TRUTH DISCOVERY FOR MOBILE WORKERS IN EDGE-ASSISTED MOBILE CROWDSENSING

BY

SYED AMIR ALI SHAH

NATIONAL UNIVERSITY OF MODERN LANGUAGES

ISLAMABAD

March 2023

TRUTH DISCOVERY FOR MOBILE WORKERS IN EDGE-ASSISTED MOBILE CROWDSENSING

By
SYED AMIR ALI SHAH
BSCS, The University of Agriculture, Peshawar, 2020

A THESIS SUBMITTED IN PARTIAL FULFILMENT OF
THE REQUIREMENTS FOR THE DEGREE OF

MASTER OF SCIENCE

In **Computer Science**

To

FACULTY OF ENGINEERING & COMPUTER SCIENCE

NATIONAL UNIVERSITY OF MODERN LANGUAGES ISLAMABAD

© Syed Amir Ali Shah, 2023

NATIONAL UNIVERSITY OF MODERN LANGUAGES FACULTY OF ENGINEERIING & COMPUTER SCIENCE

THESIS AND DEFENSE APPROVAL FORM

The undersigned certify that they have read the following thesis, examined the defense, are satisfied with overall exam performance, and recommend the thesis to the Faculty of Engineering and Computer Sciences for acceptance.

Thesis Title: Truth Discovery for Mobile Workers in Edge-Assisted Mobile Crowdsensing

Submitted By: Syed Amir Ali Shah **Registration #:** 42 MS/CS/F20

Master of Science in Computer Science (MSCS)
Title of the Degree

Computer Science
Name of Discipline

Dr. Ata Ullah
Name of Research Supervisor

Signature of Research Supervisor

Dr. Muhammad Noman Malik
Name of Dean (FE&CS)

Signature of Dean (FE&CS)

Brig. Syed Nadir Ali
Name of Pro-Rector Resources/Director General

Signature of Pro-Rector Resources/DG

15th March, 2023

AUTHOR'S DECLARATION

I Syed Amir Ali Shah

Son of Syed Jahangir Shah

Registration # 42 MS/CS/F20

Discipline Computer Science

Candidate of **Master of Science in Computer Science (MSCS)** at the National University of Modern Languages do hereby declare that the thesis **Truth Discovery for Mobile Workers In Edge-Assisted Mobile Crowdsensing** submitted by me in partial fulfillment of my MSCS degree, is my original work, and has not been submitted or published earlier. I also solemnly declare that it shall not, in the future, be submitted by me for obtaining any other degree from this or any other university or institution. I also understand that if evidence of plagiarism is found in my thesis/dissertation at any stage, even after the award of a degree, the work may be canceled and the degree revoked.

Signature of Candidate

Syed Amir Ali Shah

Name of Candidate

15th March, 2023

Date

ABSTRACT

Title: Truth Discovery for Mobile Workers in Edge-Assisted Mobile Crowdsensing

The proliferation of mobile phones has led to the rise of mobile crowdsensing systems. However, many of these systems rely on the deep cloud, which can be complex and challenging to scale. To improve the performance of crowdsensing at the edge cloud, truth-discovery methods are commonly employed. These methods typically involve updating either the truth or the weight associated with a user's task. While some edge cloud-based crowdsensing systems exist, they do not provide incentives to users based on their experience. In this report, we present a new approach to truth discovery and incentive-giving that considers both the user's experience and the accuracy of their submitted data. Our modified truth-discovery algorithm updates both the weight and truth concurrently, with greater incentives offered to users who have completed more tasks and whose submitted data is close to the estimated truth. We have conducted simulations to demonstrate the effectiveness of our proposed solution in improving the incentive mechanism for experienced users.

Keywords: incentive mechanism, mobile crowdsensing, truth discovery,

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATIONS

MCS	-	Mobile Crowdsensing
VANET	-	Vehicular ad-hoc networks
EPTD	-	Efficient and Privacy-preserving Truth Discovery
L-PTD	-	Lightweight Privacy-Preserving Truth Discovery
PPATD	-	Practical and Privacy-Aware Truth Discovery
CATD	-	Confidence-Aware Truth Discovery
MEC	-	Mobile Edge Computing

LIST OF SYMBOLS

u_i	-	Utility of user i
p_i	-	Payment of user i
c_i	-	Cost of user i
μ	-	Type of task
x_i^j	-	Observed values
x_*^j	-	Estimated truth
φ	-	Maximum number of iterations for truth discovery
T, t_j, T_i	-	Task set, task j, task set of user i

ACKNOWLEDGMENT

First of all, I wish to express my gratitude and deep appreciation to Almighty Allah, who made this study possible and successful. This study would not be accomplished unless the honest espousal was extended from several sources for which I would like to express my sincere thankfulness and gratitude. Yet, there were significant contributors to my attained success and I cannot forget their input, especially my research supervisor, Associate Prof. Dr. Ata Ullah, and Co-supervisor Asst. Prof. Dr. Fazli Subhan did not leave any stone unturned to guide me during my research journey.

I shall also acknowledge the extended assistance from the administration of the Department of Computer Sciences who supported me through my research experience and simplified the challenges I faced. For all whom I did not mention but I shall not neglect their significant contribution, thanks for everything.

DEDICATION

This thesis work is dedicated to my parents and my teachers throughout my education career who have not only loved me unconditionally but whose good examples have taught me to work hard for the things that I aspire to achieve.

CHAPTER 1

INTRODUCTION

1.1 Introduction:

These days everyone has a mobile phone. Several mobile phones have in-build sensing capabilities. Crowdsensing or Mobile Crowdsensing is a technique where a large group of people having smartphones capable of sensing and computing (tablet computers, wearables) shares data in their groups to extract the information for measuring maps, and analyzing an estimate of any task of mutual benefit. Mobile crowdsensing has great capability to acquire data in large-scale sensing applications. Mobile crowdsensing has three main types, environmental (such as monitoring pollution), infrastructure (such as locating potholes), and social (such as tracking exercise data within the community). First, the Crowdsourcer/task publisher publishes the task then the task is assigned to any crowdsensing platform, the crowdsensing platform assigns the task to participants/clients, and the platform assigns tasks to clients according to their interests. Participants complete the task and after completing the task they give the sensed data back to the platform. At last, the platform returns the data to the crowdsourcer after examining the data. This whole process is shown in figure 1.1.

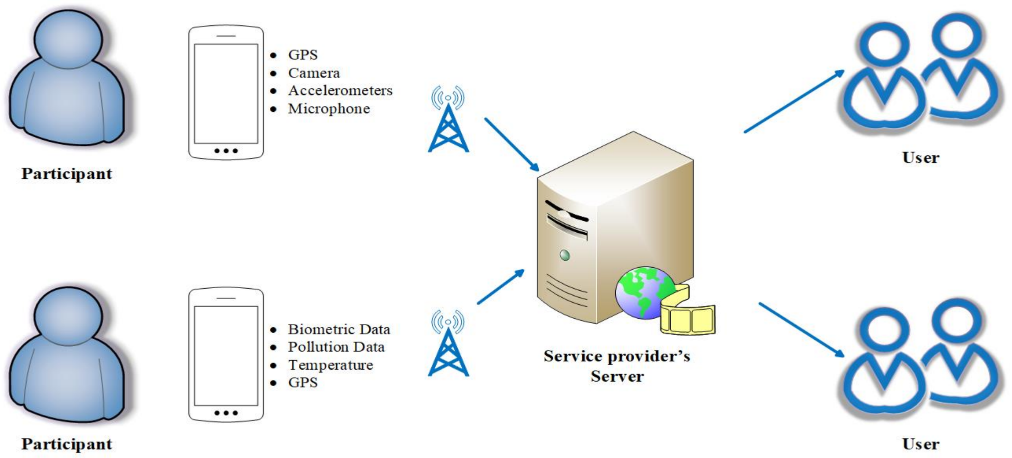

Figure 1.1: Mobile Crowdsensing

In [1] the authors discuss the incentive mechanism, which is divided into two stages, i.e., Truth discovery and cost-effective inverse auction. A family of algorithms known as "truth-discovery" has been proposed and extensively researched to extract meaningful facts from flawed data. The weights of the workers and the aggregated results are calculated together by a truth discovery algorithm, which does so without any prior knowledge of the workers' dependability. It does so on the basis that the workers whose data are more similar to the aggregated results will be given higher weights, and the data from a worker with a higher weight will be counted more in the clusters. In truth discovery, authors use an algorithm that sums the difference between each user sensing data and the estimated truth. After finding the difference, the difference for all users is found and divided by the margin in the sensing data of each user. In [2] the authors propose vehicle-assisted multi-access edge computing (VMEC) in which the tasks would be uploaded to Multi-access Edge Computing (MEC) server and Vehicle Edge Nodes (VENs) for offloading strategy. In [3] authors proposed the incentive mechanism for a multiregional scenario. In which the complete region is divided into sub-regions to accommodate users in the current large region. After this, the task is divided among all users and the result at the end is combined through an algorithm. In [4] authors proposed a security mechanism for transferring messages in Vehicular networks. In our proposed solution the algorithm for truth discovery is enhanced by considering the experience of the user and also the sensing data that is provided by the user. In multiple solutions, the authors even didn’t mention those users that have done some sensing work in the past. To collect sensing data, assign tasks, assess the truth, and then motivate mobile users, several mobile crowdsensing systems collaborate with cloud services. The truth discovery in [5] is only based on the user’s sensing data, if the sensing data is near to the estimated truth, then the user may be considered the winner of the sensing data. The authors didn’t discuss users with experience in this field, i.e., done some tasks relating to sensing work.

Versatile social event seeing is another perspective that takes advantage of unavoidable phones to conveniently accumulate data, engaging different large-scale applications. Human solidification is perhaps the standard part, and human versatility offers wonderful passageways for both distinctive ideas and data transmission. In this article, we research the deft properties of human convey ability as shown by the perspectives of both recognizing and transmission, and examine how to jump all over these expected opportunities to total data skillfully and in fact. What’s more arrangements of different open issues are acquired by human affiliation in this emerging appraisal area. Client-driven adaptive recognizing and choosing devices, such as mobile phones, music players, and in-vehicle sensors, are a new breed of Internet-connected electronics. These contraptions would drive the Internet of Things trend by providing sensor data to the Internet on a societal scale. In this article, we would examine a type of application known as flexible crowdsensing, in which people perceiving and managing gadgets communicate information and focus data to gauge and lead tasks of shared interest. We provide a concise design of existing adaptive crowdsensing applications, explain their exciting properties, depict numerous evaluation challenges, and offer new blueprints [6]. Finally, we are arguing against the need for coordinated planning.

In [7] the authors survey some of the frameworks related to systems or architecture and derived that some of the frameworks have no unity, as some have a framework that is different in some aspects and others are different in another aspect. The framework consists of the Pull and Push model, the pull is used to gather the list of active users and the pull selects the user who wishes to perform tasks. Another one is the centralized and distributed model used to collect data. The authors in [7] proposed a scheme consisting of the following concepts/techniques or parts.

a) A push-and-pull model
b) Distributed and centralized model for the gathering of data
c) The design of the sensing technique, task definition, and the complete task allocation process
d) The design of an incentive mechanism to motivate and engage users in participation in Mobile Crowdsensing tasks

By providing a study survey, authors decide on the MCS privacy issues and conduct a thorough literature review on them [8]. Sum up the distinctive features of MCS and examine its possible privacy hazard. Configure a sequence of requirements for privacy preservation in MCS. Due to this to identify that privacy preservation in MCS is necessary to avoid extra leakage of personal data, the author identifies that when scheming privacy-preserving schemes, besides privacy hazards, the accomplishment of responsibility, feasibility, and, effective protect privacy comprehensively, coherence is also necessary.

1.2 Application Areas:

In [2] authors discuss four application areas of Crowdsensing. Surrounding applications, armature applications, general applications, and etiquette applications. Surrounding applications consist of space forecast, wind contamination, and sound contamination. Armature applications consist of road traffic evaluation, parking lot accessibility, route conditions, and shipment locating. General applications include facility proposal, tour help, geo-localized cell phone, parking lot suggestion, individual/common well-being, and metropolitan general occasion while etiquette applications include society etiquette, juncture etiquette, correlation detection, metropolitan habits flexibility, and primary care. All these applications are discussed in detail later.

Figure 1.2: Application Areas of Crowdsensing

1.2.1 Surrounding Applications:

The basic aim of the surrounding applications of MCS is to preserve nature and look after the Wind, Space forecast, and sound contamination levels. In the last several years, most of the scientific studies have been conducted based on the mobile devices of volunteer contributors. This includes space forecast, wind contamination, and sound contamination.

1.2.2 Armature Applications:

The armature applications of MCS are related to the large-scale measurement of public infrastructures like road traffic evaluation, parking lot accessibility, condition of roads, real-time transit tracking, power line condition, and outages of public works (broken traffic lights). Armature applications comprise road traffic evaluation, parking lot accessibility, route condition, and shipment locating.

1.2.3 General Applications:

The general applications of MCS are human-powered and reveal multiple aspects like people sharing sensed information amongst themselves involving recommendations and opinions, life experience, and service/activity suggestions. These general applications involve facility proposals, tour help, geo-localized cell phones, parking lot suggestions, individual/common well-being, and metropolitan general occasions.

1.2.4 Etiquette Applications:

Another human-powered sensing application of Mobile crowd-sensing comes under the umbrella of the etiquette domain that explains various interesting points. One can get to know the crowd’s behavior, lifestyle, relationship, and healthcare. Etiquette applications contain society etiquette, juncture etiquette, correlation detection, metropolitan habits flexibility, and primary care.

In [9] authors proposed an architecture that consists of four layers to characterize works in MCS. All the work in MCS from the application layer to the physical layer passes through communication and data. This architecture can be used to differentiate between domain-specific Mobile Crowdsensing and general-purpose Mobile Crowdsensing.

1.3 Research Objectives:

Our main aim is to introduce such an incentive mechanism in which we consider both the user experience and the estimated truth of the sensing data. Objectives are as follows

1) To assess the Workers' truth discovery performance
2) To assess the user experience in the operational environment

1.4 Problem Statement:

The previously suggested incentive approach mechanism needs further enhancement that considers the users' experience. Existing approaches [1], [3], [5] only consider the sensing data if it is close to the estimated truth. At this point, researchers choose that user as the winner, which might disregard the skilled user who completed more tasks deemed appropriate for that user. The user can become discouraged and decide not to participate in any other mobile crowdsourcing studies or projects.

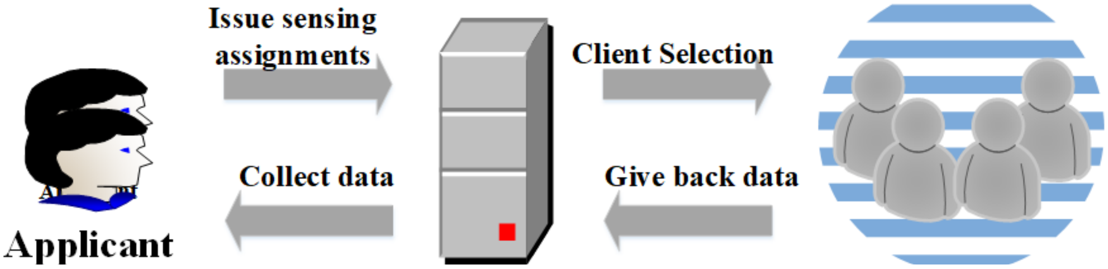

Figure 1.3: Problem Statement

.

1.5 Research Questions:

When we studied crowdsensing and when we came to the incentive mechanism, some questions arose in our minds which are as follows:

1) What evaluation methods are described in the body of existing literature for truth discovery?
2) What methods are important to obtain user experience?

1.6 Thesis Organization:

The remainder of the thesis is structured as follows: chapter 2 discusses all the literature which is somewhat in the order of the first overview of the chapter then discussed mobile crowdsensing then incentive mechanism schemes and then explains two types of incentive mechanisms after this essential of privacy preservation in mobile crowdsensing is discussed. The second last truth discovery schemes are discussed and a comparison table is created to discuss the advantages and disadvantages of the schemes. Lastly, a summary to end the second chapter. Chapter 3 discusses the overall methodology of this study. In addition, the proposed solution is also given. Finally, the evaluation metrics with the simulation framework are presented.

Chapter 4 discusses the suggested scheme and its comparison with other existing schemes with various illustrations, a flowchart, and the proposed system's algorithm. The result and discussion chapter, numbered 5, discusses the result obtained from the simulation and states how the proposed scheme differs from other existing schemes.

Chapter 6 discusses the conclusion of the thesis and proposed some future work related to mobile crowdsensing.

CHAPTER 2

LITERATURE REVIEW

2.1 Overview:

This chapter discusses literature as well as the incentive mechanism schemes. This chapter is written to explain the whole mobile crowdsensing and discusses the incentive mechanism. Also explained incentive mechanisms i.e., monetary and non-monetary incentive mechanisms. In the design of the incentive mechanism, the issues that were arising have been explained. Later, a table was created in which different methods were compared, as well as the core concept, benefit, and restriction of the scheme.

2.2 Mobile Crowdsensing:

Versatile sensors, for instance, progressed cells and vehicular developments address one more sort of geographically spread perceiving establishment that engages favorable people-driven perception According to a figure for everyday PDA shipments from 2010 to 2017, a greater number of noticeable numbers than 1.5 billion phones are depended on to be ignored all in the world. Progressed cells presently have a couple of sensors: camera, recipient, GPS, accelerometer, electronic navigator, light-dependent resistor, and Bluetooth as district sensing elements and they are soon to be integrated with flourishing and contaminating seeing sensors. Vehicle advancements now number in the hundreds of sensors, and propulsion vehicles are equipped with new types of sensors like radar and cameras [10]. Stood separated from the little, energy-obliged sensors of static sensor affiliations progressed cells and vehicular developments can stay aware of more stupefied estimations, have essential memory and cut-off, and thought direct agree to the Internet. In this way, versatile people-driven seeing can be an adaptable and savvy choice rather than conveying static far away sensor networks for thick unmistakable wire across beast areas [11]. Progressed cells have truly associated a huge load of accommodating seeing applications in gaming, sagacious circumstances, discernment, emergency response, and social affiliations. Shockingly, the development statement through adaptable perceiving and wearable sensors enjoys incited different clinical benefits applications, for instance, thriving seeing, senior thought help, and mental assistance. The broadening perceiving limits of PDAs have gone past the sensor affiliations' idea of regular development investigating where the carriers of identifying devices are currently people, the suppliers, and customers of recognized activities. Flexible people-driven perception has two major drawbacks despite its advantages.: (i) helping individuals and (ii) the faithful nature of the recognized data.

It has been suggested that the standard problem can be resolved by attending unimportant social events. A welcoming social gathering, seeing the stage recognizes a similar role to that performed by Amazon's Mechanical Turk in unrestrictedly supporting: It enables organizations and individuals (clients) to connect with a large group of people (suppliers) who are willing to perform major distinct tasks for that users are compensated. Apart from MTurk activities, which are performed on Desktop computers and consistently require human intervention, accommodating unmistakable endeavors are implemented on PDAs that meet unambiguous setting/seeing necessities (e.g., region, time, expresses sensors) and frequently do not require human intervention (i.e., re-tried perceiving tasks) [12]. Various affiliations and individuals could go probably as get-together perceiving customers. Local, state, as well as government organizations, for example, could greatly benefit from this new viewing establishment as they progress toward massive data from this current reality. Business affiliations may be astoundingly fiery about party versatility seeing data to get to acknowledge clients direct.

Experts in different areas of science and orchestrating could gather a great deal of perceived data for various evaluations. At last, we generally could go about as clients through various adaptable applications. Adaptable Crowd Sensing Applications In the going with, we present a couple of utilization spaces that can benefit from adaptable social affair recognition as well as different applications (some of them now prototyped) for each area: Intelligent Cities: Around the world, metropolitan associations with high people density and a massive number of interconnected issues make the reasonable city board a risky endeavor. Thusly, a couple of huge government and present-day assessment endeavors are in the works to exploit the most incredible limitation of the perceiving data by beginning dazzling city systems to likewise develop city support by conveying more quick cross fragments, water the trailblazer’s structures, and in the end the social advancement. From one side of the world to the next, The Songdo Business District is being constructed by South Korea's public power. It is a green, low-carbon area whose goals change in the primary full-scale declaration of a proficient city. Although these endeavors will undoubtedly have benefits, they may be expensive in large quantities. Swarm vision can reduce the expenses associated with massive expansion recognition while also providing extra modern human data. For example, our new work on Participation proposes to utilize swarm perceiving to interface with occupants in the relationship of sharp metropolitan associations; people can truly take part in distinctive endeavors to make their metropolitan locale safer and, surprisingly, more faultless. Road Mass transit: Transportation departments can collect fine grain and massive development, data about traffic plans in the country/state by using district and speed data provided by GPS sensors installed in vehicles. The organization of traffic, the improvement of new roads, etc., can then be done using these data. When comparing similar types of data collected from PDAs, drivers can obtain predictable traffic information. Drivers can in like way benefit from consistently leaving data gathered from vehicles equipped with ultrasonic sensors. To quickly repair the roads, public transportation affiliations or regions can successfully collect data on potholes using a Global positioning system and accelerometer detectors [13].

Essentially, pictures (i.e., camera sensing data) taken by individuals after and during snowfalls can be investigated in the same way to zero in there on snow cleaning and sending off. Clinical benefits and Well-being: People wearing wireless sensors for beat seeing and circulatory strain checking can give their information to the owners' PDAs. Regularly, this is done both steady and broadened length flourishing seeing of individuals [14]. Decreased seeing can utilize these current data into gigantic development clinical ideas based on constantly gathering data from various get-togethers, which can be picked pondering district, mature, etc. A specific model joins data collection from individuals who consume consistently reasonable nutritious food.

The ability of phones to confirm activity and measure peoples' levels of verifiable exercise has been shown to directly affect peoples' well-being. Based on the results of such a survey in a town, the territory may decide to build more cycling lanes to encourage individuals to participate in more proactive activities. At the most fundamental level, phones can select the degree of social correspondence for express gatherings (Using Wireless checking, a Global positioning system, or an audio sensor, for example). This paper depicts the distinct features and unique application areas of MCSC and presents a reference structure for developing human-appropriate MCSC frameworks. We go through the important regarded human and machine information in further detail and envisage the capability of vital combined human-machine structures [15]. We conclude by looking at MCSC's checks, unresolved issues, and evaluation prospects. For example, a school could track down those students (or students from express divisions) who are not helping one enough; accordingly, it could decide to figure out more gatherings close by. An equivalent part obtained in combination with data from "human sensors" can be used to monitor the spread of scourge diseases. Impelling/Advertising: Real-time region or adaptability follows/models can be used by merchants/support to zero in on unequivocal classes of people. Similarly, they can run by setting careful formats (limit of the region, time, etc.) One sale in such a summary, for instance, might inform visitors to a show about the experts they should seek out going forward. Mobile Crowd Sensing and Computing (MCSC) has evolved into a promising point of view for cross-space and massive augmentation perceiving with the growth of distant distinguishing, remote system association, and adaptable person to individual unique approaches. MCSC broadens the participatory perception vision by utilizing both disengaged participatory material knowledge from cells and client-contributed information via minimal person-to-person communication links (on the web). It also assesses the critical places and displays the combined/collaborative effort of the machine and human data in the group identification and enrolling operations.

2.3 Incentive Mechanism Schemes

As we all know the crowdsensing technique gives incentives to the participants according to the task he/she did. To give incentives to the participants there are techniques discussed by many researchers. The authors in [1] discuss the incentive mechanism, which is divided into two stages i.e., Truth discovery and cost-effective inverse auction. In truth discovery, researchers didn’t consider the user experience researchers only consider the users’ sensing truth. If one user has done some tasks in past and provides the current sensing data a bit far than the estimated truth then the user may not be considered a winner. The experience should also be considered. Before uploading, removable noise with a random variable is added to the data, and the users' secret key is used to secure sensitive information [16]. In [17], the authors proposed two tasks scenario one is to offload some burden from the network and the second one is resource allocation in Vehicle-assisted Multi-access edge computing. When the load on the server increases then it would motivate an idle vehicle to take some load off the server. In the second resource allocation, the authors use the Stackelberg game to explain the action between the MEC service provider and the user equipment UEs. [17] proposed the solution for a multi-region scenario. If one user is working in one region and the region is too large, then the region is divided into subregions and the result of each subregion is gathered into one main region. In multi-region, authors use a design that contains two algorithms i.e., weighted mean and maximin. In weighted mean maximization, the client would perform his transaction and gives incentives to all worker of micro-regions while in Minimum Maximization the client would allocate all the budget to all micro-regions and perform his tasks in all micro-regions alongside. In [6] authors proposed a security mechanism for message transfer in Vehicular networks. In this paper, authors provide a three-stage framework for Vehicular ad-hoc networks (VANET network data, Elliptic encryption, authenticated messages).

Efficient and Privacy-preserving Truth Discovery (EPTD) for mobile crowdsensing can get high accuracy and protection for both the sensor data of users and weight privacy [18]. If one user in EPTD dropped out at any stage, it would not affect the server. It is constructed on a single server setting. In [19], the authors proposed a lightweight privacy-preserving truth discovery (L-PPTD) that includes two non-colluding cloud programs and embraces extra protection of sensing data to reduce the workload on workers. Authors in [20] designed a Practical and Privacy-Aware Truth Discovery (PPATD), in which authors constructed a double security system to minimize the workload on the server as well as the employee. It is considered to be a failure if the number of online users is not up to the mark(threshold). In [21] the designer designs, Confidence-Aware Truth Discover (CATD) for encryption of the sensing data gain from unreliable sources in cloud storage, then, the data is decrypted on the requester side to make sure the data is available for the requester. In [22] authors designed the type of crowdsensing system to minimize the workload on the users and shift them to the server-side, by developing two designs for crowdsensing systems one is for a single-server setting while the second one is for a two-server to further shift most of the workload to the server. The authors used this for improved and faster results.

In [23], the authors design the type of architecture to transfer some features of Mobile Crowd Sensing (MCS) to Mobile Edge Computing (MEC). This idea is developed for reducing the threads to the privacy of the users. The MEC is established to deal with both raw sensing data and also aggregated amounts of data. In the [24] scheme authors developed the idea of the edge node. The edge node work as an assignment agent to assign work to the users and make sure the privacy protection of the participants. Privacy protection prevents an unauthorized Crowd Sensing server from accessing the users' private data. Authors used some schemes for unclear location and task allocation while considering the privacy of the user.

Crowdsourcing aggregation, a popular topic in the world of crowdsourcing, is a highly relevant field [25]. Crowdsourcing is the practice of performing certain tasks (for example, answering a series of questions) by collecting contributions from a wide number of people. One critical challenge in crowdsourcing is to consolidate the noisy answers provided by crowd workers to acquire the proper answers. Since workers may have varying levels of competence, it is critical to assess worker capabilities in the aggregation. Many crowdsourced aggregation solutions have been suggested in this direction.

Truth finding and crowdsourcing accumulation have been investigated separately and applied to several domains. These two themes, however, have a lot in common: 1) Their goals are to increase the quality of accumulation results; 2) They share the belief that dependable sources (workers) tend to produce high-quality knowledge and that information from credible sources (workers) is more likely to be correct. Differences between the two fields, on the other hand, stimulate approaches to various issues. Data generation is responsible for some significant variances. Truth discovery is generally used in online and database data integration, where the data has become available, making it a passive data-generating method.

On the contrary, crowdsourcing is proactive, and applicants have greater control over what and how data is created. Individuals frequently use their smartphones to participate in crowdsensing. This engagement has a cost for them because it consumes resources like battery, memory, bandwidth, and, on occasion, their time [26]. Furthermore, because sensed data typically comprises private location information, individuals' participation may result in revealing their position; as a result, they may be reluctant to take part in crowdsensing. As a result, user-encouraging measures should be considered. In general, prior research incentive systems can be categorized as follows:

2.3.1 Mechanisms of Monetary Incentivizing:

Participants in this category are compensated financially for their involvement in crowdsensing projects. [27] indicate that the greater the prize, the faster the activities are completed. These rewards are classified into two types: fixed and changeable. Users decide the price they want to pay for their own collected data in the changeable incentive mechanisms. This pricing strategy results in greater prices and more earnings for sellers. The system chooses whether to accept or decline the offer based on the pricing supplied by users. The program then selects users from among bids for task allocation and eliminates others. The approach for these techniques is set up in such a way that users who were eliminated in the last round have a better probability of being selected in the present round. The fixed-paying technique has already set costs for various sensing tasks, which are fixed and do not alter until the task is completed. For example, in [28], Students at a university are requested to photograph the stuff of trash cans and attach labels indicating the stuff of the trash cans to the photographs. Each person earns a specific sum in return for submitting each valid photo. The information gathered is utilized to improve the placement of trash cans.

2.3.2 Mechanisms of Non-monetary Incentivization:

This category does not provide cash or non-monetary prizes, but rather creates a positive mood. These mechanisms are further subdivided into those that amuse and those that provide social services, as follows:

Amusement: Crowdsensing activities are made into games in this subcategory, so individuals can engage in sensing while also playing the game. Such games must be suitably engaging for people to enjoy playing them.

Social Services: This type of non-monetary process comprises sensing activities to help everyone involved in crowdsensing. Contributing to pollutant sensing projects, for example, would help officials to adopt acceptable measures to regulate air pollution. In this sense, some programs [29] use air quality assessment sensors to collect data on pollution levels. This information, together with the user's location, is transferred to a database. This information would be used to create air pollution charts, which would be distributed to people who have contributed. Several systems help clients in making educated purchasing decisions. When a user who has engaged in the sensing process asks about the cost of the product, that request is transmitted to the server, along with the user's location. Following that, the server displays to the user a list of products available in nearby stores, together with their prices. The user selects her desired product after comparing prices.

2.4 Essential of Privacy Preservation in MCS

Impartiality means that no group, except the authorized one, can get or reveal the real identity of an MCS candidate. Currently, data providers give more observation to their privacy though privacy identification is considered the most important one. As impartiality is instantly related to privacy identification. It is a basic requirement to attain privacy [30]. Unattachability consult that no one can observe or notice either two connotations are from the same junction or not. Therefore, it is impractical for an intruder to detect the behavior of the junction when the unattainability requirement is fulfilled. In MCS, unattainability is required in task commencement, task release, and data compliance.

Reliability is delivered to discover disassembled or unreliable data to minimize their negative effect on the culminating task results. MCS, as a distinctive implementation methodology of IoT, needs reliability such as trust evaluation [31]. Privacy and sincerity are the fundamental requirements in the communication system. Messages convey in MCS, as well as tasks, individual details, data, etc., should be coded so that the system can hold out against listening stealthily assault. Apart from that, the system must also warranty that the notes/message gotten are identical to the original without alteration. Restriction of access consults that the assignment fulfillment, individual information, and the composed data must only be available to sanction or truthful groups on the small parallelism of an entrance strategy, which usually includes the user’s profile and additional elements.

Inherence specifies that it would not create any barrier to the standard operation of MCS even though the privacy preservation action is applied. On a sad note, some intellectuals present solutions to position privacy by using spatial-temporal cloaking techniques. Data providers are finite in ciphering(computation) ability, and battery capacity like a cellular phone. Especially, composite computation is not permitted as it affects the passion of data providers disparately even though the battery is drying out dramatically. Therefore, ciphering effectiveness is an unavoidable element when the designer designs an MCS-based protocol. Transmission cost is one more contemplation in proficiency requirements. High transmission not only raises the cost of data providers but also speeds up their power utilization. Taking into consideration that there may be quite a few mobile clients involved in the MCS scheme and a huge amount of data is transmitted, scholars are committed to reducing the reaction round amount. In short, it is notable to issue transmission proficiency privacy-preservation techniques.

2.5 Incentive Mechanism Design Issues:

After a vast study of the literature review, in [32] authors observe the succeeding design feature as obligatory conditions, for crowdsensing to be a success. Profitable Utility, Data Standard, Area Reporting, Equity, Sufficient No of Candidates, Flexible to Increased Demands, Free / Human Supervise.

For any design/plan, budget is a vital element. For programs like CS to maintain both stability and coverage, a crucial group of people is required. In any case, budgetary constraints may force such objectives. The non-monetary incentive method for CS in [32] is to attract particular user delight and hobbies to uplift their subject engagement. These kinds of incentives contain the use of games, and contests, and give in-group benefits. The idea behind these kinds of rewards is to lighten the load associated with engagement by turning it into something entertaining. These kinds of rewards allow the system to retain a large number of individuals for it to continue running. Although its development and execution require some time and knowledge of a specific domain, it is typically accomplished by specialists and experienced designers. A natural question is, how might a CrowdSensing incentive tactic improve the collection of reliable data? To answer this question the most common methods of reputation schemes are used. Normally, user position may be evaluated from previous presentations, the evaluation of peers, or by the union of both. It is complex to tackle the issue of geographical coverage of CS for mechanisms of Incentive. Consider that the goal is to calculate the temperature within a city, and the parameter of interest is temp. Buying samples from users who are evenly distributed throughout the city would be the rational decision. Addressing the problem of geographic imbalance in sample prices (i.e., lower-priced samples in some areas and too expensive in others) and their excessive use in others are among the challenges. The system only purchases the cluster with the lowest priced samples (i.e., low reportage) in the first case, and available samples are only found in certain areas of the target area in the second case.

On the other hand, regional differences in the variable's variability may exist. Regions with high variability would require more specimens to recreate the parameter, so regions with low variability would require fewer samples. Choosing the appropriate number of participants for each region to recreate the parameter in that area is challenging in both situations the regions' variance must be estimated. To achieve objectives like user retention, coverage, and financial viability, equity is essential. This is generally believed as providing a level playing field for all people involved. But in the context of variable pricing models based on inverse auctions, fairness is understood as a tactic where consumers with lower bids are more likely to be chosen than those with higher bids. This fairness (incentive approaches based on reverse auctions) may result in an imbalance in geographical coverage as well as user dropout. In the latter case, the competitors with lower bookings wages may be located in a specific region of the targeted region. The samples would always be taken from those receiving a lower reservation pay rate in the latter scenario (i.e., fairness). Nearly all of the bidders with higher amounts would give up the system after several rounds of bidding without winning.

Maintaining a reasonable level group of participants to assure quality and effectiveness is a severe issue in CS. The ultimate success of a CS framework is determined by the system's ability to retain a critical number of individuals. Variables such as test recurrence (ie, every minute or day), goal calculation the position, sort of concept to also be evaluated, the deviation of the interest variable, and detecting needs could be utilized to determine the required amount of respondents. Users chosen to respond to requests for detection are entirely dependent on incentives for their continued and active participation, though. Extensibility is defined as the extent to which a solution to a problem will work as the magnitude of the issue grows larger. In this regard, the incentive mechanism should be able to maintain the CS system's efficiency, utility, and functioning regardless of expansion packs from regional to more widespread patterns. Additional features may include the addition of new target regions along with the provision of new services. The platform must offer conceptual frameworks that encourage players to relocate from their current locations to new ones, build their confidence, and continually recruit new workers to comply with these extra criteria. Personalized and anticipated incentives should be able to draw participants with the aid of the system. In other words, it can be crucial to employ procedures that can quickly adjust to various participant expectations depending on what motivates them. The ability of the system to function autonomously in the absence of human intervention, i.e., continued user engagement in scenarios when the user acts as a static carrier of the sensors, is an equally significant design challenge. In the first case, people give up using mobile phones to freely experience their surroundings as they continue living their regular lives undisturbed. It's possible that they won't be told about when or how data collection and reporting tasks are wirelessly assigned to their mobile phones. In the latter situation, users participate in the process, such as snapping a snapshot, manually activating a sensor, or manually accepting or rejecting a request for samples.

Usually, users in MCS are heterogeneous in nature. Devices are completely responsible for associated assumptions for varied users. For example, individuals use various sorts of devices and can do different types of sensing activities [33]. Furthermore, the human aspect might have an impact on practice. Users differ in their readiness to participate, participation patterns, talents, and reputation. Due to time constraints, users may ignore the assigned sensing activities to begin engaging willingly. Users may want to execute sensing chores without changing their scheduled schedule or update sensing data without removing mobile devices from their package while using participative routines. Furthermore, the quality of skills necessary to meet the requisite ability varies across various people. These factors should be considered throughout the work allocation approach. When participants are selected for the sensing activities, participants offer a straightforward approach to obtaining pertinent data on these factors. Nonetheless, some malevolent individuals may create fraudulent assertions about their relevant content to gain additional incentives from the platform. In turn, we can acquire these details from their previous information.

The current MCS task allocation model for multitasking presupposed that activities don't truly operate independently but rather compete with one another for the resources that are shared by participants. Nonetheless, in certain cases, activities may begin to share the same sensing data or the sensing data of one task may be generated from the sensing data of another task due to temporal and spatial adjustment. For instance, two sensing jobs attempt to gather data regarding local traffic patterns at identical sensing periods. To reduce the number of users, the system in this instance simply recruits users to carry out a particular task while assuming the details of the others. To achieve this, it is necessary to examine two barriers. Study the relationship between the two activities, taking into account context sensing, as well as temporal and spatial adjustments, to start. The second concern is how to develop a task allocation approach that is efficient in collecting information for the source task and highly accurate in implying data for related activities.

As a practical and affordable alternative to large-scale sensing networks, researchers are now looking at the benefits of mobile sensor nodes. There are a few differences between the two sensing methods, though [34]. Beginning with the basics, MCS relies on portable devices such as cell phones whereas WSNs rely on a small sensor network. Due to the greater computing, memory, and energy capabilities of devices like mobile phones, this difference enables MCS to do local processing. Furthermore, because portable devices and smartphones use rechargeable batteries, MCS local processing is less power-limited than WSN local processing. The second difference is that, in contrast to WSNs, which often have hundreds, if not thousands of sensor nodes, MCS typically has a larger scale (hundreds of thousands or millions of devices, including smartphones, dispersed over a city or nation). Several thousand sensors are required to install sensors for traditional WSNs on a city-wide basis. It was demonstrated in [35] that 90,000 sensors and 1,000,000 relays are necessary to undertake citywide (approximately 900 km2) environmental monitoring to preserve full area coverage and communication connectivity. The third distinction is the presence of humans in MCS, which raises some challenges, such as concerns about private information, but also raises certain opportunities, such as those arising from utilizing human engagement in a way that makes the system smarter. Human interference raises the question of incentive systems to encourage users to participate in MCS operations. The fourth distinction is the dynamic character of MCS as a result of user mobility, power level variation, and changes in user behavior and engagement. Another difference between MCS and traditional WSNs is that sensors in MCS are mobile and move randomly and autonomously, whereas sensors in traditional WSNs are often stationary and positioned in predictable or random places.

In the MCS paradigm, there are two types of communication: formal and informal. 1) Access communication network 2) IP-based core network The telecommunications network's flexible communications network connects customers to service providers directly. It involves every piece of equipment linking the core network to the user terminal, with a diameter ranging from a few hundred meters to several kilometers. The core network uses a fiber-optic topology due to its high transmission rate. An essential component of a communications network, the core network offers a range of services to customers connected through the access network architecture. In this study, the core network consists of servers, computer equipment, and applications. The following points would provide a summary of the two media as well as a detailed breakdown of the inner and outside communication channels used in the aforementioned structure [36].

Depending on the arrival time of the participants, the selection methods are classified as offline or online. The decision of whether or not to pick participants is decided as participants arrive for online selection, and the participants are arriving continually. The choice is made in advance for offline selection based on historical or foreseeable facts, and all participants are ready.

Algorithms come in a variety of kinds, both offline and online. Graph-based algorithms, machine learning (ML)-based algorithms, and aggressive (greed-based) algorithms are the three categories into which algorithms are categorized offline. Online scenarios, however, entail that only tasks or participants come dynamically, whilst the other component is known in advance.

When many MCS system components are in constant communication with the MCS server, this is known as a multi-side online situation.

The matching issue is known as offline matching when all of the details of the tasks and participants are known in advance. The greedy selection technique is a popular strategy in numerous studies. Since it has been established that the majority of matching problems are Nanoparticles, the excessive selection technique is a useful tool for overcoming this difficulty. Choosing the best option at each stage until the objective function or constraint is met is the key component of the greedy selection technique. Machine learning algorithms are widely used in related research. Picking suitable places for the perceptron's sensing duties required the use of a greedy selection strategy in conjunction with a sorting algorithm. When there were no sensors available, Wang et al. employed the Naïve Bayes algorithm to evaluate the quality of the air [37]. The cross-validation results for the sensing zones are used to determine prior knowledge. Based on the anticipated outcomes, a participant selection algorithm was developed. The basic concept is to extract data from unsensed areas using several methods, then choose the one with the biggest difference as the sensing area. The Naïve Bayes algorithm is vital.

Graph theory is used in some matching methods. Initial participant locations are selected appropriately, after which potential routes are shown on a location graph. Maximizing the utility (universal coverage) of choosing PoIs is the goal of the matching algorithm. A dynamic programming approach was used to find the optimal path for approximation.

One-sided online matching and multi-sided online matching are the two forms of online matching difficulties. With one-sided online matching, only the tasks are entered dynamically while the participants are pre-announced, or just the participants are entered dynamically while the tasks are chosen. Most comparable activities involve users dynamically connecting to the MCS server and presuming task knowledge from the beginning. To allocate online jobs, a greedy method is utilized. First, the quality improvement of any task assignments to one or more participants is calculated. To optimize overall growth, the best tasks for each participant are chosen.

A choice on whether to select or not to pick should be made in real-time according to multi-side online matching, which suggests that more than one component of the MCS system is unclear beforehand. Particularly in the case of two-sided online matching, it is suggested that both the tasks and the participants are dynamic. Moreover, the automatic threshold greedy algorithm seeks to optimize the threshold by assessing the matching outcomes in each cycle.

The parallel approach can perform truth discovery efficiently on huge datasets, while the streaming algorithm can do truth discovery efficiently on both large data sets and data streams. They can thus aid in the successful and scalable discovery of truth in large-scale quantitative crowdsourcing applications [38]. However, algorithms must interact with servers at each step, which is wasteful and frequently leads to privacy issues. The study presents a Bayesian modeling method for concurrently learning the latent subjects of questions, quite well source dependability, and question responses (truths). Moreover, if the number of reported observations increases, the source integrity assumption may no longer apply, and it is more logical to have several origin dependabilities per origin. It mixes the prediction model and truth evaluation processes.

2.6 Truth-Discovery-Based Schemes:

Mobile Crowd-Sensing is divided into three parts: the server, the users, and the Cluster Head node. The Server, for example, is responsible for monitoring all users as well as preserving and retrieving the observational sensed data that participants input. The platform's observation sensory tasks are accepted by users, who then collect and evaluate the observation sensory data. The cluster manager is responsible for managing the cluster's nodes (users) and data processing (with a similar specific user). There are many mobile crowd-sensing schemes out there. The [38] architecture for crowd-sensing systems can protect not just users' sensory data but also individual reliability ratings produced from truth discovery methodologies. The suggested framework's central idea is to conduct weighted aggregation of users' encrypted data that used a parallel cryptographic technique. Furthermore, in crowdsensing applications, the sensory data supplied by specific individuals is typically unreliable owing to a variety of factors such as low sensor quality, a failure of sensor calibration, noise levels, partial perspectives of observations, and even deception. As a result, the potential of crowd sensing can only be unlocked by appropriately combining untrustworthy information from many participating users, who invariably contribute noisy, contradictory, and diverse data.

It is critical to capture the variance in information quality across various participating users when collecting crowd-sensing data. Some users consistently supply accurate and valuable data, whilst others may produce biased or even false data. When aggregating sensory input, an ideal solution should be able to incorporate the probability of a user supplying correct data in the form of user weight and make the aggregated results near the information supplied by trustworthy users. The difficulty here is that user dependability is frequently unknown a predetermined and must be deduced from acquired data. The concept of truth discovery was created to meet this obstacle [38-39] that is recent, several studies have been conducted to identify facts from untrustworthy user information. The basic premise held by truth discovery methodologies is that a specific user's data will be given more weight if it is closer to the pooled findings, and a specific user's data will be considered more in the aggregation operation if this user has a higher weight. Based on this idea, several truth discovery algorithms have been presented to determine user weight and aggregated outcomes together.

Although truth discovery algorithms have significantly improved aggregation accuracy, researchers fail to take into account a key practical challenge in the design of crowd-sensing systems, namely the protection of user privacy. The ultimate aggregate results in many crowd-sensing applications can be public and useful to the society or community, but the data from every individual user may comprise sensitive confidential info and so should be securely safeguarded. Gathering health data, like treatment results, for example, might lead to a better evaluation of the effects of new pharmaceuticals or medical equipment, but it may compromise the privacy of the persons involved. By aggregating participant reports, geotagging campaigns can give precise and fast localization of specified items (e.g., garbage, manhole, automated external defibrillator, etc.), but with the risk of exposing participants' sensitive location information. Even incredibly tough issues may be solved with crowd wisdom by aggregating the replies of a big population. Individual users' personal information, however, can be deduced from their responses. [40] is executed on two separate cloud platforms. This technique can secure both users' sensory data and reliable information while also achieving great efficiency and fault tolerance by relying on modular arithmetic features. In this technique, two clouds estimate the object facts collaboratively without revealing the users' private information.

Furthermore, to decrease user computing costs, sensitive information is disrupted by inserting random integers, and all sophisticated ciphertext-based processes are outsourced to cloud platforms. Furthermore, to decrease each user's computational and connection burden, appropriate for applications requiring simple sensory data protection. In this method, each user is just required to supply altered sensory data as well as encrypted random values to begin the truth-finding operation. In [41], the authors address a crucial issue in MCS systems: incentivizing user engagement. Authors create incentive structures based on reverse multiple auctions. The authors look at single-minded and multi-minded simultaneous auction models. For the former, authors offer truthful, individual rationality, and computationally efficient methods that approximate social utility maximization with a guaranteed approximation ratio. Authors devise an iterative downward mechanism for the latter that provides near-optimal societal utility while meeting individual rationality and computing efficiency.

Individual users are frequently charged a fee to participate in such crowd-sensing assignments. On the one hand, it depletes user resources like computational power, battery life, and so on. However, a significant number of sensing jobs necessitate the input of specific sorts of users' sensitive private information, resulting in information leakage for participating users. Users, for example, indicate the sorts of ailments authors suffer from by sharing images of their medical gadgets. Users frequently exchange information about their locations when authors submit air quality estimate samples. As a result, users will be hesitant to do sensing activities unless authors are rewarded in a way that compensates them for their participation costs. However, the majority of the current MCS systems are either voluntary or lack adequate incentive mechanisms.

In mobile crowd sensing, ESPPTD-based improved snipping private information truth discovery that avoids slicing [42]. The necessity for forwarding reduces computation and transmission while protecting data privacy. This method divides the sensing data into several components, hides sensitive information in distinct components, and sends each component to a different destination node to provide higher privacy protection. [43] provide an energy-efficient and efficient dynamic slice-based secure aggregation of data ASSDA approach. ASSDA limits the number of times a node's data slice may be delivered in a given time slot and slices all leaf nodes' data based on the number of receivers and the transmission distance to increase data slicing efficiency while decreasing node energy consumption, prolonging the life of the network, and safeguarding node actual information from eavesdropping.

Because the fundamental assumption that reliable data sources provide trustworthy information supports truth discovery, sensory data generated via credible users must be close to reality. Truth evaluation and weight updating are the two processes that currently distinguish known truth discovery algorithms. As previously said, there are many good crowdsensing techniques and some of them also have shortcomings, thus with that in mind, we created table 2.1 in which we included some of them, then described their core idea in a different column, an advantage in another column, and restriction in another column.

Table 2.1: Literature Comparison

Scheme	Basic Idea	Limitations	Advantages
[1] ITDELMC	Get user-sensing data. Match data Gives incentives to the user	No task allocating technique Ignored experienced user	Gives incentives to the user under budget constraints through quality function
[2] RPPTD	Discuss truth discovery under the security domain	No incentive mechanism	Data are collected and encrypted before it is sent from the user
[3] EPTDMCS	Users can drop out at any stage. Users can work offline	Can't resist a high attack from a cloud server. No incentive mechanism	Resolve the issue that the user must be online at all times during the truth discovery
[11] PEMCTD	Both of the solutions safeguard taking particular data and reliability levels throughout the truth-finding process.	the weight computation is captured by the distance function and comprises linear operations between a user's distance information and the sum of distances for all users	Security study demonstrates that the sensory data and user reliability level are properly safeguarded throughout the truth discovery operation.

[13] P2TA	Deploying edge nodes to maximize task approval rate while respecting participants' privacy	It may limit the CS-ability servers to receive users' location data directly.	authors use issue abstraction to answer the challenge of combined privacy protection and work allocation.
[25] CCS-TA	temporal and spatial correlation of data sensed in several sub-areas	The organizer may be unable to find a participant to carry out a task in the designated salient cell.	to actively choose a small number of sensory cells in each cycle while deducing missing values from the remaining cells
[27] AIMLPMS	SPECTRUM is a tool for determining the best form of incentive for any particular crowd-sensing activity	SPECTRUM misses the mark of concrete realization through empirical validation.	With SPECTRUM, smart cities would be able to efficiently incentivize users and encourage participation in vast crowdsensing applications.
[33] IWTDCC	calculate the truth for every activity based on worker reliability and precision	It can only be used for numerical crowdsourcing data or predefined possibilities. It is inadequate for crowdsourcing textual answers.	According to the Bayesian analysis, researchers compute the correlation for each pair of workers.
[35] PSTDLSQC	Identify ground truth from potentially messy, physically unclear, and inconsistent statements presented by multiple information sources.	In each step, algorithms must interact with servers, which is inefficient and readily produces privacy problems.	To capture relevant decisions from large-scale video content audience opinions, a parallel and streaming algorithm can be used.

The primary distinction between MCSC and conventional sensor networks is that the grassroots are involved in vast sensing. The following benefits are specific to grassroots involvement for MCSC: Due to the fact that MCSC makes use of already-existing infrastructure for sensing and transmission, its implementation costs are very inexpensive. In addition, users of mobile devices have inherent mobility that offers exceptional spatiotemporal coverage when compared to static sensor network deployments. Participatory sensing's key characteristic is that citizens are involved in the entire sensing process. Similar reasoning applies to MCSC, but it goes beyond participatory sensing and adds a number of new characteristics.

Data selection is frequently required to enhance data quality since human engagement in crowd sensing generates duplicate, subpar, or even fraudulent data for MCS systems. While MCS systems often employ diverse hardware, some of these components may have less computational power than others. Thus, two separate data processing techniques are produced by MCS: the centralized technique sends all collected data to a backend server for processing, whereas the self-supported technique equips the device with data processing capabilities.

2.7 Summary:

To ensure that our research could be understood properly, we added some of our materials to the second chapter of the literature review after first reviewing the papers based on our study. First and foremost, we wrote a chapter overview outlining what would be covered in the chapter as well as the structure of the chapter's literature review. Under the heading of mobile crowd sensing, what it is, and how it functions, the entire subject has been covered and explained. We had the idea to reward our employees for their work in mobile crowd-sensing because we thought that if we start giving them incentives, they would do a better job. This was a nice idea that was well received. Both monetary incentives and non-monetary incentive mechanisms are covered in this chapter. In an incentive system, the employee receives a reward or money, whereas, in a non-incentive system, the employee works voluntarily or in a volunteer capacity.

The fourth section, "Essential of Privacy Preservation in MCS," it is discussed how and what is needed to protect user privacy in MCS. The fifth category dealt with the issues that MCS presented. We arrived at some conclusions and discussed them in the paper after reading several articles on truth discovery. A table discussing truth-discovery strategies concludes.

CHAPTER 3

METHODOLOGY

3.1 Overview:

There are five sections in this chapter: the introduction, the literature review, the problem definition, the evaluation metrics, the proposed solution, and the final and fifth sections, which discuss the simulation framework. Based on the research we conducted in the literature review chapter; we would describe our research under the chapter's heading. Under the heading of problem identification, we would define our issue and let you know which article it is discussed. The evaluation metrics go on to describe the strategy we will use to solve the issue we have just outlined. Finally, we have explained the simulation that we used in the simulation framework after detailing the response that we offered in the suggested solution.

3.2 Literature Review:

After investigating a total of 43 publications linked to my research in the chapter of literature review, which included around 19 journal papers, 10 conference papers, some chapters of three books, three reports, and eight survey papers. The primary categories of the studies we investigated in the literature review chapter were truth discovery and weight updating. And we utilized Google and Google Scholar to find our material, with the search term "truth discovery in MCS." Many papers are obtained from the IEEE System Journal, some from MDPI, some from Elsevier, a few from Hindawi, several from Springer, and some from aclanthology.org are included. After reviewing all of these publications, we discovered that not all, but many, were pertinent to my research. All of my research is focused on fixing that problem, which I also mentioned in the next section. Finally, we compared the relevant research in the literature, and in the literature review, we addressed the limitations of these papers. A thorough examination of relevant schemes, including their advantages and disadvantages, has been conducted.

3.3 Problem Identification:

The initial phase in mobile crowd sensing is generally truth discovery, followed by weight updating. We collected several schemes and then recognized the advantages and limits of each scheme in the literature and described it, showing the limitations in different schemes, and then we chose a scheme [1] where the limitation is selected to provide a solution for it. All of the articles we examined on MCS, explained MCS extremely well, however after some thought, we recognized that there is one mistake that all of the papers make. If a user supplies data that is near to the estimated truth, probably, individuals who are completing such work for the first time will also give data that is close to the estimated truth, and those who have done many such tasks will also give data that is close to the estimated truth. But none of them discuss these techniques for an experienced user. In this case, there is no apparent distinction between the two users, who have done anything in the past and those who have not. The count of previously completed tasks is not taken into account in the schemes during the truth discovery computation [1], [3], [5]. So we discovered this to be an issue, and after much thought, we devised a solution, which we have included below under the heading of the proposed solution.

3.4 Proposed Solution:

In this section, we will explain the process we followed to arrive at our proposed solution. Our approach involved an in-depth review of approximately 43 relevant studies, as mentioned in the previous paragraph. We analyzed the findings of these studies and tried to devise a solution that would meet the expectations of both the applicant and the user or client. The applicant wanted their tasks completed accurately, while the users aimed to complete more tasks in less time. We encountered challenges in balancing these objectives, but after extensive research, we have identified a solution that we have outlined in Chapter 4. To provide a visual representation of our proposed solution, please refer to Figure 3.1 for additional details.

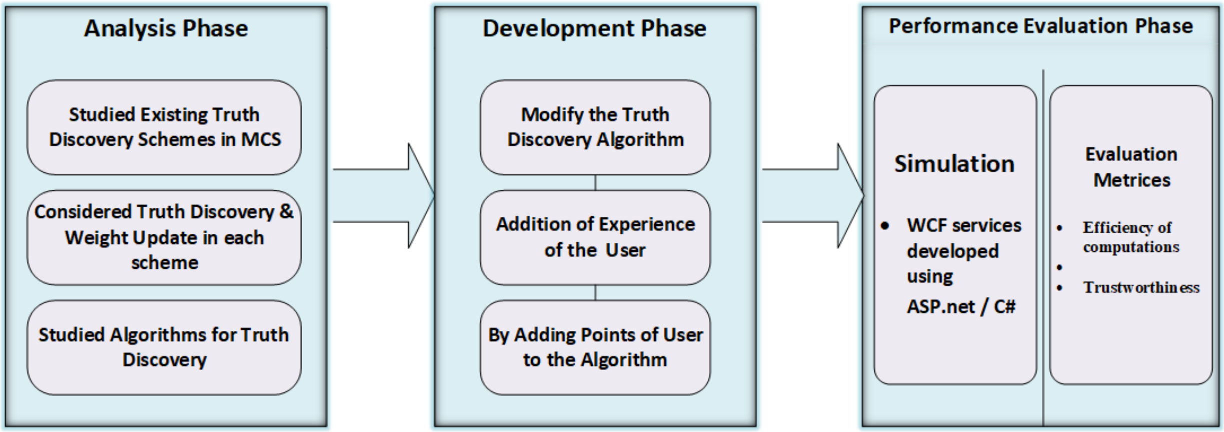

Figure 3.1: Operational Framework of TMWEMS

Some MCS user tasks are completed for enjoyment or as a social duty, while others are completed for financial gain. There is no problem if a user works as a social service, but there is a problem if it is for incentives. If the task is for an incentive, users strive to make it as comprehensive and correct as possible, whether it is true and complete or not, and they aim to provide a truth that is near to the estimated truth. So, we attempted to tackle this problem by making it our aim, and we came up with a solution in which we award users points for each activity completed, and this problem is solved.

3.5 Evaluation Metrics:

Traditionally, a variety of criteria are used to assess the efficacy and efficiency of truth discovery techniques. The effectiveness of a system may be assessed using memory costs and running time. When ground truth is provided, quality performance metrics such as recall, precision, accuracy (or error rate) for categorical data, mean of absolute error (MAE), and root of mean square error (RMSE) for continuous data are computed.

3.5.1 The Efficiency Of Computations:

If the truth and the winner set can be determined in polynomial time, an incentive mechanism is computationally more efficient.

3.5.2 Trustworthiness:

If reporting the real cost is a weakly dominant option for all users, an incentive mechanism is honest. In other words, regardless of what others provide, no user may boost its value by contributing a phony cost.

Because these are the most commonly used metrics, we will also compare performance using these metrics and with the base methods[1][3-4].

3.6 Simulation Framework:

We have used Windows Communication Foundation (WCF) for our simulation which is part of Asp.net/c# in which we have created a function through which we have run our truth discovery algorithm, and the function we have created for truth discovery is then called from the C# web application

3.7 Summary:

First, after conducting in-depth research on our subject and reporting it as a literature review. What percentage of the publications are journal articles, conference papers, and survey papers, respectively? After identifying the issue that is emerging in the pre-existing MCS system and condensing everything into a single concept known as problem identification. When the issue was identified, then a solution was needed, so put that need into words in the proposed solution. In the end, described our simulation framework after discussing where needed to make changes to the evaluation matrices. Then described our simulation setup and the function created and then called via the c# web application in the simulation framework.

CHAPTER 4

PROPOSED SOLUTION

4.1 Overview

This chapter discusses the proposed solution as well as the implications of the solution we are going to give. After the complete explanation of our scheme, we then discussed our system model, i.e. from where our base is and how it is started, and explained the system model through a diagram, after the system model, the flowchart has been discussed and In the flowchart, all the modules are explained through a diagram. After the flowchart, the algorithm has been explained through equations, and then the truth discovery algorithm which is our proposed scheme has been written, and finally, the summary of this chapter has also been written.

4.2 PROPOSED SOLUTION

In our proposed solution, we improve the incentive mechanism for the experienced user. As the previous mechanism doesn't consider the experience of a user. We would improve the truth discovery algorithm by adding the experience of the user i.e., how many tasks a user performed so far. If the experienced user provides sensing truth that is a little far from the estimated truth and a new user provides sensing truth that is near to the estimated truth, it may be by chance that a new user provides sensing truth that is near to the estimated truth. So, we would consider both the users' experience and the sensing truth.

Truth discovery is an approach used in mobile crowdsensing to reconcile discrepancies in sensory data collected from various sources and produce truthful data. For example, a medical research Centre may seek to undertake truth discovery over health data (that is, treatment results and daily health situations) obtained from various individuals equipped with wearables to gain an improved assessment of illness facts. The main concept behind truth discovery methodologies is to repeatedly estimate the dependability level of each data source in the form of weights, and then estimate the ground truths of target objects by weighted aggregation of individual sensory inputs, until some convergence condition is met. More specifically, a truth discovery technique typically begins with randomly generated ground facts and then runs three subroutines iteratively: (i) weight assessment, (ii) truth assessment, and (iii) convergence measurement.

In this study, we would, as in previous designs [44], use the representative algorithm CRH to actualize the aforementioned iterative technique of truth finding, which was presented by [45]. For the sake of clarity, we discuss the three subroutines using the scenario of truth discovery for a single object, although support for many objects may be easily added. Remember that in the first round of weight estimation, the predicted ground facts are initialized randomly [45].

In [1], it assigns tasks with the budget from the deep cloud to the edge cloud after this step the tasks and budget are given to the users from the edge cloud. After sensing the data, the user creates a bid on that data and gives the data to the edge cloud. In the discovered cloud the truth of sensing data is d and the edge cloud gives the estimated truth to the deep cloud to match the truth between tasks and the sensing data. The edge cloud then performs a reverse auction and gives payments to the user according to their estimated truth.

If one user has some experience with 30 tasks and he makes a bid with data and the bid is under the budget. Still, his truth is slightly changed from the estimated truth and another user who has no experience creates a bid with data the bid is above the budget but his truth is correct to the estimated truth. Then, that user would be considered a winner of the bid, which is unfair to the experienced user. We introduce a mechanism in which the user would get incentives according to their bid with data and truth discovery.

The DATE is the truth discovery mechanism in crowdsensing with copiers [46]. DATE, on the other hand, can only be used for numerical crowdsourcing data or choices from predetermined possibilities. DATE is useless for text-based crowdsourced answers. It's because multiple text answers may have identical meanings and hence be assigned the same weight. After all, DATE cannot recognize the semantics of text answers. In [46] DATE's textual data processing capacity is expanded. To support crowdsourced textual answers, researchers create and cluster content linear interpretations of sampled data.

When aggregating sensory input, an ideal solution should be able to incorporate the probability of a user supplying correct data in the form of user weight and make the aggregated results near the information supplied by reliable users. The difficulty in this situation, though, is that user dependability is typically unknown a priori and must be inferred from gathered data. The fundamental tenet of truth discovery methods is that a user's data will be given more weight if it is more closely aligned with the aggregated findings and vice versa. If a user is given more weight, their data will also be given more consideration throughout the aggregation process. The calculation of user weight and aggregated outcomes in a combined way has been proposed using many truth discovery techniques based on this idea.

4.3 System Model:

Over the past few years, significant progress has been made in mobile crowd sensing. Two popular architecture styles that have emerged in this field are server-centric, as seen in both [47] and [48]. In this approach, mobile devices are used to collect relevant information, which is then uploaded and stored on a server. The processing of this data is performed on a backend cloud server using sophisticated computing techniques. Unfortunately, this approach underutilizes the power of the crowd by not leveraging the processing power of mobile devices. According to the authors of [48], smartphones are used to gather information about the user's surroundings through mobile sensing, while maintaining user privacy. This information is retained, processed when necessary, and provided to third-party applications. This architecture primarily focuses on mobile devices. In our work, we have utilized a service provider server in conjunction with the mobile phones of the users performing the tasks and the data of the applicant who wants to perform the task. Figure 4.1 provides a system model that illustrates our approach.

The clients who wish to perform tasks first send their tasks to the crowdsensing server, which then forwards the tasks to the users who are willing to perform them or distributes them among them. After the users have completed the tasks, the task data is sent to the mini or small server. The mini server calculates the users' experience and adds experience points to the task weight, which is then submitted to the crowdsensing server. Finally, the data is returned to the client through the crowdsensing server. Our approach leverages the processing power of mobile devices and ensures the efficient completion of tasks.

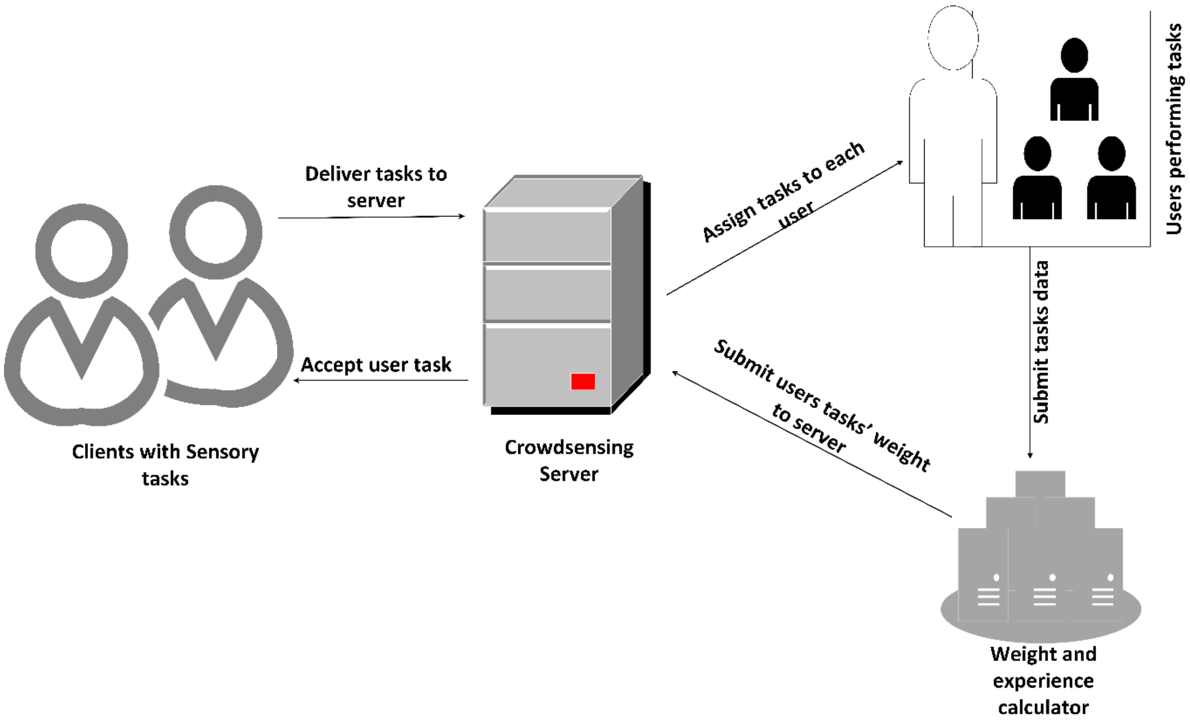

Figure 4.1: System Model

4.4 Flowchart:

The first stage starts, and we are told that our algorithm begins here. The second phase is tasks with a budget, in which we are instructed about the tasks and budget, which are the

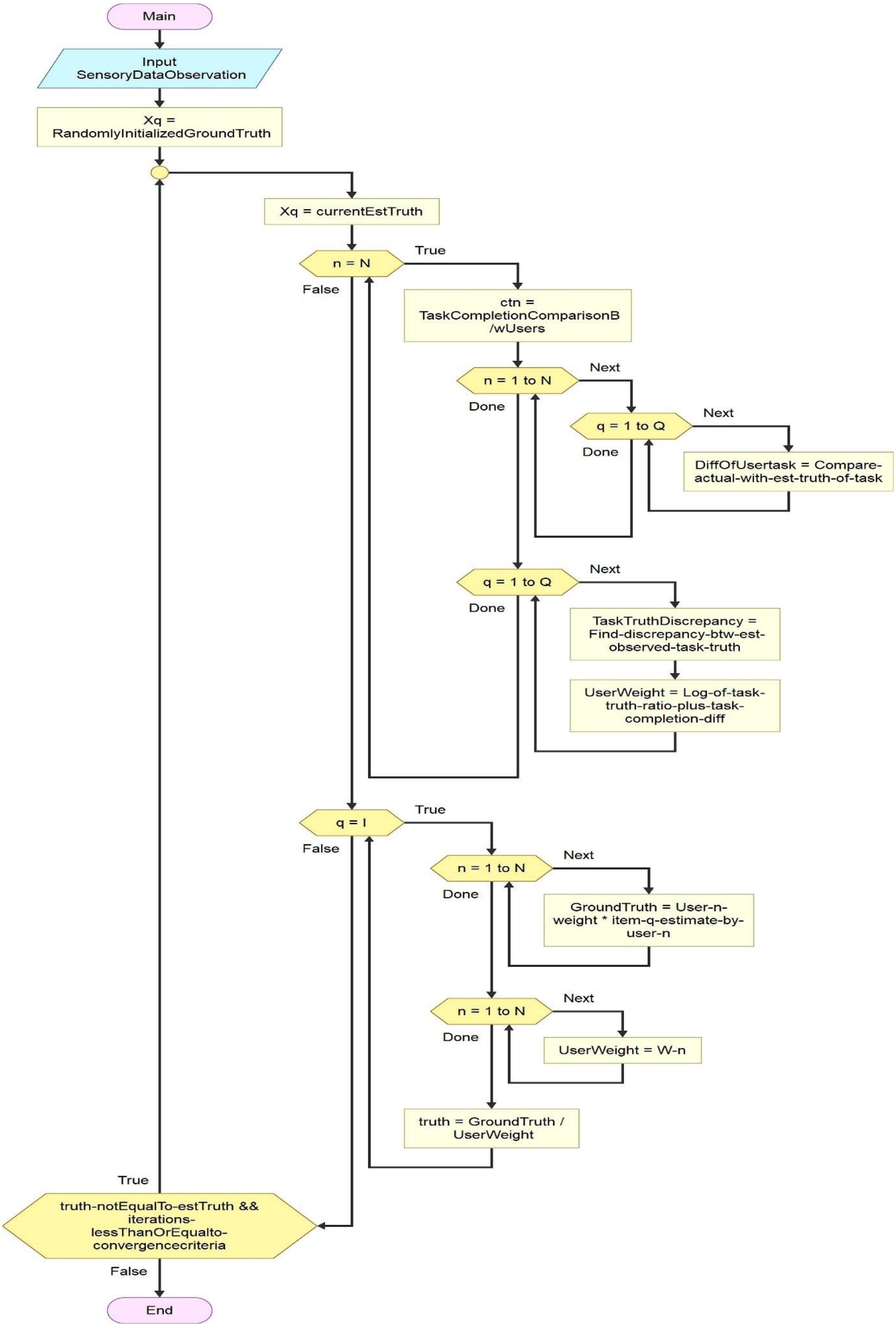

Figure 4.2: Flow chart for the proposed scheme

the user, the user completes these tasks and submits his bid. The user's truth is discovered when he submits his bid. Following the revelation of the truth, it is determined if the truth is near the estimated truth or not. Whether the truth is close to the estimated truth, the second condition is tested to see if the user has the experience, and if the estimated truth is not close, the user is given another assignment if he wishes.

If the user has no experience, he will have to wait until all of the jobs have been completed, and then a condition will be verified to ensure that no other user has offered any truth that is near to the estimated truth, and he will be rewarded. If provided, another task should be assigned to him. Our algorithm is finished after payment.

4.5 Proposed Scheme Algorithm

Let's say there are Q items in the given sensing task, denoted as $I = \{i_1, i_1, \dots, i_Q\}$, and these items will be scrutinized by N participating workers as $U = \{u_1, u_2, \dots, u_N\}$. We refer to these workers' weights (i.e., reliability) as $W = \{w_1, w_2, \dots, w_N\}$. Let x_q^n indicate the worker's sensory data of the worker u_n for item i_q. For every item $i_q \epsilon I$, in this framework, there is a ground truth that none of the participants are aware of. Our objective is to determine the estimated values $\{x_q\}_{q=1}^{Q}$ of the ground truths for all the items while also looking for workers' experience and adding points for experience to its task weight.

4.5.1 Truth Discovery:

Typically, the truth discovery methods follow an iterative two-step process. Weight update and truth update.

To estimate the truth, we employed the CRH truth discovery technique [39,44]. In CRH, there are two critical steps: weight update and truth update.

4.5.1.1 Weight Update:

This stage will include estimating each worker's weight based on the discrepancy between their sensory data as well as the estimated truths. Most often, a worker's weight u_n is determined using the formula used by [1]:

$$w_n = f\left(\sum_{q=1}^{Q} d\left(x_q^n, x_q\right)\right) \tag{1}$$

Where f is a monotonically decreasing function, and $d\left(x_q^n, x_q\right)$ is the distance function, which calculates the variance between the estimated truths and the workers' sensory data. We have made a small adjustment to Equation 1 by including user experience in the context of the proposed scheme. The following equation counts user experience.

$$ct_n = \sum_{n=1}^{N} d\ (e_n, e_{n-1}) \tag{2}$$

Where ct_n is a variable that stores the users' experience, and $d\ (e_n, e_{n-1})$ is the difference function which calculates the difference between the experience of user n and user $n-1$. After calculating the user experience in Equation 2, it will be added to Equation 1 and It's get Equation 3.

$$w_n = f\left(\sum_{q=1}^{Q} d\left(x_q^n, x_q\right)\right) + ct_n \tag{3}$$

Equation 4 represents the distance function.

$$d\left(x_q^n, x_q\right) = \frac{\left(x_q^n - x_q\right)^2}{std_q} \tag{4}$$

Where std_q indicates the standard deviation of observational sensory data depending on the object q.

$$w_n = \log\left(\frac{\sum_{n'=1}^{N}\sum_{q=1}^{Q} d\left(x_q^{n'}, x_q\right)}{\left(\sum_{q=1}^{Q} d\left(x_q^{n}, x_q\right)\right) + ct_n}\right) \tag{5}$$

We adopted the logarithmic function as $f(.)$ for any user $u \in U_n$ because of its strong operational performance.

4.5.1.2 Truth Update:

In this phase, we assumed that each user's weight is constant. Following the calculation of worker weights, the ground truth for each item i_q may be approximated as given in Equation 6:

$$x_q = \frac{\sum_{n=1}^{N} w_n\, x_q^n}{\sum_{n=1}^{N} w_n} \tag{6}$$

When sensory data is continuous, this number represents the weighted average of the workers' observations of the item i_q. However, when the data is categorical, x_q is a vector where each element denotes the probability that a certain candidate outcome or response is the correct one. The outcome or response that has the highest value in the vector x_q will be the estimated truth of an object i_q. The Equation 5 and Equation 6 will be applied repeatedly in truth discovery algorithms until a convergence requirement is achieved. The threshold of the change in the estimated truths across two successive iterations, or a predetermined number of iterations, might serve as the convergence criteria. The processing steps are shown in the Algorithm. Equations are cited from [44].

Table 4.1 summarizes the notations used in a specific context, including Q and i_Q for tasks and specific items, N and W for workers and their weights, x_q^n and x_q for sensory data and estimated truth, and e_n for worker experience. These notations simplify complex concepts and variables and aid communication and understanding.

Table 4.1: Notations summary

Notation	Description
Q, i_Q	items in given tasks, item i
N, W	number of workers, weight of workers
x_q^n, x_q	sensory data of worker n for item q, estimated truth for item q
e_n	experience of worker n

ALGORITHM 1: TRUTH DISCOVERY ALGORITHM

Input: observation of sensory data for N participants $\{x_q^n\}_{n,q=1}^{N,Q}$

Output: estimated truth $\boldsymbol{x_q}$, weights $\boldsymbol{w_n}$

Randomly initialize the ground truth $\boldsymbol{x_q}$

$q \leftarrow 0;$

$\boldsymbol{do}$

$\boldsymbol{x'_q} \leftarrow \boldsymbol{x_q}$

$\boldsymbol{while}(\boldsymbol{n} \rightarrow \boldsymbol{N});$

$ct_n = d\,(e_n, e_{n-1})$

$(\boldsymbol{for\ n' = 1};\ \boldsymbol{n'} \rightarrow \boldsymbol{N};\ \boldsymbol{n'}++)$

$(\boldsymbol{for\ q = 1};\ \boldsymbol{q} \rightarrow \boldsymbol{Q};\ \boldsymbol{q}++)$

$\boldsymbol{Diff\ of\ user's\ task} = truth\ of\ q\ provided\ by\ n - estimated\ truth\ of\ q$

$(\boldsymbol{for\ q = 1};\ \boldsymbol{q} \rightarrow \boldsymbol{Q};\ \boldsymbol{q}++)$

$\boldsymbol{Diff\ of\ est\ and\ obs\ task\ truth} = d(x_q^n, x_q)$

$weight\ of\ user\ n = \log\left(\frac{Diff\ of\ user's\ task}{Diff\ of\ est\ and\ obs\ task} + ct_n\right)$

$n++:$

$\boldsymbol{while}(\boldsymbol{q} \rightarrow \boldsymbol{I});$

$(\boldsymbol{for\ n = 1};\ \boldsymbol{n} \rightarrow \boldsymbol{N};\ \boldsymbol{n}++)$

Ground truth = weight of user n × est truth of item q by user n)

$(\boldsymbol{for\ n = 1};\ \boldsymbol{n} \rightarrow \boldsymbol{N};\ \boldsymbol{n}++)$

Weight of user $= w_n$;

$truth = \dfrac{Ground\ truth}{weight\ of\ user}$;

$q{+}{=}\ 1$;

$\boldsymbol{while}\ truth\ \neq est\ truth\ \boldsymbol{and}\ q \leq convergence\ criterian$;

4.6 Summary

This is a summary of our fourth chapter, which we have also designated as our proposed scheme. We stated what we did in this chapter and how we discussed it in the title of the overview at the beginning of this chapter. In the second heading, we have fully detailed our proposed system, including its name and function, and we have compared it to current schemes. Two graphs have been created to demonstrate how our approach differs from others. The first graph compares running time and the number of users, while the second graph compares (MAPE) and the number of iterations. The system model was then discussed, along with its many components that pertain to our research. A graphic was used to clarify each component's relationship to the other components in the system model. The flowchart follows. As you are aware, the algorithm is explained in the flowchart, beginning with where it begins and progressing through the stages to the decision; if the choice is correct, which step will follow; if it is incorrect, which step will follow? Finally, we reviewed the algorithm of our proposed scheme, which we illustrated using equations.

CHAPTER 5

RESULTS AND DISCUSSION

5.1 Overview

In this chapter, we go over all of the results and keep altering the data for different results so that the varied results can be read in full detail and the overall solution can be clearly understood. To do this, we've graphed the schemes we compared and put them alongside the graphed lines of our own and earlier designs. Additionally, in this chapter, we have discussed the probable effects of our plan. Then, in the conclusion, a summary is written.

5.2 Performance Matrices

The table 5.1 lists various parameters used for a simulation study, including the number of tasks and participants, the type of sensory data, the convergence criteria, and the programming language and communication protocol used. The notation column provides a shorthand reference to the mathematical symbols used in the study

Metrics were utilized, as they were in the earlier schemes, and several other schemes also used the same metrics. Along with these, we employed metrics such as running time vs. number of users and mean absolute percentage error (MAPE) vs. Number of iterations, among others.

In the evaluation of various schemes, metrics played a crucial role, just as they did in earlier schemes. In addition to commonly used metrics, such as precision, recall, and accuracy, we also incorporated other metrics, such as running time versus the number of users and mean absolute percentage error (MAPE) versus the number of iterations. These metrics helped us to evaluate the performance of different recommendation systems comprehensively. For example, running time versus the number of users helped us to understand the scalability of the systems, while MAPE versus the number of iterations gave us insights into how quickly the algorithms converged. Overall, utilizing various metrics allowed us to gain a more nuanced understanding of the strengths and weaknesses of the recommendation systems under consideration

Table 5.1: Simulation parameter

Parameter	Value	Notation
Number of tasks	1000	Q
Number of participants	200	N
Type of sensory data	Surroundings data for a map	x_q^n, x_q
Convergence criteria	100	-
Performance metrics	Accuracy, precision, recall, F1 score	-
Sampling method	Random sampling	-
Experiment design	Cross-sectional study	-
Data analysis method	Statistical analysis	-
Programming language	C#	-
Communication protocol	WCF	-
Data transfer format	XML or JSON	-
Security	Transport security or message security	-
Fault handling	Exception handling or fault contracts	-
Performance optimization	Message compression or message caching	-
Development environment	Visual Studio	-
Testing framework	NUnit or MSTest	-
Data visualization tool	Microsoft Excel or Tableau	-

5.3 MAPE vs Number of iterations:

In Figure 5.1, we compare the Mean absolute percentage error (MAPE) and the number of iterations between our scheme and the first scheme IMTEC. As can be seen, the bigger the number of iterations, the lower our mean absolute percentage error (MAPE). The number of iterations in the first scheme is decreasing, as shown by the graph, however, the number of iterations in our proposed scheme is more. As a result, we collect various forms of data from

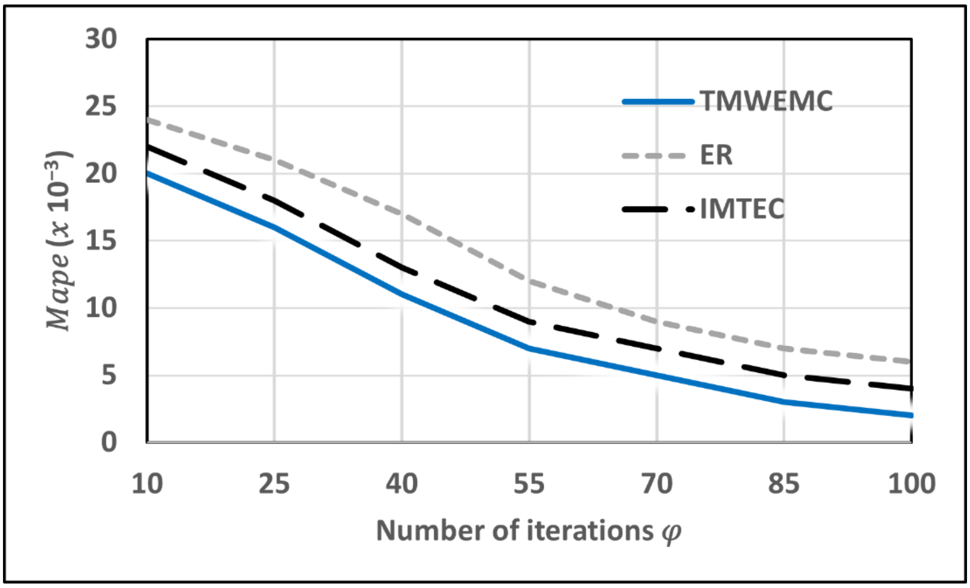

Figure 5.1: MAPE vs Number of iterations

consumers to increase the number of iterations in our system. (MAPE) falls as the iteration number grows, and the system becomes effective.

For equal reliability, we utilized the ER in this graph. ER believes that each edge cloud has the same reliability. The reliability of the assumed truth of the edge cloud is thought to be dependent on the number of users within it. If you look at the graph again, you will notice that the graph line of our scheme is at the bottom, which implies that the (MAPE) of our scheme is all coming smaller than, and if (MAPE) is the least, then our scheme is the best.

5.4 Running time vs Number of users

As shown in Figure 5.2, we made a comparison between the running time and the number of users as specified in the preceding schemes, thus we compared our scheme as well as existing schemes. so that we can determine how dependable our scheme is and how much load it can manage. As you can see from the graph line of the first scheme, IMTEC, the first point is 0.24, the second point is 0.35, and so on for the third, fourth, and fifth points. By inspecting the graph, you can see for yourself that the values of the first scheme, IMTEC, are shown on the graph, as well as the graph line, and you can also see the graph line of our suggested scheme. As you can see in Figure 5.3, we evaluated the effectiveness of our plan by

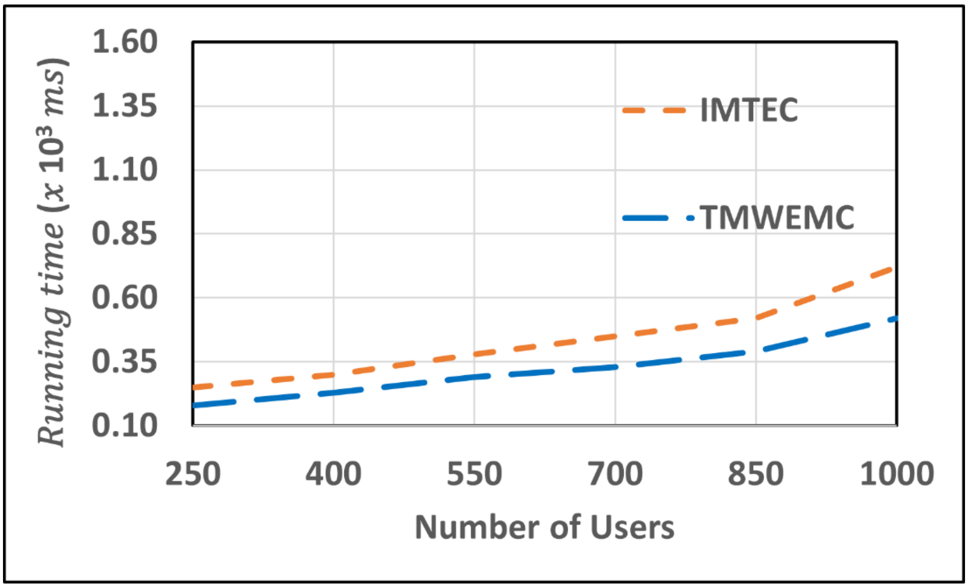

Figure 5.2: Running time vs number of users

comparing it to other plans already in place, and we then presented the results to you in the form of a graph. [48] have utilized the MSensing scheme in their paper where authors discuss platform-centric and user-centric, where one pays the user from the platform directly while the other pays the user according to the job. It is a good scheme, but researchers also neglected the experience as a result of which the platform's effectiveness is somewhat declining.

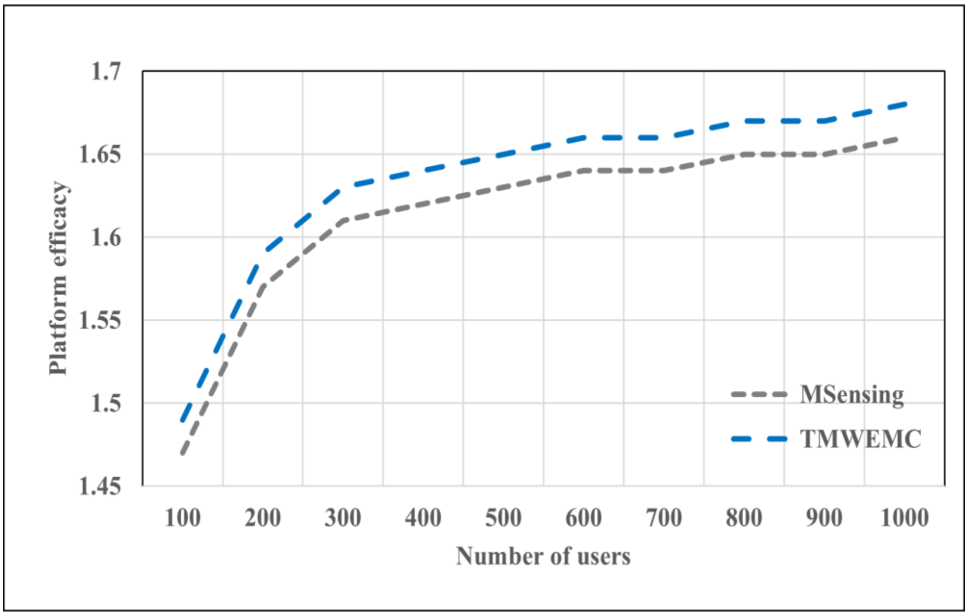

Figure 5.3: Platform efficacy vs Number of users

This means some users might get disappointed and may not participate in an upcoming project crowdsensing. So, keeping in mind that issue we introduce such a mechanism to look after such users who have some experience. That would also encourage the new user get to participate in such projects. If we ignore this issue, we may lose users because researchers would not work to get experience, researchers would work to get great incentives by just providing the data with a bid and researchers may not look at their previous performance. Our research would boost the users to extend their experience to get great incentives. And the user would look to their previous work and would try hard to improve from the last experience. And the new users would try to perform more tasks to gain more experience.

In the previous solution, the number of users is very high i.e., the high number of users slows down the process of crowdsensing. Dealing with numerous users makes the algorithm work slowly and the number of iterations increases. The number of users increases the Mean Absolute Percentage Error (MAPE) decreases but while decreasing the MAPE the number of iterations increases. When the number of iterations increases algorithm's efficiency decreases.

5.5 Platform efficacy vs Cost Range:

Similarly, when the number of jobs in this scheme grows, so does the running time in comparison to our system as you can see in Figure 5.4. The previous research on the same topic

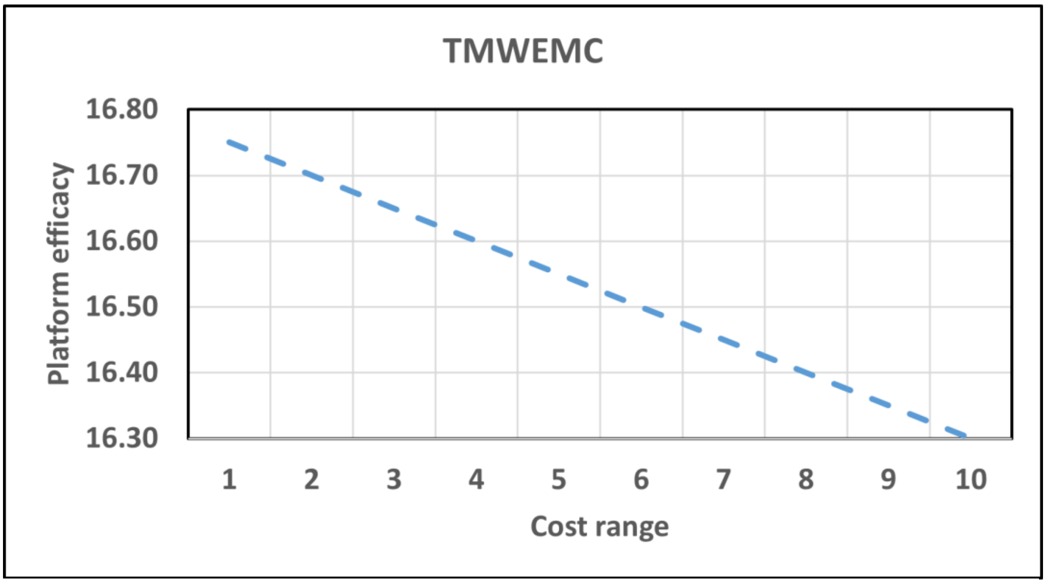

Figure 5.4: Platform efficacy vs cost range

doesn't include any such mechanism which would consider both the users' experience and the estimated truth.

5.6 Running time vs Number of tasks:

Comparisons are made between the Equal Reliability (ER), Square Root Distance (SRD), and Goad Mechanism for fact Detection in Edge-assisted Vast Portable Crowdsensing techniques (IMTEC). ER assumes that all edge clouds are equally reliable. The distance function is implemented by SRD. The MAPE for all three techniques decreased dramatically as the number of users rose. The accuracy of IMTEC was the greatest of the three algorithms studied. Because IMTEC takes into account the number of users in the edge clouds, edge clouds with more users play a bigger role in truth discovery. Since IMTEC took into account the standard deviation of each task's provided data and tasks with huge data fluctuations had less of an influence on changing users' weights, we can observe that IMTEC used a more appropriate distance function than SRD. Figure 5.5 compares our method with the MSensing approach, and you can see from the graph that our strategy takes less time as the number of tasks rises than the MSensing scheme does. If so, many tasks that were in the old system are completed, therefore our scheme is doing far better than before. As you can see from the graph, M Sensing measured the task's completion time in seconds, and we followed suit by measuring the task's completion time in seconds to compare the two.

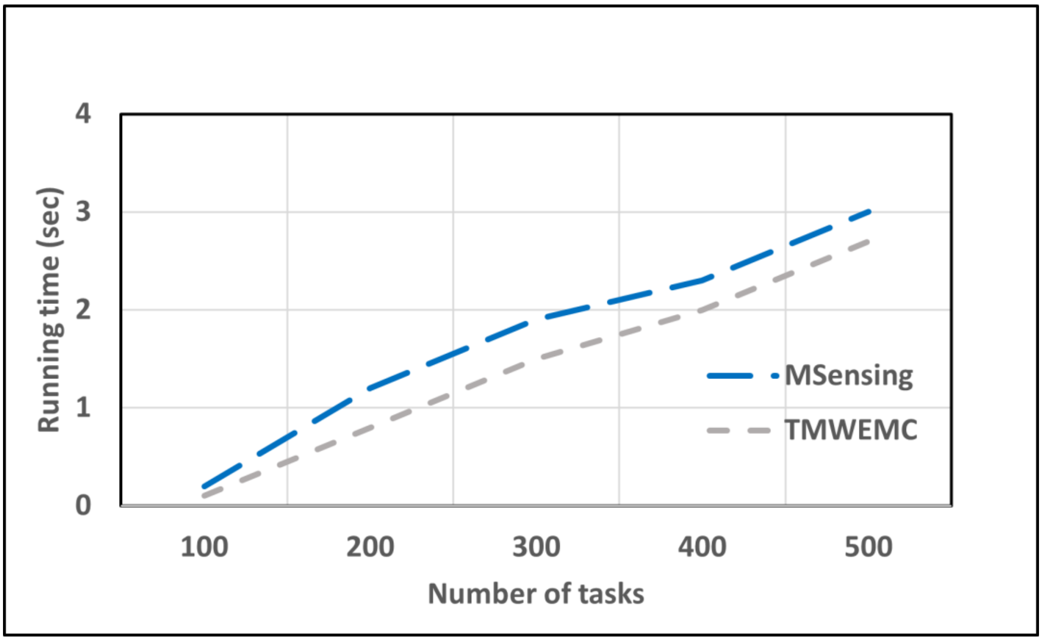

Figure 5.5: Running time (sec) vs Number of tasks

If we closely examine the currently implemented schemes, particularly the MSensing scheme, we see that as the number of tasks rises, the time (which we have measured in seconds) increases. The time will also grow as the number of tasks rises; therefore, we have taken this into account and addressed all of these problems in our proposed scheme. When workers' GPS data is combined, for instance, it might compromise their private location data while yet providing crucial insights for smart mobility applications. Expense, or privacy cost, results from the possible privacy leaking. Consequently, it is vital to develop a strong incentive system to make up for the numerous expenses that employees incur in order to encourage worker involvement.

An MCS application must balance two competing goals: data reliability and user privacy protection. The reliability of the data may be impacted by robust privacy protections. The procedures for maintaining privacy are, however, countered by safeguarding the validity of the data. Therefore, it is vital to make a trade-off between preserving user privacy and guaranteeing the reliability of the data. Future work should concentrate on figuring out how to use certain straightforward cryptographic operations to transmit a reputation value (which is a proxy for evaluating data trustworthiness) between anonymous contributions without the participation of any reliable third party and with little communication cost.

5.7 Summary

After comparing our strategy to other plans in this chapter and went over the outcomes. Then, discuss the findings of our research, any potential outcomes, and what makes it special and effective in comparison to other approaches. Graphs were used to illustrate its effects and the effectiveness of our plan. This graph was made after a simulation, and since the simulation relied on data, we collected it and plotted it. We then talked about these graphs in terms of graph metrics. After comparing our scheme with the other schemes in the graph matrices, and then discussed the graphs created for this comparison one by one, explaining why our scheme is superior to the others and why it should be chosen.

CHAPTER 6

CONCLUSION AND FUTURE WORK

6.1 Overview:

In this chapter, we have explained our research conclusion and future work. We have concluded our research, and it has been stated that this research has been advantageous to us. It has also been stated under two heads in the future work where and where work may be done in crowdsensing, then we wrote a summary of our research and described how it would assist others. Finally, it has been stated regarding future work that if someone wants to work in crowdsensing, which module may be used and there is still work to be done in crowdsourcing.

6.2 Contribution to Research Work

In this thesis, we proposed a new truth discovery scheme for mobile workers in edge-assisted mobile crowdsensing. We found that existing schemes such as IMTEC and MSensing had limitations in terms of scalability and execution time when the number of users or tasks increased.

To address these limitations, we added an extra point to our scheme by incorporating worker experience into the calculation of worker weights. We then compared our proposed scheme to IMTEC and MSensing in terms of execution time.

When the number of tasks was 500, MSensing took 3 seconds while our proposed scheme took only 2.65 seconds, representing a 12.5% improvement in execution time. This demonstrates the effectiveness and efficiency of our proposed scheme.

Overall, our contribution to research work lies in proposing an extension to existing truth discovery that takes into account worker experience and addresses scalability and execution time issues in existing schemes. By improving execution time and reducing processing overhead, our proposed scheme can facilitate more efficient and effective mobile crowdsensing applications.

6.3 Future Research Directions:

In addition to these areas, another potential area for future work is exploring the effectiveness of different truth discovery algorithms for different types of data. While the algorithms evaluated in this thesis demonstrated promising results for the types of data used, it's possible that different algorithms may be more effective for different types of data or applications. For instance, some algorithms may be more effective for structured data, while others may be better suited for unstructured data. Evaluating the effectiveness of different algorithms for different types of data can help identify the most suitable approach for a particular application or context.

Another potential area for future work is exploring the impact of user bias on truth discovery algorithms. Users may have biases or preconceptions that could impact the accuracy of the results, particularly in cases where subjective information is being provided. Therefore, it would be valuable to investigate methods for identifying and mitigating user bias in truth discovery algorithms. For instance, algorithms could be designed to account for user bias or to weight the contributions of different users based on their level of bias.

Overall, these potential areas for future work demonstrate the ongoing need for continued research and development in the field of truth discovery. By addressing these challenges and exploring new approaches, progress can be made towards more accurate and reliable results, particularly in sensory-based applications.

6.4 Limitaions:

One limitation of the current study is the focus on initial user authentication, without extensively exploring the potential for authenticated users to later become intruders or attackers to the system. While regular security checks and user monitoring were implemented, the potential for authenticated users to later become threats remains a limitation. This highlights the need for ongoing monitoring and evaluation of user behavior to identify potential threats and address them in a timely manner.

To address this limitation, future studies could explore the use of more sophisticated user monitoring and authentication mechanisms. For instance, algorithms could be designed to continually monitor user behavior and adjust their trustworthiness scores over time. Additionally, user behavior could be monitored to identify potential warning signs that a user may be turning malicious, such as a sudden increase in the number of false statements they make. By implementing more robust and adaptable authentication and monitoring mechanisms, the risk of authenticated users turning into intruders or attackers can be reduced, improving the reliability and accuracy of the system.

REFERENCES

[1] J. Xu, S. Yang, W. Lu, L. Xu, and D. Yang, "Incentivizing for truth discovery in edge-assisted large-scale mobile crowdsensing," *Sensors (Switzerland)*, vol. 20, no. 3, Feb. 2020, doi: 10.3390/s20030805.

[2] Paul. Bennett and Association for Computing Machinery. Special Interest Group on Knowledge Discovery & Data Mining., *Proceedings of the ACM SIGKDD Workshop on Human Computation : 2009, Paris, France, June 28-28, 2009*. 2009.

[3] F. Laws, C. Scheible, H. Schütze, and S. Schütze, "Active Learning with Amazon Mechanical Turk," in *Proceedings of the 2011 Conference on Empirical Methods in Natural Language Processing*, Jul. 2011, pp. 1546–1556. [Online]. Available: http://crowdflower.com/

[4] H. D. Ma, "Internet of things: Objectives and scientific challenges," *J Comput Sci Technol*, vol. 26, no. 6, pp. 919–924, Nov. 2011, doi: 10.1007/s11390-011-1189-5.

[5] D Yang, G Xue, X Fang, J Tang, and ACM SIGMOBILE., "Proceedings of the 18th annual international conference on Mobile computing and networking," in *Crowdsourcing to Smartphones: Incentive Mechanism Design for Mobile Phone Sensing*, Aug. 2012, Mobicom 12., pp. 173–184. doi: 10.1145/2348543.2348567.

[6] IEEE Computer Society., Institute of Electrical and Electronics Engineers., IEEE Computer Society. Technical Committee on Computer Communications., and University of Texas at Austin., *2013 IEEE 14th International Symposium and Workshops on a World of Wireless, Mobile and Multimedia Networks (WoWMoM) : digital proceedings : Madrid, Spain, June 4-7, 2013*. IEEE, 2013.

[7] D. Zhao, H. Ma, and S. Tang, "COUPON: Cooperatively building sensing maps in mobile opportunistic networks," in *Proceedings - IEEE 10th International Conference*

on Mobile Ad-Hoc and Sensor Systems, MASS 2013, 2013, pp. 295–303. doi: 10.1109/MASS.2013.68.

[8] X. Hu, T. H. S. Chu, H. C. B. Chan, and V. C. M. Leung, "Vita: A crowdsensing-oriented mobile cyber-physical system," *IEEE Trans Emerg Top Comput*, vol. 1, no. 1, pp. 148–165, Jun. 2013, doi: 10.1109/TETC.2013.2273359.

[9] Huadong Ma, Dong Zhao, and Peiyan Yuan, "Opportunities in mobile crowd sensing," *IEEE Communications Magazine*, vol. 52, no. 8, pp. 29–35, Aug. 2014, doi: 10.1109/MCOM.2014.6871666.

[10] M. J. Morón, R. Luque, and E. Casilari, "On the capability of smartphones to perform as communication gateways in medical wireless personal area networks," *Sensors (Switzerland)*, vol. 14, no. 1, pp. 575–594, Jan. 2014, doi: 10.3390/s140100575.

[11] Q. Li *et al.*, "A Confidence-Aware Approach for Truth Discovery on Long-Tail Data," *Proc. VLDB Endow.*, vol. 8, no. 4, pp. 425–436, Dec. 2014, doi: 10.14778/2735496.2735505.

[12] Y. Li *et al.*, "Conflicts to Harmony: A Framework for Resolving Conflicts in Heterogeneous Data by Truth Discovery," *IEEE Trans Knowl Data Eng*, vol. 28, no. 8, pp. 1986–1999, Aug. 2016, doi: 10.1109/TKDE.2016.2559481.

[13] J. Guillén, J. Miranda, J. Berrocal, J. García-Alonso, and J. M. Murillo, "IEEE People as a Service: A Mobile-centric Model for Providing Collective Sociological Profiles," *IEEE Softw*, vol. 31, no. 2, pp. 48–53, Apr. 2014, doi: 10.1109/MS.2013.140.

[14] L. G. Jaimes, I. J. Vergara-Laurens, and A. Raij, "A Survey of Incentive Techniques for Mobile Crowd Sensing," *IEEE Internet of Things Journal*, vol. 2, no. 5. Institute of Electrical and Electronics Engineers Inc., pp. 370–380, Oct. 01, 2015. doi: 10.1109/JIOT.2015.2409151.

[15] L. Wang *et al.*, "CCS-TA: Quality-guaranteed online task allocation in compressive crowdsensing," in *UbiComp 2015 - Proceedings of the 2015 ACM International Joint*

Conference on Pervasive and Ubiquitous Computing, Sep. 2015, pp. 683–694. doi: 10.1145/2750858.2807513.

[16] J. Gao, Q. Li, B. Zhao, W. Fan, and J. Han, "Truth Discovery and Crowdsourcing Aggregation: A Unified Perspective," *Proc. VLDB Endow.*, vol. 8, no. 12, pp. 2048–2049, Aug. 2015, doi: 10.14778/2824032.2824136.

[17] F. Ma *et al.*, "FaitCrowd: Fine grained truth discovery for crowdsourced data aggregation," in *Proceedings of the ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, Aug. 2015, vol. 2015-August, pp. 745–754. doi: 10.1145/2783258.2783314.

[18] C. Miao *et al.*, "Cloud-enabled privacy-preserving truth discovery in crowd sensing systems," in *SenSys 2015 - Proceedings of the 13th ACM Conference on Embedded Networked Sensor Systems*, Nov. 2015, pp. 183–196. doi: 10.1145/2809695.2809719.

[19] H. Jin, L. Su, D. Chen, K. Nahrstedt, and J. Xu, "Quality of information aware incentive mechanisms for mobile crowd sensing systems," in *Proceedings of the International Symposium on Mobile Ad Hoc Networking and Computing (MobiHoc)*, Jun. 2015, vol. 2015-June, pp. 167–176. doi: 10.1145/2746285.2746310.

[20] Cheng *et al.*, "Deco: False data detection and correction framework for participatory sensing," in *2015 IEEE 23rd International Symposium on Quality of Service (IWQoS)*, IEEE, 2015, pp. 213–218. doi: 10.1109/IWQoS.2015.7404736.

[21] M. Louta, K. Mpanti, G. Karetsos, and T. Lagkas, "Mobile Crowd Sensing Architectural Frameworks: A Comprehensive Survey," in *2016 7th International Conference on Information, Intelligence, Systems & Applications (IISA)*, Jul. 2016, pp. 1–7. doi: 10.1109/IISA.2016.7785385.

[22] R. I. Ogie, "Adopting incentive mechanisms for large-scale participation in mobile crowdsensing: from literature review to a conceptual framework," *Human-centric Computing and Information Sciences*, vol. 6, no. 1, Dec. 2016, doi: 10.1186/s13673-016-0080-3.

[23] R. W. Ouyang, L. M. Kaplan, A. Toniolo, M. Srivastava, and T. J. Norman, "Parallel and Streaming Truth Discovery in Large-Scale Quantitative Crowdsourcing," *IEEE Transactions on Parallel and Distributed Systems*, vol. 27, no. 10, pp. 2984–2997, Oct. 2016, doi: 10.1109/TPDS.2016.2515092.

[24] Q. Li, Y. Li, J. Gao, B. Zhao, W. Fan, and J. Han, "Resolving conflicts in heterogeneous data by truth discovery and source reliability estimation," in *Proceedings of the ACM SIGMOD International Conference on Management of Data*, 2014, pp. 1187–1198. doi: 10.1145/2588555.2610509.

[25] G. Cardone, A. Corradi, L. Foschini, and R. Ianniello, "ParticipAct: A Large-Scale Crowdsensing Platform," *IEEE Trans Emerg Top Comput*, vol. 4, no. 1, pp. 21–32, Jan. 2016, doi: 10.1109/TETC.2015.2433835.

[26] Miao, C and Su, Lu and Jiang, Wen and Li, Y and Tian, and Miaomiao, *A lightweight privacy-preserving truth discovery framework for mobile crowd sensing systems*. Atlanta, GA, USA: IEEE, 2017. doi: 10.1109/INFOCOM.2017.8057114.

[27] L. Cheng *et al.*, "Compressive sensing based data quality improvement for crowd-sensing applications," *Journal of Network and Computer Applications*, vol. 77, pp. 123–134, Jan. 2017, doi: 10.1016/j.jnca.2016.10.004.

[28] G. Xu, H. Li, and R. Lu, "Practical and Privacy-Aware Truth Discovery in Mobile Crowd Sensing Systems," in *Proceedings of the 2018 ACM SIGSAC Conference on Computer and Communications Security*, Oct. 2018, pp. 2312–2314. doi: 10.1145/3243734.3278529.

[29] Y. Zheng, H. Duan, and C. Wang, "Learning the Truth Privately and Confidently: Encrypted Confidence-Aware Truth Discovery in Mobile Crowdsensing," *IEEE Transactions on Information Forensics and Security*, vol. 13, no. 10, pp. 2475–2489, Oct. 2018, doi: 10.1109/TIFS.2018.2819134.

[30] M. Marjanovic, A. Antonic, and I. P. Zarko, "Edge computing architecture for mobile crowdsensing," *IEEE Access*, vol. 6, pp. 10662–10674, Jan. 2018, doi: 10.1109/ACCESS.2018.2799707.

[31] Y. Y. Chen, P. Lv, D. K. Guo, T. Q. Zhou, and M. Xu, "A Survey on Task and Participant Matching in Mobile Crowd Sensing," *J Comput Sci Technol*, vol. 33, no. 4, pp. 768–791, Jul. 2018, doi: 10.1007/s11390-018-1855-y.

[32] P. Hua, X. Liu, J. Yu, N. Dang, and X. Zhang, "Energy-efficient adaptive slice-based secure data aggregation scheme in WSN," in *Procedia Computer Science*, 2018, vol. 129, pp. 188–193. doi: 10.1016/j.procs.2018.03.033.

[33] B. Wang, L. Kong, L. He, F. Wu, J. Yu, and G. Chen, "I(TS, CS): Detecting faulty location data in mobile crowdsensing," in *Proceedings - International Conference on Distributed Computing Systems*, Jul. 2018, vol. 2018-July, pp. 808–817. doi: 10.1109/ICDCS.2018.00083.

[34] D. E. Boubiche, M. Imran, A. Maqsood, and M. Shoaib, "Mobile crowd sensing – Taxonomy, applications, challenges, and solutions," *Comput Human Behav*, vol. 101, pp. 352–370, Dec. 2019, doi: 10.1016/j.chb.2018.10.028.

[35] G. Xu, H. Li, S. Liu, M. Wen, and R. Lu, "Efficient and Privacy-Preserving Truth Discovery in Mobile Crowd Sensing Systems," *IEEE Trans Veh Technol*, vol. 68, no. 4, pp. 3854–3865, Apr. 2019, doi: 10.1109/TVT.2019.2895834.

[36] H. Shen, G. Bai, Y. Hu, and T. Wang, "P2TA: Privacy-preserving task allocation for edge computing enhanced mobile crowdsensing," *Journal of Systems Architecture*, vol. 97, pp. 130–141, Aug. 2019, doi: 10.1016/j.sysarc.2019.01.005.

[37] A. Capponi, C. Fiandrino, B. Kantarci, L. Foschini, D. Kliazovich, and P. Bouvry, "A Survey on Mobile Crowdsensing Systems: Challenges, Solutions, and Opportunities," *IEEE Communications Surveys and Tutorials*, vol. 21, no. 3, pp. 2419–2465, 2019, doi: 10.1109/COMST.2019.2914030.

[38] Y. Wu, Z. Yan, K. K. R. Choo, and L. T. Yang, "IEEE Access Special Section Editorial: Internet-of-Things Big Data Trust Management," *IEEE Access*, vol. 7. Institute of Electrical and Electronics Engineers Inc., pp. 65223–65227, 2019. doi: 10.1109/ACCESS.2019.2915489.

[39] W. Guo, W. Zhu, Z. Yu, J. Wang, and B. Guo, “A Survey of Task Allocation: Contrastive Perspectives from Wireless Sensor Networks and Mobile Crowdsensing,” *IEEE Access*, vol. 7, pp. 78406–78420, 2019, doi: 10.1109/ACCESS.2019.2896226.

[40] K. AbualSaud *et al.*, “A Survey on Mobile Crowd-Sensing and Its Applications in the IoT Era,” *IEEE Access*, vol. 7. Institute of Electrical and Electronics Engineers Inc., pp. 3855–3881, 2019. doi: 10.1109/ACCESS.2018.2885918.

[41] H. Vahdat-Nejad, E. Asani, Z. Mahmoodian, and M. H. Mohseni, “Context-aware computing for mobile crowd sensing: A survey,” *Future Generation Computer Systems*, vol. 99, pp. 321–332, Oct. 2019, doi: 10.1016/j.future.2019.04.052.

[42] L. Jiang, X. Niu, J. Xu, D. Yang, and L. Xu, “Incentivizing the workers for truth discovery in crowdsourcing with copiers,” in *Proceedings - International Conference on Distributed Computing Systems*, Jul. 2019, vol. 2019-July, pp. 1286–1295. doi: 10.1109/ICDCS.2019.00129.

[43] C. Zhang, L. Zhu, C. Xu, K. Sharif, and X. Liu, “PPTDS: A privacy-preserving truth discovery scheme in crowd sensing systems,” *Inf Sci (N Y)*, vol. 484, pp. 183–196, May 2019, doi: 10.1016/j.ins.2019.01.068.

[44] Y. Qiao, J. Wu, H. Cheng, Z. Huang, Q. He, and C. Wang, “Truthful Mechanism Design for Multiregion Mobile Crowdsensing,” *Wirel Commun Mob Comput*, vol. 2020, 2020, doi: 10.1155/2020/8834983.

[45] R. Sharma and L. Hourany, “Cloud computing-based Elliptic Curve Augmented Encryption framework for Vehicular Ad-Hoc Networks,” in *2020 5th International Conference on Innovative Technologies in Intelligent Systems and Industrial Applications (CITISIA)*, Nov. 2020, pp. 1–11. doi: 10.1109/CITISIA50690.2020.9371811.

[46] Y. Zheng, H. Duan, X. Yuan, and C. Wang, “Privacy-Aware and Efficient Mobile Crowdsensing with Truth Discovery,” *IEEE Trans Dependable Secure Comput*, vol. 17, no. 1, pp. 121–133, Jan. 2020, doi: 10.1109/TDSC.2017.2753245.

[47] Y. Wang, Z. Yan, W. Feng, and S. Liu, “Privacy protection in mobile crowd sensing: a survey,” *World Wide Web*, vol. 23, no. 1, pp. 421–452, Jan. 2020, doi: 10.1007/s11280-019-00745-2.

[48] J. Chen, Y. Liu, Y. Xiang, and K. Sood, “RPPTD: Robust Privacy-Preserving Truth Discovery Scheme,” *IEEE Syst J*, pp. 1–8, Aug. 2021, doi: 10.1109/jsyst.2021.3099103.

[49] J. Xue, Q. Hu, Y. An, and L. Wang, “Joint task offloading and resource allocation in vehicle-assisted multi-access edge computing,” *Comput Commun*, vol. 177, pp. 77–85, Sep. 2021, doi: 10.1016/j.comcom.2021.06.014.

[50] C. Lv, T. Wang, C. Wang, F. Chen, and C. Zhao, “ESPPTD:An efficient slicing-based privacy-preserving truth discovery in mobile crowd sensing,” *Knowl Based Syst*, vol. 229, Oct. 2021, doi: 10.1016/j.knosys.2021.107349.

[51] L. Jiang, X. Niu, J. Xu, D. Yang, and L. Xu, “Incentive Mechanism Design for Truth Discovery in Crowdsourcing with Copiers,” *IEEE Trans Serv Comput*, pp. 1–3, Apr. 2021, doi: 10.1109/TSC.2021.3075741.