from 0.6 to 0.8 there is a huge loss in platform utility due to high selectiveness attitude shown by the users. This may result in users being idle that incurring to a loss for crowd-sensing recruiters. While reputation un-aware selectiveness performs better than SRLI-ϱ=0.8 because of the recruitment of a comparatively better number of participants in SNR. This phenomenon explains the very high selectiveness has a severe impact than recruiting users in an arbitrary manner.

$$\varrho[1 - \frac{r_i}{r_{max}^{\Gamma_T}}] \leq [1 - \frac{r_{\Gamma_T}^{avg}}{r_{max}^{\Gamma_T}}] \wedge SRLI\ mode \tag{3.15}$$

$$\varrho[1 - \frac{r_i}{r_{max}^{\Gamma_T}}] > [1 - \frac{r_{\Gamma_T}^{avg}}{r_{max}^{\Gamma_T}}] \wedge SRHI\ mode \tag{3.16}$$

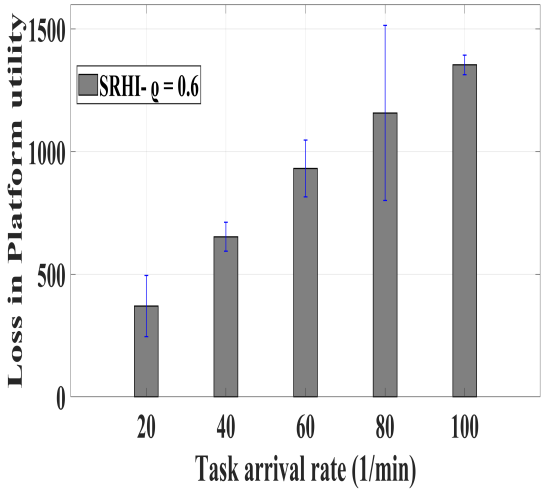


Figure 3.3: Loss in Platform utility for SRHI

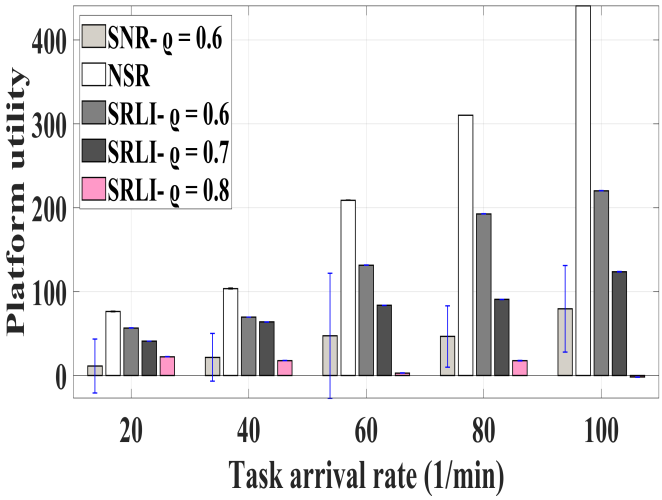


Figure 3.4: Platform utility under SRLI, SNR and NSR.

Fig 3.3 illustrates the effect of income based selectiveness approach on the utility of system (platform utility) under given simulation settings. In order to investigate the impact of SRHI, where users select to join the communities that have higher incomes, we evaluate the platform utility and user utility performance of SRHI. Due to joining a high-income community, a user may not be selected as the users in the community are supposed to be receiving high-income as a result of providing trustworthy sensing data. It eventually helps to improve their reliability with the showcase of high reputation scores to the MCS platform. Due to the selectiveness of joining in only high income groups, participants may end up not being involved in any sensing group. That indeed leads to a significantly low amount of valuable sensing tasks received by the platform. This phenomenon may result in a significant reduction / loss in platform utility. The loss increases with the increase in task arrival rate due to the increase in the unsensed high valued tasks.

Moreover, the formed communities/groups demand high income for delivering required information, causing no profit to the platform. Another major reason that contributes to loss of platform utility is that the mobility of users causes in-completion of task after the payment phase. This phenomenon can be highly perceivable when the total number of participants selected are very less when has happened in case of the SRHI recruitment scheme.

Average User Utility:

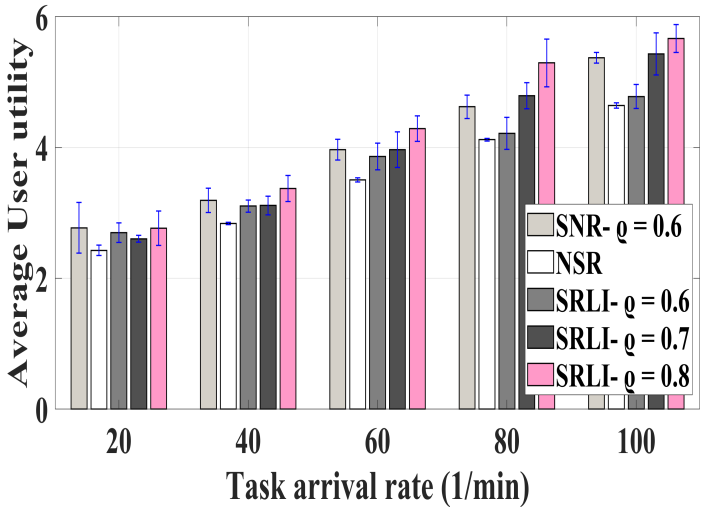


Figure 3.5: Average user utility under SRLI, SNR with NSR

In Fig 3.5 illustrates the Average user utility under all the presented recruitment schemes. Since it is average user utility of selected participants, it is clear that if users' opt their sensing community based on its average income (see Eq 3.15) then users' gain higher utility than the utility obtained in other recruitment schemes. The reason we avoided SRHI in comparison is the fact that it causes substantial loss to the platform/recruiter, which is completely unacceptable. Comparatively, high selectiveness (i.e. ϱ=0.8) among participants might improve their income. However, it also reduces the chance of sensing more number of tasks. On the other hand, NSR ends up in causing low user utility because of 1) More users being recruited by which more tasks are covered, 2) Does not allow user selectiveness to form reputation-based groups for improving average user utility.

Rewards to Malicious and Disinformation Ratio:

Rewards to Malicious represents the total payments received by malicious users'. These false payments can be reduced with reputation-based user recruitment. This is exactly the reason for the rise in false payments under reputation un-aware selection (SNR). Furthermore, it is clearly evident in Fig 3.7 that out of all reputation aware schemes NSR does better under almost all task arrival rates. With the low selective margin (such as ϱ=0.6 and 0.7 malicious users find it easier to damage the platform because of the flexibility to

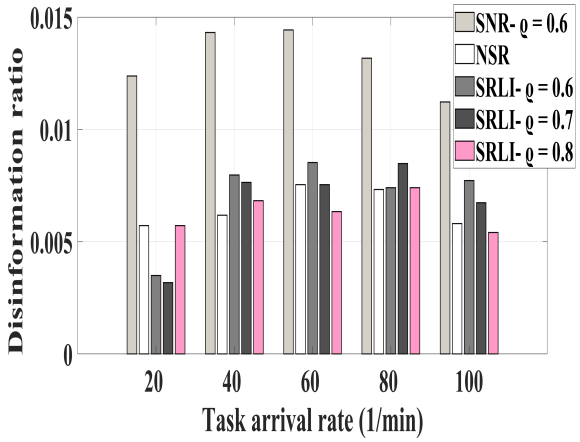


Figure 3.6: Disinformation ratio under SRLI, SNR and NSR.

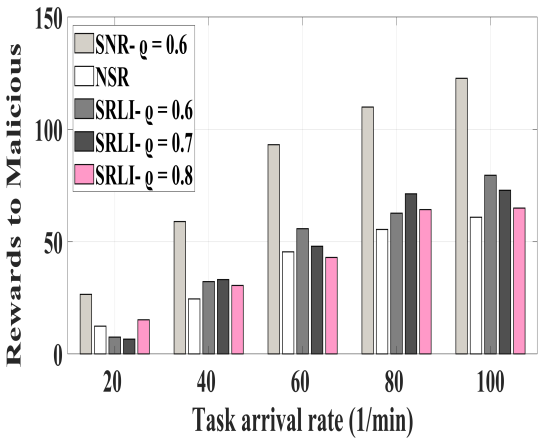


Figure 3.7: Rewards for Malicious users under SRLI, SNR and NSR.

join in various communities and get selected for the first time in crowd-sensing. There can be two ways that a malicious user can attack 1)Since all the users will have the same initial reputation, there is a fair chance of recruiting malicious user. 2) Another way is malicious users may behave trustworthy until they improve their reputation above the threshold and attack subsequently [15].

Disinformation ratio is another considerable metric to analyze the ratio of tasks attacked by malicious users. Disinformation ratio has a direct correlation with rewards for malicious as a percentage of selected malicious users increase, and false information obtained also increases. Having said that, malicious users who try to be trustworthy for a period of time doesn't affect the platform for that instance.

3.3.2 Simulation Results for Reputation-Based Selectiveness (SRR)

We consider the same performance metrics that are used for evaluation of income based reputation mechanism. In the previous simulations, we kept malicious user probability as 0.05. For SRR, we move a step forward and consider 0.07 malicious along with the 0.05 and also we are eager to analyze the performance for less selectiveness, i.e. ϱ=0.5 along with 0.6, 0.7, 0.8. Moreover, Eq 3.17 represents the criteria every user follow to join in community to contribute for a task.

$$\varrho[R_i] \leq [\frac{R_{total}}{\Gamma_T}] \wedge SRR\ mode \tag{3.17}$$

Additional Performance Metrics

- **Total Energy consumption:** It is the total sum of energy consumption values for sensors that are associated with the tasks which are bid by total selected users in entire 30min. It is worthwhile mentioning that G.P.S energy consumption is added each time user i contributes data for the platform (see Eq 3.18).

$$T.E = \sum_{t} \sum_{T} \sum_{i} \left(x_i^{T^t} . \beta_{T^t c} + G.P.S \right) \tag{3.18}$$

  We do not consider any stand-by energy drain for built-in sensors. So it is clear that in order to save residual energy of their device, users need to avoid sensing energy hunger tasks by switching off the respective sensor.

- **Average Energy Consumption:** The ratio of total energy consumption by the selected users to the number of selected as evaluate in Eq 3.19 . The energy consumption values w.r.t sensors used in this chapter are presented in 3.3.

$$A.E = \frac{\sum_{t} \sum_{T} \sum_{i} \left(x_i^{T^t} . \beta_{T^t c} + G.P.S \right)}{Totalselected(|W|^{total})} \tag{3.19}$$

Table 3.3: Energy consumption values

Sensor	Energy Consumption Value (Joule)
Global Positioning System	0.125
Accelerometer	0.05
Camera	0.1
Microphone	0.075
LightSensor	0.025

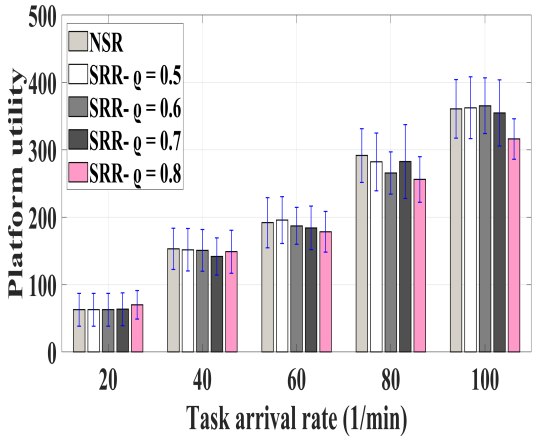


Figure 3.8: Platform utility under SRR and NSR for 5% malicious probability

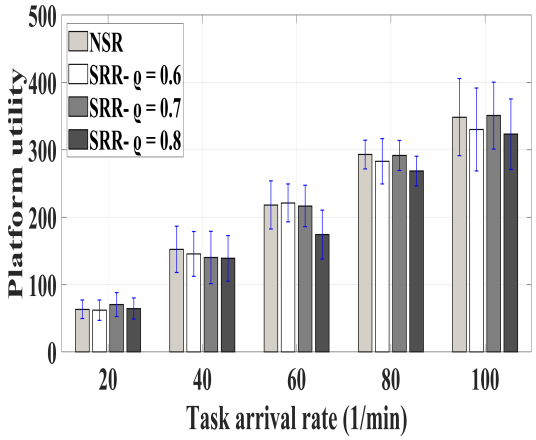


Figure 3.9: Platform utility under SRR and NSR for 7% malicious probability

Platform Utility

Fig 3.9, 3.8 represents the results of platform utility for the NSR, SRR models under 5% and 7% malicious probability. With the 5% probability of false users, there is an almost similar performance for NSR and SRR. While as at a higher number of tasks/min,there is a mixed behavior among different selective margin with SRR. This pattern of behaviour is evident that, finding an appropriate community/group of users may not always be successful. However, it is transparent that SRR performs better than NSR under certain conditions.

To investigate that we further increase the malicious probability to 7%, as average user utility has started to decline (see Fig 3.11) by lesser selectiveness(i.e. ϱ=0.5) we stick with $\varrho > 0.5$ since we target for a methodology that resembles NSR in all the performance metrics. As Fig 3.8 demonstrate, SRR with closeness margin (ϱ=0.7) gains similar utility for the platform. This strengthens the proposal of decentralized selective data acquisition at mobile edges.

Average User Utility

In Fig 3.11, 3.10 average user, utility is illustrated for selective and non-selective user selection methods. As shown in the figure, reduction in selectiveness enables users to join for more number of tasks and as a result, gaining higher utility compared to very high

selectiveness(ϱ=0.8). With the joining of high reputation community, lowers the chance of winning the auction, since system considers reputation along with the bid of the user in the winner selection phase. Meanwhile, users that join fewer reputation groups have a higher chance to find a place in winner set,which eventually increases their winning percentage in the reverse auction. This phenomenon can be clearly observed under both, malicious probabilities presented. Since it is the average utility of selected participants malicious probability does not have a direct impact on the utility of user. That being said, there is a possibility that malicious user may grab the chance of co-operative/reliable user if a reputation for both is equal to 0.7 (initially each user is assigned 0.7 as reputation. Later it completely depend on their performance).

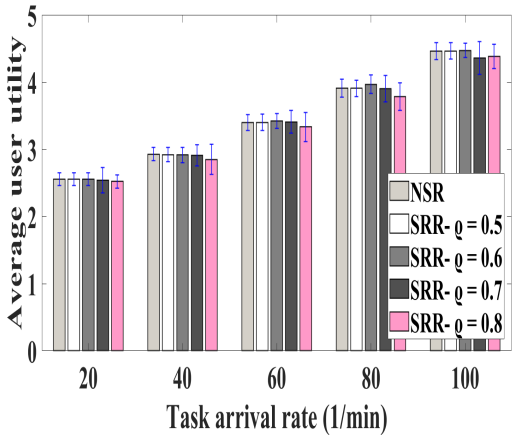


Figure 3.10: Average user utility under SRR and NSR for 5% malicious probability.

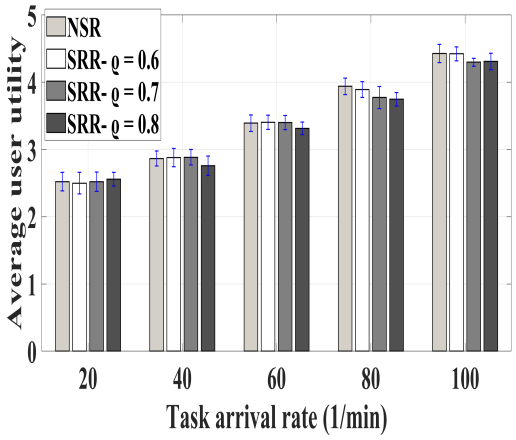


Figure 3.11: Average user utility under SRR and NSR for 7% malicious probability.

Rewards for Malicious and Disinformation Ratio

As shown in Fig 3.12 Non-selective / centralized approach(NSR) outperforms less and/or moderate selective decentralized data acquisition (when ϱ=0.5,0.6,0.7) in terms of restricting payments to malicious users. However, high selectiveness makes malicious users unsuccessful in attacking crowd-sensing system due to the fact that the system doesn't allow malicious users reputation to grow unless he/she acts reliably. This process reduces the chances of malicious users being selected and/or get paid by the recruiter.

Simultaneously with the increase in malicious probability has shown the similar outcomes that are illustrated as such in Fig 3.13. The payments for malicious users have a

major impact on the utility of platform,which can be observed in Fig 3.8. We can draw a conclusion here that NSR is the best to mitigate malicious behaviour under any malicious probabilities where as selective mobile edge data acquisition depends on various factors. One of the major drawbacks of selective data acquisition is the increase in the percentage of attacked,tasks as seen in Fig 3.15, 3.14 which is indeed higher under high malicious percentages.

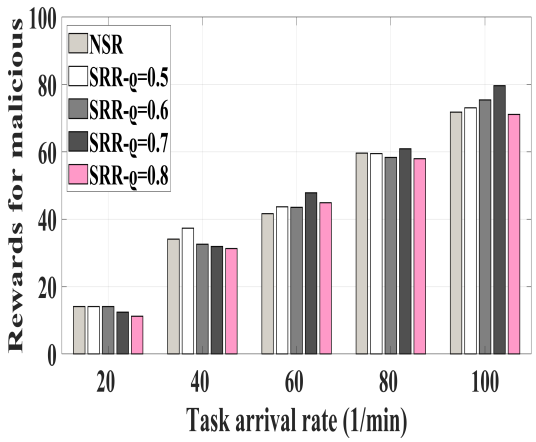


Figure 3.12: Rewards for malicious under SRR and NSR for 5% malicious probability

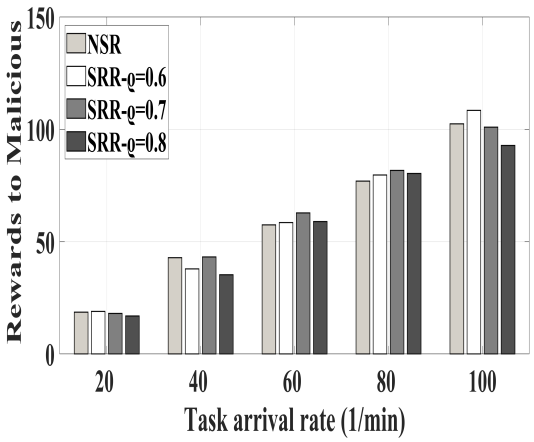


Figure 3.13: Rewards for malicious under SRR and NSR for 7% malicious probability

Total Energy Consumption

As we have seen in Fig 3.17, 3.16 total energy consumption for acquiring data is very similar to each scheme. This behaviour is exhibited since users may show affinity to contribute to specific tasks, however eventually one or the other user might show interest in sensing the remaining tasks. Moreover, the difference in energy consumption due to unsensed tasks is negligible since GPS is high energy consuming compared to other sensors.

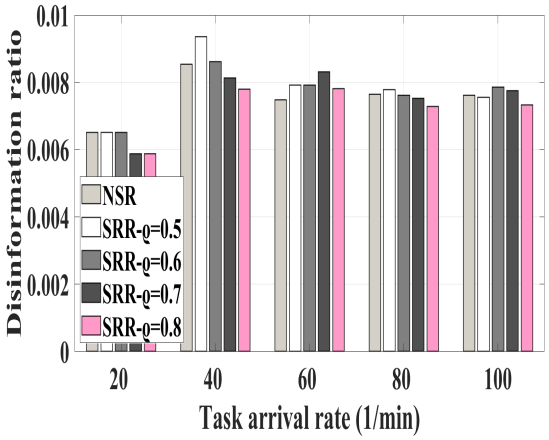


Figure 3.14: Disinformation ratio under SRR and NSR for 5% malicious probability

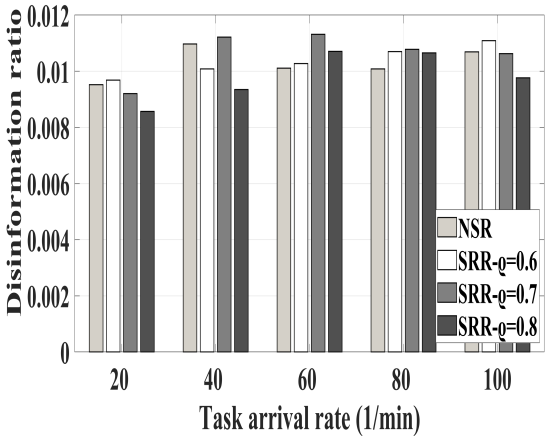


Figure 3.15: Disinformation ratio under SRR and NSR for 7% malicious probability

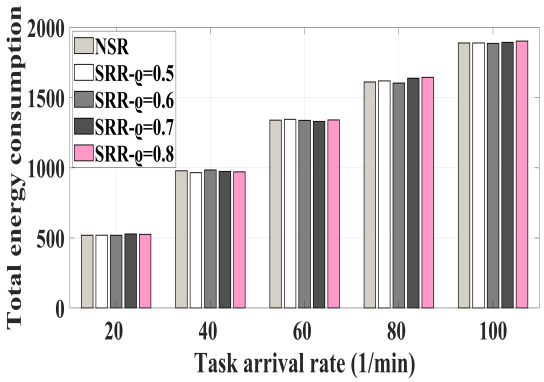


Figure 3.16: Total energy consumption under SRR and NSR for 5% malicious probability

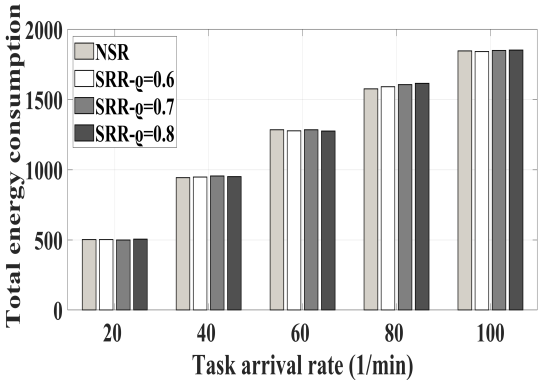


Figure 3.17: Total energy consumption under SRR and NSR for 7% malicious probability

3.4 Summary

With the advent of Network function virtualization (NFV), mobile edge computing could be a viable tool to support crowd-sensing to perform decentralized data acquisition. Data service function (DSF) can be embedded with virtual network functions (VNF) to monitor

and operate data collection and winner selection at the nearest point, thus reducing the overall latency. In this chapter, we combined decentralized data acquisition with the user selectiveness while ensuring data trustworthiness. The results have shown that income-based reputation aware selectiveness (SRLI, SRHI) diminish the platform utility by 25%-50% with the loss of 14%-28% user utility over the different task arrival rates. However, selective reputation unawareness led to higher payments for malicious users resulting in accumulation of unreliable data. On the other hand, simulation results for SRR shown promising progress over income based selectiveness in terms of all the considered metrics. Besides that, our feasibility study reveals that proper selection of closeness margin (ϱ) with SRR resembles benchmark (NSR)in platform utility with the loss of negligibly small user utility. That being said the model is successful in encouraging users with providing the opportunity to opt for the desired group to join. To this end, we conclude that SRR-ϱ=0.7 perform consistent with 5% malicious probability, whereas with 7% it inherited Non-selective data solicitation. In this chapter, users are open to any type of data sensing. In the next chapter we consider user unwillingness to sense the task by totally turning-off the sensor according to their convenience. Having said that platform has to select a participant according to their availability of sensors which shows its impact in various performance metrics that are presented in this chapter.

Chapter 4

Comfort-Aware Participant Recruitment for Mobile Crowd-Sensing

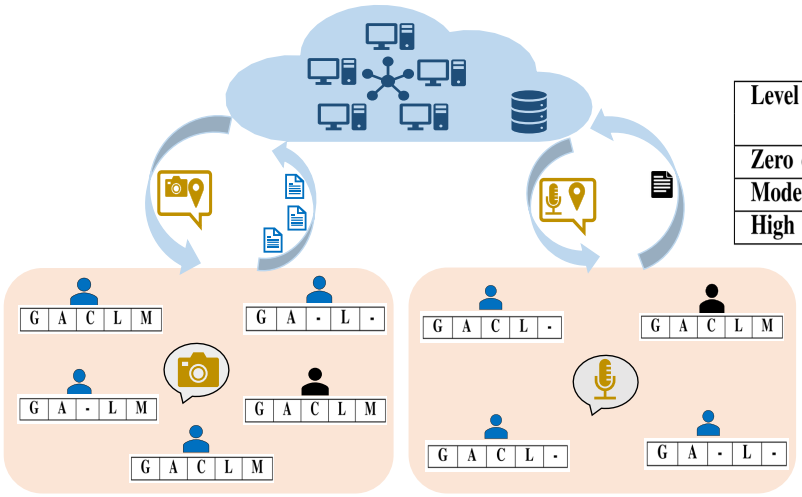
4.1 Introduction

To maintain balance in MCS, it is equally important to consider user preference along with other attributes of users in the process of recruitment. The reason for preference may vary such, as the user might feel difficulties/discomfort in participating for a particular type of tasks or the user might have privacy concerns due to the possibility of sensitive information leakage. Since each built-in sensor can reveal certain personal information, for instance, microphone and camera might compromise the confidentiality of not only participant but also surrounding people [53, 124]. As a result of these concerns, data acquisition might be affected by user unwillingness to contribute [78].

In previous chapter, we formulated user willingness by criterion based community formations. In this chapter, we move a step forward to make crowd-sensing comfortable to the participants while ensuring the acceptable level of user utility. Having said that, we model participants such that they can restrict access to the sensor in their device based on their level of comfort. More specifically, users might not participate in data acquisition for tasks that require specific sensory data. The purpose of this chapter is to gain user comfort (reduce user discomfort) while maintaining the trustworthiness of data with the presence of vulnerable participants in the sensing environment.

4.2 Proposed Solution

In the direction of trustworthy data acquisition, we propose the Reputation and Comfort Level Aware Participant Selection (RA-CLAPS) framework [20] that is built on top of previously presented Non-Selective reputation based recruitment strategy (NSR). That being said, RA-CLAPS do not authorize users to have the choice of forming communities on the basis of reputation and/or income unlike selective data acquistion.



Level of comfort	Sensor configuration (G A C L M)
Zero comfort level	1 1 1 1 1
Moderate comfort level	1 1 0 1 1
High comfort level	1 1 0 1 0

G	Global Positioning System
A	Accelerometer
C	Camera
L	Light sensor
M	Microphone

Figure 4.1: A schematic representation of the reputation and comfort level-aware participant recruitment scheme. Here co-operative and malicious participants co-exist. Two malicious participants are coloured differently for illustrative purposes. While G,A,C,L,M represents built-in sensors and 1 indicates ON, 0 indicates OFF for respective sensors.

In this chapter, we provide users with the freedom of modifying their sensor configuration before entering into the data acquisition process. We introduce the term user comfort, which increases with the switching of more sensors. We divide the complete set of participants into groups based on their comfort levels. We formulate this user comfort in the presence of malicious users, who maintains the complete set of sensors available all the time to increase their chance of selecting among comfort hungry users. Our model stands out among the other works because of the provision of user comfort with consideration of the vulnerable environment and simultaneously achieving acceptable platform utility, user utility. Fig 4.1 describes the scenario of users advertising their sensor configuration and take part in the auction. We considered a sensor set that often exists with every

smart device such as GPS, accelerometer, camera, Light sensor, Microphone. As we see in the figure, malicious participants are collocated alongside others. Malicious users are represented with dark colour, and it is clearly seen that all the attackers keep a complete set of sensors as available. We assume participant is interested in providing data if the corresponding sensor is kept ON, i.e. $\kappa_i^T = 1$, where $T_{\mathcal{C}}$ represented the sensor associated with the task T and $\mathcal{C}_i$ denotes the sensor configuration set of user i (refer to Eq 4.1).

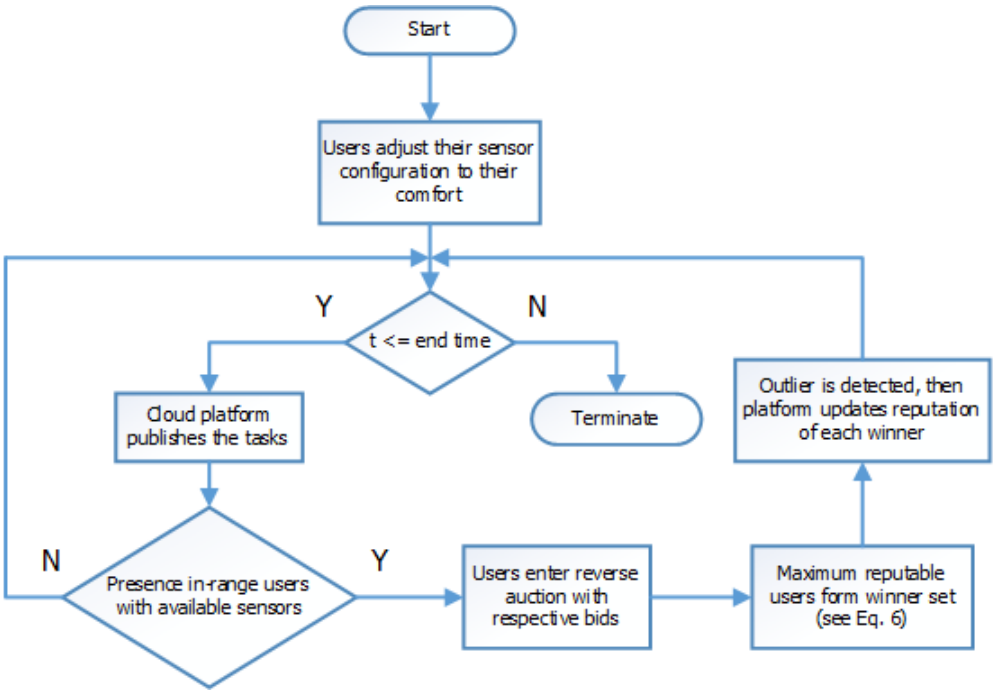


Figure 4.2: Flow chart to describe the proposed RA-CLAPS

$$\kappa_i^T = \begin{Bmatrix} 1, & [T_{\mathcal{C}} \subseteq \mathcal{C}_i \wedge RA - CLAPS\ scheme] \\ & \vee [\varrho[R_i] \leq [\frac{R_{total}^T}{\Gamma_T}] \wedge SRR\ scheme] \\ & \vee\ NSR\ scheme \\ 0, & Otherwise \end{Bmatrix} \tag{4.1}$$

As we see in the Fig 4.2 users adjust their sensor configuration at the start of the entire process of crowd-sensing (i.e. when time/campaign index t=0). In order to avoid any false payments, RA-CLAPS make sure that in-range participants having their required sensor turned ON and also can only participate for the bidding process. The user recruitment phase and rewarding phase is similar to the Non-selective reputation based recruitment

Table 4.1: Notations and equations that they appear

Notation	Equations	Description
$\mathcal{C}_i$	4.1	Set of sensors corresponding to the comfort group of user/smartphone i
$T_{\mathcal{C}}$	4.1	Sensor associated task T
a_i	4.2	Binary value that represents the availability of *Accelerometer* in sensor set of user i
c_i	4.2	Binary value that represents the availability of *Camera* in sensor set of user i
l_i	4.2	Binary value that represents the availability of *Lightsensor* in sensor set of user i
m_i	4.2	Binary value that represents the availability of *Microphone* in sensor set of user i
w_a	4.2	weight of *Accelerometer* or sensitivity of data that could be disclosed with use of it.
w_c	4.2	weight of *Camera* or sensitivity of data that could be disclosed with use of it.
w_l	4.2	weight of *Lightsensor* or sensitivity of data that could be disclosed with use of it.
w_m	4.2	weight of *Microphone* or sensitivity of data that could be disclosed with use of it.

presented in the previous chapter. After sorting out the winner set, platform computes their comfort level along with other performance metrics. Table 4.1 represent the notations used in the rest of the thesis along with the equations that use these variables.

4.3 Performance Evaluation Metrics

The simulation environment is similar to the previously mentioned 1000m x 1000m terrain with few changes in user bids as well as closeness margin as presented in Table 4.2. To evaluate the impact of our proposed comfort level aware user selection scheme, we introduce a new metric called '$userdiscomfort$' consider the following performance metrics such as

Table 4.2: Simulation settings for implementing user comfort

Parameter	Value
Size of terrain	$1000\,\text{m} \times 1000\,\text{m}$
Total number of Users	1000
Range of task	30 m
Task Arrival Rates	$20; 40; 60; 80; 100/\text{min}$
Initial reputation of a participant	0.7
Closeness margin (ϱ)	0.7
Probability of malicious users	0.05
Probability of invalid reading submitted by a user	$0.02, 0.03$
Weight factor δ	0.6
Task Value	$\{1, 2, 3, 4, 5\}$
Bid value/Sensing costs	Varied according to sensitivity of data
Total number of sensing campaigns	30
Simulation Duration	30 min
Average walking speed($\vec{V}$)	$2m/s$
Mobility awareness	False
Upper reputation threshold for malicious users to turn on attacking	0.8

- ***User Discomfort level:*** Denotes the average weight of available sensors for all participants per campaign as formulated in Eq. 4.2. Where a_i, c_i, l_i, m_i are binary variables for user i that takes zero if the corresponding sensor is unavailability in the user sensor configuration. Weights of sensor correspond to the sensitivity of data that can be disclosed through that particular sensor.

$$Discomfort = \frac{\sum_{t}\sum_{i}\Big(a_i.w_a + c_i.w_c + l_i.w_l + m_i.w_m)/|W^t|\Big)}{t_{total}} \tag{4.2}$$

  Moreover, we consider maximum discomfort for all the participants under Non-selective, selective recruitment schemes presented in chapter 3. According to the mentioned formula, maximum discomfort level can be 10 and is obtained when all the sensors are ON.

Table 4.3: User bid arrangement based on the sensitivity of data

Task-associated sensor	Value of the task	Bid value of user	Weight of the sensor
Camera	5	3, 4	$w_c = 4$
Microphone	4	2, 3	$w_m = 3$
Accelerometer	3	1, 2	$w_a = 2$
Light sensor	1,2	1	$w_l = 1$

- **Platform Utility:** It is the difference of total reputable value obtained for the crowd-sensed tasks and incentives allotted to the winners. Where t denotes the index of campaign and r_i^t denotes rewards for user i in campaign t.

$$U_{platform} = \sum_t \left(v^R(W^t) - \sum_i {\mathrm{r}_i}^t \right), \tag{4.3}$$

  Given that, $v^R(W^t)$ denotes the total reputable value obtained from the winner set W at t time slot.

- **Average user Utility:** *User Utility* indicates the average profit gained by a user per campaign. It denotes the average difference between total received rewards and sensing costs of all selected participants normalized by the total number of sensing campaigns (t_{total}) as formulated in Eq. 4.4.

$$U_{user} = \frac{\sum_t \left(\left(\sum_i {\mathrm{r}_i}^t - \sum_i {b_i}^t \right) / |W^t| \right)}{t_{total}}. \tag{4.4}$$

  Where r_i^t denotes the rewards obtained and b_i^t denotes the sensing cost for user i at time t and/or campaign t.

- **Dis-information ratio:** The ratio of tasks affected by at least one malicious user to the total number of tasks. Affected tasks can be stated as the tasks for which malicious users' are recruited.

- **Rewards to malicious:** It is the total payments given to the malicious in entire simulation period.

- **Total Energy consumption:** It is the total sum of energy consumption values for sensors that are associated with the tasks which are bid by total selected users in entire 30min.

$$T.E = \sum_{t} \sum_{T} \sum_{i} \left(x_i^{T^t} . \beta_{T^t c} + G.P.S \right) \tag{4.5}$$

We do not consider any stand-by energy drain for built-in sensors. So it is clear that in order to save residual energy of their device, users need to avoid sensing energy hunger tasks by switching off the respective sensor. The energy values used in this chapter, are derived from the power values presented in [14].

- **Average Energy Consumption:** The ratio of total energy consumption by the selected users to the number of selected. the table 4.4 represents the energy consumption values for 50sec usage of that particular sensor except for GPS which is assumed to be used only for 10sec.

$$A.E = \frac{\sum_{t} \sum_{T} \sum_{i} \left(x_i^{T^t} . \beta_{T^t c} + G.P.S \right)}{Total selected(|W|^{total})} \tag{4.6}$$

Table 4.4: Energy consumption values

Sensor	Energy Consumption Value (Joule)
Global Positioning System	4.0
Accelerometer	0.25
Camera	3.8
Microphone	1.25
LightSensor	0.15

4.4 Simulation Results

Average User Utility

Fig 4.3, 4.4 presents the average benefit gained by the selected users that is calculated from Eq 4.4 under each scheme. The illustration shows that the proposed comfort aware

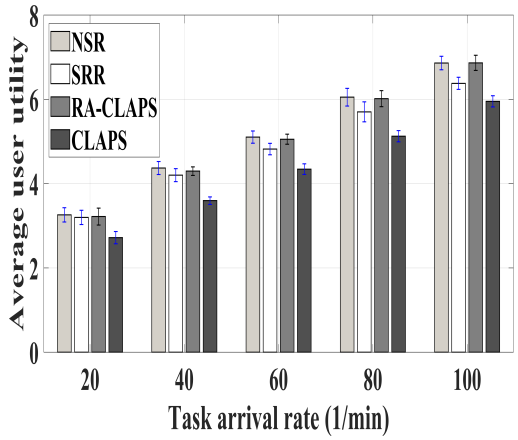


Figure 4.3: Average user utility under RA-CLAPS, CLAPS, SRR and NSR with 5% malicious probability

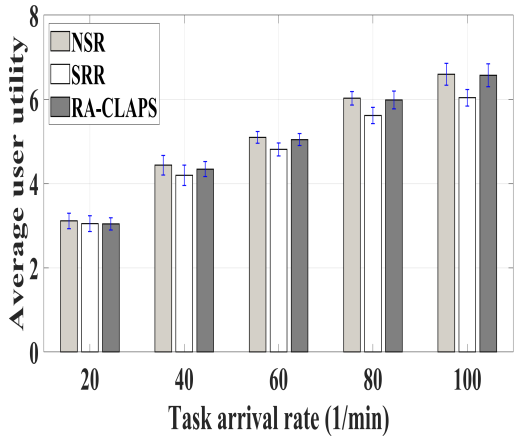


Figure 4.4: Average user utility under RA-CLAPS, CLAPS, SRR and NSR with 7% malicious probability

RA-CLAPS has maintained almost equal or negligibly lesser user utility than the baseline NSR under every task arrival rate. With enabling selectiveness for users has diminished the chance of electing a participant for multiple tasks, thereby reducing average user utility. While in the case of CLAPS, malicious users have every chance of seizing the opportunities of legitimate users.

Platform Utility

Fig 4.5 represents the utility obtained by the platform under two non-comfort aware (NSR, SRR) and two comfort ware (RA-CLAPS, CLAPS) recruitment schemes. The motivation behind comparing the proposed scheme with reputation un-aware approach (CLAPS) is to exhibit the sustainability of trustworthiness with user comfort. As seen in fig, NSR remains as the top performer in fetching utility for the platform, whereas RA-CLAPS shows improvement than SRR under 100 tasks/min. This phenomenon is because of the fact that NSR, SRR does not allow users to turn-off sensors; as a result, availability of acceptable users is comparatively more. That being said, the chance of finding incomplete tasks is high with comfort aware user recruitment. Moreover, it is possible to see the high reputation users not willing to perform the task resulting in a reduction of the level of reliable data. In addition to that, mobility of participants might cause when users move out of the sensing rang before completing the task.

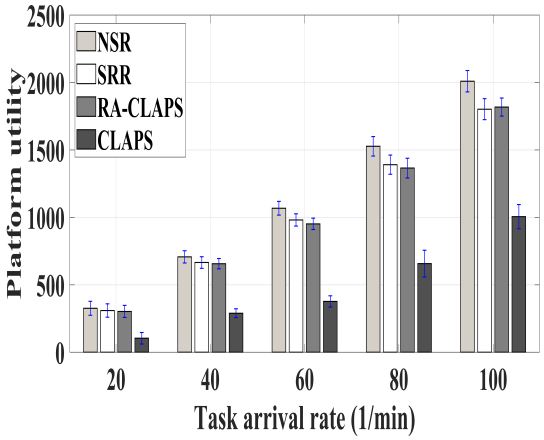


Figure 4.5: Platform utility under RA-CLAPS, CLAPS, SRR and NSR with 5% malicious probability

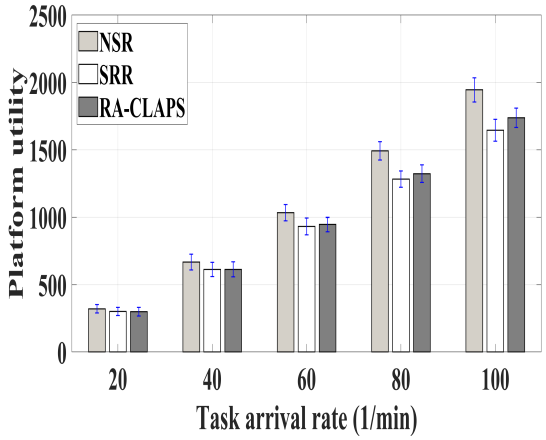


Figure 4.6: Platform utility under RA-CLAPS, CLAPS, SRR and NSR with 7% malicious probability

While under 7% malicious probability, RA-CLAPS has shown better performance compared to community selectiveness strategy. The reason for the decrease of utility under SRR is this simulation settings users bid according to sensor associated with the task as shown in Table 4.3. Since users are modelled to choose arbitrary bid out of very limited options, it turns out to be quite challenging for the platform to differentiate the most reliable user. This phenomenon affects for SRR due to the fact that users with similar reliability score form a community thus producing similar output for $v_i^R(W) - c_i/R_i$ (see Eq 3.7).

Rewards to Malicious and Disinformation Ratio

Total payments made to illegitimate users is illustrated in Fig 4.7. As one would expect the reputation unawareness has shown the highest payments for malicious users, whereas Non-selective reputation aware data acquisition (NSR) is successful in minimizing payments to malicious users. It is noticeable that RA-CLAPS mitigated the possibility of malicious users receiving payments from the platform. This slight increase compared to NSR is due to the fact that malicious users present in zero comfort zone in order to leverage their possibility of being selected. As we discussed before community selectiveness has failed to restrict unreliable users due to the degree of similarity in terms of trust score. This effect is highly perceivable with the increase of malicious probability, as seen in fig 4.8.

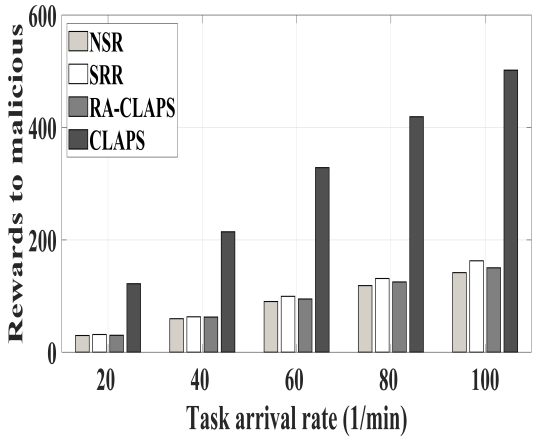


Figure 4.7: Rewards to malicious RA-CLAPS, CLAPS, SRR and NSR with 5% malicious probability

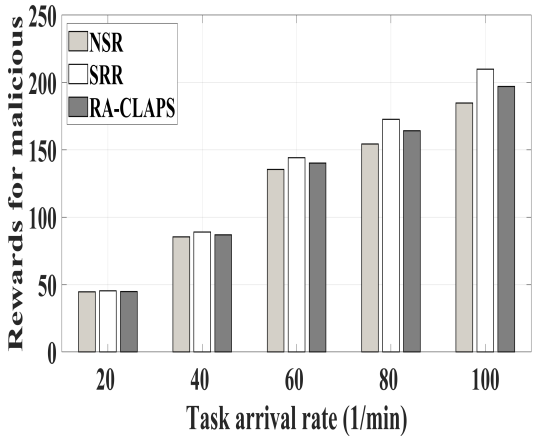


Figure 4.8: Rewards to malicious RA-CLAPS, CLAPS, SRR and NSR with 7% malicious probability

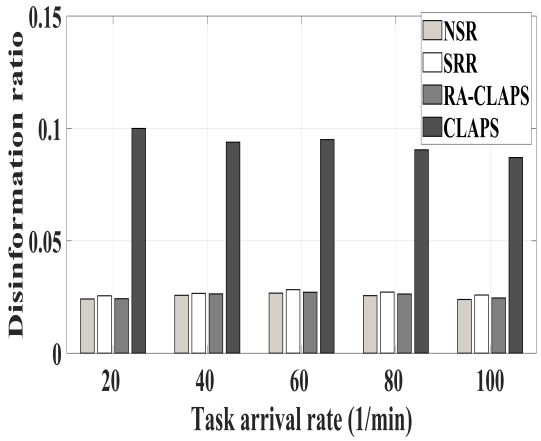


Figure 4.9: Disinformation ratio under RA-CLAPS, CLAPS, SRR and NSR with 5% malicious probability

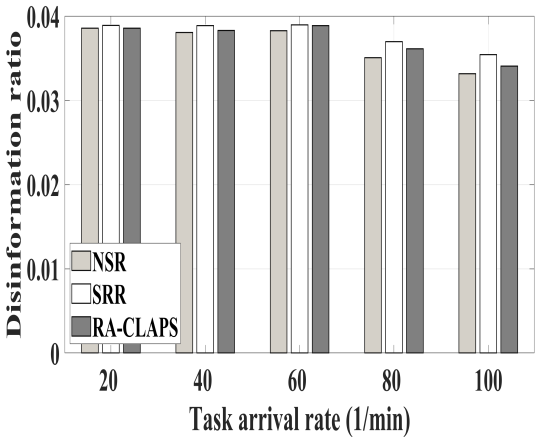


Figure 4.10: Disinformation ratio under RA-CLAPS, CLAPS, SRR and NSR with 7% malicious probability

Synchronously, disinformation ratio shows a similar impact as seen in fig 4.9, 4.10. Rewards for malicious users increases with the increase in tasks arrival rate since the total number of selected illegitimate participants is more. Results show that RA-CLAPS performed remained consistent performer for higher task requests irrespective of malicious

percentage. While under lower task arrival rate RA-CLAPS experienced challenges to find high reputable user within the range of task and having the required sensor available.

Total Energy Consumption and Average Energy Consumption

Total energy consumption is defined as total energy usage by all the selected participants for their contribution to data acquisition. We do not consider energy during transmission of data in our simulations. It is certainly true that the total number of selected users presented in Fig 4.17, 4.18 has direct impact on total energy consumption presented in Fig 4.13, 4.14. As we observe, reputation unaware (CLAPS) recruitment scheme selects more users which is evident in Fig 4.17. While SRR has selected just more than NSR, which supports the statement of users not being assigned multiple tasks due to selectiveness.

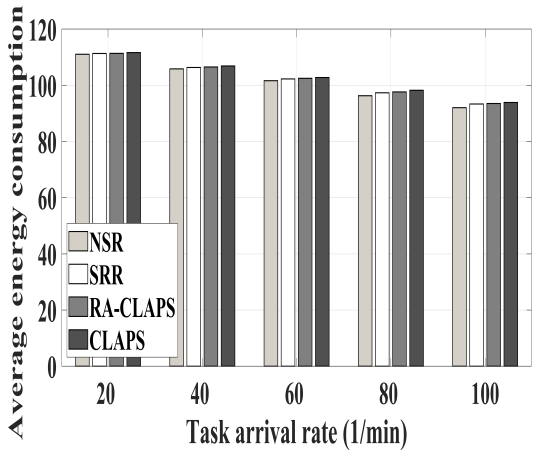


Figure 4.11: Average energy consumption under RA-CLAPS, CLAPS, SRR and NSR with 5% malicious probability

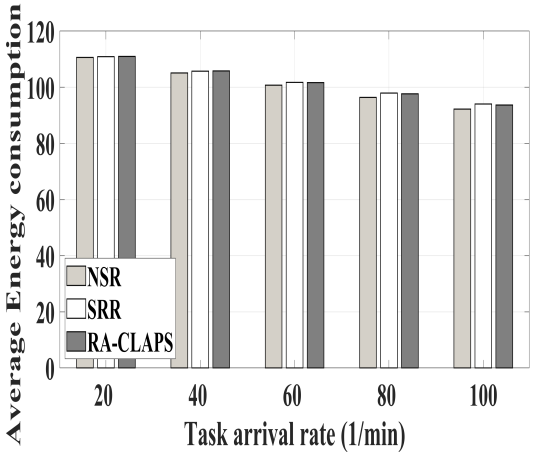


Figure 4.12: Average energy consumption under RA-CLAPS, CLAPS, SRR and NSR with 7% malicious probability

Likewise, average energy consumption presented in fig 4.11, 4.12 depends on the probability of user being bid for multiple tasks (directly proportional). Besides that, it also depends on the type of tasks the selected users perform as energy consumption values vary for different sensors. Since the G.P.S consumes a higher portion of energy the difference of average energy consumption observed is very less. Since the proposed scheme deliberately allows users to turn-off some sensors to participate with comfort, it is important to show

the percentage of tasks (T'_U) that are bid by at least one user (see Eq 3.3). Fig 4.15, 4.16 represents the total percentage of tasks that are surrounded by no less than one user.

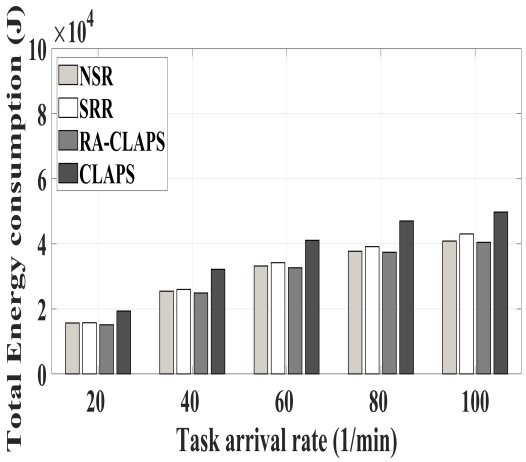


Figure 4.13: Total energy consumption under RA-CLAPS, CLAPS, SRR and NSR with 5% malicious probability

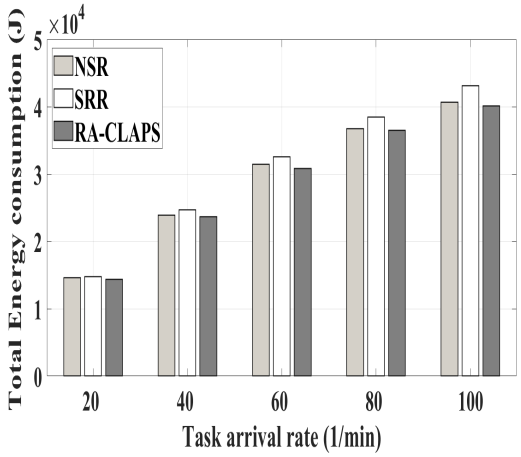


Figure 4.14: Total energy consumption under RA-CLAPS, CLAPS, SRR and NSR with 7% malicious probability

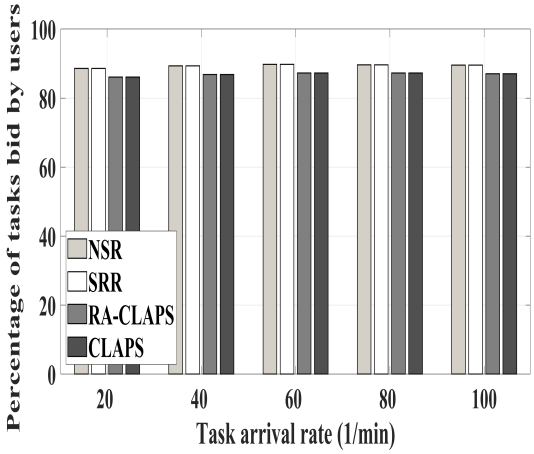


Figure 4.15: Total percentage of sensed tasks under RA-CLAPS, CLAPS, SRR and NSR with 5% malicious probability

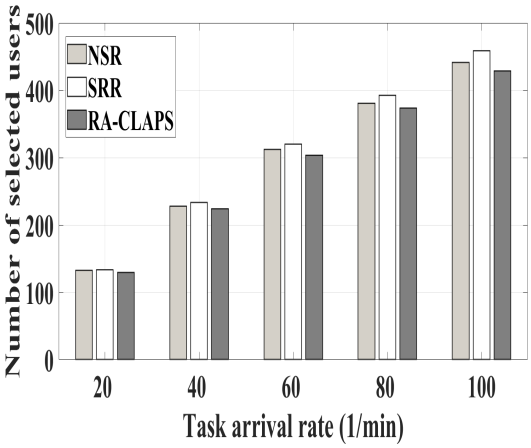


Figure 4.16: Total percentage of sensed tasks under RA-CLAPS, CLAPS, SRR and NSR with 7% malicious probability

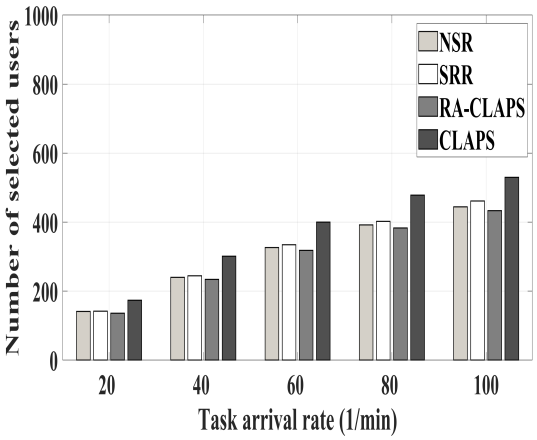


Figure 4.17: Total selected users under RA-CLAPS, CLAPS, SRR and NSR with 5% malicious users

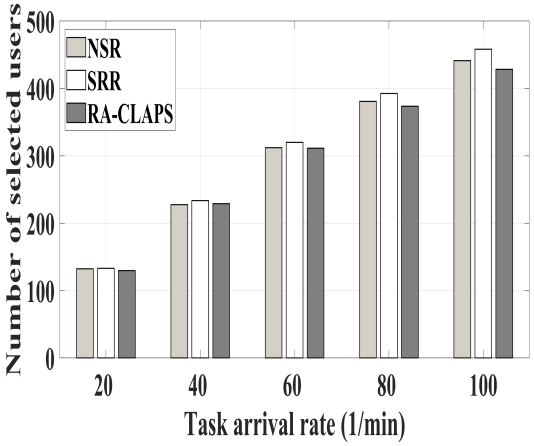


Figure 4.18: Total selected users under RA-CLAPS, CLAPS, SRR and NSR with 7% malicious users

Average Comfort Gain of Selected Participants

Tables 4.5, 4.6 denotes the average gain of user comfort for the selected participants under RA-CLAPS. Since we consider zero comfort for the users under previously presented Non-selective and selective data acquisition schemes, we assume average comfort of each user is zero. However, comfort aware data solicitation gain user comfort based on the weight of sensitive information that they restrict to contribute. Despite the poor performance due to reputation unaware, CLAPS has been productive in providing user comfort as illustrated in tab 4.7. When requested tasks are less the impact of selecting a high comfort users make huge difference to overall comfort gain that makes the normalized percentage error comparatively high.

Table 4.5: Achieved improvement in Average user comfort under RA-CLAPS (70-20-10) w.r.t Non-comfort aware schemes with 5% malicious probability

Tasks/min	**Comfort gain**
20	(6.08 ± 0.89) %
40	(6.1 ± 0.57) %
60	(5.84 ± 0.51) %
80	(6.08 ± 0.43) %
100	(6.40 ± 0.20)%

Table 4.6: Achieved improvement in Average user comfort under RA-CLAPS (70-20-10) w.r.t Non-comfort aware schemes with 7% malicious probability

Tasks/min	**Comfort gain**
20	(5.53 ± 0.81) %
40	(5.33 ± 0.46) %
60	(6.2 ± 0.52) %
80	(5.87 ± 0.59) %
100	(5.78 ± 0.22)%

Table 4.7: Achieved improvement in Average user comfort under CLAPS (70-20-10) w.r.t Non-comfort aware schemes with 5% malicious probability

Tasks/min	**Comfort gain**
20	(5.5 ± 0.82) %
40	(5.61 ± 0.47) %
60	(5.54 ± 0.40) %
80	(5.96 ± 0.33) %
100	(6.06 ± 0.22)%

4.5 Summary

In this chapter, we commence introducing comfort aware mobile crowd sensing in order to provide significance for user preference/constraints to take part in data acquisition. We consider different malicious user probabilities who do not turn-off sensors at any point in time through the simulations so as to leverage their attack opportunities. In order to enrich the discussion, we compare the proposed Reputation and comfort aware participant selection scheme (RA-CLAPS) with the previously presented Selective and Non-selective user recruitment methodologies (NSR, SRR) and also with reputation unaware version of the comfort-aware framework. A thorough simulation study has shown that under these simulation settings RA-CLAPS has improved user comfort by at most 6.4% with the loss of 7%- 10% platform utility under 5% malicious probability. While under higher vulnerable scenario (7% malicious) proposed scheme has shown comparatively lower platform utility. It is worth noting that RA-CLAPS is successful in keeping the negligible difference in user utility, the ratio of attacked tasks compared to benchmark NSR. Moreover, simulation

results for CLAPS have clearly shown the impact of reputation based recruitment. In the next chapter, we aim to reduce the loss of platform utility while authorizing users to dynamically change the pre-defined available sensor set. Moreover, we consider different combinations of comfort level settings to attain a better overall outcome. Since the increase of malicious probability has a decrease in obtained user comfort, we continue with 5% malicious probability for the further research study.

Chapter 5

An adaptive approach to Trustworthiness and Comfort-Aware Participant Recruitment in Mobile Crowd-Sensing

5.1 Introduction

In the previous feasibility studies, we considered users with selectiveness as well as with the opportunity to switch off the sensors that might collect sensitive data. In the earlier chapters, we formulated RA-CLAPS that assume 70% of *zero comfort level*: participants provides access to all the available sensors, 20% of *moderate level*: denies access to most sensitive data collection sensor such as camera and 10% of *high comfort level*: participants allow access to only sensors that would not reveal any personal identifiers(e.g. accelerometer, light sensor). Simulation results have shown that 6% comfort gain costs almost 10% loss in platform utility. However, users are assumed to modify their sensor set before registering for crowd-sensing. In this chapter, we aim to reduce the loss of platform utility and simultaneously gain user comfort with maintaining the reliability of data. Moreover, we propose Adaptive Reputation and Comfort-Aware Participant Selection strategy (Adaptive-RA-CLAPS) that allows users to modify their available sensor set in between the transition of sensing campaigns. During this process, we attempt to investigate the behaviour of RA-CLAPS under varying comfort level settings.

The decision of users adaptive modification of sensor set is based on their reputation.

The motivation behind choosing the reputation of the user as the criterion for authorizing adaptive modification of sensor set is to emphasize the prominence of high reputation score which indeed can improve the trustworthiness of data obtained from the crowd-sensing participants.

5.2 System Overview

We propose an Adaptive Reputation and Comfort-Aware Participant Selection scheme to dynamically revise their sensor configuration set in the course of involvement in data acquisition. Our focus to the proposal of Adaptive RA-CLAPS is to improve the flexibility and comfort of the participants and decrease the loss in platform utility. It is also essential to mention that the simulation results of RA-CLAPS showed the loss of platform utility for allowing users to turn-off the access to some of the built-in sensors. Moreover, it is not idealistic to make participants to choose their sensor set only at the beginning of the process of data collection. As mentioned in [67], users may be uncomfortable to share certain information from specific locations which may lead to participants unwillingness for that sensor associated tasks for the entire duration of data solicitation. That could eventually effect platform due to unavailability of reliable participants.

In addition to three phases (i.e. Comfort level adjustment, user recruitment, rewarding phase) presented in previous chapters, the proposed model includes the sensor adjustment phase succeeding rewarding phase as seen in the flowchart 5.4. Since the previous work has considered only one set of variation in comfort level, in this chapter we simulate multiple variations for the proposed three comfort zones. After the selection of appropriate combination, then we proceed build an adaptive phase for that particular set of arrangement.

5.2.1 Simulation Study

We considered 80 tasks/min as the standard arrival rate which is a moderate load level to draw a conclusion. We use the same performance metrics as well as simulation settings that are detailed in the previous chapter (see section 4.3).

Platform utility

The fig 5.1 illustrates the performance of RA-CLAPS under various combinations of comfort levels. The simulation results show that the higher the percentage of zero comfort

participants, lower the loss of utility for the platform. This phenomenon is because of the unavailability of higher reputable users around the sensing task in many instances. It is intuitive to see the decrease but we attempt to quantify the relation between comfort level and platform utility. It is worth noting that with the appropriate setting of comfort level platform utility is well maintained comparatively as it is evident in the case of 90-5-5, 80-15-5, 70-2-10, 70-25-5, 70-15-15, 60-30-10 combinations (where x-y-z: x indicated percentage of zero comfort participants, y indicates moderate, z indicated high comfort participants). Moreover, it is visible to the considerable decrease of utility when percentage of high comfort users is more than medium comfort participants. Since the results have shown almost 30% decrease in platform utility for combinations with 60% of zero comfort participants, we consider not to further decrease the percentage of zero comfort. In addition to that, it is evident that platform utility has exhibited an inverse relationship with the high comfort levels.

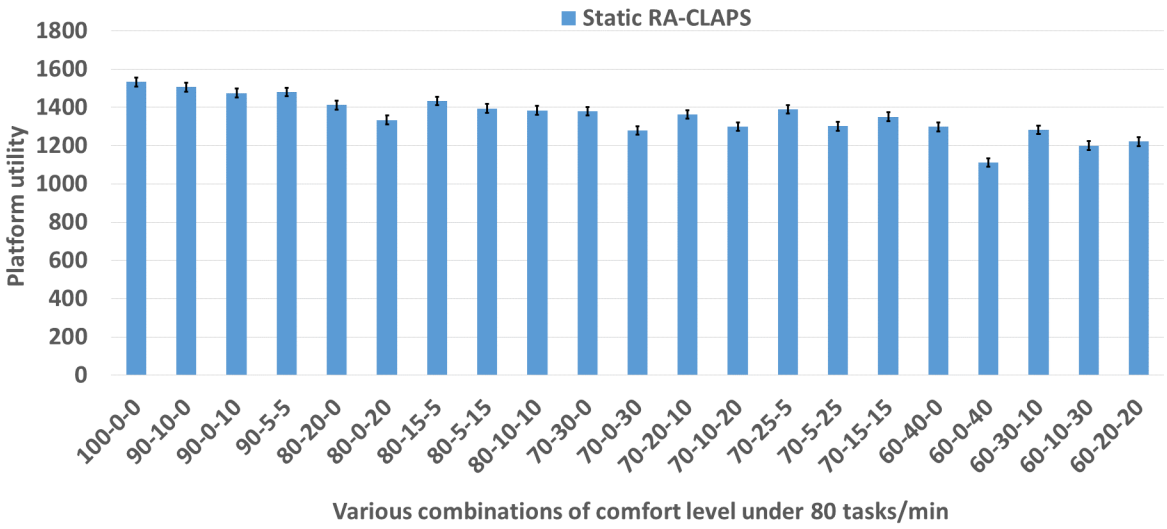


Figure 5.1: Platform utility under static RA-CLAPS for 80 tasks/min.

Average User Utility

Fig 5.2 represents the average user utility for various combinations of comfort levels under 80 tasks/min. It represents the average user utility for the selected participants. As we

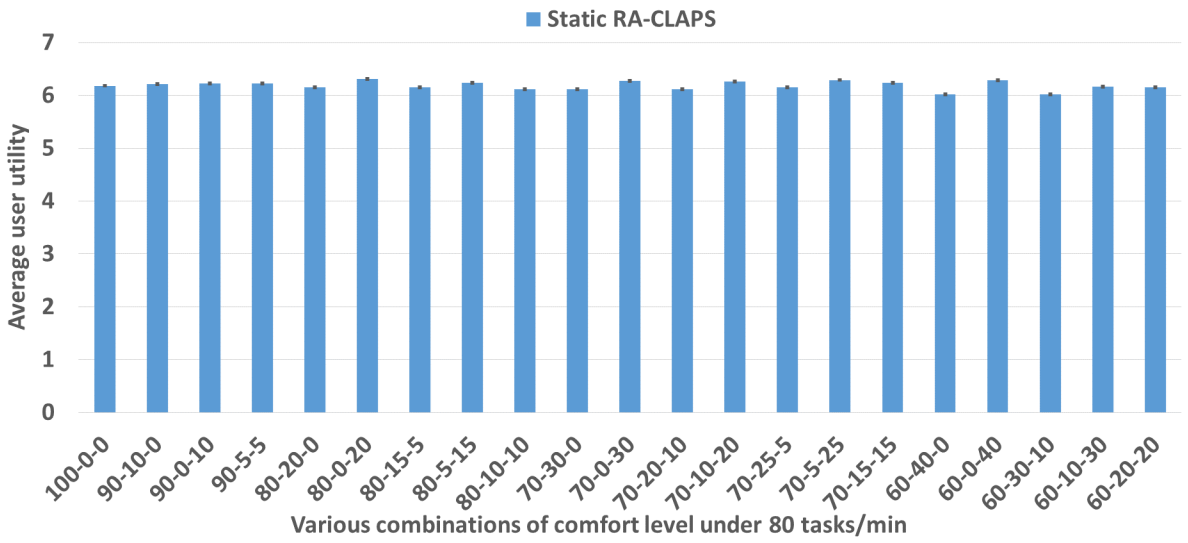


Figure 5.2: Average user utility under static RA-CLAPS for 80 tasks/min.

can see user utility has shown increment when the portion of higher comfort users is more comparative to medium comfort participants such as 80-0-20, 80-5-25, 70-0-30, 70-10-20, 70-5-25, 60-0-40. This effect is exactly opposite to the behaviour shown with platform utility. The reason for this performance is explained as, with the effect of sensor turn-off, platform settles with the available truthful user most probably from zero comfort group, since they participate in comparatively more tasks than moderate, high comfort users. Thus eventually improve their income and indeed user utility.

Case Study on Average User Discomfort

Fig 5.3 provides the overview of average discomfort level of users under various combinations (maximum possible discomfort level is 10). It is computed for the selected users as we do in case of other performance metrics. It can be clearly visualized that, comparatively higher proportion of medium comfort participants has maintained a considerable level of user discomfort. One would expect to see more comfort gain with a greater percentage of high comfort users, but results have shown the other. This phenomenon can be justified as the proportion of selected participants are more from the other two comfort groups (i.e. zero and moderate). This statement also strengthens the argument that, users from zero comfort group tend to be selected more often than high comfort level users.

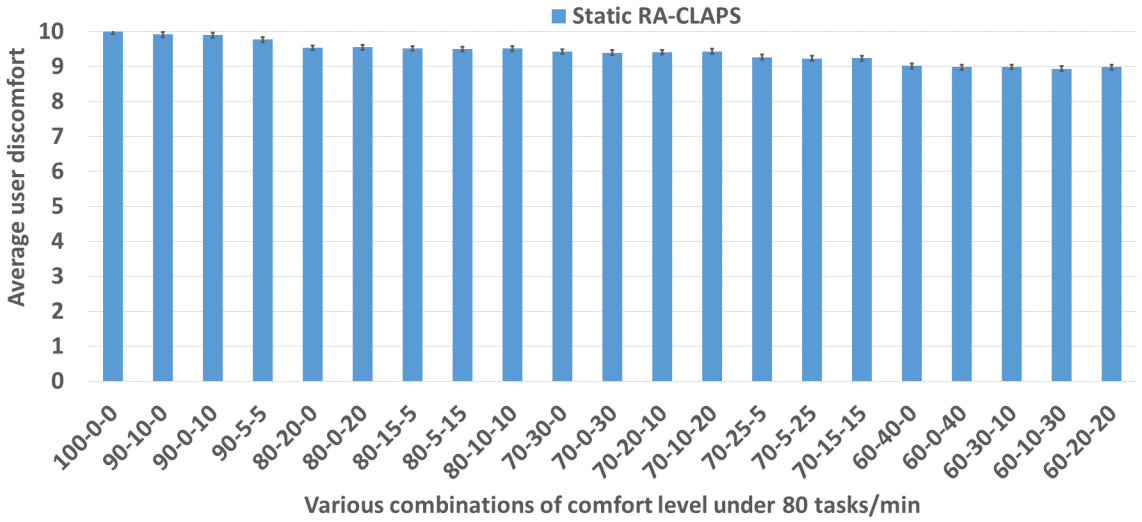


Figure 5.3: Average User Discomfort under static RA-CLAPS for 80 tasks/min.

The thorough investigation has shown the relationship among user comfort, utility and platform utility that authorizing more participants to turn off both camera and microphone (High comfort level users) have shown an extreme decrease in utility of the platform due to the presence of more unsensed. Though considering large number of high comfort users improves the average user comfort of participants and user utility by miniature, platform loss is comparatively large. To this end, we conclude by choosing the appropriate presetting as 90% zero comfort, 5% medium comfort and 5% high comfort users. The reason for choosing out of other better performances is, it enables users to be in all three comfort zones alongside reaching better platform utility and user comfort with substantial user utility.

5.2.2 Sensor Adjustment Phase

While stated in [67], user constraints may vary from location to location; we move forward to initiate sensor adjustment phase after certain period. With the completion of reverse auction winner set W is announced. We model the modification step for completion of every three campaigns or 3 min. The justification behind this assumption is that every user has mobility with an average speed of 2 m/s. Considering that fact user travels a

considerable distance away from the previous location in 3min, we insist users to keep the complete set of sensors ON after completion every 3min.

As we mentioned previously, sensor adjustment in the midst of 30min simulation time is enabled based on users reputation. That being said, we breakdown the reputation levels into four levels and each level offer certain levels of comfort. They are as follows.

- High reliability level : $R_w \geq Th_{High}$
- Moderate reliability level : $Th_{Moderate} \leq R_w < Th_{High}$
- Naive reliability level : $Th_{Naive} < R_w < Th_{Moderate}$
- Low reliability level : $R_w < Th_{Naive}$

where R_w represents the reputation of winner $w \in W$.

In accordance with the considered system model, we add sensor modification step as a choice to the users, which is represented by the probability P($\mathcal{C}_w$). Moreover, we assume users always aim to increase their comfort levels as they increase their reputation score/reliability level. To do so, each user intends to turn-on comparatively more sensors that increase the probability of being selected for multiple tasks. We impose this behaviour in the algorithm as participants in the low-reliability level exhibit all the available set as ON in order to get notified by the platform. While with the intention to provide choice of comfort-based participation, we allow 25% of users (who belong to winner set) to restrict their contribution for microphone associated task. Simultaneously, 33% of moderate comfort users aim at deliberately reject to perform camera related tasks. In addition to that, 50% of high-reliability users tend to restrict their contribution to less sensitive information collection. However, other users remain with their previous sensor configuration and the above-mentioned modification lasts until the system suggests them to turn ON the complete sensor set (after every n^{th}).

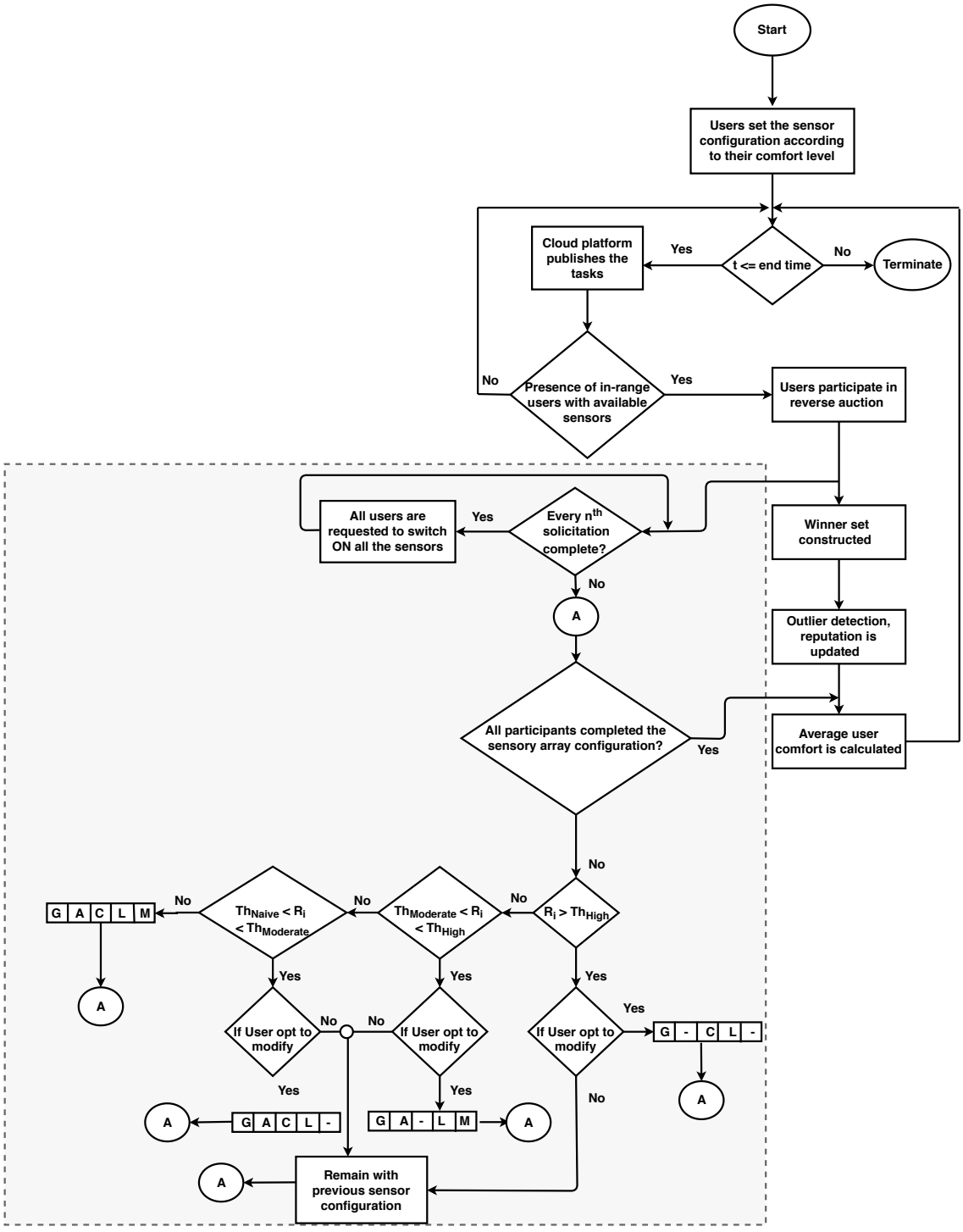


Figure 5.4: Flow chart to describe the proposed RA-CLAPS

5.3 Additional Simulation Setting for Implementing Adaptive RA-CLAPS

We evaluate the performance of the Adaptive RA-CLAPS under the performance metrics, simulation settings mentioned in 4.3. Additional settings implemented in this chapter are presents in Tab 5.1.

Table 5.1: Simulation settings

Parameter	Value
Th_{High}	0.95
$Th_{Moderate}$	0.8
Th_{Naive}	0.7
n	3

Additional Performance Metrics

Overall Efficiency is the difference in percentage of gain in user comfort to the average loss percentage in user and platform utilities when compared with benchmark scheme NSR as shown in Eq. 5.1, 5.2, 5.3. The justification behind the consideration of overall efficiency is to evaluate the impact of user comfort over the utilities of user, platform as determined by Eq. 5.4. Positive overall efficiency indicates the improvement of the user comfort level is more than the average loss occurred in user and platform utilities.

$$E_{U_{user}} = \frac{\left(U_{user}^{RA-CLAPS} - U_{user}^{NSR}\right) * 100}{U_{user}^{NSR}} \tag{5.1}$$

$$E_{U_{platform}} = \frac{\left(U_{platform}^{RA-CLAPS} - U_{user}^{NSR}\right) * 100}{U_{platform}^{NSR}} \tag{5.2}$$

$$E_{C_{user}} = \frac{\left(C_{user}^{RA-CLAPS} - C_{user}^{NSR}\right) * 100}{C_{user}{}^{NSR}} \tag{5.3}$$

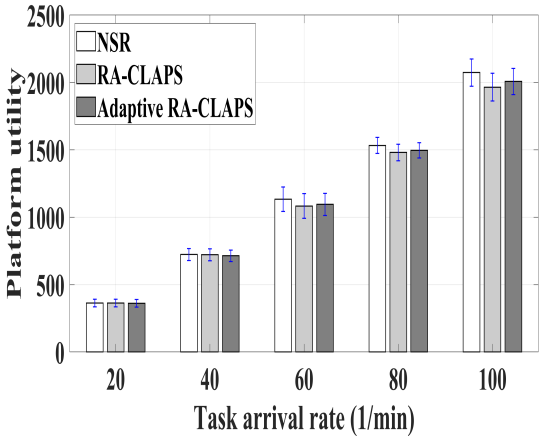


Figure 5.5: Platform utility under RA-CLAPS, Adaptive RA-CLAPS and NSR

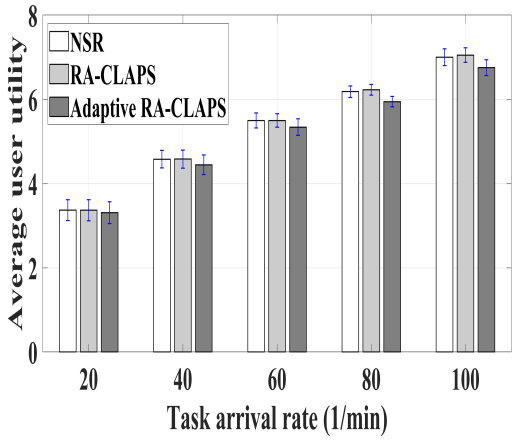


Figure 5.6: Average User utility under RA-CLAPS, Adaptive RA-CLAPS and NSR

$$E_{overall} = \frac{\left(E_{U_{user}} + E_{U_{platform}}\right)}{2} + E_{C_{user}} \tag{5.4}$$

Platform Utility

As illustrated in Fig 5.5 the proposed adaptive RA-CLAPS has shown the considerable improvement in the utility of platform compared to previous proposed RA-CLAPS [20]. In addition to that, under lower task request rate both adaptive and static approaches performed similar to benchmark scheme NSR. However, the reason for the considerable decrease exhibited by comfort aware schemes is due to the unavailability of participants around the task as discussed before. Compared to benchmark adaptive comfort aware scheme has reduced platform utility by 0.7% - 3.2%, whereas compared with RA-CLAPS there is an increase up to 2.2% under different task arrival rates.

Average User Utility

Comparison of the difference between total rewards gained to the total bids under NSR, RA-CLAPS, Adaptive RA-CLAPS can be seen in Fig 5.6. Amongst the three recruitment strategies, the previously proposed RA-CLAPS scheme under the chosen comfort

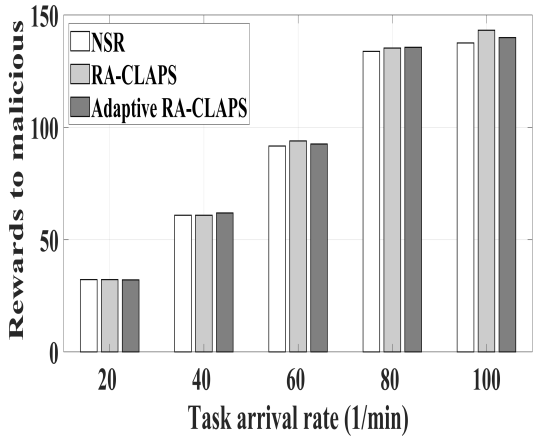


Figure 5.7: Rewards to malicious under RA-CLAPS, Adaptive RA-CLAPS and NSR

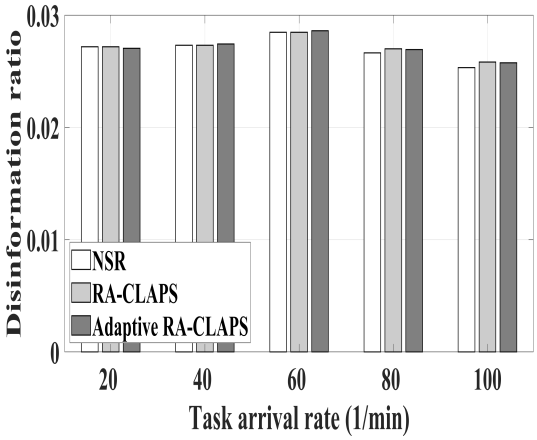


Figure 5.8: Disinformation ratio under RA-CLAPS, Adaptive RA-CLAPS and NSR

level combination has shown the best average user utility figures. Contrary to that, under Adaptive version average user utility is dropped by 3.6% in accordance with the benchmark NSR and 4.2% compared to RA-CLAPS. This behaviour is because of the reset of sensor configuration (all sensors ON) after every certain period of time / certain number of campaigns. This helps every participant to get equal opportunity to bid for reverse auction, hence increasing the chance of recruiting the new participant. As a result, it reduces the average utility of participant but in turn, increases the total number of participants recruited, as shown in Fig 5.11.

Rewards to Malicious and Disinformation ratio

As we mentioned in performance metrics, Rewards to malicious denotes the total payments received by malicious users or total false payments made by the cloud platform. In Fig 5.7 total false payments under NSR, RA-CLAPS, Adaptive RA-CLAPS are illustrated with varying task arrival rate. Since Non-selective reputation-based recruitment strategy do not favour users with comfort-aware sensing, it shows the least payments for malicious users. However, adaptive version of RA-CLAPS has shown the comparative deduction in payments to adversaries under 100 tasks/min, which is reflected in terms of gain in platform utility.

In addition to the rewards to illegitimate users, another metric that is related to the

number of selected adversaries is Disinformation ratio. It denotes the ratio of misinformed tasks in the overall 30min crowd-sensing. Fig 5.8 shows the negligible difference among the three discussed recruitment schemes. Since the number of attacked tasks is the same in three schemes, the trustworthiness of data is successfully sustained by comfort-ware user selection frameworks (RA-CLAPS, Adaptive RA-CLAPS) in a vulnerable environment.

Total Energy Consumption and Average Energy Consumption

Average energy consumption value of user has shown a gradual decrease with the increase in the number of tasks/min as plotted in Fig 5.9. It depends on the multiple tasks that are bid by the selected participant and the energy value associated with that task. Fig 5.10 represents the total energy consumption of selected participants. That being said it depends on the total number of selected participants that is presented in Fig 5.11 . Having said that it is essential to present the total selected participants in the entire simulation time(see Fig 5.12).

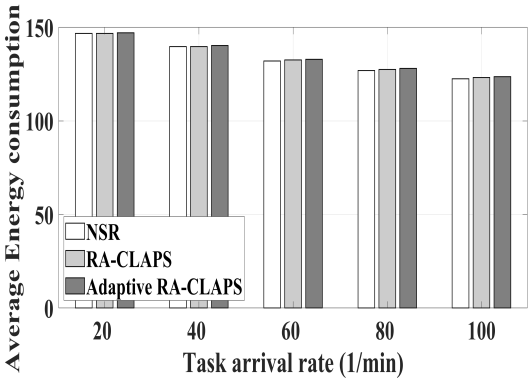


Figure 5.9: Average Energy consumption under RA-CLAPS, Adaptive RA-CLAPS and NSR

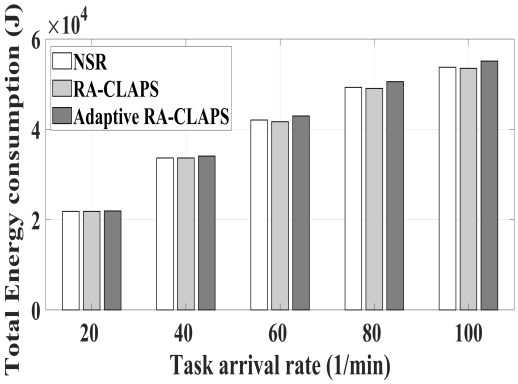


Figure 5.10: Total Energy consumption under RA-CLAPS, Adaptive RA-CLAPS and NSR

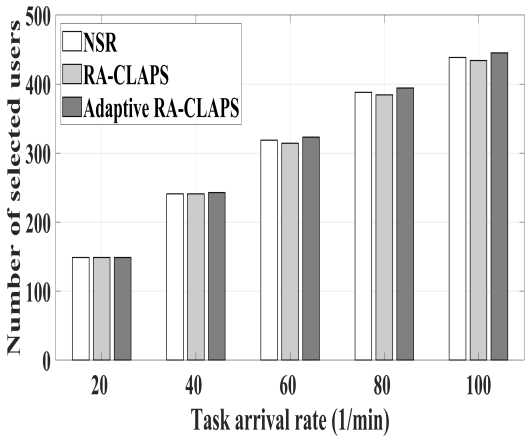


Figure 5.11: Total number of selected users under RA-CLAPS, Adaptive RA-CLAPS and NSR

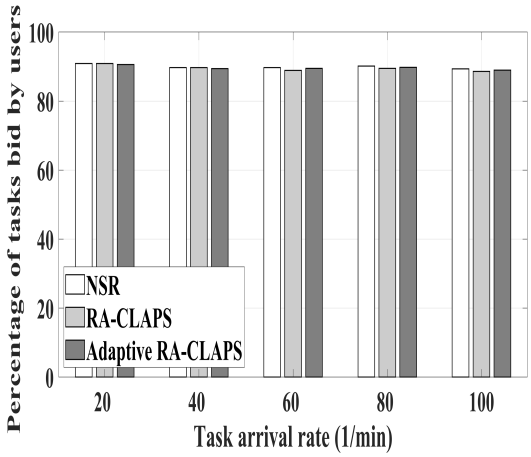


Figure 5.12: Percentage of tasks bid by users under RA-CLAPS, Adaptive RA-CLAPS and NSR

Average User Discomfort

The most essential metric in order to show the effect of the proposed comfort-aware scheme is average user discomfort, as presented in Table 5.2,5.3. Since the adaptive RA-CLAPS authorize users to switch-off sensors in between sensing campaigns based on the reliability level of users, as explained in flowchart 5.4, user discomfort has shown a considerable decrease. In other words, user comfort has increased due to the process of turning off their respective inconvenient sensors. The percentage gain in user comfort is represented along with the normalized percentage error to validate the obtained results.

Table 5.2: Achieved improvement in Average user comfort under Static RA-CLAPS (90-5-5) w.r.t Non-comfort aware schemes with 5% malicious probability

Tasks/min	**Comfort gain**
20	(0.06 ± 0.002) %
40	(0.08 ± 0.002) %
60	(2.14 ± 0.001) %
80	(2.29 ± 0.0004) %
100	(2.67 ± 0.0009)%

Table 5.3: Achieved improvement in Average user comfort under Adaptive RA-CLAPS (90-5-5) w.r.t Non-comfort aware schemes with 5% malicious probability

Tasks/min	**Comfort gain**
20	(2.26 ± 0.0008) %
40	(3.29 ± 0.0006) %
60	(3.53 ± 0.0005) %
80	(4.41 ± 0.0007) %
100	(5.03 ± 0.001)%

Efficiency

Overall efficiency of the system is presented in table 5.4 that is calculated based on the Eq. 5.4. It is the compared efficiency improvement in both static and adaptive RA-CLAPS with benchmark approach. Simulation results proved that overall efficiency of system has improved under adaptive approach compared to static RA-CLAPS, indeed supporting adaptive sensor modification.

Table 5.4: Percentage difference in various parameters under Static RA-CLAPS (90-5-5), Adaptive RA-CLAPS (90-5-5) w.r.t Non-Selective Reputation-aware scheme under same settings.

Tasks/min	*Static* **RA-CLAPS**				*Adaptive* **RA-CLAPS**			
	$E_{U_{user}}$	$E_{U_{platform}}$	$E_{C_{user}}$	$E_{overall}$	$E_{U_{user}}$	$E_{U_{platform}}$	$E_{C_{user}}$	$E_{overall}$
20	-0.04%	-0.05%	0.06%	0.01%	-1.70%	-0.74%	2.26%	1.03%
40	-0.02%	-0.30%	0.08%	-0.05%	-3.00%	-1.40%	3.30%	1.14%
60	-0.04%	-4.50%	2.14%	-0.12%	-2.90%	-3.4%	3.50%	0.38%
80	0.70%	-3.40%	2.30%	0.92%	-3.86%	-2.44%	4.40%	1.26%
100	0.70%	-5.20%	2.70%	0.42%	-3.6%	-3.20%	5.03%	1.66%

5.4 Summary

The objective behind the adaptive sensor set modification is to allow users the flexibility and comfort during the process of data collection. In order to simultaneously facilitate the trustworthiness of data and gain user comfort, we proposed Adaptive RA-CLAPS. Moreover, with the intention to decrease the loss of platform utility obtained under RA-CLAPS, we choose the proper combination of comfort level groups and then applied adaptive modification phase on top of that. The simulation results have shown the improvement of platform utility by almost 2.5% and average user comfort by 1.4%- 3.2% compared to previous presented RA-CLAPS. However, there is a considerable decrease in user utility by 1.7%- 4.2% due to the increase in total selected users by at most 2.8%. In addition to that, compared to benchmark NSR, it showed the gain in average user comfort by 5% at the cost of 3.1% loss in platform utility and a 3.5% loss in user utility, however overall efficiency is high than the predecessors.

The proposed adaptive RA-CLAPS has facilitated platform utility, whereas the RA-CLAPS (non-adaptive) has promoted user utility, user comfort. To this end, we state that the comfort aware approach could achieve considerable platform utility in an environment where sensors could be limited than expected.

Chapter 6

Conclusion and Future Directions

Advancements in mobile technology has facilitated data acquisition through non-dedicated sensors. That being said, Mobile Crowd-Sensing can be a viable tool for the collection of enormous, heterogeneous data that could be requested of a variety of IoT applications.Due to the fact of anonymous participants being the primary source of information, data trustworthiness is one of the major challenges faced by MCS. There may be participants who accidentally send in-appropriate data due to misplacing of the sensor while collection of data, sensor malfunction or adversaries that deliberately send false data. In addition to that, the developed crowd-sensing system should be attracting/ motivating and sustaining a large number of participants in order to uphold the process of data collection. Many user-centric incentive mechanisms, context specific solutions have been proposed for motivating participants and to address user drop-off problems.

We have investigated the impact of users intention to improve their income by forming groups based on the comparison of income with the average group income. We implement this in the presence of 5% population of adversaries to study the robustness of the system. Simulations results have shown that promoting income based selectiveness could degrade the platform utility as well as user utility by a considerably higher margin. With this in mind, we formulated reputation-based community formations under two malicious probability settings(5%, 7%). Through simulations results, it is evident that authorizing community formation w.r.t reputation scores can facilitate platform but obstruct the improvement of user utility [21]

With the motive to investigate the consequences of participants turning off some of the available built-in sensors that could cause privacy breach, we introduced reputation and comfort-aware participant selection scheme(RA-CLAPS) [20]. This allows users to modify

their sensor set according to their level of comfort. We arbitrarily chose the combination of three comfort levels as 70% zero comfort level users, 20% moderate comfort level users and 10% users in high comfort level. Thorough simulation results have shown that the proposed model could entirely favour users with the loss of platform utility by almost 10% while achieving at most 6.4% gain in user comfort.

Since RA-CLAPS has considered only one combination of comfort levels, we continued with other possible combinations to better understand the behaviour of the proposed model. We then introduced an adaptive version of RA-CLAPS that facilitated users to dynamically modify sensor set in between switching to a new task. Our study has shown that adaptive version of reputation and comfort-aware scheme attempted to decrease the loss in platform utility for the gain of user comfort, however it led to decrease in considerable average user utility and increase in the total number of users profited.

The aforementioned milestones are obtained by simultaneously considering user and platform utilities, and mobility of users. In addition to that, the trustworthiness of data is never compromised while considering malicious users collocated with rational users. In addition to the that overall efficiency of the system is improved compared to other considered models. Moreover, one of the considerable limitation of this work includes that if the density of users in the considered terrain decreases, there is a possibility that proposed selective, comfort aware recruitment schemes might not show positive impact over platform and user utilities due to the unavailability of participants. Besides, we assume each user can perform unlimited number of tasks in entire simulation period which may be unfeasible in practice.

Furthermore, implementing reputation and comfort-aware recruitment scheme for tackling user drop-off problem could be another interesting part of future work. We considered Random-way point mobility model with mobility unawareness, which can be extended with other mobility models such as social mobility and also successful estimation of next location of users might decrease the loss in platform utility. User location privacy is considered only for specific period during the entire process, since we have considered user location as reference while accepting the data. However, implementation of user location anonymizing could be a possible future study. This thesis work is confined to user recruitment and data collection process, besides challenges and issues during the data transmission would be a considerable extension to this work. Consideration of multiple cloud servers alongside untrusted servers that cause security breaches could be another challenging aspect. Studying the impact of the proposed user recruitment strategies in a budget constrained environment is a way forward to extend this work. Applying game theory on interactions among users, between users and platform would be an interesting, challenging work.

References

[1] Cisco visual networking index: Global mobile data trac forecast update 2015-2020 white paper. Cisco, 2016.

[2] Tarek Abdelzaher, Yaw Anokwa, Peter Boda, Jeff Burke, Deborah Estrin, Leonidas Guibas, Aman Kansal, Samuel Madden, and Jim Reich. Mobiscopes for human spaces. *IEEE pervasive computing*, 6(2):20–29, 2007.

[3] Alessandro Acquisti and Jens Grossklags. Privacy and rationality in individual decision making. *IEEE security & privacy*, 3(1):26–33, 2005.

[4] Saurabh Amin, Steve Andrews, Saneesh Apte, Jed Arnold, Jeff Ban, Marika Benko, Re M Bayen, Benson Chiou, Christian Claudel, Coralie Claudel, et al. Mobile century using gps mobile phones as traffic sensors: A field experiment. 2008.

[5] Haleh Amintoosi and Salil S Kanhere. A reputation framework for social participatory sensing systems. *Mobile Networks and Applications*, 19(1):88–100, 2014.

[6] Jian An, Xiaolin Gui, Zhehao Wang, Jianwei Yang, and Xin He. A crowdsourcing assignment model based on mobile crowd sensing in the internet of things. *IEEE Internet of Things Journal*, 2(5):358–369, 2015.

[7] Fazel Anjomshoa and Burak Kantarci. Sober-mcs: Sociability-oriented and battery efficient recruitment for mobile crowd-sensing. *Sensors*, 18(5):1593, 2018.

[8] Paolo Bellavista, Dimitri Belli, Stefano Chessa, and Luca Foschini. A social-driven edge computing architecture for mobile crowd sensing management. *IEEE Communications Magazine*, 57(4):68–73, 2019.

[9] Dimitri Belli, Stefano Chessa, Luca Foschini, and Michele Girolami. A social-based approach to mobile edge computing. In *2018 IEEE Symposium on Computers and Communications (ISCC)*, pages 00292–00297. IEEE, 2018.

[10] Jeffrey A Burke, Deborah Estrin, Mark Hansen, Andrew Parker, Nithya Ramanathan, Sasank Reddy, and Mani B Srivastava. Participatory sensing. 2006.

[11] Tracy Camp, Jeff Boleng, and Vanessa Davies. A survey of mobility models for ad hoc network research. *Wireless communications and mobile computing*, 2(5):483–502, 2002.

[12] Andrew T Campbell, Shane B Eisenman, Nicholas D Lane, Emiliano Miluzzo, and Ronald A Peterson. People-centric urban sensing. In *Proceedings of the 2nd annual international workshop on Wireless internet*, page 18. ACM, 2006.

[13] Andrew T Campbell, Shane B Eisenman, Nicholas D Lane, Emiliano Miluzzo, Ronald A Peterson, Hong Lu, Xiao Zheng, Mirco Musolesi, Gahng-Seop Ahn, et al. The rise of people-centric sensing. *IEEE Internet Computing*, (4):12–21, 2008.

[14] Aaron Carroll and Gernot Heiser. The systems hacker's guide to the galaxy. *Proc. of APSYS*, 2013.

[15] Younghun Chae, Lisa Cingiser DiPippo, and Yan Lindsay Sun. Trust management for defending on-off attacks. *IEEE Transactions on Parallel and Distributed Systems*, 26(4):1178–1191, 2014.

[16] Craig H Chapman, Kush G Parikh, Oliver B Downs, Robert C Cahn, and Jesse S Hersch. Determining road traffic conditions using data from multiple data sources, March 22 2011. US Patent 7,912,628.

[17] Man Hon Cheung, Richard Southwell, Fen Hou, and Jianwei Huang. Distributed time-sensitive task selection in mobile crowdsensing. In *Proceedings of the 16th ACM International Symposium on Mobile Ad Hoc Networking and Computing*, pages 157–166. ACM, 2015.

[18] Cory Cornelius, Apu Kapadia, David Kotz, Dan Peebles, Minho Shin, and Nikos Triandopoulos. Anonysense: privacy-aware people-centric sensing. In *Proceedings of the 6th international conference on Mobile systems, applications, and services*, pages 211–224. ACM, 2008.

[19] Chenyun Dai, Dan Lin, Elisa Bertino, and Murat Kantarcioglu. An approach to evaluate data trustworthiness based on data provenance. In *Workshop on Secure Data Management*, pages 82–98. Springer, 2008.

[20] Venkat Surya Dasari, Burak Kantarci, and Murat Simsek. Trustworthiness and comfort-aware participant recruitment for mobile crowd-sensing in smart environments. In *2019 IEEE Symposium on Computers and Communications (ISCC) Barcelona, Spain(accepted)*, pages 1–6. IEEE, 2019.

[21] Venkat Surya Dasari, Maryam Pouryazdan, and Burak Kantarci. Selective versus non-selective acquisition of crowd-solicited iot data and its dependability. In *2018 IEEE International Conference on Communications Workshops (ICC Workshops)*, pages 1–6. IEEE, 2018.

[22] Chrysanthos Dellarocas. Analyzing the economic efficiency of ebay-like online reputation reporting mechanisms. In *Proceedings of the 3rd ACM Conference on Electronic Commerce*, pages 171–179. ACM, 2001.

[23] Akshay Dua, Nirupama Bulusu, Wu-Chang Feng, and Wen Hu. Towards trustworthy participatory sensing. In *Proceedings of the 4th USENIX conference on Hot topics in security*, pages 8–8, 2009.

[24] Akshay Dua, Nirupama Bulusu, Wu-Chang Feng, and Wen Hu. Combating software and sybil attacks to data integrity in crowd-sourced embedded systems. *ACM Transactions on Embedded Computing Systems (TECS)*, 13(5s):154, 2014.

[25] Lingjie Duan, Takeshi Kubo, Kohei Sugiyama, Jianwei Huang, Teruyuki Hasegawa, and Jean Walrand. Incentive mechanisms for smartphone collaboration in data acquisition and distributed computing. In *2012 Proceedings IEEE INFOCOM*, pages 1701–1709. IEEE, 2012.

[26] Jingyao Fan, Qinghua Li, and Guohong Cao. Privacy-aware and trustworthy data aggregation in mobile sensing. In *2015 IEEE Conference on Communications and Network Security (CNS)*, pages 31–39. IEEE, 2015.

[27] Xinxin Fan, Ling Liu, Mingchu Li, and Zhiyuan Su. Grouptrust: dependable trust management. *IEEE Transactions on Parallel and Distributed Systems*, 28(4):1076–1090, 2017.

[28] Wei Feng, Zheng Yan, Hengrun Zhang, Kai Zeng, Yu Xiao, and Y Thomas Hou. A survey on security, privacy, and trust in mobile crowdsourcing. *IEEE Internet of Things Journal*, 5(4):2971–2992, 2018.

[29] Olga Galinina, Konstantin Mikhaylov, Kaibin Huang, Sergey Andreev, and Yevgeni Koucheryavy. Wirelessly powered urban crowd sensing over wearables: Trading energy for data. *IEEE Wireless Communications*, 25(2):140–149, 2018.

[30] Saurabh Ganeriwal, Laura K Balzano, and Mani B Srivastava. Reputation-based framework for high integrity sensor networks. *ACM Transactions on Sensor Networks (TOSN)*, 4(3):15, 2008.

[31] Raghu K Ganti, Fan Ye, and Hui Lei. Mobile crowdsensing: current state and future challenges. *IEEE Communications Magazine*, 49(11):32–39, 2011.

[32] Hugo Gascon, Sebastian Uellenbeck, Christopher Wolf, and Konrad Rieck. Continuous authentication on mobile devices by analysis of typing motion behavior. *Sicherheit 2014–Sicherheit, Schutz und Zuverlässigkeit*, 2014.

[33] Bugra Gedik and Ling Liu. Location privacy in mobile systems: A personalized anonymization model. In *25th IEEE International Conference on Distributed Computing Systems (ICDCS'05)*, pages 620–629. IEEE, 2005.

[34] Stylianos Gisdakis, Thanassis Giannetsos, and Panagiotis Papadimitratos. Security, privacy, and incentive provision for mobile crowd sensing systems. *IEEE Internet of Things Journal*, 3(5):839–853, 2016.

[35] Vipul Goyal, Omkant Pandey, Amit Sahai, and Brent Waters. Attribute-based encryption for fine-grained access control of encrypted data. In *Proceedings of the 13th ACM conference on Computer and communications security*, pages 89–98. Acm, 2006.

[36] Bin Guo, Yan Liu, Wenle Wu, Zhiwen Yu, and Qi Han. Activecrowd: A framework for optimized multitask allocation in mobile crowdsensing systems. *IEEE Transactions on Human-Machine Systems*, 47(3):392–403, 2017.

[37] Bin Guo, Zhiwen Yu, Liming Chen, Xingshe Zhou, and Xiaojuan Ma. Mobigroup: Enabling lifecycle support to social activity organization and suggestion with mobile crowd sensing. *IEEE Transactions on Human-Machine Systems*, 46(3):390–402, 2015.

[38] Bin Guo, Zhiwen Yu, Daqing Zhang, and Xingshe Zhou. Cross-community sensing and mining. *IEEE Communications Magazine*, 52(8):144–152, 2014.

[39] Bin Guo, Zhiwen Yu, Xingshe Zhou, and Daqing Zhang. From participatory sensing to mobile crowd sensing. In *2014 IEEE International Conference on Pervasive Computing and Communication Workshops (PERCOM WORKSHOPS)*, pages 593–598. IEEE, 2014.

[40] Daojing He, Sammy Chan, and Mohsen Guizani. User privacy and data trustworthiness in mobile crowd sensing. *IEEE Wireless Communications*, 22(1):28–34, 2015.

[41] Shibo He, Dong-Hoon Shin, Junshan Zhang, and Jiming Chen. Toward optimal allocation of location dependent tasks in crowdsensing. In *IEEE INFOCOM 2014-IEEE Conference on Computer Communications*, pages 745–753. IEEE, 2014.

[42] Baik Hoh, Marco Gruteser, Hui Xiong, and Ansaf Alrabady. Enhancing security and privacy in traffic-monitoring systems. *IEEE Pervasive Computing*, 5(4):38–46, 2006.

[43] Jia Hu, Hui Lin, Xuancheng Guo, and Ji Yang. Dtcs: An integrated strategy for enhancing data trustworthiness in mobile crowdsourcing. *IEEE Internet of Things Journal*, 5(6):4663–4671, 2018.

[44] Chao Huang and Dong Wang. Spatial-temporal aware truth finding in big data social sensing applications. In *2015 IEEE Trustcom/BigDataSE/ISPA*, volume 2, pages 72–79. IEEE, 2015.

[45] Kuan Lun Huang, Salil S Kanhere, and Wen Hu. Are you contributing trustworthy data?: the case for a reputation system in participatory sensing. In *Proceedings of the 13th ACM international conference on Modeling, analysis, and simulation of wireless and mobile systems*, pages 14–22. ACM, 2010.

[46] Panagiotis G Ipeirotis, Foster Provost, and Jing Wang. Quality management on amazon mechanical turk. In *Proceedings of the ACM SIGKDD workshop on human computation*, pages 64–67. ACM, 2010.

[47] Roslan Ismail, Colin Boyd, Audun Jøsang, and Selywn Russel. Strong privacy in reputation systems. In *Proceedings of the 4th International Workshop on Information Security Applications (WISA)*, 2003.

[48] Luis G Jaimes and Juan M Calderon. Location-based social networks data for mobile crowdsensing. In *2018 IEEE 8th Annual Computing and Communication Workshop and Conference (CCWC)*, pages 690–694. IEEE, 2018.

[49] Luis G Jaimes, Idalides J Vergara-Laurens, and Andrew Raij. A survey of incentive techniques for mobile crowd sensing. *IEEE Internet of Things Journal*, 2(5):370–380, 2015.

[50] Haiming Jin, Hongpeng Guo, Lu Su, Klara Nahrstedt, and Xinbing Wang. Dynamic task pricing in multi-requester mobile crowd sensing with markov correlated equilibrium. In *IEEE INFOCOM 2019-IEEE Conference on Computer Communications*, pages 1063–1071. IEEE, 2019.

[51] Haiming Jin, Lu Su, Danyang Chen, Klara Nahrstedt, and Jinhui Xu. Quality of information aware incentive mechanisms for mobile crowd sensing systems. In *Proceedings of the 16th ACM International Symposium on Mobile Ad Hoc Networking and Computing*, pages 167–176. ACM, 2015.

[52] Haiming Jin, Lu Su, and Klara Nahrstedt. Theseus: Incentivizing truth discovery in mobile crowd sensing systems. In *Proceedings of the 18th ACM International Symposium on Mobile Ad Hoc Networking and Computing*, page 1. ACM, 2017.

[53] Prajakta Joglekar and Vrushali Kulkarni. Privacy issues in urban computing using mobile crowdsensing. *International Journal of Computer Applications*, 168(3), 2017.

[54] Lampros A Kalogiros, Kostas Lagouvardos, Sotiris Nikoletseas, Nikos Papadopoulos, and Pantelis Tzamalis. Allergymap: A hybrid mhealth mobile crowdsensing system for allergic diseases epidemiology: a multidisciplinary case study. In *2018 IEEE International Conference on Pervasive Computing and Communications Workshops (PerCom Workshops)*, pages 597–602. IEEE, 2018.

[55] Burak Kantarci, Kevin G Carr, and Connor D Pearsall. Sonata: Social network assisted trustworthiness assurance in smart city crowdsensing. *International Journal of Distributed Systems and Technologies (IJDST)*, 7(1):59–78, 2016.

[56] Burak Kantarci and Hussein T Mouftah. Mobility-aware trustworthy crowdsourcing in cloud-centric internet of things. In *2014 IEEE Symposium on Computers and Communications (ISCC)*, pages 1–6. IEEE, 2014.

[57] Burak Kantarci and Hussein T Mouftah. Trustworthy sensing for public safety in cloud-centric internet of things. *IEEE Internet of Things Journal*, 1(4):360–368, 2014.

[58] Apu Kapadia, Nikos Triandopoulos, Cory Cornelius, Daniel Peebles, and David Kotz. Anonysense: Opportunistic and privacy-preserving context collection. In *International Conference on Pervasive Computing*, pages 280–297. Springer, 2008.

[59] Merkourios Karaliopoulos, Iordanis Koutsopoulos, and Michalis Titsias. First learn then earn: Optimizing mobile crowdsensing campaigns through data-driven user

profiling. In *Proceedings of the 17th ACM International Symposium on Mobile Ad Hoc Networking and Computing*, pages 271–280. ACM, 2016.

[60] Sunyoung Kim and Eric Paulos. Inair: sharing indoor air quality measurements and visualizations. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, pages 1861–1870. ACM, 2010.

[61] Michael Kinateder and Siani Pearson. A privacy-enhanced peer-to-peer reputation system. In *International Conference on Electronic Commerce and Web Technologies*, pages 206–215. Springer, 2003.

[62] Ioannis Krontiris, Marc Langheinrich, and Katie Shilton. Trust and privacy in mobile experience sharing: future challenges and avenues for research. *IEEE Communications Magazine*, 52(8):50–55, 2014.

[63] Nicholas D Lane. Community-aware smartphone sensing systems. *IEEE Internet Computing*, 16(3):60–64, 2012.

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

[65] Ting Li, Taeho Jung, Zhijin Qiu, Hanshang Li, Lijuan Cao, and Yu Wang. Scalable privacy-preserving participant selection for mobile crowdsensing systems: Participant grouping and secure group bidding. *IEEE Transactions on Network Science and Engineering*, 2018.

[66] Xiaohui Li and Qi Zhu. Social incentive mechanism based multi-user sensing time optimization in co-operative spectrum sensing with mobile crowd sensing. *Sensors*, 18(1):250, 2018.

[67] Guangchun Luo, Ke Yan, Xu Zheng, Ling Tian, and Zhipeng Cai. Preserving adjustable path privacy for task acquisition in mobile crowdsensing systems. *Information Sciences*, 2018.

[68] Tie Luo, Salil S Kanhere, Sajal K Das, and Hwee-Pink Tan. Incentive mechanism design for heterogeneous crowdsourcing using all-pay contests. *IEEE transactions on mobile computing*, 15(9):2234–2246, 2016.

[69] Tie Luo and Chen-Khong Tham. Fairness and social welfare in incentivizing participatory sensing. In *2012 9th Annual IEEE Communications Society Conference on*

Sensor, Mesh and Ad Hoc Communications and Networks (SECON), pages 425–433. IEEE, 2012.

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

[71] Daniel Miessler. Hp study reveals 70 percent of internet of things devices vulnerable to attack. *Retrieved June*, 30:2015, 2014.

[72] Jianbing Ni, Kuan Zhang, Qi Xia, Xiaodong Lin, and Xuemin Sherman Shen. Enabling strong privacy preservation and accurate task allocation for mobile crowdsensing. *IEEE Transactions on Mobile Computing*, 2019.

[73] Xiaoguang Niu, Qiongzan Ye, Yihao Zhang, and Dengpan Ye. A privacy-preserving identification mechanism for mobile sensing systems. *IEEE Access*, 6:15457–15467, 2018.

[74] Nsikak P Owoh and M Mahinderjit Singh. Security analysis of mobile crowd sensing applications. *Applied Computing and Informatics*, 2018.

[75] Alex Page, Shurouq Hijazi, Dogan Askan, Burak Kantarci, and Tolga Soyata. Research directions in cloud-based decision support systems for health monitoring using internet-of-things driven data acquisition. *Int. J. Services Comput.*, 4(4):18–34, 2016.

[76] Ivana Podnar Zarko, Aleksandar Antonic, and Krešimir Pripužic. Publish/subscribe middleware for energy-efficient mobile crowdsensing. In *Proceedings of the 2013 ACM conference on Pervasive and ubiquitous computing adjunct publication*, pages 1099–1110. ACM, 2013.

[77] Layla Pournajaf, Daniel A. Garcia-Ulloa, Li Xiong, and Vaidy Sunderam. Participant privacy in mobile crowd sensing task management: A survey of methods and challenges. *SIGMOD Rec.*, 44(4):23–34, May 2016.

[78] Layla Pournajaf, Li Xiong, Daniel A Garcia-Ulloa, and Vaidy Sunderam. A survey on privacy in mobile crowd sensing task management. *Dept. Math. Comput. Sci., Emory Univ., Atlanta, GA, USA, Tech. Rep. TR-2014-002*, 2014.

[79] Evangelos Pournaras, Izabela Moise, and Dirk Helbing. Privacy-preserving ubiquitous social mining via modular and compositional virtual sensors. In *2015 IEEE 29th International Conference on Advanced Information Networking and Applications*, pages 332–338. IEEE, 2015.

[80] Maryam Pouryazdan and Burak Kantarci. Ta-crocs: Trustworthiness-aware coalitional recruitment of crowd-sensors. In *2018 IEEE Global Communications Conference (GLOBECOM)*, pages 1–7. IEEE, 2018.

[81] Maryam Pouryazdan, Burak Kantarci, Tolga Soyata, Luca Foschini, and Houbing Song. Quantifying user reputation scores, data trustworthiness, and user incentives in mobile crowd-sensing. *IEEE Access*, 5:1382–1397, 2017.

[82] Maryam Pouryazdan, Burak Kantarci, Tolga Soyata, and Houbing Song. Anchor-assisted and vote-based trustworthiness assurance in smart city crowdsensing. *IEEE Access*, 4:529–541, 2016.

[83] Fudong Qiu, Fan Wu, and Guihai Chen. Privacy and quality preserving multimedia data aggregation for participatory sensing systems. *IEEE Transactions on Mobile Computing*, 14(6):1287–1300, 2015.

[84] Rajib Kumar Rana, Chun Tung Chou, Salil S Kanhere, Nirupama Bulusu, and Wen Hu. Ear-phone: an end-to-end participatory urban noise mapping system. In *Proceedings of the 9th ACM/IEEE international conference on information processing in sensor networks*, pages 105–116. ACM, 2010.

[85] Sasank Reddy, Andrew Parker, Josh Hyman, Jeff Burke, Deborah Estrin, and Mark Hansen. Image browsing, processing, and clustering for participatory sensing: lessons from a dietsense prototype. In *Proceedings of the 4th workshop on Embedded networked sensors*, pages 13–17. ACM, 2007.

[86] Ju Ren, Yaoxue Zhang, Kuan Zhang, and Xuemin Sherman Shen. Sacrm: social aware crowdsourcing with reputation management in mobile sensing. *Computer Communications*, 65:55–65, 2015.

[87] Francesco Restuccia, Nirnay Ghosh, Shameek Bhattacharjee, Sajal K Das, and Tommaso Melodia. Quality of information in mobile crowdsensing: Survey and research challenges. *ACM Transactions on Sensor Networks (TOSN)*, 13(4):34, 2017.

[88] Haggai Roitman, Jonathan Mamou, Sameep Mehta, Aharon Satt, and LV Subramaniam. Harnessing the crowds for smart city sensing. In *Proceedings of the 1st international workshop on Multimodal crowd sensing*, pages 17–18. ACM, 2012.

[89] Johannes Schobel, Rüdiger Pryss, and Manfred Reichert. Using smart mobile devices for collecting structured data in clinical trials: Results from a large-scale case study.

In *2015 IEEE 28th International Symposium on Computer-Based Medical Systems*, pages 13–18. IEEE, 2015.

[90] Adi Shamir. How to share a secret. *Communications of the ACM*, 22(11):612–613, 1979.

[91] Venkat SuryaDasari, Maryam Pouryazdan, and Burak Kantarci. On the impact of selective data acquisition in mobile crowd-sensing performance. In *2018 IEEE Canadian Conference on Electrical & Computer Engineering (CCECE)*, pages 1–4. IEEE, 2018.

[92] Astarita Vittorio, Vaiana Rosolino, Iuele Teresa, Caruso Maria Vittoria, P Giofrè Vincenzo, et al. Automated sensing system for monitoring of road surface quality by mobile devices. *Procedia-Social and Behavioral Sciences*, 111:242–251, 2014.

[93] Marco Voss, Andreas Heinemann, and Max Muhlhauser. A privacy preserving reputation system for mobile information dissemination networks. In *First International Conference on Security and Privacy for Emerging Areas in Communications Networks (SECURECOMM'05)*, pages 171–181. IEEE, 2005.

[94] Jiangtao Wang, Feng Wang, Yasha Wang, Leye Wang, Zhaopeng Qiu, Daqing Zhang, Bin Guo, and Qin Lv. Hytasker: Hybrid task allocation in mobile crowd sensing. *IEEE Transactions on Mobile Computing*, 2019.

[95] Jiangtao Wang, Feng Wang, Yasha Wang, Daqing Zhang, Brian Y Lim, and Leye Wang. Allocating heterogeneous tasks in participatory sensing with diverse participant-side factors. *IEEE Transactions on Mobile Computing*, 2018.

[96] Leye Wang, Dingqi Yang, Xiao Han, Tianben Wang, Daqing Zhang, and Xiaojuan Ma. Location privacy-preserving task allocation for mobile crowdsensing with differential geo-obfuscation. In *Proceedings of the 26th International Conference on World Wide Web*, pages 627–636. International World Wide Web Conferences Steering Committee, 2017.

[97] Leye Wang, Daqing Zhang, Animesh Pathak, Chao Chen, Haoyi Xiong, Dingqi Yang, and Yasha Wang. Ccs-ta: quality-guaranteed online task allocation in compressive crowdsensing. In *Proceedings of the 2015 ACM international joint conference on pervasive and ubiquitous computing*, pages 683–694. ACM, 2015.

[98] Leye Wang, Daqing Zhang, and Haoyi Xiong. effsense: energy-efficient and cost-effective data uploading in mobile crowdsensing. In *Proceedings of the 2013 ACM*

conference on Pervasive and ubiquitous computing adjunct publication, pages 1075–1086. ACM, 2013.

[99] Shiguang Wang, Dong Wang, Lu Su, Lance Kaplan, and Tarek F Abdelzaher. Towards cyber-physical systems in social spaces: The data reliability challenge. In *2014 IEEE Real-Time Systems Symposium*, pages 74–85. IEEE, 2014.

[100] Wendong Wang, Hui Gao, Chi Harold Liu, and Kin K Leung. Credible and energy-aware participant selection with limited task budget for mobile crowd sensing. *Ad Hoc Networks*, 43:56–70, 2016.

[101] Xiaojie Wang, Zhaolong Ning, Xiping Hu, Edith C-H Ngai, Lei Wang, Bin Hu, and Ricky YK Kwok. A city-wide real-time traffic management system: Enabling crowd-sensing in social internet of vehicles. *IEEE Communications Magazine*, 56(9):19–25, 2018.

[102] Xinlei Wang, Kannan Govindan, and Prasant Mohapatra. Collusion-resilient quality of information evaluation based on information provenance. In *2011 8th Annual IEEE Communications Society Conference on Sensor, Mesh and Ad Hoc Communications and Networks*, pages 395–403. IEEE, 2011.

[103] Zhibo Wang, Jiahui Hu, Qian Wang, Ruizhao Lv, Jian Wei, Honglong Chen, and Xiaoguang Niu. Task-bundling-based incentive for location-dependent mobile crowd-sourcing. *IEEE Communications Magazine*, 57(2):54–59, 2019.

[104] Yutian Wen, Jinyu Shi, Qi Zhang, Xiaohua Tian, Zhengyong Huang, Hui Yu, Yu Cheng, and Xuemin Shen. Quality-driven auction-based incentive mechanism for mobile crowd sensing. *IEEE Transactions on Vehicular Technology*, 64(9):4203–4214, 2015.

[105] Dapeng Wu, Lei Fan, Chenlu Zhang, Honggang Wang, and Ruyan Wang. Dynamical credibility assessment of privacy-preserving strategy for opportunistic mobile crowd sensing. *IEEE Access*, 6:37430–37443, 2018.

[106] Guanlin Wu, Junjie Chen, Weidong Bao, Xiaomin Zhu, Wenhua Xiao, and Ji Wang. Towards collaborative storage scheduling using alternating direction method of multipliers for mobile edge cloud. *Journal of Systems and Software*, 134:29–43, 2017.

[107] Hai-Qin Wu, Liangmin Wang, and Guoliang Xue. Privacy-aware task allocation and data aggregation in fog-assisted spatial crowdsourcing. *IEEE Transactions on Network Science and Engineering*, 2019.

[108] Yu Xiao, Pieter Simoens, Padmanabhan Pillai, Kiryong Ha, and Mahadev Satyanarayanan. Lowering the barriers to large-scale mobile crowdsensing. In *Proceedings of the 14th Workshop on Mobile Computing Systems and Applications*, page 9. ACM, 2013.

[109] Haoyi Xiong, Daqing Zhang, Guanling Chen, Leye Wang, Vincent Gauthier, and Laura E Barnes. icrowd: Near-optimal task allocation for piggyback crowdsensing. *IEEE Transactions on Mobile Computing*, 15(8):2010–2022, 2016.

[110] Ke Yan, Guangchun Luo, Xu Zheng, Ling Tian, and Akshita Maradapu Vera Venkata Sai. A comprehensive location-privacy-awareness task selection mechanism in mobile crowd-sensing. *IEEE Access*, 7:77541–77554, 2019.

[111] D. Yang, G. Xue, X. Fang, and J. Tang. Incentive mechanisms for crowdsensing: Crowdsourcing with smartphones. *IEEE/ACM Transactions on Networking*, PP(99):1–13, 2015.

[112] Dejun Yang, Guoliang Xue, Xi Fang, and Jian Tang. Crowdsourcing to smartphones: Incentive mechanism design for mobile phone sensing. In *Proceedings of the 18th annual international conference on Mobile computing and networking*, pages 173–184. ACM, 2012.

[113] Kan Yang, Kuan Zhang, Ju Ren, and Xuemin Shen. Security and privacy in mobile crowdsourcing networks: challenges and opportunities. *IEEE communications magazine*, 53(8):75–81, 2015.

[114] S. Yang, F. Wu, S. Tang, X. Gao, B. Yang, and G. Chen. On designing data quality-aware truth estimation and surplus sharing method for mobile crowdsensing. *IEEE Journal on Selected Areas in Communications*, 35(4):832–847, April 2017.

[115] Shuo Yang, Fan Wu, Shaojie Tang, Xiaofeng Gao, Bo Yang, and Guihai Chen. On designing data quality-aware truth estimation and surplus sharing method for mobile crowdsensing. *IEEE Journal on Selected Areas in Communications*, 35(4):832–847, 2017.

[116] Hui Zang and Jean Bolot. Anonymization of location data does not work: A large-scale measurement study. In *Proceedings of the 17th annual international conference on Mobile computing and networking*, pages 145–156. ACM, 2011.

[117] Yufeng Zhan, Yuanqing Xia, and Jinhui Zhang. Incentive mechanism in platform-centric mobile crowdsensing: A one-to-many bargaining approach. *Computer Networks*, 132:40–52, 2018.

[118] Bo Zhang, Chi Harold Liu, Jianyu Lu, Zheng Song, Ziyu Ren, Jian Ma, and Wendong Wang. Privacy-preserving qoi-aware participant coordination for mobile crowdsourcing. *Computer Networks*, 101:29–41, 2016.

[119] Maotian Zhang, Panlong Yang, Chang Tian, Shaojie Tang, Xiaofeng Gao, Baowei Wang, and Fu Xiao. Quality-aware sensing coverage in budget-constrained mobile crowdsensing networks. *IEEE Transactions on Vehicular Technology*, 65(9):7698–7707, 2016.

[120] Xinglin Zhang, Zheng Yang, Yunhao Liu, Jianqiang Li, and Zhong Ming. Toward efficient mechanisms for mobile crowdsensing. *IEEE Transactions on Vehicular Technology*, 66(2):1760–1771, 2016.

[121] Y. Zhang, N. Meratnia, and P. Havinga. Outlier detection techniques for wireless sensor networks: A survey. *IEEE Communications Surveys Tutorials*, 12(2):159–170, Second 2010.

[122] Yu Zhang and Mihaela van der Schaar. Reputation-based incentive protocols in crowdsourcing applications. In *2012 Proceedings IEEE INFOCOM*, pages 2140–2148. IEEE, 2012.

[123] Yueqian Zhang, Murat Simsek, and Burak Kantarci. Machine learning-based prevention of battery-oriented illegitimate task injection in mobile crowdsensing. In *Proceedings of the ACM Workshop on Wireless Security and Machine Learning*, pages 31–36. ACM, 2019.

[124] Sheng Zhong, Hong Zhong, Xinyi Huang, Panlong Yang, Jin Shi, Lei Xie, and Kun Wang. Connecting human to cyber-world: Security and privacy issues in mobile crowdsourcing networks. In *Security and Privacy for Next-Generation Wireless Networks*, pages 65–100. Springer, 2019.

[125] Pan Zhou, Wenbo Chen, Shouling Ji, Hao Jiang, Li Yu, and Dapeng Wu. Privacy-preserving online task allocation in edge-computing-enabled massive crowdsensing. *IEEE Internet of Things Journal*, 2019.

[126] Tongqing Zhou, Zhiping Cai, Kui Wu, Yueyue Chen, and Ming Xu. Fidc: A framework for improving data credibility in mobile crowdsensing. *Computer Networks*, 120:157–169, 2017.

[127] Christoph Ziegler and E von Zezschwitz. Implicit authentication 2.0: Behavioural biometrics in smart environments. In *Human Computer Interaction in the Internet of Things Era*, pages 100–107. Citeseer, 2015.

APPENDICES

Through out the thesis, we run each simulation setting under 10 seeds and we ensure the reliability of results by calculating marginal error at 95% confidence interval as shown in 6.1 where $\bar{x}$ is mean and SD denotes standard deviation and N denotes number of samples.

$$error = \bar{x} + \frac{t * SD}{\sqrt{N}} \tag{6.1}$$

We acknowledge that entire results presented in the thesis are influenced by random number generators that are used to choose the location of user, tasks and also other initial conditions which we tried to reduce the impact of randomness by choosing 10 different seeds. Since the total sample size is less than 30, t distribution table is used instead of z-table. We plotted the marginal error bars for crucial metrics that determine the efficiency of overall system such as platform utility, user utility and user comfort. Th remaining performance metrics support and justify the obtained overall efficiency of system and trustworthiness which are related to the user, platform utility and user comfort.